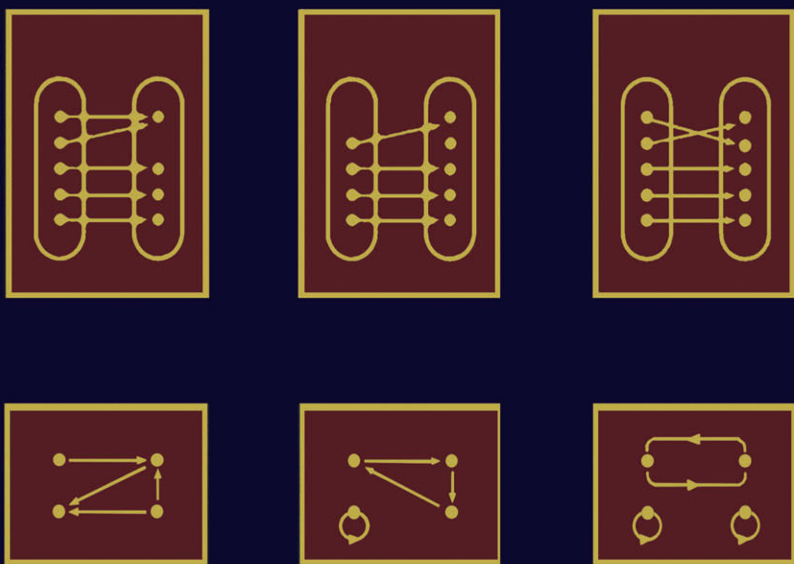
Handbook of Analysis and Its Foundations



Eric Schechter

Handbook of Analysis and Its Foundations

This Page Intentionally Left Blank

Handbook of Analysis and Its Foundations

Eric Schechter
Vanderbilt University

ACADEMIC PRESS
San Diego London Boston New York Sydney Tokyo Toronto

This book is printed on acid-free paper. ∞

ACADEMIC PRESS, INC.
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA
1300 Boylston Street, Chestnut Hill, MA 02167, USA
http://www.apnet.com

ACADEMIC PRESS LIMITED
24–28 Oval Road, London NW1 7DX, UK
http://www.hbuk.co.uk/ap/

Library of Congress Cataloging-in-Publication Data

Schechter, Eric, 1950–
Handbook of analysis and its foundations / Eric Schechter.
p. cm.
Includes bibliographical references and index.
ISBN 0-12-622760-8 (alk. paper)
1. Mathematical analysis. I. Title.
QA300.S339 1997
515—dc20 96-32226
CIP

Printed in the United States of America
96 97 98 99 00 IP 9 8 7 6 5 4 3 2 1

In dealing with mathematical problems, specialization plays, as I believe, a still more important part than generalization. Perhaps in most cases where we seek in vain the answer to a question, the cause of the failure lies in the fact that problems simpler and easier than the one in hand have been either not at all or incompletely solved. All depends, then, on finding out these easier problems, and on solving them by means of devices as perfect as possible and of concepts capable of generalization. — *David Hilbert*

Logic sometimes makes monsters. During half a century we have seen the rise of a crowd of bizarre functions which seem to try to resemble as little as possible the honest functions which serve some purpose. No longer continuity, or perhaps continuity but no derivatives, etc. Nay, more: from the logical point of view, it is these strange functions which are the most general. Those which one meets without seeking, no longer appear except as a particular case. — *Henri Poincaré*

Mathematics belongs to man, not to God. We are not interested in properties of the positive integers that have no descriptive meaning for finite man. When a man proves a positive integer to exist, he should show how to find it. If God has mathematics of his own that needs to be done, let him do it himself. — *Errett Bishop*

He considered, perhaps in his moments of less lucidity, that it is possible to achieve happiness on earth when it is not very hot, and this idea made him a little confused. He liked to wander through metaphysical obstacle courses. That was what he was doing when he used to sit in the bedroom every morning with the door ajar, his eyes closed and his muscles tensed. However, he himself did not realize that he had become so subtle in his thinking that for at least three years in his meditative moments he was no longer thinking about anything. — *Gabriel García-Márquez* (novelist)

This Page Intentionally Left Blank

Contents

This Page Intentionally Left Blank

Preface

About the Choice of Topics

Handbook of Analysis and its Foundations — hereafter abbreviated *HAF* — is a self-study guide, intended for advanced undergraduates or beginning graduate students in mathematics. It will also be useful as a reference tool for more advanced mathematicians. *HAF* surveys analysis and related topics, with particular attention to existence proofs.

HAF progresses from elementary notions — sets, functions, products of sets — through intermediate topics — uniform completions, Tychonov's Theorem — all the way to a few advanced results — the Eberlein-Smulian-Grothendieck Theorem, the Crandall-Liggett Theorem, and others. The book is self-contained and thus is well suited for self-directed study. It will help to compensate for the differences between students who, coming into a single graduate class from different undergraduate schools, have different backgrounds. I believe that the reading of part or all of this book would be a good project for the summer vacation before one begins graduate school in mathematics. At least, this is the book I wish *I* had had before I began *my* graduate studies.

HAF introduces and shows the connections between many topics that are customarily taught separately in greater depth:

> set theory, metric spaces, abstract algebra, formal logic, general topology, real analysis, and linear and nonlinear functional analysis, plus a small amount of Baire category theory, Mac Lane-Eilenberg category theory, nonstandard analysis, and differential equations.

Included in these customary topics are the usual nonconstructive proofs of existence of pathological objects. Unlike most analysis books, however, *HAF* also includes some chapters on set theory and logic, to explain why many of those classical pathological objects are presented without examples.

HAF contains the most fundamental parts of an entire shelf of conventional textbooks. In his "automathography," Halmos [1985] said that one good way to learn a lot of mathematics is by reading the first chapters of many books. I have tried to improve upon that collection of first chapters by eliminating the overlap between separate books, adhering to consistent notation, and inserting frequent cross-referencing between the different topics. *HAF*'s integrated approach shows connections between topics and thus partially counteracts the fragmentation into specialized little bits that has become commonplace in mathematics in recent decades. *HAF*'s integrated approach also supports the development of interdisciplinary topics, such as the "intangibles" discussed later in this preface.

The content is biased toward the interests of analysts. For instance, our treatment of algebra devotes much attention to convexity but little attention to finite or noncommutative groups; our treatment of general topology emphasizes distances and meager sets but omits manifolds and homology. *HAF* will not transform the reader into a researcher in algebra, topology, or logic, but it will provide analysts with useful tools from those fields.

HAF includes a few "hard analysis" results: Clarkson's Inequalities, the Kobayashi-Rasmussen Inequalities, maximal inequalities for martingales and for Lebesgue measure, etc. However, the book leans more toward "soft analysis" — i.e., existence theorems and other qualitative results. Preference is given to theorems that have short or elegant or intuitive proofs and that mesh well with the main themes of the book. A few long proofs — e.g., Brouwer's Theorem, James's Theorem — are included when they are sufficiently important for the themes of the book.

As much as possible, I have tried to make this book current. Most mathematical papers published each year are on advanced and specialized material, not appropriate for an introductory work. Only occasionally does a paper strengthen, simplify, or clarify some basic, classical ideas. I have combed the literature for these insightful papers as well as I could, but some of them are not well known; that is evident from their infrequent mentions in the *Science Citation Index.* Following are a few of *HAF*'s unusual features:

- A thorough introduction to filters in Chapters 5 and 6, and nets in Chapter 7. Those tools are used extensively in later chapters. Included are ideas of Aarnes and Andenæs [1972] on the interchangeability of subnets and superfilters, making available the advantages of both theories of convergence. Also included, in 15.10, is Gherman's [1980] characterization of topological convergences, which simplifies slightly the classic characterization of Kelley [1955/1975].

- an introduction to symmetric and preregular spaces, filling the conceptual gaps that are left in most introductions to T_0, T_1, T_2, and T_3 spaces — see the table in 16.1.

- a unified treatment of topological spaces, uniform spaces, topological Abelian groups, topological vector spaces, locally convex spaces, Fréchet spaces, Banach spaces, and Banach lattices, explaining these spaces in terms of increasingly specialized kinds of "distances" — see the table in 26.1.

- converses to Banach's Contraction Fixed Point Theorem, due to Bessaga [1959] and Meyers [1967], in Chapter 19. These converses show that, although Banach's theorem is quite easy to prove, a longer proof cannot yield stronger results.

- the Brouwer Fixed Point Theorem, proved via van Maaren's geometry-free version of Sperner's Lemma. This approach is particularly intuitive and elementary in that it involves neither Jacobians nor triangulations. It decomposes the proof of Brouwer's Theorem into a purely combinatorial argument (in 3.28) and a compactness argument (in 27.19).

- introductions to both the Lebesgue and Henstock integrals and a proof of their equivalence in Chapter 24. (More precisely, a Banach-space-valued function is Lebesgue integrable if and only if it is almost separably valued and absolutely Henstock integrable.)

- pathological examples due to Nedoma, Kottman, Gordon, Dieudonné, and others, which illustrate very vividly some of the differences between $\mathbb{R}^n$ and infinite-dimensional Banach spaces.
- an introduction to set theory, including the most interesting equivalents of the Axiom of Choice, Dependent Choice, the Ultrafilter Principle, and the Hahn-Banach Theorem. (For lists of equivalents of these principles, see the index.)
- an introduction to formal logic following the substitution rules of Rasiowa and Sikorski [1963], which are simpler and — in this author's opinion — more natural than the substitution rules used in most logic textbooks. This is discussed in 14.20.
- a discussion of model theory and consistency results, including a summary of some results of Solovay, Pincus, Shelah, et al. Those results can be used to prove the nonconstructibility of many classical pathological objects of analysis; see especially the discussions in 14.76 and 14.77.
- Neumann's [1985] nonlinear Closed Graph Theorem.
- the automatic continuity theorems of Garnir [1974] and Wright [1977]. These results are similar to Neumann's, but instead of assuming a closed graph, they replace conventional set theory with ZF + DC + BP. Their result explains in part why a Banach space in applied math has a "usual norm;" see 14.77.

In compiling this book I have acted primarily as a reporter, not an inventor or discoverer. Nearly all the theorems and proofs in *HAF* can be found in earlier books or in research journal articles — but in many cases those books or articles are hard to find or hard to read. This book's goal is to enhance classical results by modernizing the exposition, arranging separate topics into a unified whole, and occasionally incorporating some recent developments.

I have tried to give credit where it is due, but that is sometimes difficult or impossible. Historical inaccuracies tend to propagate through the literature. I have tried to weed out the inaccuracies by reading widely, but I'm sure I have not caught them all. Moreover, I have not always distinguished between primary and secondary sources. In many cases I have cited a textbook or other secondary source, to give credit for an exposition that I have modified in the present work.

Existence, Examples, and Intangibles

Most existence proofs use either compactness, completeness, or the Axiom of Choice; those topics receive extra attention in this book. (In fact, *Choice, Completeness, Compactness* was the title of an earlier, prepublication version of this book; papers that mention that title are actually citing this book.) Although those three approaches to existence are usually quite different, they are not entirely unrelated — AC has many equivalent forms, some of which are concerned with compactness or completeness (see 17.16 and 19.13).

The term "foundations" has two meanings; both are intended in the title of this book:

(i) In nonmathematical, everyday English, "foundations" refers to any basic or elementary or prerequisite material. For instance, this book contains much elementary set theory, algebra, and topology. Those subjects are not part of analysis, but are prerequisites for some parts of analysis.

(ii) "Foundations" also has a more specialized and technical meaning. It refers to more advanced topics in set theory (such as the Axiom of Choice) and to formal logic. Many mathematicians consider these topics to be the basis for all of mathematics.

Conventional analysis books include only a page or so concerning (ii); this book contains much more. We are led to (ii) when we look for examples of pathological objects.

Students and researchers need examples; it is a basic precept of pedagogy that every abstract idea should be accompanied by one or more concrete examples. Therefore, when I began writing this book (originally a conventional analysis book), I resolved to give examples of everything. However, as I searched through the literature, I was unable to find explicit examples of several important pathological objects, which I now call **intangibles**:

- finitely additive probabilities that are not countably additive,
- elements of $(\ell_\infty)^* \setminus \ell_1$ (a customary corollary of the Hahn-Banach Theorem),
- universal nets that are not eventually constant,
- free ultrafilters (used very freely in nonstandard analysis!),
- well orderings for $\mathbb{R}$,
- inequivalent complete norms on a vector space,

etc. In analysis books it has been customary to prove the existence of these and other pathological objects without constructing any explicit examples, without explaining the omission of examples, and without even mentioning that anything has been omitted. Typically, the student does not consciously notice the omission, but is left with a vague uneasiness about these unillustrated objects that are so difficult to visualize.

I could not understand the dearth of examples until I accidentally ventured beyond the traditional confines of analysis. I was surprised to learn that the examples of these mysterious objects are omitted from the literature because they *must* be omitted: Although the objects exist, it can also be proved that explicit constructions do *not* exist. That may sound paradoxical, but it merely reflects a peculiarity in our language: The customary requirements for an "explicit construction" are more stringent than the customary requirements for an "existence proof." In an existence proof we are permitted to postulate arbitrary choices, but in an explicit construction we are expected to make choices in an algorithmic fashion. (To make this observation more precise requires some definitions, which are given in 14.76 and 14.77.)

Though existence without examples has puzzled some analysts, the relevant concepts have been a part of logic for many years. The nonconstructive nature of the Axiom of Choice was controversial when set theory was born about a century ago, but our understanding and acceptance of it has gradually grown. An account of its history is given by Moore [1982]. It is now easy to observe that nonconstructive techniques are used in many of the classical existence proofs for pathological objects of analysis. It can also be shown, though less easily, that many of those existence theorems cannot be proved by other, constructive techniques. Thus, the pathological objects in question are *inherently* unconstructible.

The paradox of existence without examples has become a part of the logicians' folklore, which is not easily accessible to nonlogicians. Most modern books and papers on logic are written in a specialized, technical language that is unfamiliar and nonintuitive to outsiders: Symbols are used where other mathematicians are accustomed to seeing words, and distinctions are made which other mathematicians are accustomed to blurring — e.g., the distinction between first-order and higher-order languages. Moreover, those books and papers of logic generally do not focus on the intangibles of analysis.

On the other hand, analysis books and papers invoke nonconstructive principles like magical incantations, without much accompanying explanation and — in some cases — without much understanding. One recent analysis book asserts that analysts would gain little from questioning the Axiom of Choice. I disagree. The present work was motivated in part by my feeling that students deserve a more "honest" explanation of some of the non-examples of analysis — especially of some of the consequences of the Hahn-Banach Theorem. When we cannot construct an explicit example, we should say so. The student who cannot visualize some object should be reassured that no one else can visualize it either. Because examples are so important in the learning process, the lack of examples should be discussed at least briefly when that lack is first encountered; it should not be postponed until some more advanced course or ignored altogether.

Though most of *HAF* relies only on conventional reasoning — i.e., the kind of set theory and logic that most mathematicians use without noticing they are using it — we come to a better understanding of the idiosyncrasies of conventional reasoning by contrasting it with unconventional systems, such as ZF + DC + BP or Bishop's constructivism. *HAF* explains the relevant foundational concepts in brief, informal, intuitive terms that should be easily understood by analysts and other nonlogicians.

To better understand the role played by the Axiom of Choice, we shall keep track of its uses and the uses of certain weakened forms of AC, especially

> the Principle of Dependent Choices (DC), which is constructive and is equivalent to several principles about complete metric spaces;
>
> the Ultrafilter Principle (UF), which is nonconstructive and is equivalent to the Completeness and Compactness Principles of logic, as well as dozens of other important principles involving topological compactness; and
>
> the Hahn-Banach Theorem (HB), also nonconstructive, which has many important equivalent forms in functional analysis.

Most analysts are not accustomed to viewing HB as a weakened form of AC, but that viewpoint makes the Hahn-Banach Theorem's nonconstructive nature much easier to understand.

This book's sketch of logic omits many proofs and even some definitions. It is intended not to make the reader into a logician, but only to show analysts the relevance of some parts of logic. The introduction to foundations for analysts is *HAF*'s most unusual feature, but it is not an overriding feature — it takes up only a small portion of the book and can be skipped over by mathematicians who have picked up this book for its treatment of nonfoundational topics such as nets, F-spaces, or integration.

Abstract versus Concrete

I have attempted to present each set of ideas at a natural level of generality and abstraction — i.e., a level that conveys the ideas in a simple form and permits several examples and applications. Of course, the level of generality of any part of the book is partly dictated by the needs of later parts of the book.

Usually, I lean toward more abstract and general approaches when they are available. By omitting unnecessary, irrelevant, or distracting hypotheses, we trim a concept down to reveal its essential parts. In many cases, omitting unnecessary hypotheses does not lengthen a proof, and it may make the proof easier to understand because the reader's attention is then focused on the few possible lines of reasoning that still remain available. For instance, every metric space can be embedded isometrically in a Banach space (see 22.14), but the "more concrete" setting of Banach spaces does not improve our understanding of metric space results such as the Contraction Fixed Point Theorem in 19.39.

Here is another example of my preference for abstraction: Some textbooks build Hausdorffness into their definition of "uniform space" or "topological vector space" or "locally convex space" because most spaces used in applications are in fact Hausdorff. This may shorten the statements of several theorems by a word or two, but it does not shorten the proofs of those theorems. Moreover, it may confuse beginners by entangling concepts that are not inherently related: The basic ideas of Hausdorff spaces are independent from the other basic ideas of uniform spaces, topological spaces, and locally convex spaces; neither set of ideas actually requires the other. In *HAF*, Hausdorffness is a separate property; it is not built into our definitions of those other spaces. Our not-necessarily-Hausdorff approach has several benefits, of which the greatest probably is this:

> The weak topology of an infinite-dimensional Banach space is an important nonmetrizable Hausdorff topology that is best explained as the supremum of a collection of pseudometrizable, non-Hausdorff topologies.

(If the reader is accustomed to working only in Hausdorff spaces, *HAF*'s not-necessarily-Hausdorff approach may take a little getting used to, but only a little. Mostly, one replaces "metric" with "pseudometric" or with the neutral notion of "distance;" one replaces "the limit" with "a limit" or with the neutral notion of "converges to.")

However, a more general approach to a topic is not necessarily a simpler approach. Every idea in mathematics can be made more general and more abstract by making the hypotheses weaker and more complicated and by introducing more definitions, but I have tried to avoid the weakly upper hemisemidemicontinuous quasipseudospaces of baroque mathematics. It is unavoidable that the beginning graduate student of mathematics must wade through a large collection of new definitions, but that collection should not be made larger than necessary. Thus we sometimes accept slightly stronger hypotheses for a theorem in order to avoid introducing more definitions. Of course, ultimately the difference between important distinctions and excessive hair-splitting is a matter of an individual mathematician's own personal taste.

Converses to main implications are included in *HAF* whenever this can be managed conveniently, as well as in a few inconvenient cases that I deemed sufficiently important. Lists of *dissimilar but equivalent* definitions are collected into one long definintion-and-theorem, even though that one theorem may have a painfully long proof. The single portmanteau theorem is convenient for reference, and moreover it clearly displays the importance of a concept. For instance, the notion of "ultrabarrelled spaces" seemed too advanced and specialized for this book until I saw the long list of dissimilar but equivalent definitions that now appears in 27.26. To prevent confusion, I have called the student's attention to contrasts between *similar but inequivalent concepts*, either by juxtaposing them (as in the case of barrels and ultrabarrels) or by including cross-referencing remarks (as in the case of Bishop's constructivism and Gödel's constructivism).

Although the content is chosen for analysts, the writing style has been influenced by algebraists. Whenever possible, I have made degenerate objects such as the empty set into a special case of a rule, rather than an exception to the rule. For instance, in this book and in algebra books, $\{S : S \subseteq X\}$ is an "improper filter" on X, though it is not a filter at all according to the definition used by many books on general topology.

ORDER OF TOPICS

I have followed a Bourbaki-like order of topics, first introducing simple fundamentals and later building upon them to develop more specialized ideas. The topics are ordered to suit pedagogy rather than to emphasize applications. For instance, convexity is commonly introduced in functional analysis courses in the setting of Banach spaces or topological vector spaces, but I have found it expedient to introduce convexity as a purely algebraic notion, and then add topological considerations much later in the book. Most topological vector spaces used in applications are locally convex, but *HAF* first studies topological vector spaces without the additional assumption of local convexity.

Topics covered within a single chapter are closely related to each other. However, in many cases the end of a chapter covers more advanced and specialized material that can be postponed; it will not be needed until much later in the book, if at all. Most of Part C (on topological and uniform spaces) can be read without Part B (logic and algebra). However, most readers should skim through Chapters 5, 6, and 7. Those chapters introduce filters and nets — tools that are used more extensively in this book than in most analysis books.

I have felt justified in violating logical sequencing in one important instance. The real number system is, in some sense, *the* foundation of analysis, so it must be used in examples quite early in the book. Examples given in early chapters assume an informal understanding of the real numbers, such as might be acquired in calculus and other early undergraduate courses. A more precise definition of the reals is neither needed nor attainable until Chapter 10. Much conceptual machinery must be built before we can understand and prove a statement such as this one:

> There exists a Dedekind complete, chain ordered field, called the *real numbers.* It is unique up to isomorphism if we use the conventional reasoning methods of analysts. (It is not unique if we restrict our reasoning methods to first-order languages and permit the use of nonstandard models.)

The existence and uniqueness of the complete ordered field justify the usual definition of $\mathbb{R}$. I am surprised that these algebraic results are not proved (or even mentioned!) in many introductory textbooks on analysis.

A traditional course on measure and integration would correspond roughly to part of Chapter 11, all of Chapter 21, and parts of Chapters 22–25 and 29. Integration theory is commonly introduced separately from functional analysis, but I have mixed the two topics together because I feel that each supports the other in essential ways. All of the usual definitions of the Lebesgue space $L^1[0,1]$ (e.g., in 19.38, 22.28, or 24.36) are quite involved; these definitions cannot be properly appreciated without some of the abstract theory of completions or Banach spaces or convergent nets. Conversely, an introduction to Banach spaces is narrow or distorted if it omits or postpones the rather important example of L^p spaces; the remaining elementary examples of Banach spaces are not diverse enough to give a proper feel for the subject.

How to Use This Book

Because students' backgrounds differ greatly, I have tried to assume *very few prerequisites.* The book is intended for students who have finished calculus plus at least four other college math courses. *HAF* will rely on those four additional courses, not for specific content, but only for mathematical maturity — i.e., for the student's ability to learn new material at a certain pace and a certain level of abstraction, and to fill in a few omitted details to make an exercise into a proof. Students with that amount of preparation will find *HAF* self-contained; they will not need to refer to other books to read this one. Students with sufficient mathematical maturity may not even need to refer to their college calculus textbooks; Chapters 24 and 25 reintroduce calculus in the more general setting of Banach spaces. Proofs are included, or at least sketched, for all the main results of this book except a few consistency results of formal logic. For those consistency results we give references in lieu of proofs, but the conclusions are explained in sufficient detail to make them clear to beginners.

Parts of *HAF* might be used as a classroom textbook, but *HAF* was written primarily for individual use. My intended reader will skip back and forth from one part of the book to another; different readers will follow different paths through the book. The reader should begin by skimming the table of contents to get acquainted with the ordering of topics. To facilitate skipping around in the book, I have included a large index and many cross-referencing remarks. Newly defined terms are generally given in **boldface** to make them easy to find. These definitions are followed by alternate terminology in *italics* if the literature uses other terms for the same concept or by cautionary remarks if the literature also uses the same term for other concepts. The first few pages of the first chapter introduce many of the symbols and typographical conventions used throughout the book; the index ends with a list of symbols. A list of charts, tables, diagrams, and figures is included in the index under "charts."

Mathematics textbooks usually postpone exercises until the end of each subchapter or each chapter, but *HAF* mixes exercises into the main text. In fact, *HAF* does not always distinguish sharply between "discussions," "theorems," "examples," and "exercises." All such assertions are *true statements*, with varying degrees of importance, generality, or difficulty, and with varying amounts of hints provided. Every student knows that reading through any proof in any math book is a challenge, whether that proof is marked "exercise" or not. Some computations and deductions are easier or more instructive to do than to watch, so for brevity I have intentionally given some proofs as sketches. All the "exercises" are actually part of the text; most of them will serve as essential examples or as steps in proofs of later theorems. Thus, in each chapter that is studied, *the reader should work through, or at least READ through, every exercise;* no exercise should be skipped.

ACKNOWLEDGMENTS

I am especially grateful to Isidore Fleischer, Mai Gehrke, Paul Howard, and Constantine Tsinakis, who helped with innumerable questions about algebra and logic. I am also grateful to many other mathematicians who helped or tried to help with many different questions: Richard Ball, Howard Becker, Lamar Bentley, Dan Biles, Andreas Blass, Douglas Bridges, Norbert Brunner, Gerard Buskes, Chris Ciesielski, John Cook, Matthew Foreman, Doug Hardin, Peter Johnstone, Bjarni Jónsson, William Julian, Keith Kearnes, Darrell Kent, Menachem Kojman, Ralph Kopperman, Wilhelmus Luxemburg, Hans van Maaren, Roman Mańka, Peter Massopust, Ralph McKenzie, Charles Megibben, Norm Megill, Michael Mihalik, Zuhair Nashed, Neil Nelson, Michael Neumann, Jeffrey Norden, Simeon Reich, Fred Richman, Saharon Shelah, Stephen Simons, Steve Tschantz, Stan Wagon, and others too numerous to list here. I am also grateful to many students who read through earlier versions of parts of this book. Of course, any mistakes that remain in this book are my own.

This work was supported in part by a Summer Award from the Vanderbilt University Research Council. I would also like to thank John Cook, Mark Ellingham, Martin Fryd, Bob Messer, Ruby Moore, Steve Tschantz, John Williams, and others for their help with TEX. This book was composed using several different computers and wordprocessors. It was typeset using LATEX, with some fonts and symbols imported from AMS-TEX.

I am also grateful to my family for their support of this project.

To Contact Me

I've surveyed the literature as well as I could, but it's enormous; I'm sure there is much that I've overlooked. I would be grateful for comments from readers, particularly regarding errors or other suggested alterations for a possible later edition. I will post the errata and other insights on the book's World Wide Web page on the internet.

Eric Schechter, August 16, 1996
`http://math.vanderbilt.edu/~schectex/ccc/`

Part A

SETS AND ORDERINGS

This Page Intentionally Left Blank

Chapter 1

Sets

Mathematical Language and Informal Logic

1.1. *A few typographical conventions.* Certain kinds of mathematical objects are most often represented by certain kinds of letters. For instance, mathematicians often represent a point by "x" and a function by "f," and very seldom the other way around. This book will usually adhere to the following guidelines, which are consistent with much (but not all!) of the literature of algebra, topology, and analysis. The reader is cautioned that there is no *standard* usage, in the literature or even in this book. The guidelines in the following list will be helpful, but the guidelines will have exceptions (which should be clear from the context). There is even some overlap between the categories listed above. For instance, in atomless set theory, discussed in 1.46, all sets are sets of sets.

$i,\ j,\ k,\ m,\ n,\ p,\ \ldots$	integers
$p,\ q,\ r,\ s,\ t,\ \ldots$	real numbers
$\mathbb{C},\ \mathbb{N},\ \mathbb{Q},\ \mathbb{R},\ \mathbb{Z}$	sets of numbers
$W,\ X,\ Y,\ Z,\ \Omega,\ \Gamma,\ \Lambda,\ \ldots$	main sets — e.g., linear spaces
$A,\ B,\ C,\ L,\ S,\ T,\ \ldots$	subsets of main sets
$a,\ b,\ y,\ z,\ \alpha,\ \beta,\ \lambda,\ \mu,\ \omega,\ \ldots$	elements of sets
$\mathcal{A},\ \mathcal{B},\ \mathcal{C},\ \ldots$	sets of sets — e.g., filters, topologies
$f,\ g,\ p,\ q,\ \alpha,\ \beta,\ \lambda,\ \mu,\ \pi,\ \ldots$	functions
$\Gamma,\ \Delta,\ \Phi,\ \Psi,\ \ldots$	collections of functions

1.2. All letters are variables, but some letters are more variable than others (as George Orwell might have put it). Every high school student has understood at least one example of this:

$$\text{the solutions of} \qquad ax^2+bx+c=0 \qquad \text{are} \qquad x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

Here the letters a, b, c are treated as real constants, but they can be *any* real constants;

they vary only slightly less than x does. Usually it should be clear from the context just which letters are varying more than others.

1.3. *Notes on "and" and "or."* Although mathematicians base their language on English or other "natural" languages, mathematicians alter the language slightly to make it more precise or to make it fit their purposes better. Some of the differences between English and mathematics may confuse the beginner.

For instance, there are two different meanings for the English word "or:"

$\sqcup$	vel	inclusive or	A or B or both
$+$	aut	exclusive or	A or B but not both.

Latin distinguishes between these two meanings by using two different words: "**vel**" and "**aut**;" see Rosser [1953/1978]. In everyday English, the term "or" is ambiguous; it could have either meaning. For clarification in English, "vel" is sometimes called "and/or," and "aut" is sometimes called "either/or." *In mathematics, "or" generally means "vel," unless specified otherwise.*

Undergraduate mathematics students sometimes confuse "and" and "or" in the following fashion: What is the solution set of $x^2 - 4x + 3 > 0$, in the real line? It is

$$\{x \in \mathbb{R} \,:\, x < 1\} \;\cup\; \{x \in \mathbb{R} \,:\, x > 3\} \quad = \quad \{x \in \mathbb{R} \,:\, x < 1 \text{ or } x > 3\}.$$

Thus, the appropriate word is "or." However, some calculus students write the solution as "$x < 1$ and $x > 3$," by which they mean "the points x that satisfy $x < 1$, and also the points x that satisfy $x > 3$" — thus they are using "and" for $\cup$ (union). Though such students may think that they know what they mean, this usage is not standard in mathematics and should be discontinued by students who wish to proceed in higher mathematics.

Another word for "or" is **disjunction**; the most commonly used symbol for it is $\vee$. Another word for "and" is **conjunction**; the most commonly used symbol for it is $\wedge$. However, we shall use $\sqcup$ and $\sqcap$ for "or" and "and," in order to reserve the symbols $\vee$ and $\wedge$ for use in some related lattices.

We shall use "not-A" or "$\neg A$" as abbreviations for the statement that "statement A is not true;" some mathematicians use other symbols such as $\sim A$. The symbol $\neg$, meaning "not," is also called **negation**. In conventional (ordinary) logic, used throughout most of this book, $\neg\neg A = A$; that is, not-not-A is equal to A. That equality fails in constructivist or intuitionist logic, which is discussed very briefly in Chapters 6 and 13.

1.4. The statement "A **implies** B" will sometimes be abbreviated as "$A \Rightarrow B$" or "$A \rightarrow B$;" the latter expression will be used in our chapter on logic. Either of these expressions means "**if** A is true **then** B is true" — or more precisely, "whenever A is true, then B is also true." The usage of "if ... then" in mathematics differs from the usage in English, because the mathematical statement $A \Rightarrow B$ makes no prediction about B in the case where A is false. For instance, in everyday English the statement "If it rains, then I will take my umbrella" is ambiguous — it could have either of the following meanings:

(i) If it rains, then I will take my umbrella. If it doesn't rain, then I won't take my umbrella.

(ii) If it rains, then I will take my umbrella. If it doesn't rain, then I might or might not take my umbrella.

In mathematics, however, (ii) is the only customary interpretation of "if ... then."

The mathematicians' implication also differs from the nonmathematicians' implication in this respect: we may have $A \Rightarrow B$ even if A and B are not causally related. For instance, "if ice is hot then grass is green" is true in mathematics, but it is nonsense in ordinary English, since there is no apparent connection between the temperature of ice and the color of grass. The mathematicians' implication is sometimes referred to as **material implication**, to distinguish it from certain other kinds of implications not commonly used in mathematics but sometimes studied by philosophers and specialized logicians.

The **converse** of the statement "$A \Rightarrow B$" is the statement "$B \Rightarrow A$." These two statements are *not* equivalent; the beginner must be careful not to confuse them. For instance, "$x = 3$" implies "x is a prime number," but "x is a prime number" does not imply "$x = 3$."

The statement "A if and only if B" may be abbreviated "A **iff** B;" it is also written "$A \iff B$." This statement means that both $A \Rightarrow B$ and the converse implication $B \Rightarrow A$ are true.

Statement A is **stronger** than statement B if $A \Rightarrow B$; then we may say B is **weaker** than A. More generally, a property P of objects is stronger than a property Q if every object that has property P also must have property Q — i.e., if the statement "X has property P" is stronger than the statement "X has property Q." (A related but slightly different meaning of "stronger than" is introduced in 9.4.) The mathematical usage of the terms "stronger" and "weaker" (and of other comparative adjectives such as coarser, finer, higher, lower) differs from the common nonmathematical English usage in this important respect: *In English, two objects cannot be "stronger" than each other, but in mathematics they can.* Thus, when $A \iff B$, each statement is stronger than the other. In particular, a statement is always stronger than itself. To say that

$$A \text{ implies } B \qquad \text{and} \qquad B \text{ does } \textit{not} \text{ imply } A,$$

we could say that A is **strictly stronger** than B. For instance, the property of being equal to 3 is strictly stronger than the property of being a prime number.

In general, "if ... then" is quite different from "if and only if." However, in mathematical *definitions* the words "and only if" generally are omitted and are understood implicitly, particularly when the defined word or phrase is displayed in boldface or italics. For instance, in our earlier sentence

> Statement A is **stronger** than statement B if $A \Rightarrow B$; then we may say B is **weaker** than A.

the "if" is really understood to be "if and only if."

1.5. When A and B are variables taking the values "true" or "false," then an expression such as "A and B" is a function of those variables — that is, the value of "A and B" depends on the values of A and B. The **truth table** below shows how several functions of

A and B depend on the values of A and B. In the table, "T" and "F" stand for "true" and "false," respectively.

A	B	not-A	A or B	A and B	$A \Rightarrow B$	$A \Longleftrightarrow B$
T	T	F	T	T	T	T
T	F	F	T	F	F	F
F	T	T	T	F	T	F
F	F	T	F	F	T	T

If a statement A is known to be always false, then the statement "$A \Rightarrow B$" is true, regardless of what we know or do not know about B; under these circumstances we may say that the implication "$A \Rightarrow B$" is **vacuously true**, or **trivially true**. The term "trivially true" can also be used to describe the implication "$A \Rightarrow B$" if B is known to be always true, since in that case the validity of A need not be considered.

1.6. *Exercises.*

a. The statement "$A \Rightarrow B$" is equivalent to the statement "B or not-A." Explain.

b. The **contrapositive** of "$A \Rightarrow B$" is the statement "not-$B \Rightarrow$ not-A." Show that an implication and its contrapositive *are* equivalent. We shall use them interchangeably.

c. **(De Morgan's Laws for logic.)** Explain:

$$\text{(not-}A\text{) and (not-}B\text{)} \qquad \text{is equivalent to} \qquad \text{not-(}A \text{ or } B\text{);}$$

$$\text{(not-}A\text{) or (not-}B\text{)} \qquad \text{is equivalent to} \qquad \text{not-(}A \text{ and } B\text{).}$$

1.7. Duality arguments. Some concepts in mathematics occur in pairs; each member of the pair is said to be *dual* to the other. A few examples are listed in the table below; these examples and others are developed in more detail in later chapters. The statements about these concepts occur in pairs. In some cases, one of the two statements is preferred, because it is more relevant to applications or is simpler in appearance.

A concept	and	$\wedge$	$\cap$	min	inf	open	int	ideal	$\leq$	$\subseteq$
Its dual	or	$\vee$	$\cup$	max	sup	closed	cl	filter	$\geq$	$\supseteq$

Generally there is a simple and mechanical method for transforming a statement into its dual statement and for transforming the proof of a statement into the proof of the dual statement. For instance, De Morgan's Laws for logic (given in 1.6.c) can be used to convert between *and*s and *or*s, by inserting a few *not*s. Other such conversion rules will be given in later chapters. In some cases, for brevity, we state and/or prove only one of the two statements in the pair. The other statement is left unstated and/or unproved, but the reader should be able to fill in the missing details without any difficulty.

1.8. *On parsing strings of symbols.* In this book, we generally read set-theoretical operations ($\cap, \cup, \complement$, etc.) first, then set-theoretical relations ($=, \subseteq, \subsetneqq$, etc.), then logical relations between statements. For instance,

$$A \quad = \quad B \qquad \Longleftrightarrow \qquad C \quad \cap \quad D \quad \supseteq \quad E \qquad\qquad (*)$$

should be interpreted as

$$\Big[A \quad = \quad B\Big] \qquad \Longleftrightarrow \qquad \Big[(C \cap D) \quad \supseteq \quad E\Big].$$

Generally we omit the parentheses, but we may sometimes use extra spacing to make the correct interpretation more obvious:

$$A \quad = \quad B \qquad\qquad \Longleftrightarrow \qquad\qquad C \cap D \quad \supseteq \quad E\,.$$

We emphasize that this order of precedence depends on the context — i.e., the present book is concerned with abstract analysis. In a different context, the expression $(*)$ could be read in an entirely different order. For instance, in some books on logic, $\cap$ means "and" and $\supseteq$ means "is implied by." Hence all four of the symbols $=$, $\Leftrightarrow$, $\cap$, and $\supseteq$ are *binary operations on statements* — i.e., they are operators $\Box$ with the syntax that if P and Q are statements, then $P \Box Q$ is a statement. Therefore, in a logic book, the displayed equation $(*)$ could make sense with *any* arrangement of parentheses, and it would have different meanings with different arrangements of parentheses. In that context, $(*)$ would be highly ambiguous; some parentheses would be needed for clarification.

1.9. Proof by contradiction is a nonconstructive technique of logic, so widely used in mainstream mathematics that it generally goes unremarked. It may be confusing to beginning mathematicians who have never seen it explained. The technique is this:

> If we wish to prove $A \Rightarrow B$, we can assume the truth of both A and not-B. From those two assumptions we deduce a contradiction; the contradiction demonstrates that indeed $A \Rightarrow B$.

The justification of this technique is 1.6.a.

Proof by contradiction has this advantage: We work from two assumptions (both A and not-B) rather than just the one assumption of A; thus we have more statements on which to build. Consequently, proofs by contradiction are often easier to *discover* than direct proofs.

Proofs by contradiction also have a couple of disadvantages:

- Proofs by contradiction are often harder to *read* than direct proofs because they are conceptually more complicated. Proofs by contradiction are conceptually complicated. A beginning student of mathematics may prefer to assume that A is *true* and try to discover what else is then true — a sort of one-directional approach. But a proof by contradiction works simultaneously in two directions, mixing together statements (such as A and its consequences) that we take to be true with statements (such as not-B) that we temporarily pretend are true but shall eventually decide are false. This scheme must seem diabolical, or at least amoral, to beginners: It is not concerned so much with "what is true," but rather with "what implies what."

- A proof by contradiction is often *nonconstructive*: It may prove the existence of some mathematical object without producing any explicit example of that object. For a very vivid example of this lack of examples, see 6.5. The availability or unavailability of explicit examples is one of the main themes of this book. A proof by contradiction may convince us that a statement is *true*, but it may not give us as much *intuitive understanding* of that statement as a direct proof would.

1.10. The phrase "**we may assume**" is often used in the literature in ways that may bewilder the novice. For instance, consider a proposition of this form:

> *Proposition A.* Let X be a mathematical object satisfying hypothesis $H(X)$. Then X satisfies conclusion $C(X)$.

A published proof of Proposition A might begin something like this:

(!) We may assume that X also satisfies property $P(X)$.

The reasoning step (!) has several possible meanings; we shall describe three of them below. The simplest meaning of (!) would be that

> (1) Hypothesis $H(X)$ actually *implies* property $P(X)$, by some reasoning that should be evident to a sufficiently advanced reader.

Readers who are not so advanced may spend many hours trying to fill in that reasoning. However, (!) may not mean (1) after all. Indeed, if (1) were true then (!) would probably be worded a bit differently — e.g., the proof might have begun by saying "We first observe that, obviously, $H(X) \Rightarrow P(C)$." A more likely meaning of (!) is this:

> (2) $H(X)$ and not-$P(X)$ together imply $C(X)$, by some reasoning that should be evident to the reader. Hence, in trying to prove $H(X) \Rightarrow C(X)$, we may concentrate on the case where $P(X)$ holds.

That is harder but still manageable. Alas, (!) has yet a third meaning, and this one is much too subtle for some beginners:

> (3) The text will now give the details of a proof of a slightly easier proposition. After reading the proof provided for the easier proposition, the reader is expected to figure out the details of *how to use* that easier proposition to prove Proposition A. The easier proposition is as follows:
>
> *Proposition B.* Let Y be a mathematical object satisfying hypotheses $H(Y)$ and $P(Y)$. Then Y satisfies conclusion $C(Y)$.

The missing details might go as follows: Let any object X be given, satisfying hypothesis $H(X)$ but not necessarily property $P(X)$. By some clever method (which the reader must figure out), we now construct a collection of related objects $Y_1, Y_2, Y_3, \ldots$, with each Y_k satisfying both hypothesis $H(Y_k)$ and property $P(Y_k)$. Then Proposition B is applicable to the Y_k's, and so we can draw conclusions $C(Y_1), C(Y_2), C(Y_3), \ldots$. By some clever method (which, again, the reader must figure out), we may then use that information to help us prove $C(X)$.

In such an argument, object X does *not* necessarily satisfy $P(X)$, despite the wording of statement (!). The effect of statement (!) is to discard the original object X, replace it with the new object Y_k, and **relabel** Y_k to call it X now. Some other relabeling arguments will be discussed and used in 2.19, 7.21, and 16.5.

1.11. *How much formalism do we need?* It is not necessary to learn the definitions of "noun" and "verb" to become a fluent speaker of English (or any other natural language). One can learn the language quite well just by studying examples; this is the method by which toddlers learn their native tongue.

Similarly, most mathematicians use logic properly without ever knowing its formal rules. This book is intended for "most mathematicians," and we shall discuss logic and formal set theory as little as possible. The few concepts from logic and set theory that we shall need will be developed briefly and informally. For a more complete and formal development, the interested reader is referred to more advanced and specialized books and papers.

Informal reasoning is not always reliable, in part because informal language is not always reliable. Natural languages (such as English) evolved to suit the mundane, ordinary, real world, but mathematicians often find themselves considering extraordinary ideas.

For instance, a self-referencing statement such as

This statement is false

cannot be true *or* false. (This is the simplest form of the **Paradox of the Liar**, also known as the **Paradox of Epimenides**.) Such statements do not arise in "ordinary" reality, but such statements show mathematicians a need for careful rules about language and reasoning.

The simplest way to deal with self-referencing statements is to simply prohibit them and avoid the confusion. We shall follow that policy in this book. However, we remark that self-referencing recently has been analyzed in a meaningful and useful way by Aczel [1988] and Barwise and Etchemendy [1987]. Such analyses are especially useful in the theory of computer programs. A computer program may operate on data files that are stored in memory; one of those files may be the program that is operating.

1.12. We should mention one more type of self-referencing before we leave the topic. The self-referencing in Epimenides's Paradox is very direct: The word "this" in the sentence "This sentence is false" points directly to the sentence in which that word is located. But Quine's Paradox, below, involves a more indirect type of self-referencing, which has some important uses in logic.

A typical sentence in English consists of a subject followed by a predicate. For instance, in each of the sentences

Jane is a girl.

Jane runs with the ball.

the subject is "Jane" and the predicate is the remainder of the sentence. The subject is some "thing" that is being discussed; the predicate says that the subject "is" something or "does" something.

Mathematicians often wish to discuss mathematical objects, so in a mathematics text the subject of a sentence can be a mathematical symbol or formula. For instance,

> $\Box$ is a box symbol.
>
> "$\Box$" is a box symbol.
>
> x is a variable.
>
> "x" is a variable.
>
> "$x = y$" is an equation.

are all acceptable sentences in a mathematics book or paper. Whether we include or omit the quotation marks is generally a matter of taste; our main rule is that the intended meaning should be clear. In this author's opinion, the last example would become confusing if the quotation marks were omitted, but the quotation marks are optional in the other examples. (Of course, in a book or paper on logic, the quotation marks may have a more technical meaning, and then their use or omission is no longer a matter of taste.)

We shall now consider sentences that follow the format described above, but in these sentences the subject will be some phrase of the English language — i.e., a sentence fragment. Thus, we shall consider sentences that discuss certain sentence fragments. In each case, the sentence fragment will consist of a sentence whose subject has been omitted.

> "is a girl" is a sentence fragment composed of three words.
>
> "runs with the ball" is a sentence fragment composed of four words.
>
> "is a sentence fragment" is a sentence fragment.
>
> "is composed of five words" is composed of five words.

Each of those four sentences is true. The last two sentences have a peculiar structure: they consist of a sentence fragment in quotes, followed by the same sentence fragment without quotes, followed by a period. In Hofstadter [1979], the process of forming such a sentence from such a fragment is called **quining**. Thus, the last sentence displayed above is the result of starting from the fragment

is composed of five words

and then quining that fragment.

Now, **Quine's Paradox** consists of the peculiar sentence

> "yields a falsehood when preceded by its quotation" yields a falsehood when preceded by its quotation.

or, in Hofstadter's terminology,

> "yields a falsehood when quined" yields a falsehood when quined.

These sentences are paradoxical: they are false if true, and true if false. (Think about it for a moment.) These sentences do not involve direct self-referencing of the sort found in Epimenides's Paradox; there is no "this" that points to itself. However, in either formulation, Quine's peculiar sentence discusses another sentence that would be formed as the result of a quining. Just by coincidence (not really), the sentence being discussed happens to be identical to the sentence doing the discussing. Quine formed this paradox in order to explain Gödel's Proof; see 14.62.

Basic Notations for Sets

1.13. A **set** is a collection of objects. This is not really a definition, since we do not state what a "collection" is; we shall rely on the reader's intuition about these terms. A more formal approach will be introduced in 1.44 and the sections thereafter.

Three common ways to specify a set are by listing the objects in the set, by specifying a larger set and a property that determines the subset in question, and by listing a parameter set and a way to form some object from each value of the parameter. For instance, the set of odd positive integers can be represented in any of these ways:

$$\{1,3,5,7,\ldots\} \quad = \quad \{n \in \mathbb{N} \,:\, n \text{ is odd}\} \quad = \quad \{2m+1 \,:\, m \in \mathbb{N}\}.$$

In the last expression, $\mathbb{N}$ is used as a **index set**, or **parameter set**. (Some mathematicians would write that last expression as $\{2m+1 \mid m \in \mathbb{N}\}$, but this book will have too many other uses for vertical bars.)

The order of the elements of a set is not relevant, and repetitions are ignored; for instance, $\{1,2,3,4\} = \{4,3,1,2\} = \{1,2,3,1,4\}$. To emphasize this we may occasionally refer to a set as an *unordered set* to contrast it with ordered sets, such as those in 1.32. Two sets A and B are defined to be **equal** (as sets) if they contain the same elements — i.e., if they satisfy $x \in A \Leftrightarrow x \in B$.

Two mathematical objects may be equal as sets even though they have different additional structures associated with them. For instance, the real number system with its usual topology is different from the real number system with the discrete topology — i.e., these are different topological spaces. But these topological spaces are equal *as sets*, since they have the same members.

The term "collection" will usually mean the same thing as "set," but occasionally "collection" may have the more general meaning of "class," discussed in 1.44.

1.14. Here are the two most basic notions of sets:

> "$x \in S$" is read as: x **belongs** to S, or x is an **element** of S, or x is a **member** of S. It is occasionally written as "$S \ni x$."

> "$A \subseteq B$" means $x \in A \Rightarrow x \in B$; that is, each element of A is also an element of B. It is read as: A is a **subset** of B, or B is a **superset** of A. It is also written as "$B \supseteq A$."

Unfortunately, the terms "include" and "contain" are ambiguous. As they are commonly used in the mathematical literature,

> either of the statements "U **includes** V" or "U **contains** V" can have either of the meanings "$U \ni V$" or "$U \supseteq V$."

When the words "include" or "contain" are used, the reader must determine the intended meaning from context.

The statement "x is not an element of S" can be written $x \notin S$; the statement "A is not a subset of B" is occasionally written as $A \not\subseteq B$. When $S \subseteq X$ and $S \neq X$, we say S is a **proper subset** of X, or X is a **proper superset** of S; this is sometimes written $S \subsetneqq X$ or $X \supsetneqq S$.

The symbols $\subset$ and $\supset$ are ambiguous: They are used for $\subseteq$ and $\supseteq$ by some mathematicians, and for $\subsetneqq$ and $\supsetneqq$ by other mathematicians. We shall not use $\subset$ or $\supset$ in this book.

1.15. *Some sets of numbers.* Numbers are the basis of what most analysts consider to be "analysis." The list below shows some of the most commonly used sets of numbers.

$\mathbb{N}$	positive integers (also known as natural numbers)
$\mathbb{Z}$	integers
$\mathbb{Q}$	rational numbers (quotients of integers)
$\mathbb{R}$	real numbers (introduced formally in Chapter 10)
$[-\infty, +\infty]$	extended reals (introduced in 1.17)
$\mathbb{C}$	complex numbers (introduced in Chapter 10)
$\mathbb{T}$	the circle group (introduced in 10.32.f)
$\mathbb{F}$	unspecified field — generally understood to be $\mathbb{R}$ or $\mathbb{C}$
$\mathbb{A}, \mathbb{B}, \mathbb{D}$	directed sets ("generalized numbers;" see 7.3)

We assume an informal acquaintance with $\mathbb{Q}$ and $\mathbb{R}$ — e.g., techniques of computation, such as in college calculus. Relying on that informal acquaintance only for some illustrative examples, in later chapters we shall carefully develop basic ideas of orderings, groups, and fields, leading up to formal definitions of $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ in Chapter 10.

1.16. In this book, $\mathbb{N}$ and $\mathbb{Z}$ have their usual, classical meanings, and we assume a familiarity with the elementary properties of those sets of numbers. *Caution*: Many mathematicians agree with our definition that $\mathbb{N} = \{1, 2, 3, \ldots\}$, but many others instead use the symbol $\mathbb{N}$ to represent the set $\{0, 1, 2, 3, \ldots\}$.

Set theorists often find it useful to define the integers (and everything else) in terms of sets — see 1.46. Zermelo defined the nonnegative integers $0, 1, 2, 3, \ldots$ to be the sets $\varnothing$,

$\{\varnothing\}$, $\{\{\varnothing\}\}$, $\{\{\{\varnothing\}\}\}$, and so on. Later, von Neumann defined the nonnegative integers to be the sets

$$\varnothing, \qquad \{\varnothing\}, \qquad \{\varnothing, \{\varnothing\}\}, \qquad \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\},$$

and so on, as described in 5.44. Either of these definitions is manageable, but von Neumann's more complicated definition has a few advantages for purposes of set theory, and so it is now widely used in that field.

For purposes outside of set theory, however, it is conceptually simpler not to attach the labels "0," "1," "2," etc., to particular sets. (This point is discussed further by Hirsch [1995].) Instead we usually view the integers as indivisible objects, for which we define algebraic operations in the usual fashion. Thus $\mathbb{N} \cup \{0\}$ is a *monoid* and $\mathbb{Z}$ is a *ring*; these notions are discussed in Chapter 8.

In dealing with the integers, we shall rely on the reader's intuition, not on a precise definition and list of properties. Although it is possible to specify the positive integers uniquely by Peano's Axioms (see 14.52), that specification is nontrivial and rests on an understanding of conventional language. In nonstandard analysis, language is sometimes used in a different fashion, and then the "integers" take on a new meaning, as described in 14.68 and 14.69. The reader of this book does *not* need to be familiar with the nonstandard integers; we have mentioned them here only to emphasize our reliance on some shared intuition about the standard (i.e., conventional) integers.

1.17. Let $+\infty$ and $-\infty$ be the names given to some two objects that are not real numbers; the object $+\infty$ may also be abbreviated as ∞. The **extended real line**, denoted $[-\infty, +\infty]$, is the set $\mathbb{R} \cup \{-\infty, +\infty\}$ — that is, the real number system $\mathbb{R}$ plus these two additional points. We extend the ordering of $\mathbb{R}$ to this larger set by defining $-\infty < r < +\infty$ for all real numbers r. Addition and multiplication are usually extended to this larger set of numbers by the rules indicated in the following tables. In the tables, "undef.," "pos.," and "neg." are abbreviations for "undefined," "positive real," and "negative real." The product of 0 and $\pm\infty$ is sometimes left undefined, but more often it is defined to be 0. That product may come as a surprise to some students and is discussed further in 15.28.c.

PLUS	$-\infty$	real	$+\infty$
$-\infty$	$-\infty$	$-\infty$	undef.
real	$-\infty$	real	$+\infty$
$+\infty$	undef.	$+\infty$	$+\infty$

TIMES	$-\infty$	neg.	0	pos.	$+\infty$
$-\infty$	$+\infty$	$+\infty$	0	$-\infty$	$-\infty$
neg.	$+\infty$	pos.	0	neg.	$-\infty$
0	0	0	0	0	0
neg.	$-\infty$	neg.	0	pos.	$+\infty$
$+\infty$	$-\infty$	$-\infty$	0	$+\infty$	$+\infty$

1.18. *Preview of assorted infinities.* The term "infinity" has several different meanings in mathematics, and it is important not to confuse these with each other.

Some older mathematics books sometimes refer to a **potential infinity** such as $\lim_{x \downarrow 0} \frac{1}{x}$; this is a very large finite number that gets larger without bound. Our dealings with potential infinities may be simplified if we adjoin to $\mathbb{R}$ some **ideal points** $-\infty$ and $+\infty$, as discussed in the preceding section. The resulting number system $[-\infty, +\infty]$ is algebraically somewhat awkward — unlike $\mathbb{R}$, it is not a field; indeed, it is not even an additive monoid.

By adding many more ideal points, we obtain a more satisfactory algebraic system. The **hyperreal line**, $^*\mathbb{R}$, is an ordered field strictly larger than $\mathbb{R}$; it is discussed in 10.18. Among other things, it contains some numbers that are infinitely large and some numbers (besides zero) that are infinitely small. An infinitely large number is a constant that plays a role similar to the role played by finite variable such as the number x in the expression "$\lim_{x\uparrow\infty} f(x)$." Similarly, an infinitely small but positive constant number plays a role similar to that played by the finite variable x in the expression "$\lim_{x\downarrow 0} f(x)$."

Yet another kind of infinity is the "number" of elements in an infinite set such as $\mathbb{N}$ or $\mathbb{R}$; that "number" is called the **cardinality** of the set. Some older mathematics books refer to it as an **actual infinity**, in contrast to the potential infinity mentioned above. There are several different infinite cardinalities — for instance, the cardinalities of the sets $\mathbb{N}$ and $\mathbb{Q}$ are equal (see 2.20.f), but the cardinality of the set $\mathbb{R}$ is larger (see 10.44.f). In fact, there are infinitely many different sizes of infinities; that follows from 2.20.l. Some arithmetic of cardinalities is possible — for instance, in later chapters we shall see that $\text{card}(S \times T) = \max\{\text{card}(S), \text{card}(T)\}$ when S and T are infinite sets, and $\text{card}(2^X) > \text{card}(X)$ for any set X. However, this arithmetic should not be confused with the arithmetic of the hyperreal numbers; cardinalities do not form a field. Infinite cardinal numbers are sometimes denoted by $\aleph_n$; we consider this notation briefly in 5.48. Also related are the infinite ordinals, introduced in 5.44; the first infinite ordinal is often denoted ω.

Our several notions of the "infinite" are only distantly related. To avoid confusion, think of them as entirely unrelated uses of the same strings of letters.

Yet another unrelated use of "infinity" is that in theology. The beginner is urged to put aside any spiritual notions of infinity, for mathematicians have tamed infinity and made it entirely a secular matter. (On the other hand, mathematics is not devoid of spiritual questions; see particularly 6.8 and 14.71.)

1.19. The set with no elements is called the **empty set** (or *null set*); it is denoted by $\varnothing$ (or by $\{\ \}$ in some books).

The word "nothing" is used in different ways in English. For instance, if we order things by our preferences, then "a ham sandwich is better than nothing" can be written

$$\text{ham sandwich} \quad > \quad \varnothing.$$

However, "nothing is better than true love" should *not* be written as "$\varnothing >$ true love." Rather, it should be written as

$$\{x \ : \ x > \text{true love}\} \quad = \quad \varnothing.$$

Thus, we *cannot* conclude that "a ham sandwich is better than true love."

1.20. *A few more sizes of sets.* A **singleton** is a set $\{x\}$ containing exactly one element. The objects x and $\{x\}$ can never be equal (see 1.49), and in some contexts the distinction between x and $\{x\}$ is crucial. In some other contexts, however, x and $\{x\}$ are used in substantially different ways, so that no confusion is possible if we find it convenient to write x and $\{x\}$ interchangeably. (For instance, the unique solution of $u^2 + 2u + 1 = 0$ is $u = -1$, and the solution set is $\{-1\}$; generally these two answers are interchangeable.)

Examples for beginners to think about. $\{x,y\}$ is a singleton if and only if $x = y$, and $\{\{x,y\}\}$ is a singleton in any case. The set $\{\varnothing\}$ is a singleton.

A set S is **finite** if the number of elements in S is a finite number — i.e., if S can be written in the form $\{x_1, x_2, x_3, \ldots, x_n\}$ for some $n \in \mathbb{N} \cup \{0\}$. (We permit $n = 0$; thus, the empty set is a finite set.) A set that cannot be so written is **infinite**.

A set S is **countable** if it is empty or can be written in the form $\{x_1, x_2, x_3, \ldots\}$. We emphasize that repetitions are permitted; thus by our definition any finite set is also a countable set. A set that is not countable is **uncountable**. A set S is **countably infinite** if it is countable and infinite — or, equivalently, if and only if it can be written in the form $\{x_1, x_2, x_3, \ldots\}$ without repetitions. *Caution*: Some mathematicians use these terms a little differently and apply the term "countable" only to the sets of the form $\{x_1, x_2, x_3, \ldots\}$ without repetitions — i.e., only to the sets that we have called "countably infinite."

Other sizes of sets will be discussed in 2.16.

Ways to Combine Sets

1.21. The **power set** of a given set X is $\{S : S \subseteq X\}$, the set of all subsets of X. We shall usually write the power set of X as $\mathcal{P}(X)$. It is also denoted 2^X, for reasons discussed in 2.20.k.

For instance, the power set of $\{0,1\}$ is the set $\mathcal{P}(\{0,1\}) = \Big\{\varnothing, \{0\}, \{1\}, \{0,1\}\Big\}$. The power set of the empty set is $\mathcal{P}(\varnothing) = \{\varnothing\}$, which is a singleton — i.e., it has one element, so it is not empty.

1.22. Suppose that $\mathcal{S} = \{S_\lambda : \lambda \in \Lambda\}$ is a set of sets — i.e., Λ is a set, and S_λ is a set for each $\lambda \in \Lambda$. Then the **union** of the S_λ's is the set $\{x : x \in S_\lambda \text{ for at least one } \lambda\}$. It can be denoted by any of these expressions:

$$\operatorname{Un}(\mathcal{S}), \qquad \bigcup \mathcal{S}, \qquad \bigcup \{S_\lambda : \lambda \in \Lambda\}, \qquad \bigcup_{\lambda \in \Lambda} S_\lambda.$$

Other notations are available in certain special cases: The union of finitely many sets $S_1, S_2, \ldots, S_n$ may be written as $\bigcup_{k=1}^n S_k$ or as $S_1 \cup S_2 \cup \cdots \cup S_n$. The union of a sequence of sets $S_1, S_2, S_3, \ldots$ may be written as $\bigcup_{k=1}^\infty S_k$ or as $S_1 \cup S_2 \cup S_3 \cup \cdots$. Note that if $L \subseteq \Lambda$, then $\bigcup_{\lambda \in L} S_\lambda \subseteq \bigcup_{\lambda \in \Lambda} S_\lambda$. In particular, $\bigcup_{\lambda \in \varnothing} S_\lambda$ is just the empty set.

1.23. Again suppose that $\mathcal{S} = \{S_\lambda : \lambda \in \Lambda\}$ is a set of sets. Then the **intersection** of the S_λ's is the set $\{x : x \in S_\lambda \text{ for every } \lambda\}$. It can be denoted by any of these expressions:

$$\operatorname{Int}(\mathcal{S}), \qquad \bigcap \mathcal{S}, \qquad \bigcap \{S_\lambda : \lambda \in \Lambda\}, \qquad \bigcap_{\lambda \in \Lambda} S_\lambda.$$

The expressions

$$\bigcap_{k=1}^{n} S_k = S_1 \cap S_2 \cap \cdots \cap S_n \qquad \text{and} \qquad \bigcap_{k=1}^{\infty} S_k = S_1 \cap S_2 \cap S_3 \cap \cdots$$

are interpreted in a fashion analogous to that for unions.

If $L \subseteq \Lambda$ and $L \neq \varnothing$, then $\bigcap_{\lambda \in L} S_\lambda \supseteq \bigcap_{\lambda \in \Lambda} S_\lambda$. The expression $\bigcap_{\lambda \in \varnothing} S_\lambda$ is not meaningful without further specification, but the following convention is often useful: If the S_λ's are all subsets of a fixed set X whose choice is understood, then we may agree to let $\bigcap_{\lambda \in \varnothing} S_\lambda$ mean X.

1.24. If S and X are sets, then the **complement** (or *relative complement*) of S in X is the set

$$X \backslash S \quad = \quad \{x \in X : x \notin S\}.$$

For example, $\{a, b, c\} \backslash \{c, d\} = \{a, b\}$. We emphasize that S is not necessarily a subset of X. *Caution*: Some mathematicians write the set $X \setminus S$ instead as $X - S$. However, that also can be interpreted as $\{x - s : x \in X, s \in S\}$ in contexts where subtraction is meaningful — e.g., if S and X are subsets of $\mathbb{R}$.

If the choice of X is clear and/or does not need to be mentioned explicitly, and S is a subset of X, then the relative complement of S in X may be written more briefly as $\complement S$. (Some mathematicians write this as $\mathcal{C}S$ or S^c or $\overline{S}$.) The $\complement$ notation is useful especially when we are considering many subsets R, S, T, U, etc., of a single set X; note then $S \backslash T = S \cap \complement T$. The $\complement$ notation simplifies the appearance of many results — for instance, $\complement\complement S = S$. More generally, $\complement^n S = S$ when $n = 0, 2, 4, 6, \ldots$. Here we adopt the convention that $\complement^0 S = S$, and $\complement^{n+1} S = \complement\complement^n S$; this exponential notation will be particularly helpful in 13.11.

The symbol $\complement$ will be given a more general meaning in 13.1; see the discussion in 13.3.

1.25. Also simplified by the $\complement$ notation are **De Morgan's Laws for sets**:

$$\complement\left(\bigcup_{\lambda \in \Lambda} S_\lambda\right) = \bigcap_{\lambda \in \Lambda} (\complement S_\lambda), \qquad \complement\left(\bigcap_{\lambda \in \Lambda} S_\lambda\right) = \bigcup_{\lambda \in \Lambda} (\complement S_\lambda).$$

That is: The complement of a union is the intersection of complements, and vice versa. The proofs are an easy exercise.

There is a *duality* (as in 1.7) between statements about any collection of sets and statements about the complements of those sets. This duality is order-reversing; i.e.,

$$S \subseteq T \quad \Longleftrightarrow \quad \complement S \supseteq \complement T.$$

By De Morgan's Laws, the duality transforms unions to intersections, and vice versa.

1.26. We say that two sets **meet** if their intersection is nonempty; otherwise the sets are **disjoint**. Note that $\varnothing$ and any set are disjoint. A collection of sets is *disjoint* (or for emphasis, **pairwise disjoint**) if each pair of distinct sets in the collection is disjoint. A **partition** of a set X is a collection of pairwise disjoint sets that have union equal to X.

A collection of sets $\{S_\lambda : \lambda \in \Lambda\}$ is **fixed** if their intersection $\bigcap_{\lambda \in \Lambda} S_\lambda$ is nonempty; the collection is **free** if its intersection is empty. We emphasize that this does not refer to the pairwise intersection. For instance, if a, b, c are distinct objects, then the collection

$$\mathcal{S} \quad = \quad \Big\{\{a, b\},\ \{b, c\},\ \{c, a\}\Big\}$$

is free but not disjoint.

A collection $\{S_\lambda : \lambda \in \Lambda\}$ of subsets of a set X is said to be a **cover**, or **covering**, of X if $\bigcup_{\lambda \in \Lambda} S_\lambda = X$. Note that this condition is satisfied if and only if $\{\complement S_\lambda : \lambda \in \Lambda\}$ is free (where $\complement$ denotes complement in X). Thus, "cover" and "free" are dual concepts, in the sense of 1.7. Also note that a partition is the same thing as a disjoint covering.

Examples. Let $X = \{a, b, c, d, e\}$ consist of five distinct elements. Then the collection of sets $\{\{a, b\}, \{c\}, \{d, e\}\}$ is a partition of X; $\{\{a, b\}, \{c\}, \{d\}\}$ is a disjoint collection but not a partition or a cover; $\{\{a, b\}, \{b, c\}, \{c, d, e\}\}$ is a free cover. Any disjoint collection of two or more sets is free. However, a collection consisting of just one nonempty set is disjoint and not free.

Further definitions. Suppose that $\mathcal{S} = \{S_\lambda : \lambda \in \Lambda\}$ is a cover of a set X. Then

- A **subcover** is a cover of X that is of the form $\{S_\lambda : \lambda \in \Lambda_0\}$ for some set $\Lambda_0 \subseteq \Lambda$.
- A **refinement** of $\mathcal{S}$ is a cover $\mathcal{T} = \{T_\mu : \mu \in M\}$ of X with the property that each T_μ is contained in some S_λ.
- A **precise refinement** of $\mathcal{S}$ is a cover of X of the form $\mathcal{T} = \{T_\lambda : \lambda \in \Lambda\}$ (with the same set Λ), such that $T_\lambda \subseteq S_\lambda$ for each λ.

Of course, any precise refinement is a refinement.

1.27. The **symmetric difference** of two sets S and T is the set

$$\begin{aligned} S \bigtriangleup T &= (S \backslash T) \cup (T \backslash S) &&= (S \cap \complement T) \cup (\complement S \cap T) \\ &= (S \cup T) \setminus (S \cap T) &&= \{x \ : \ x \text{ is in } S \text{ or in } T \text{ but not both}\}. \end{aligned}$$

For instance, $\{a, b, c\} \bigtriangleup \{b, c, d\} = \{a, d\}$. Note that $S \bigtriangleup T = T \bigtriangleup S$, that $S \bigtriangleup S = \varnothing$, and that $S \bigtriangleup \varnothing = S$. Also, for subsets of a given set X, we have $\complement(S \bigtriangleup T) = S \bigtriangleup (\complement T) = (\complement S) \bigtriangleup T$.

Exercise. Show that

$$\begin{aligned} (R \bigtriangleup S) \bigtriangleup T &= R \bigtriangleup (S \bigtriangleup T) \\ &= \{x : x \text{ is in exactly one or three of the sets } R, S, T\}. \end{aligned}$$

More generally, use induction to show that

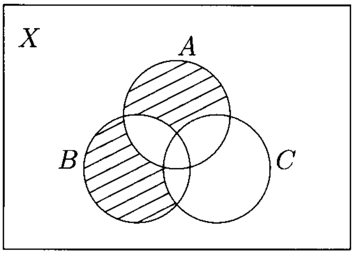
$$S_1 \bigtriangleup S_2 \bigtriangleup \cdots \bigtriangleup S_n = \{x : x \text{ is in an odd number of the sets } S_j\}.$$

1.28. A **Venn diagram** is used to indicate the unions, intersections, and complements of several sets. Two of these diagrams are shown below. Typically, a Venn diagram is used for two or three subsets of a larger set X (which is sometimes called "the universe," or "the universal set," in this context — see 1.44). The set X may be represented by a rectangle, and its subsets A, B, and C are represented as disks contained in that rectangle. If no assumptions are made about the sets A, B, C, then they are drawn "in general position" — i.e., so that each of the eight sets

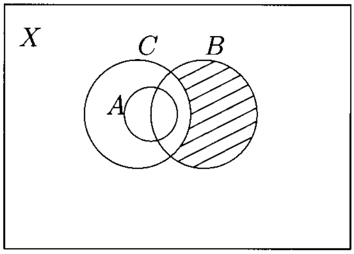
$$A \cap B \cap C, \qquad A \cap B \cap \complement C, \qquad A \cap \complement B \cap C, \qquad A \cap \complement B \cap \complement C,$$

$$\complement A\cap B\cap C,\qquad \complement A\cap B\cap \complement C,\qquad \complement A\cap \complement B\cap C,\qquad \complement A\cap \complement B\cap \complement C$$

is represented by a single nonempty region in the rectangle. (See the first diagram.) If some relationship between the sets is known, then this may be reflected in the diagram; for instance, if we know that $A \subseteq C$, then we may draw the disk for A inside the disk for C. (See the second diagram.)



Shaded region
is $(A \triangle B)\backslash C$.



If $A \subseteq C$, then
$(A \triangle B)\backslash C = B\backslash C$.

Do not rely too heavily on Venn diagrams or other figures — particularly complicated ones — for they can be erroneous in subtle ways. A common error is to attribute to a figure more generality than it truly possesses and thus to overlook certain special cases not explained by the figure. (In 15.19 are some further remarks about the limitations of diagrams.) However, simple diagrams can be trusted if constructed carefully, and, in any case, diagrams can be used to help us *find* other proofs that do not rely on diagrams.

1.29. *Distributive laws.* The following equations occur in dual pairs. In each case it is only necessary to prove one equation; the other then follows by duality using De Morgan's Laws (1.25).

a. Intersection and union distribute over each other. That is:

$$S \cap (T \cup U) \;=\; (S \cap T) \cup (S \cap U)$$

and

$$S \cup (T \cap U) \;=\; (S \cup T) \cap (S \cup U)$$

for all sets S, T, U.

b. In fact, intersection and union are *infinitely distributive* over each other:

$$\left(\bigcup_{\alpha\in A} S_\alpha\right) \cap \left(\bigcup_{\beta\in B} T_\beta\right) \;=\; \bigcup_{\alpha\in A}\,\bigcup_{\beta\in B} (S_\alpha \cap T_\beta)$$

and

$$\left(\bigcap_{\alpha\in A} S_\alpha\right) \cup \left(\bigcap_{\beta\in B} T_\beta\right) \;=\; \bigcap_{\alpha\in A}\,\bigcap_{\beta\in B} (S_\alpha \cup T_\beta)$$

for any index sets A, B and any sets S_α, T_β. This notion is generalized in 4.23.

1.30. *Closure under operations.* Let $\mathcal{S}$ be a collection of subsets of a set X. We say that $\mathcal{S}$ is closed under some set operation if performing that operation on members of $\mathcal{S}$ yields another member of $\mathcal{S}$. For instance, $\mathcal{S}$ is

closed under finite union if $S_1, S_2 \in \mathcal{S} \;\Rightarrow\; S_1 \cup S_2 \in \mathcal{S}$

or equivalently, if $S_1, S_2, \ldots, S_n \in \mathcal{S} \;\Rightarrow\; S_1 \cup S_2 \cup \cdots \cup S_n \in \mathcal{S}$ for each positive integer n. Similarly, $\mathcal{S}$ is

closed under countable union if $S_1, S_2, S_3, \ldots \in \mathcal{S} \;\Rightarrow\; \bigcup_{j=1}^{\infty} S_j \in \mathcal{S}$;

closed under arbitrary union if $\{S_\lambda : \lambda \in \Lambda\} \subseteq \mathcal{S} \;\Rightarrow\; \bigcup_{\lambda \in \Lambda} S_\lambda \in \mathcal{S}$.

We define closures under intersections analogously. A collection $\mathcal{S}$ is **closed under complementation** if $S \in \mathcal{S} \;\Rightarrow\; \complement S \in \mathcal{S}$. These "closures" are special cases of *Moore closures*; see 4.4.d.

Functions and Products of Sets

1.31. A **function** (or *map* or *mapping* or *operator* or *operation*) from a set X into a set Y is a *rule* that assigns to each **argument** $x \in X$ a unique **value** $f(x) \in Y$. This is not really a definition; we rely on the reader's intuition about what a "rule" is. However, in 1.36 we give an alternate definition that is less intuitive but more precise, in terms of subsets of products of sets.

We write $f : X \to Y$ to abbreviate the statement that f is a function from X into Y. The set X is the **domain** of f, often abbreviated Domain(f) or Dom(f). We say f is a *function on* X, or f is *defined on* X.

The set Y is the **codomain** of f. It should not be confused with the **range** of f, which is the set $\{f(x) : x \in X\}$, often abbreviated Range(f) or Ran(f). The function f is called **surjective** (or *onto*, or a *surjection*) if the range is equal to the codomain. (See also 2.7.)

The distinction between range and codomain may confuse some beginners. The range is a very specific set — it is the set of all the values taken on by the function. The codomain of the function is a somewhat arbitrary or nominal set; the codomain is any convenient set large enough to contain the range of the function. We may choose to describe the function in terms of the codomain instead of the range because we do not actually know the range. Another reason is so that we can compare several functions that have different ranges. For instance, the functions $f(x) = x^2$ and $g(x) = x^3$ (both defined for real numbers x) have different ranges, but they can both be viewed as having codomain $\mathbb{R}$, thus permitting us to ask such questions as: Is $f(x)$ always less than $g(x)$?

The concept of "function" evolved over several centuries; some earlier definitions are listed by Rüthing [1984].

1.32. An **ordered pair** is an ordered list (y_1, y_2) consisting of two mathematical objects y_1, y_2, which may or may not be different from each other. The ordered pair is then a new mathematical object. For some purposes in set theory (discussed in 1.46), it is convenient to view an ordered pair as a special kind of set; the ordered pair (y, z) can be represented by the set $\{\{y\}, \{y, z\}\}$. This representation (which is not used in most branches of mathematics outside of set theory) preserves the essential property of ordered pairs: Two ordered pairs (y_1, y_2) and (z_1, z_2) are considered to be *equal* (i.e., to be representations of the same mathematical object) if and only if $y_1 = z_1$ and $y_2 = z_2$.

More generally, for any nonnegative integer n, an **ordered n-tuple** (or *finite sequence*, with *length* n) is a list of n objects — i.e., an object expressed in any of the forms

$$(y_1, y_2, \ldots, y_n) \quad = \quad (y_j \ : \ j = 1, 2, \ldots, n) \quad = \quad (y_j)_{j=1}^{n} \quad = \quad (y_j),$$

where $y_1, y_2, \ldots, y_n$ are any mathematical objects. The notation (y_j) can only be used if the value of n is understood. There are several ways to represent n-tuples in terms of other objects (and we usually do not need to concern ourselves about which of these representations is being used). One representation is as an iteration of ordered pairs: $(x_1, x_2, \ldots, x_n) = ((x_1, x_2, \ldots, x_{n-1}), x_n)$.

Another way to view an ordered n-tuple is as a function with domain $\{1, 2, 3, \ldots, n\}$; its value at the argument j is y_j. Thus, if we represent the same function by f, we will have $f(j) = y_j$. In particular, an ordered pair may be viewed as a function with domain $\{1, 2\}$.

An ordered n-tuple can also be written as a column, as in $\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$. This notation is used chiefly when the y_j's are numbers, but it may be used for other y_j's as well.

A **sequence** (or *infinite sequence*) is an object of the form

$$(y_1, y_2, y_3, \ldots) \quad = \quad (y_j : j \in \mathbb{N}) \quad = \quad (y_j)_{j=1}^{\infty} \quad = \quad (y_j).$$

A sequence is a function with domain $\mathbb{N}$. Again, the notation (y_j) can only be used if the choice of the domain is understood. A sequence (y_k) is a **subsequence** of a sequence (x_j) if $y_k = x_{\varphi_k}$ for some positive integers $\varphi_1 < \varphi_2 < \varphi_3 < \cdots$. For instance, $(1, 3, 9, 27, 81, \ldots)$ is a subsequence of $(1, 3, 5, 7, 9, \ldots)$.

For finite and infinite sequences, it is understood that the order of the objects is being noted. Thus, two finite or infinite sequences (x_j) and (y_j) are considered to be equal if and only if they have the same length and satisfy $x_j = y_j$ for all j. The unordered sets $\{1, 2\}$ and $\{2, 1\}$ are considered to be the same, but the ordered pairs $(1, 2)$ and $(2, 1)$ are different.

We now generalize. Any function with domain Λ may be viewed as a "Λ-tuple"

$$(y_\lambda : \lambda \in \Lambda) \quad = \quad (y_\lambda)_{\lambda \in \Lambda} \quad = \quad (y_\lambda),$$

where y_λ is the value of the function at the argument λ. Again, the notation (y_λ) can only be used if the choice of Λ is understood. The notation of Λ-tuples is used mainly when Λ is equipped with some sort of ordering (see especially 7.6), but that is not a requirement. The object y_λ is called the λth **component** (or **element** or **entry** or **value**) of the Λ-tuple. In particular, in the ordered pair (x, y), the objects x and y are the first and second component, respectively. We may sometimes refer to (y_λ) as a **parametrized set**; then Λ is the **parameter set**.

We may occasionally write $\Lambda = \{\alpha, \beta, \gamma, \ldots\}$ and $(y_\lambda) = (y_\alpha, y_\beta, y_\gamma, \ldots)$, where the indices $\alpha, \beta, \gamma, \ldots$ are intended to represent typical elements of Λ. If interpreted properly, this notation is occasionally useful, because it emphasizes the conceptual similarity between n-tuples and more general functions. However, this notation is not standard and should only be used with caution. It may give some readers the impression that the parameter set Λ is a sequence, but that meaning is not intended.

Throughout this book, we use braces $\{\ \}$ for unordered sets and parentheses $(\ )$ for sequences or other parametrized sets. Note that the mathematical literature does not always observe this notational convention.

1.33. The **product** of n sets S_1, S_2, ..., S_n is the set of ordered n-tuples

$$\prod_{j=1}^{n} S_j \quad = \quad S_1 \times S_2 \times \cdots \times S_n \quad = \quad \{(x_1, x_2, \ldots, x_n) \ : \ x_j \in S_j \text{ for all } j\}.$$

The product of a sequence of sets $S_1, S_2, S_3, \ldots$ is the set of sequences

$$\prod_{j=1}^{\infty} S_j \quad = \quad S_1 \times S_2 \times S_3 \times \cdots \quad = \quad \{(x_1, x_2, x_3, \ldots) \ : \ x_j \in S_j \text{ for all } j\}.$$

The product of an arbitrary collection of sets $(S_\lambda : \lambda \in \Lambda)$ is the set

$$\prod_{\lambda \in \Lambda} S_\lambda \quad = \quad \{(y_\lambda)_{\lambda \in \Lambda} \ : \ y_\lambda \in S_\lambda \text{ for all } \lambda\}.$$

In other words, it is the collection of all functions $f : \Lambda \to \bigcup_{\lambda \in \Lambda} S_\lambda$ that satisfy $f(\lambda) \in S_\lambda$ for each $\lambda \in \Lambda$. This collection of functions may also be viewed as a collection of Λ-tuples; if we write $\Lambda = \{\alpha, \beta, \gamma, \ldots\}$, then the product $\prod_{\lambda \in \Lambda} S_\lambda$ may be written as

$$S_\alpha \times S_\beta \times S_\gamma \times \cdots \quad = \quad \{(x_\alpha, x_\beta, x_\gamma, \ldots) \ : \ x_\alpha \in S_\alpha,\ x_\beta \in S_\beta,\ x_\gamma \in S_\gamma,\ \ldots\}.$$

This representation should only be used with caution, as noted in 1.32.

We emphasize that an ordering on Λ may or not be present; it may be stated explicitly or may be implied in a particular context; it may be of great or small importance, in any particular context. The set $A \times B$ is not the same as $B \times A$, but for some purposes $A \times B$ is "essentially the same" as $B \times A$, and a rearrangement of ordering may be clear in some contexts. It is convenient to be able to say that

$$\prod_{\lambda \in \Lambda} S_\lambda \qquad \text{is "essentially the same" as} \qquad \left(\prod_{\lambda \in \Lambda_1} S_\lambda\right) \times \left(\prod_{\lambda \in \Lambda_2} S_\lambda\right)$$

for some purposes, whenever $\{\Lambda_1, \Lambda_2\}$ is a partition of Λ, but this equation is only valid after an obvious rearrangement of the ordering of Λ and removal of some parentheses.

1.34. Associated with any product of sets $P = \prod_{\lambda \in \Lambda} S_\lambda$ is another collection of mappings. For each $\lambda \in \Lambda$, the λth **coordinate projection** is the surjective mapping $\pi_\lambda : P \to S_\lambda$ given by

$$\pi_\lambda(f) \;=\; f(\lambda) \qquad \text{or} \qquad \pi_\lambda(x_\alpha, x_\beta, x_\gamma, \ldots) \;=\; x_\lambda$$

depending on whether we view $\prod_{\lambda \in \Lambda} S_\lambda$ as a collection of functions f with domain Λ or as a collection of Λ-tuples.

Notations for this mapping vary throughout the literature; the notation π_λ will be used for coordinate projections throughout most of this book. If Λ is the set $\{1, 2, \ldots, n\}$ or $\mathbb{N}$, then the jth coordinate projection will be denoted by π_j; it is the map that takes the n-tuple $(x_1, x_2, \ldots, x_n)$ or the sequence $(x_1, x_2, x_3, \ldots)$ to x_j.

The term "projection" has other meanings; see for instance 8.12 and 22.45.

1.35. When all the S_λ's are equal to one set S, then their product $\prod_{\lambda \in \Lambda} S_\lambda$ may also be written as

$$S^\Lambda \;=\; \{f \;:\; f \text{ is a function from } \Lambda \text{ into } S\}.$$

It is called the **Λth power of S**. It is related to, but should not be confused with, the *power set* of S; see 1.21 and 2.20.k.

If Λ contains just n elements for some positive integer n, then S^Λ may also be written as S^n.

1.36. The **graph** of a function $f : X \to Y$ is the set of ordered pairs

$$\mathrm{Graph}(f) \quad = \quad \mathrm{Gr}(f) \quad = \quad \{(x, f(x)) : x \in X\} \quad \subseteq \quad X \times Y.$$

We sometimes identify a function with its graph; with this viewpoint, a **function** from X into Y is simply a subset of $X \times Y$ with the property that each element of X is the first component of one and only one of the ordered pairs that are members of $X \times Y$.

Thus, a function may be viewed as a set of ordered pairs. On the other hand, we noted in 1.32 that an ordered pair may be viewed as a function with domain $\{1, 2\}$. To avoid confusion or circular reasoning, generally we do not adopt both of these viewpoints simultaneously.

1.37. *Degenerate examples.* The **empty function** is the rule that makes no assignments; its domain and graph are both the empty set.

If S is any set, then $S^\varnothing = \{\varnothing\}$, since the *only* rule assigning to each element of $\varnothing$ a corresponding element of S is the empty function. The set $S^\varnothing$ is also denoted S^0.

If S is any *nonempty* set, then $\varnothing^S = \varnothing$, since there is *no* rule that assigns to each element of S a corresponding element of $\varnothing$.

We emphasize that $\varnothing \neq \{\varnothing\}$.

1.38. *An example using products.* Any intersection of unions can be expressed as a union of intersections, and conversely: For any sets C and A_γ ($\gamma \in C$) and $S_{\gamma,\alpha}$ ($\alpha \in A_\gamma$), we have

$$\bigcap_{\gamma \in C} \; \bigcup_{\alpha \in A_\gamma} S_{\gamma,\alpha} \quad = \quad \bigcup_{f \in \prod_{\gamma \in C} A_\gamma} \; \bigcap_{\gamma \in C} S_{\gamma, f(\gamma)}$$

and

$$\bigcup_{\gamma \in C} \bigcap_{\alpha \in A_\gamma} S_{\gamma,\alpha} = \bigcap_{f \in \Pi_{\gamma \in C} A_\gamma} \bigcup_{\gamma \in C} S_{\gamma, f(\gamma)}.$$

Note that if C and the A_γ's are finite sets, then $\prod_{\gamma \in C} A_\gamma$ is also a finite set; hence

> any finite intersection of finite unions can be represented as a finite union of finite intersections, and conversely.

An analogous statement is not true for countable intersections and unions; see the remarks in 13.10.

1.39. *More notations for functions.*

a. The symbol $\mapsto$ stands for "**maps to,**" and is sometimes used to indicate by individual values the rule that defines a function. For instance, the function defined by $f(x) = x^2$ could also be written as $x \mapsto x^2$.

b. In some cases the rule defining a function f is given explicitly (as by $x \mapsto x^2$). In other cases the rule is only given implicitly or indirectly, and it may be necessary to verify that the function is **well defined**, i.e., that the description of f does in fact determine a function. In some cases we do not specify the rule, but simply postulate its existence; this is the effect of the *Axiom of Choice* — see (AC3) in 6.12.

c. A function f may also be denoted by the expression "$f(\cdot)$," with the raised dot showing where the argument should be inserted. This notation is useful in more complicated expressions. For instance, if Γ is a mapping from $X \times Y$ into Z, then

- for each fixed $x \in X$ we obtain a mapping $\varphi_x = \Gamma(x, \cdot) : Y \to Z$, defined by $y \mapsto \Gamma(x, y)$, and
- for each fixed $y \in Y$ we obtain a mapping $\psi_y = \Gamma(\cdot, y) : X \to Z$, defined by $x \mapsto \Gamma(x, y)$.

Thus the one function Γ determines two families of functions, $\{\varphi_x : x \in X\}$ and $\{\psi_y : y \in Y\}$. The functions in one family have the other family as their domain. We may say that these two families are **dual** to one another since each family determines the other.

It should be noted that the different functions $\varphi_x, \varphi_{x'}$ are not necessarily *distinct*; it is possible for two different points x, x' in X to have the same action on Y, and yet differ in some other respect that is not under consideration. Thus the mapping $x \mapsto \varphi_x$, from X to Z^Y, is not necessarily injective.

When two families of functions are dual to each other, notations such as $\langle\ ,\ \rangle$ or $(\ ,\ )$ are often used. Thus, $\Gamma(x, y)$ might be written as $\langle \varphi_x, \psi_y \rangle$, or perhaps as $\langle x, y \rangle$.

d. Let f be a function. Some other notations for $f(x)$ are

- f_x — used as in 1.32, especially when $X = \mathbb{N}$. However, be aware that f_x has other meanings as well; for instance, it can mean $\partial f / \partial x$, the partial derivative of f with respect to x. The reader must interpret "f_x" from the context.

- fx — used especially when f is *linear* (see Chapter 11). However, this simple juxtaposition of symbols also often means composition or multiplication. Again, the reader must interpret "fx" from the context.

1.40. A **binary operation** on a set X is a mapping from $X \times X$ into X. It is often written in the form $(x, y) \mapsto x \Box y$, for some symbol $\Box$. Familiar examples are addition ($+$) and multiplication ($\cdot$) on $\mathbb{R}$, and intersection ($\cap$), union ($\cup$), and symmetric difference ($\triangle$) on $\mathcal{P}(\Omega)$ for any set Ω.

A binary operation $\Box$ is **associative** if

$$x \Box (y \Box z) \;=\; (x \Box y) \Box z \qquad \text{for all } x, y, z \in X.$$

When this condition is satisfied, then parentheses are not needed; both sides of the equation above can be represented more simply as $x \Box y \Box z$. By repeated uses of this rule, we find that parentheses are not needed in an expression such as $x_1 \Box x_2 \Box \cdots \Box x_n$ for any positive integer n.

A binary operation $\Box$ is **commutative**, or **Abelian**, if

$$x \Box y \;=\; y \Box x \qquad \text{for all } x, y \in X.$$

If the operation $\Box$ is both commutative and associative, then the value of an expression such as $x_1 \Box x_2 \Box \cdots \Box x_n$ does not depend on the order of the x_j's.

In some instances, when the meaning is clear, the symbol $\Box$ may be omitted altogether — i.e., a binary operation is indicated by juxtaposition of the arguments, thus: $(x, y) \mapsto xy$. A function written this way is often called *multiplication*, although its behavior may differ significantly from the behavior of multiplication of real numbers — e.g., it need not be commutative; see for instance 2.3.

Another symbol commonly used for a binary operation is "+," called *addition*. Algebraists occasionally use + for a noncommutative operation, but analysts generally do not. *In this book addition* ($+$) *will always denote a commutative operation.*

When addition or subtraction are used for binary operations, they customarily are applied *last*, i.e., after any other operations in the expression. For instance, $ax + b - c/d$ is generally interpreted to mean $(ax) + b - (c/d)$.

Let $\Box$ be a binary operation on a set X, let S be a set, and let $f : S \times X \to X$ be some function. We say that f **distributes** (or is *distributive*) over $\Box$ if

$$f(s, x \Box y) \;=\; f(s, x) \Box f(s, y) \qquad \text{for all } s \in S \text{ and } x, y \in X.$$

A familiar example is that, in ordinary arithmetic of real numbers, multiplication distributes over addition — that is, $s(x + y) = (sx) + (sy)$. Another example was given in 1.29.a; further examples will be given in later chapters.

1.41. Let X and Λ be sets. By a **Λ-ary operation** on X we shall mean any mapping from X^Λ into X. Such a function may be written as $f = f(x_\alpha, x_\beta, x_\gamma, \ldots)$, where $\Lambda = \{\alpha, \beta, \gamma, \ldots\}$. We may consider the point $(x_\alpha, x_\beta, x_\gamma, \ldots) \in X^\Lambda$ as the single argument of f, but alternatively we may view f as having many arguments $x_\alpha, x_\beta, x_\gamma, \ldots \in X$. We may refer to the αth argument of f, the βth argument of f, etc. The set Λ (or the number of elements in Λ, if it is finite) is called the **arity** of the operation.

A binary operation (defined in 1.40) is the same thing as a 2-ary operation. We note a few other important cases:

- When Λ is a finite set, with n elements, then a Λ-ary operation is also called an ***n*-ary operation**; it may be viewed as a mapping from $X \times X \times \cdots \times X$ (n factors) into X. Typically, it is written in the form $y = f(x_1, x_2, \ldots, x_n)$. We may consider it to be a function with n arguments in X — the first argument, the second argument, etc. Operations that are n-ary for finite n are also called **finitary** operations.

- A 1-ary operation on X is a mapping from X into X. It is also called a **unary operation**. Typical examples are $x \mapsto -x$ (for numbers) or $S \mapsto \complement S$ (for subsets of a given set).

- It is occasionally useful to view a specially selected member of a set X (such as the number 0, in $\mathbb{R}$) as an "operation." Since $X^0 = X^\varnothing = \{\varnothing\}$ is a singleton, a 0-ary operation on X is a function from a singleton into X — i.e., it is a constant member of X. In effect, it is a function with 0 arguments. It is also called a **nullary operation.**

- The set Λ does not have to be finite. For instance, an $\mathbb{N}$-ary operation on X is a mapping from $X^{\mathbb{N}}$ into X — that is, a mapping that takes each sequence of elements of X to an element of X. Let $X = \mathcal{P}(\Omega)$ for some set Ω; then
$$(S_1, S_2, S_3, \ldots) \mapsto \bigcup_{j=1}^{\infty} S_j \qquad \text{and} \qquad (S_1, S_2, S_3, \ldots) \mapsto \bigcap_{j=1}^{\infty} S_j$$
are two $\mathbb{N}$-ary operations on X that are important in measure theory.

ZF Set Theory

1.42. *Remark.* This subchapter can be postponed; it will not be needed until much later in the book.

1.43. *How big can sets be?* As we remarked in 1.11, we will attempt to avoid self-referencing statements. We must also avoid certain kinds of self-referencing definitions of sets. Defining sets in a self-referencing way can lead to sets that are too "big" to be meaningful. This is evident in the following paradox.

> **Russell's Paradox.** It seems that some sets are members of themselves. For instance, *the collection of all sets that can be described in fewer than 100 words of English* is a set that has just been so described; and *the collection of all sets that are mentioned in this book* is a set that has just been mentioned. Let us call such sets "self-inclusive."

On the other hand, some sets do not include themselves. For instance, *the collection of all pages in this book* is a set of pages; it is not a page. We shall call such sets "non-self-inclusive."

But now what about *the collection of all non-self-inclusive sets*? Is it a member of itself? It is if it isn't, and it isn't if it is — a contradiction either way.

1.44. *A few ways to avoid paradoxical sets.* In 1.13 we defined a *set* to be "a collection of objects," but that definition leads to Russell's Paradox. A slightly better definition is: A set is a collection of *already fixed* objects — i.e., the objects must already be fixed before the collection is formed (Scott [1974], Shoenfield [1977]). This precludes self-referencing. With this definition, the collection of all sets is not a set. Unfortunately, this "definition" is not very precise; we shall give it some precision in 5.53.

A safer but more restrictive method for avoiding excessively large collections is by specifying in advance some manageable collection V of sets, and then prohibiting the use of any sets outside that collection. The collection V may be smaller than what one ordinarily thinks of as "all sets," but it may still be large enough for all the applications one is interested in. Thus, without substantial inconvenience, one may replace the term "set" with "member of V." The collection V is then called the **universe** (or *universal set*, if it is a set). In many contexts in mathematics, a universe is not specified explicitly or even mentioned; one simply assumes that the universe being used is large enough for one's applications. In other contexts it is useful to discuss the choice of the universe and even to specify it explicitly; see for instance 1.48, 5.53, 5.54, and 9.39. One small, easily manageable universe is the "superstructure" over $\mathbb{R}$, which is commonly used in nonstandard analysis; it is described in 14.65.

The most commonly used universe is the one described by the axioms of conventional set theory, **ZF + AC**. This stands for *Zermelo-Fraenkel set theory, as modified by Skolem, plus the Axiom of Choice.* We shall list the axioms of ZF in 1.47; we shall introduce AC in 6.12. Conventional set theory does not permit Russell's Paradox, and it *apparently* does not lead to any other contradictions either. But we can only say "apparently;" the uncertainty of this is discussed further in 14.71 and the sections thereafter.

1.45. Although we shall only apply the term "set" to members of our universe V, it is grammatically convenient to be able to discuss other, much bigger collections, at least informally — e.g., to discuss as a "collection" those sets that satisfy a certain property. Any collection of objects will be called a **class**, but we shall distinguish between two types of classes:

- A **set** is a member of the universe V. Intuitively, it is a class of ordinary size.
- A **proper class** is a collection that is not a member of V. Intuitively, it is a much bigger class, one that is too big for us to safely apply to it the rules for sets.

A set can be a member of something; a proper class cannot. We refuse to consider proper classes as members of anything.

Thus, V is the collection of all sets. It is a proper class, but not a set. We cannot form the "set of all sets" or the "class of all classes," so Russell's Paradox does not arise.

In set theory, it is easy to give examples of proper classes — e.g., the class of all singletons or the class of all ordinals (investigated later in this book). Outside of set theory, examples are harder to produce. The class of all linear spaces and the class of all topological spaces are proper classes, but these examples are somewhat contrived — for instance, a theorem about topological spaces usually only involves a few topological spaces at a time; it can be formulated so that it does not require us to simultaneously consider *all* topological spaces. However, occasionally a proper class really is needed outside set theory. For instance, a convergence structure on a set X (see Chapter 7) can be described by a "limit" function defined on the collection of all proper filters on X, or defined on the collection of all nets on X. The net approach has some intuitive advantages — nets are very much like sequences — but the net approach must be used with some caution: The collection of all nets on X is a proper class, not a set.

We will need proper classes only a few times in this book, so we shall not develop a systematic theory for them. We shall simply use them in an ad hoc fashion, in ways that obviously make sense; it should be clear in each context that we are avoiding self-referencing arguments such as Russell's Paradox. The usual operations of sets make sense for classes sometimes, but not always, and it is sometimes convenient to apply the terminology of sets to a few classes. For instance, it usually makes sense to consider the intersection of two classes. By a **function of classes** we shall mean a mapping $f : \mathcal{M} \to \mathcal{N}$ from one class into another — i.e., a rule that assigns to each $M \in \mathcal{M}$ some particular $f(M) \in \mathcal{N}$; see 1.50, 5.51, 5.53, and 6.23. Generally the graph of such a function f is not a set of ordered pairs, but rather a class of ordered pairs.

1.46. *What are sets made of?* A typical set is $\{0, 1, 2, -5/3, \pi\}$; this set has five elements. For most purposes in "ordinary" mathematics — i.e., outside of set theory and logic — we do not think of the individual numbers $0, 1, 2, -5/3, \pi$ as sets that may contain other objects. Instead we think of these numbers as *indivisible*; for this reason we may refer to them as **atoms** (or *urelements* or *individuals* or *primitive objects*). Likewise, we generally do not think of an ordered pair $(-5/3, \pi)$ as a set.

A set need not contain just atoms — it may contain other sets for its members. For instance, $\{\varnothing, \{1, 2\}, \{-5/3, \pi\}\}$ is a set whose members are sets. Such sets arise naturally in analysis; for instance, a topology or a σ-algebra on a set X is a collection of subsets of X (see 5.12 and 5.25). The set of all topologies on X is a set of sets of sets. "Ordinary" mathematicians — i.e., those not involved in logic or set theory — seldom need to go to any levels deeper than this. However, logicians and set theorists quite commonly have sets nested arbitrarily deep; for instance, consider the ordinals $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \ldots$ described in 5.44.

For most purposes in most branches of mathematics, it does not matter whether "3" is an indivisible object or a set containing three objects. What matters is how we *use* 3. We may define "3" in any way we wish, provided we define "+" so that $3 + 3 = 6$. For most mathematicians, it is simpler to view "3" as an indivisible object, and our language reflects that viewpoint. The **elements** of a set are also called its **points**, whether those elements are known to be indivisible or not. The points of the set $\{0, \{0, 1\}\}$ are 0 and $\{0, 1\}$.

Although atoms seem natural to most mathematicians, they are not really needed, and in some studies of set theory it is customary to dispense with atoms altogether. All familiar objects can be represented *solely in terms of sets*, without any other basic building blocks. Some details of this representation will be worked out in later chapters, but we can outline it now:

- The nonnegative integers can be built up from the empty set, as in 5.44.
- An ordered pair (x, y) can be represented in terms of sets and an ordered n-tuple in terms of ordered pairs, as discussed in 1.32.
- Functions and relations can be represented as sets of ordered pairs; see 3.2 and 1.36.
- A product of sets is a set of functions, and a finite or infinite sequence may be viewed as a function; see 1.33.
- A negative integer can be represented by an ordered pair involving a positive integer.
- Rational numbers may be represented in terms of pairs of integers (see 8.22).
- Real numbers may be represented in terms of sets of rationals (see 10.15.d and the constructions used in that proof).

Thus, all *familiar* objects can be represented as sets.

However, to assert that

all objects can be represented as sets

is to make an additional assumption about our universe of "objects." This is one of the assumptions of conventional set theory: All of its "objects" are sets, and all the members of sets are sets, so conventional set theory is atomless. If we omit this assumption and permit the existence of objects that cannot be represented as sets, we obtain a slightly weaker system of axioms, known as **set theory with atoms** (or *set theory with urelements*); it is occasionally useful in model theory.

This metaphor may be helpful: Atomless (conventional) set theory is like a great collection of transparent bags, some of which are empty and some of which contain other bags — which may in turn contain other bags, and so on; there is nothing in the system except bags, and nothing to distinguish between the bags except the different combinations of bags-within-bags that they contain. In set theory with atoms, the bags may also contain beads. The beads do not contain anything, but can be distinguished by their markings.

Whether we view certain objects — e.g., the nonnegative integers — as sets or as atoms depends on our viewpoint; different branches of mathematics find different viewpoints advantageous. It is sometimes convenient to label the real numbers or other familiar objects as "atoms" and treat them as indivisible, even though those objects could instead be represented as sets. Then the assumptions that underlie our work are really the assumptions of conventional, atomless set theory, although "atoms" may enter into our terminology. An example of this is given in 14.65. (In our metaphor, there is still nothing in the system but bags, but we seal some of the bags shut and mark them on the outside, and agree to treat them as beads.)

1.47. Zermelo-Fraenkel Set Theory. Most of the axioms of ZF are just formal statements that correspond to our informal intuition about sets. (An exception is the Axiom of Regularity, which will appear at the end of our list of axioms; it is somewhat nonintuitive and nonconstructive.) In the next few paragraphs we shall list the axioms of set theory to give a general impression, but later in this book we shall usually rely on the reader's intuition rather than on the list of axioms. Further discussions of ZF set theory can be found in books on set theory; for elementary treatments see, for instance, Halmos [1960] or Stoll [1963].

In some later chapters we shall briefly consider modifications of conventional set theory. For that reason, the reader is encouraged to put aside the usual intuitive meaning of "set" and view the axioms below as a self-contained theory that does not refer to anything familiar. To help put aside familiar intuitive notions of sets, some readers may find it helpful to glance ahead to the peculiar example in 1.48.

We assume that we are given some collection of objects, which we call "sets." We assume that some pairs of these "sets" are related by a relationship, called "is a member of," denoted $\in$. That is, when S and T are "sets," then the statement "$S \in T$" is either true or false. We assume that this collection of "sets" and this relationship "$\in$" satisfy certain axioms, listed below. We may then explore the consequences of those axioms.

The statement $A \subseteq B$ is an abbreviation for the statement that every member of A is also a member of B. That is, $A \subseteq B$ is defined to mean $x \in A \;\Rightarrow\; x \in B$.

There are at least two different ways to deal with equality of sets. Some books take equality ($=$) to be a logical symbol with its customary properties, as listed in 14.27.a. Two objects are equal if and only if they are not distinct; thus equals can be substituted for equals. In such books, equality ($=$) and membership ($\in$) are already meaningful before we get to the relation between them. The relation between them is taken to be the first axiom:

> *Axiom (called "Extensionality" in some books).* Two sets are the same if and only if they have the same members. That is, $(A = B) \iff ((x \in A) \iff (x \in B))$.

Other books follow a slightly different approach, which we shall follow here. Equality of sets is taken, not as a primitive notion, but as a defined notion. We *define* two sets to be equal when they have the same members. Thus, $A = B$ means $x \in A \iff x \in B$. With this definition, we cannot automatically assume that equality ($=$) has all of its usual properties; we must not be misled by the fact that the symbol we are using ($=$) is a familiar symbol. From our definition of equality of sets (and our understanding of the logical symbol "$\Leftrightarrow$"), it is easy to prove that (i) $A = A$, (ii) $A = B \;\Rightarrow\; B = A$, and (iii) $A = B$ and $B = C$ imply $A = C$. However, the last two axioms in 14.27.a — which state that "equal" quantities can be substituted for one another in any expression — do *not* follow directly from our definition of equality of sets, and so they will require some assumption about sets. One particular instance of the substitution principle is:

> **Axiom of Extensionality.** If $A = B$ and $A \in C$, then $B \in C$.

It turns out that this is all we need — the general substitution principle (described in the last two axioms of 14.27.a) can be proved from our Axiom of Extensionality, by induction on the length of the formulas involved. We shall omit the proof; it can be found in Takeuti and Zaring [1982].

The next few axioms are formal restatements of some of our basic rules about permitted methods for forming sets, discussed informally earlier in this chapter.

> **Axiom of the Power Set.** If S is a set, then there exists a set whose members consist precisely of the subsets of S. That set is denoted by $\mathcal{P}(S)$.

> **Axiom of Replacement.** Let f be a function defined on a set X — that is, a rule that assigns to each element $x \in X$ some set $f(x)$. Then $\{f(x) : x \in X\}$ is a set.

> **Axiom of Comprehension (or Separation).** If X is a set, and $P(x)$ is a property that is true or false for each $x \in X$, then $\{x \in X : P(x) \text{ is true}\}$ is a set. That is, there exists a set Y such that $x \in Y \iff [x \in X$ and $P(x)$ is true]. This can also be stated as: The intersection of a set and a class is a set.

Though the Axiom of Replacement and the Axiom of Comprehension are usually presented as separate axioms, some mathematicians formulate their language in a slightly different fashion, so that one of these axioms becomes a consequence of the other; see Bell and Machover [1977]. Actually, we have skipped over some of the complexity of the last two axioms. We shall not explain at this point precisely what is meant by "function" or "property;" those terms mean approximately what one would expect them to mean, but for axiomatic set theory the function f or property P must be expressed in a formal first-order language. (First-order languages are introduced in 14.15 and thereafter; see also the related comments in 14.67.) This actually gives infinite *schemes* of axioms — one for each function f and one for each property P.

> **Axiom of the Empty Set.** There exists a set that has no members; it is denoted $\varnothing$ and called the *empty set.*

We could replace this axiom with the assumption that there exists *some* set, for then the existence of the empty set follows from the Axiom of Comprehension by taking $P(x)$ to be the property $x \neq x$. Note that the empty set is unique, since two sets with the same members are equal; thus we are justified in introducing a symbol " $\varnothing$ " for it.

> **Axiom of Unions.** If S is a set, then there exists a set $\mathrm{Un}(S)$ whose members are precisely the same as the members of the members of S. That is, $\mathrm{Un}(S)$ has the property that $[A \in \mathrm{Un}(S)] \iff$ [there exists some $B \in S$ with $A \in B$].

We call $\mathrm{Un}(S)$ the *union* of the members of S — or more briefly, the union of S. To understand this axiom, keep in mind that all the elements of S are *sets.* Here are a few examples: If $S = \{A\}$ is a singleton, then $\mathrm{Un}(S) = A$; if $S = \{A_1, A_2, A_3, \ldots\}$, then $\mathrm{Un}(S) = A_1 \cup A_2 \cup A_3 \cup \cdots$. If we use either Zermelo's or von Neumann's definition of the integers (see 1.16), then $\mathrm{Un}(n) = n - 1$ for positive integers n, and $\mathrm{Un}(0) = 0$ also.

> **Axiom of Pairing.** If S and T are sets, then there exists a set whose only members are S and T; it is denoted $\{S, T\}$.

Most books on set theory present the Axiom of Pairing as a separate axiom, but in fact we can make it a consequence of the previous axioms, as follows: First, following von Neumann, define 0 to be the empty set; define 1 to be the power set of 0; define 2 to be the power set of 1. Thus

$$0 = \varnothing, \qquad 1 = \{\varnothing\}, \qquad 2 = \{\varnothing, \{\varnothing\}\}.$$

Now "2" is the name of a set that contains two elements; those elements are "0" and "1." Define a function f on 2 by taking $f(0) = S$ and $f(1) = T$. By the Axiom of Replacement, $\{S, T\}$ is a set.

The Axiom of Pairing can be used repeatedly, to define the remaining nonnegative integers 3, 4, 5, ..., with Zermelo's definition or von Neumann's definition. However, this procedure only yields finitely many nonnegative integers, since the ellipsis (...) of informal mathematics is not permitted in formal set theory. To get all of the nonnegative integers at once — i.e., to get the set $\omega = \{0, 1, 2, 3, \ldots\}$ — requires something more. The set of nonnegative integers can be constructed using

> **Axiom of Infinity.** There exists a set S with $\varnothing \in S$, and such that $A \in S \Rightarrow \{A\} \in S$.

The set S given in the axiom is not quite the set of integers that we're after. However, we can construct $\omega = \mathbb{N} \cup \{0\}$ this way: Call a set S "infinity-like" if it satisfies $\varnothing \in S$ and also satisfies $A \in S \Rightarrow \{A\} \in S$. The Axiom of Infinity guarantees the existence of at least one infinity-like set, S_0. Let

$$P(x) \quad = \quad \text{"for every } y\text{, if } y \text{ is infinity-like, then } x \in y\text{"}.$$

The Axiom of Comprehension guarantees that $S_1 = \{x \in S_0 : P(x)\}$ is a set. Clearly, S_1 is the intersection of all infinity-like sets. It can be shown that S_1 is the desired set ω; we omit the details.

The preceding axioms merely formalize the intuition about sets that we may have obtained from experience with finite sets. The one remaining axiom of ZF set theory is not just a formalization of our intuition, however:

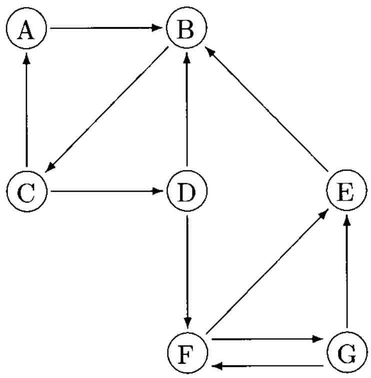
> **Axiom of Regularity (or Foundation, or Restriction).** If X is a nonempty set, then X has a member that does not meet X — i.e., there exists at least one set $A \in X$ that satisfies $A \cap X = \varnothing$.

The set A whose existence is postulated by the Axiom of Regularity is sometimes called an **$\in$-minimal element** of X (or more simply, a *minimal* element), for this reason: A is a member of X, and there does not exist another set B, also a member of X, that satisfies $B \in A$.

The Axiom of Regularity precludes the possibility of certain counterintuitive sets; see 1.49. It will be clear from the reformulations in 1.50 and 6.31 that the Axiom of Regularity is concerned with sets of sets of sets of sets of .... Since such deep nesting does not occur outside of set theory, the Axiom of Regularity has little effect on "ordinary" mathematics; it is merely a technical convenience that helps set theory work properly. It can be replaced with alternative axioms; see Aczel [1988] and Barwise and Etchemendy [1987].

1.48. *Pathological example.* Since ZF's axioms are mostly in agreement with our intuition, we would not come to understand those axioms better by looking at examples that satisfy the axioms. Instead, we shall now present a peculiar little universe that *violates* most of the axioms. This example is a modification of one by Krivine [1971].

There are seven "sets" in our peculiar little universe, denoted by A, B, C, D, E, F, G. The membership relation $\in$ is represented by an arrow in the diagram below — we say $S \in T$ if there is an arrow from S to T. Thus, the only memberships in our miniature universe are those listed beside the diagram.



memberships:		
C	$\in$	A
A, D, E	$\in$	B
B	$\in$	C
C	$\in$	D
F, G	$\in$	E
D, G	$\in$	F
F	$\in$	G

As usual, we define $S \subseteq T$ to mean that $X \in S \;\Rightarrow\; X \in T$. With this definition, the only subset relations are:

$$A \subseteq D, \qquad D \subseteq A, \qquad G \subseteq E,$$

plus the fact that each set is a subset of itself. Most of the axioms of ZF are violated:

- "Equality" doesn't mean what we would expect — the sets A and D are distinct, yet they have the same members, so they are "equal."
- The Axiom of Extensionality is violated: The sets A and D are "equal," yet they are not members of the same sets — we have $D \in F$ but not $A \in F$.
- The Axiom of the Empty Set is violated: Each of our "sets" has at least one member.
- The Axiom of Pairing is violated: There is no set whose only members are C and G.
- The Axiom of the Power Set is violated: The sets that are subsets of D are the sets A and D, yet there is no set whose only members are A and D.
- The Axiom of Unions is violated: The members of B are A, D, E; the sets that are members of A, D, or E are the sets C, F, G; yet there is no set whose only members are C, F, G.

- The Axiom of Comprehension is violated: Let $P(X)$ be the statement "$C \in X$." Then the class $\{S \in B : P(S)\}$ is not a set.
- The Axiom of Regularity is violated: The only members of E are F and G, yet neither of those sets is disjoint from E.
- The Axiom of Infinity, as stated in 1.47, only makes sense if we have already assumed the Axiom of the Empty Set. However, it is clear that our universe of seven "sets" does not yield an infinite set or an infinite collection of sets.

1.49. *Consequences of regularity.* In ZF set theory, none of the following can occur:

(i) $T \in T$ for some set T.

(ii) $T_n \in T_{n-1} \in T_{n-2} \in \cdots \in T_1 \in T_0 = T_n$ for some positive integer n and sets $T_1, T_2, \ldots, T_n$.

(iii) $\cdots \in T_3 \in T_2 \in T_1 \in T_0$ for some infinite sequence of sets $T_0, T_1, T_2, \ldots$.

Proof. Either of conditions (i) and (ii) implies (iii), since the sequence in (iii) may repeat itself. If $T_0, T_1, T_2, \ldots$ is a sequence as in (iii), let $X = \{T_0, T_1, T_2, \ldots\}$; by the Axiom of Regularity some member of X does not meet X — a contradiction.

1.50. (*Optional.*) We state without proof two more interesting consequences of the Axiom of Regularity.

> **Principle of Membership Induction.** Suppose $P(\cdot)$ is a property of sets that is $\in$-*inductive* — i.e., that has this property:
>
> Whenever $P(\cdot)$ is true for all members of a set X, then $P(X)$ is also true.
>
> Then $P(\cdot)$ is true for all sets.

> **Principle of Membership Recursion.** Let ρ be a function of classes, from $\{\text{sets}\} \times \{\text{sets}\}$ into $\{\text{sets}\}$. (That is, for any sets S, T suppose some set $\rho(S,T)$ is specified.) Then there exists a unique map $F : \{\text{sets}\} \to \{\text{sets}\}$ satisfying
>
> $$F(X) \quad = \quad \rho\Big(X,\ \{F(A) : A \in X\}\Big) \qquad \text{for each set } X.$$
>
> (In other words, it is possible to define $F(X)$ for all sets X, using a rule that specifies $F(X)$ in terms of the values of F on the members of X.)

We omit the proofs. Proofs can be found, for instance, in Johnstone [1987] and Kunen [1980]. Actually, Johnstone shows that the Principle of Membership Induction is *equivalent* to the Axiom of Regularity.

Chapter 2

Functions

2.1. Functions were defined in 1.31 and in 1.36; they will be studied in greater depth in this chapter.

Some Special Functions

2.2. *A few numerical functions.* We assume the reader has at least an informal familiarity with $\mathbb{R}$. A formal introduction to $\mathbb{R}$ is given in Chapter 10; we shall not use any of the deeper properties of $\mathbb{R}$ until then.

a. If $u_m, u_{m+1}, u_{m+2}, \ldots u_n$ are real numbers parametrized by consecutive integers $m, m+1, m+2, \ldots, n$, then their **sum** is

$$\sum_{j=m}^{n} u_j \quad = \quad u_m \;+\; u_{m+1} \;+\; u_{m+2} \;+\; \cdots \;+\; u_n\,,$$

and their **product** is

$$\prod_{j=m}^{n} u_j \quad = \quad u_m \;\; u_{m+1} \;\; u_{m+2} \;\; \cdots \;\; u_n\,.$$

(The letter j may be replaced by any other letter not already in use.) These notations will later be applied more generally — not just to sums and products of real numbers, but also to sums and products of complex numbers or members of any ring. The summation notation will also apply to sums of vectors or sums of members of any additive monoid.

b. Let X be some set. For each subset $S \subseteq X$ we define $1_S : X \to \{0,1\}$ by

$$1_S(x) \;=\; \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \in X \backslash S. \end{cases}$$

We shall call this the **characteristic function** of S. Our notation "1_S" does not reflect the choice of X, which must be understood from context. Note that

$$1_{S \cap T} \;=\; 1_S \cdot 1_T \;=\; \min\{1_S, 1_T\}.$$

Similarly, $\max\{1_S, 1_T\}$ is the characteristic function of $S \cup T$, and $1 - 1_S$ is the characteristic function of $\complement S$. *Caution*: Some mathematicians call 1_S the *indicator function* of S or denote it by χ_S, i_S, I_S, or other symbols. Another meaning for the term "indicator function" is given in 12.18.

c. The **sign function**, $\operatorname{sgn} : \mathbb{R} \to \{-1, 0, 1\}$, is defined by

$$\operatorname{sgn}(x) \;=\; \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases}$$

It may also be written as $\operatorname{sign}(x)$. (We mention it again in 15.20.)

d. For any set X, the **Kronecker delta** is the characteristic function of the diagonal set $\{(x, x) : x \in X\}$, considered as a subset of $X \times X$. It is usually written with its arguments as subscripts. Thus, it is the function $\delta : X \times X \to \{0, 1\}$ defined by

$$\delta_{xy} \;=\; \begin{cases} 0 & \text{when } x \neq y \\ 1 & \text{when } x = y. \end{cases}$$

e. Let $r_1, r_2, \ldots, r_n$ be distinct real numbers (or, more generally, distinct elements of any field — see 8.18). For $k = 1, 2, 3, \ldots, n$, let

$$L_k(t) \;=\; \prod_{j \neq k} \frac{t - r_j}{r_k - r_j}\,.$$

Show that $L_1, L_2, \ldots, L_n$ are polynomials of degree $n - 1$ that satisfy $L_k(r_j) = \delta_{jk}$ (where δ is the Kronecker delta). These are the **Lagrange polynomials**. We shall use them for a result about linear independence in 11.15, which in turn will be used for a cardinality proof in 11.35.

The Lagrange polynomials are commonly used in numerical analysis in the following fashion: Let $f(t)$ be any function defined on a set that includes the numbers $r_1, r_2, \ldots, r_n$. Let

$$p(t) \;=\; \sum_{k=1}^{n} f(r_k) L_k(t).$$

Then $p(t)$ is the unique polynomial of degree at most n that agrees with f on the set $\{r_1, r_2, \ldots, r_n\}$. It is called the **interpolating polynomial** and is used to approximate f in various ways.

2.3. The **composition** of two functions $f : X \to Y$ and $g : Y \to Z$ is the function $g \circ f : X \to Z$ defined by $(g \circ f)(x) = g(f(x))$. The symbol "$\circ$" may be included for emphasis or clarification, but it may be omitted otherwise: $g \circ f$ may be written "multiplicatively" as gf. However, the beginner is cautioned not to assume too much just on the basis of notation. For instance, unlike multiplication of real numbers, composition of functions does not satisfy the *commutative law* $gf = fg$.

2.4. A **self-mapping** of a set X is a mapping from X into X.

One particularly important self-mapping of X is the **identity mapping**, denoted $i : X \to X$, defined by $i(x) = x$ for all x. Of course, the identity maps of different sets X and Y are different functions; we may write i_X and i_Y if the distinction needs to be displayed. *Caution*: Some texts — especially on category theory — write the identity map as 1_X, but it should not be confused with the *characteristic function*, defined in 2.2.b.

For a function $f : X \to X$ we may write $f^2 = f \circ f$, $f^3 = f \circ f \circ f$, etc.; these are the **iterates** of f. It will sometimes be convenient to also write $f^1 = f$ and $f^0 = i_X$; then $f^{m+n} = f^m \circ f^n$ for any nonnegative integers m, n.

If f is a self-mapping of X, then a **fixed point** of f is any point $x \in X$ such that $f(x) = x$. Note that x is then a fixed point of all the iterates of f, too. *Preview*: Many problems that do not appear to involve fixed points can be reformulated as problems about fixed points. For instance, if we are given y and a function f and wish to find a solution x of the equation $y = f(x)$, we may rewrite the equation as $g(x) = x$, where we define $g(u) = f(u) - y + u$. This may seem rather contrived, but some problems yield to solution in this fashion. Several theorems about fixed points will be developed in later chapters.

A function $f : X \to X$ is **idempotent** if $f^2 = f$, or equivalently if $f(x) = x$ for all $x \in \text{Range}(f)$. Some elementary examples: The absolute value function, the greatest integer function, and sgn are idempotent maps from $\mathbb{R}$ into itself.

An **involution** of a set X is a function $f : X \to X$ that satisfies $f^2 = i_X$. Here are some examples. The reader should already be familiar with the first few of these; the last few are a preview of material in later chapters.

- $f(x) = -x$ is an involution on $\mathbb{R}$ (or on any additive group).
- $f(x) = 1/x$ is an involution on $\mathbb{R} \setminus \{0\}$ or on $(0, +\infty)$.
- $S \mapsto \complement S$ is an involution on $\mathcal{P}(X)$, for any set X. (Later we will generalize this to Boolean algebras.)
- $A \mapsto A^\top$ is an involution on certain collections of matrices.
- $\alpha \mapsto \overline{\alpha}$ is an involution on $\mathbb{C}$; here $\overline{\alpha}$ denotes the complex conjugate of α.

2.5. *Functions that agree.*

a. If X is a set and $S \subseteq X$, then the **inclusion map** $i : S \to X$ is the map given by $i(s) = s$ for each $s \in S$. This arrangement is sometimes abbreviated as $i : S \subseteq X$; for emphasis or clarification we may occasionally write it as $i : S \xrightarrow{\subseteq} X$. Of course, when $S = X$, then i is simply the identity map, defined in 2.4.

b. If $f : X \to Y$ and $S \subseteq X$, then the **restriction** of f to S is the function $f\big|_S : S \to Y$ that takes the value $f(s)$ at each point $s \in S$. A function f is an **extension** of a function g if g is a restriction of f; note that this occurs if and only if $g = f \circ i$ for some inclusion i.

c. Two functions $f_1 : X_1 \to Y_1$ and $f_2 : X_2 \to Y_2$ with overlapping domains are said to **agree** at a point $x_0 \in X_1 \cap X_2$ if $f_1(x_0) = f_2(x_0)$. They are said to agree on a set $S \subseteq X_1 \cap X_2$ if they agree at every point in S. Two functions **differ** at a point if they do not agree there.

Let $f_1 : X \to Y_1$ and $f_2 : X \to Y_2$ be two functions with the same domain and different codomains. (See the following diagram.) Then f_1 and f_2 agree on all of X if and only if $i_1 \circ f_1 = i_2 \circ f_2$ for some inclusions $i_1 : Y_1 \xrightarrow{\subseteq} Y$ and $i_2 : Y_2 \xrightarrow{\subseteq} Y$; here Y can be any set that contains $Y_1 \cup Y_2$. When that is the case, then we usually disregard formality and consider f_1 and f_2 to be the "same" function; but occasionally the formal distinction between two such functions is useful. See the related discussions in 9.4 and 9.20.

$$\begin{tikzcd} & Y_1 \arrow[dr, "i_1"] & \\ X \arrow[ur, "f_1"] \arrow[dr, "f_2"'] & & Y \\ & Y_2 \arrow[ur, "i_2"'] & \end{tikzcd}$$

Two functions f_1, f_2 that agree on their domain X

d. A function **vanishes** at a point or on a set if that function agrees there with the constant function 0, where "0" has any of its usual meanings — i.e., the empty set, the real number 0, the vector 0 in some linear space, etc.

2.6. Recall from 1.31 that a function $f : X \to Y$ is *surjective* if its codomain Y is equal to its range $f(X)$.

A function $f : X \to Y$ is **injective** (or *one-to-one*, or an *injection*) if it has the property that $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$. (See the following diagram.) More generally, a collection Φ of mappings defined on X (possibly with different codomains) is said to **separate the points** of X if for each pair of distinct points x_1, x_2 in X there exists at least one $f \in \Phi$ satisfying $f(x_1) \neq f(x_2)$.

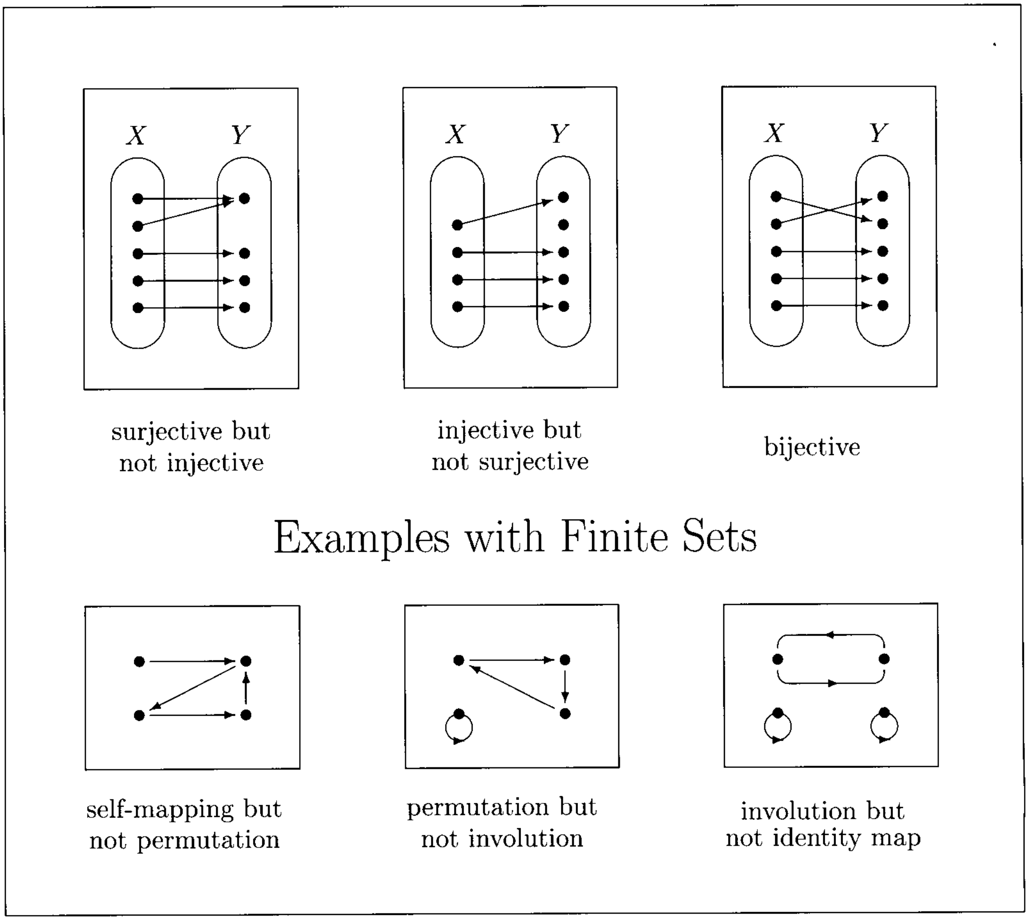
If a function $f : X \to Y$ is injective, then we may define its **inverse**, a function $f^{-1} : \text{Range}(f) \to X$, as follows: for each $y \in \text{Range}(f)$, let $f^{-1}(y)$ be the unique $x \in X$ that satisfies $f(x) = y$.

We say a function $f : X \to Y$ is **bijective** (or a *bijection* of X onto Y, or a *one-to-one correspondence between* X and Y) if it is both injective and surjective. It then follows that f^{-1} is also a bijection from Y onto X. A set $S \subseteq X \times Y$ is the graph of a bijection from X onto Y if and only if S is a set of ordered pairs such that each $x \in X$ is the first coordinate of exactly one of the pairs and each $y \in Y$ is the second coordinate of exactly one of the pairs. Of course, whenever $f : X \to Y$ is injective, then f acts as a bijection from X onto $\text{Range}(f)$.

A bijection from a set X onto itself is called a **permutation** of X.

Exercise. Any involution (defined in 2.4) is a permutation.

2.7. Let $f : X \to Y$ be a function. The **image** (or *forward image*) under f of any set $S \subseteq X$ is the set $f(S) = \{f(x) : x \in S\} \subseteq Y$. Thus the same symbol "$f$" is also used for a mapping from $\mathcal{P}(X)$ into $\mathcal{P}(Y)$; in general this does not lead to any confusion. (In a few



surjective but not injective

injective but not surjective

bijective

Examples with Finite Sets

self-mapping but not permutation

permutation but not involution

involution but not identity map

unusual contexts it can cause difficulty, however; see 14.65. Some mathematicians use a slightly different notation, such as $f[S]$ or $f :: S$, for the forward image; we shall not follow that practice here.)

The forward image map preserves some of the basic set operations:

$$\bigcap_{\lambda\in\Lambda} f(S_\lambda) \supseteq f\left(\bigcap_{\lambda\in\Lambda} S_\lambda\right) \qquad \text{and} \qquad \bigcup_{\lambda\in\Lambda} F(S_\lambda) = F\left(\bigcup_{\lambda\in\Lambda} S_\lambda\right)$$

The forward image map extends the given mapping $f : X \to Y$ if we identify each singleton in X or Y with its unique member. The *range* of f, defined in 1.31, is just the set $f(X)$ — i.e., the image of the domain.

The notation of forward images can also be applied in one or more arguments of a function of several variables. Thus, for a function $f : X \times Y \to Z$ we may write

$$f(x,T) = f(\{x\} \times T) = f(\{x\},T) = \{f(x,t) : t \in T\}$$

and similarly

$$f(S,T) \quad = \quad f(S \times T) \quad = \quad \{f(s,t) \ : \ s \in S, \ t \in T\}$$

for sets $S \subseteq X$ and $T \subseteq Y$. This notation can also be combined with the notation of binary operators. Thus we may write $S \square T = \{s \square t \ : \ s \in S, \ t \in T\}$. In particular, $S - T = \{s - t : s \in S, \ t \in T\}$, as noted in 1.24.

2.8. Let $f : X \to Y$ be any function. The **inverse image** (or *preimage*) under f of any set $T \subseteq Y$ is the set $f^{-1}(T) = \{x \in X : f(x) \in T\}$. If f is injective, then $f^{-1}(T)$ is also equal to $\{f^{-1}(t) : t \in T\}$ — that is, the forward image of $T \cap \text{Range}(f)$ under the mapping $f^{-1} : \text{Range}(f) \to X$.

Whether f is injective or not, f^{-1} is a mapping from $\mathcal{P}(Y)$ into $\mathcal{P}(X)$. It is somewhat better behaved than the forward image — it preserves *all* the basic set operations:

$$f^{-1}\left(\bigcup_{\lambda \in \Lambda} T_\lambda\right) \ = \ \bigcup_{\lambda \in \Lambda} f^{-1}(T_\lambda), \qquad f^{-1}\left(\bigcap_{\lambda \in \Lambda} T_\lambda\right) \ = \ \bigcap_{\lambda \in \Lambda} f^{-1}(T_\lambda),$$

$$f^{-1}(Y \backslash T) \ = \ X \backslash f^{-1}(T), \qquad f^{-1}(Y) \ = \ X, \qquad f^{-1}(\varnothing) = \varnothing.$$

For any point $y \in Y$, the set $f^{-1}(\{y\})$ can also be abbreviated as $f^{-1}(y)$.

2.9. *Further properties of forward and inverse images.* Let $g : X \to Y$ be some function. Then for all sets $S, T \subseteq X$ and $A, B \subseteq Y$ we have:

- **a.** $g^{-1}(g(S)) \supseteq S$.
- **b.** $g(g^{-1}(A)) = A \cap \text{Range}(g) \subseteq A$.
- **c.** Suppose g is *surjective*. Then $g(g^{-1}(A)) = A$. Also, $A \subseteq B \iff g^{-1}(A) \subseteq g^{-1}(B)$; hence $A = B \iff g^{-1}(A) = g^{-1}(B)$.

Distances

2.10. For later reference we note this form of the **Cauchy-Bunyakovskiĭ-Schwarz inequality**:

$$x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \quad \leq \quad \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \ \sqrt{y_1^2 + y_2^2 + \cdots + y_n^2}$$

for any real numbers $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$. *Hint for the proof*: $0 \leq \sum_{i \neq j}(x_i y_j - x_j y_i)^2$. A more general form of the CBS inequality will be given in 22.33.

An important consequence (which will be used in 2.12.a) is:

$$\sqrt{(x_1 + y_1)^2 + \cdots + (x_n + y_n)^2} \quad \leq \quad \sqrt{x_1^2 + \cdots + x_n^2} \ + \ \sqrt{y_1^2 + \cdots + y_n^2}\,.$$

To prove this inequality, multiply both sides of the CBS inequality by 2, then add $\sum_{j=1}^n (x_j^2 + y_j^2)$ to both sides, then take square roots on both sides.

2.11. *Definitions.* A **quasipseudometric** on a set X is a mapping $d : X \times X \to [0, +\infty)$ that satisfies $d(x,x) = 0$ and

$$d(x,y) \;\leq\; d(x,u) + d(u,y) \qquad \textbf{(triangle inequality)}$$

for all $x, y, u \in X$. The number $d(x,y)$ is called the **distance** from x to y. The name "triangle inequality" stems from the fact that in Euclidean geometry, the length of one side of a triangle is less than or equal to the sum of the lengths of the other two sides.

Note that many different distance functions can be defined on any one set X. For real-world examples, note that distance as the crow flies is different from distance as the taxicab drives. For an asymmetric example, consider a taxicab in a city that has some one-way streets; the distance from x to y is not necessarily equal to the distance from y to x.

Except for a few brief remarks in 5.15.i, all of the quasipseudometrics which we shall consider in this book also satisfy

$$d(x,y) \;=\; d(y,x); \qquad \textbf{(symmetry)}$$

they are then called **pseudometrics**. A pseudometric that also satisfies

$$d(x,y) \;>\; 0 \quad \text{when} \quad x \neq y \qquad \textbf{(positive-definiteness)}$$

is called a **metric**.

In this book we shall sometimes discuss the positive-definite case and the not-necessarily-positive-definite case simultaneously, by writing "pseudo" in parentheses. Any such discussion should be read once with the "pseudo" and once without it. This convention also applies to G-(semi)norms, F-(semi)norms, and (semi)norms, which are special types of (pseudo)metrics introduced in later chapters.

A **(pseudo)metric space** is a pair (X, d) consisting of a set X and a (pseudo)metric d on X. We may refer to X itself as a (pseudo)metric space, if d does not need to be mentioned explicitly.

A map $p : X \to Y$, from one (pseudo)metric space (X, d) into another (pseudo)metric space (Y, e), is called **distance-preserving** or **isometric** if it satisfies $e\Big(p(x_1), p(x_2)\Big) = d(x_1, x_2)$ for all $x_1, x_2 \in X$. It then preserves all the (pseudo)metric structure of X. If it is also injective (always true for metric spaces), then by a change of notation we may view p as an inclusion map that makes X a subset of Y.

2.12. *Basic examples and properties of metrics and pseudometrics.*

a. The **usual metric** on $\mathbb{R}$ is $d(x,y) = |x - y|$, where $|\ \ |$ is the usual absolute value function.

For any positive integer n, the most commonly used metrics on $\mathbb{R}^n$ (or on $\mathbb{C}^n$) are

$$\begin{aligned} d_1(x,y) &= |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n|, \\ d_2(x,y) &= \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + \cdots + |x_n - y_n|^2}, \\ d_\infty(x,y) &= \max\left\{|x_1 - y_1|,\ |x_2 - y_2|,\ \ldots,\ |x_n - y_n|\right\}, \end{aligned}$$

for any points $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$. It is easy to verify that these are indeed metrics; for d_2 use 2.10.

The metrics d_1 and d_2 are special cases of the metric

$$d_p(x,y) \quad = \quad \sqrt[p]{|x_1 - y_1|^p + |x_2 - y_2|^p + \cdots + |x_n - y_n|^p},$$

where $p \in [1, \infty)$; that d_p is a metric will be proved in 22.11. The metrics d_p for $1 \leq p \leq \infty$ may be referred to as the **usual metrics** on $\mathbb{R}^n$. They are equivalent, in a sense discussed in 22.5, and therefore they are interchangeable for most purposes.

b. A point x in a metric space (X, d) is **isolated** if there is some number $r > 0$ (which may depend on x) such that all other points have distance from x at least equal to r. A **discrete metric** on a set X is a metric that makes every point isolated. There are many such metrics on a set X, and some of them have substantially different properties — see 5.34.a and 19.11.e. The simplest discrete metric is the following one, which we shall call the **Kronecker metric**:

$$d(x,y) \quad = \quad 1 - \delta_{xy} \quad = \quad \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y \end{cases}$$

(where δ is the Kronecker delta). Some mathematicians call this *the* discrete metric.

c. (This example assumes some familiarity with calculus.) Let X be the set of all Riemann integrable (or more generally, Henstock or Lebesgue integrable) real-valued functions defined on some interval $J \subseteq \mathbb{R}$ — say on $[0,1]$, for instance. Define

$$d(p,q) \quad = \quad \int_J |p(t) - q(t)| \, dt.$$

This is a pseudometric on X, but it is not a metric. For instance, $d(p,q) = 0$ if $\{t \in J : p(t) \neq q(t)\}$ is a finite set. The pseudometric d becomes a metric if we restrict it to the continuous functions on J.

d. We may define a metric on the extended real line $[-\infty, +\infty]$ by taking $d(x,y) = |f(x) - f(y)|$ where f is some injective function from $[-\infty, +\infty]$ into $\mathbb{R}$. Three such functions $f(u)$ are given by

$$\arctan(u), \qquad \tanh(u), \qquad \frac{u}{1 + |u|}$$

with values of $f(u)$ at $u = \pm\infty$ defined by taking limits in the obvious fashion. Assorted other functions f will also suffice for this purpose. We do not think of $|f(x) - f(y)|$ as actually being the "distance" between x and y, but the metric is nevertheless useful for defining convergent sequences and other metric concepts. The three choices of f given above yield metrics that are *equivalent* in the sense that they yield the same topologies and the same uniformities, and consequently agree on many other structures — e.g., they have the same convergent sequences; these notions are discussed in later chapters. (See 18.24.d.) Any of these metrics, or any other equivalent metric, may be referred to as the **usual metric** on $[-\infty, +\infty]$.

e. $|d(x,y) - d(u,v)| \leq d(x,u) + d(y,v)$ for $x, y, u, v \in X$, if d is a pseudometric on X.

f. (*Optional.*) An **ultrametric** is a metric that satisfies the following strengthened version of the triangle inequality:

$$d(x,y) \quad \leq \quad \max\{d(x,u),\ d(u,y)\}.$$

Show that this inequality implies the triangle inequality. Show that the Kronecker metric (in 2.12.b) is an ultrametric.

2.13. *Definitions.* By a **gauge** on a set X we shall mean a collection of pseudometrics on X. A **gauge space** is a pair (X, D) consisting of a set X and a gauge D on X. We may refer to X itself as a gauge space, if D does not need to be mentioned explicitly.

As we develop the theory of gauges, we shall devote special attention to the case of a gauge $D = \{d\}$ consisting of just one pseudometric d. We shall often write "$\{d\}$" and "d" interchangeably, discuss d itself as a gauge, and develop some properties of pseudometrics as a special case of properties of gauges. Conversely, a gauge space (X, D) often can be analyzed in terms of the simpler pseudometric spaces $\{(X,d) : d \in D\}$.

A gauge D on a set X is **separating** if it has the property that

> for each pair of distinct points x and y in X, there exists at least one $d \in D$ satisfying $d(x,y) > 0$.

Most gauges used in applications are separating. Some mathematicians make the separation condition a part of their definition of "gauge," but we shall not follow that practice.

Note that a singleton gauge $D = \{d\}$ is separating precisely when the pseudometric d is a metric. Thus, most pseudometrics used *by themselves* in applications are metrics. However, some important separating gauges D used in applications consist of large collections of pseudometrics that are *not* metrics; for instance, see 28.11.b.

We caution that the term "gauge" is used in a wide variety of inequivalent ways in the literature. Our own usage follows that of Reilly [1973]; that usage works particularly well with the concepts in this book. Another, entirely unrelated meaning of the term "gauge" is given in 24.6.

2.14. *Remarks/example.* A single pseudometric is not adequate to describe the structure of some spaces; sometimes large collections of pseudometrics are needed. For instance, let $X = \mathbb{R}^{\mathbb{R}} = \{\text{functions from } \mathbb{R} \text{ into } \mathbb{R}\}$. For each $t \in \mathbb{R}$, define a pseudometric d_t on X by: $d_t(p,q) = |p(t) - q(t)|$. The resulting gauge $D = \{d_t \ : \ t \in \mathbb{R}\}$ is separating, but no proper subset of it is separating. We shall see in 9.18 and 18.9.f that the gauge D yields the product topology and product uniformity on $\mathbb{R}^{\mathbb{R}}$.

2.15. *Examples and exercises about separation.*

a. If $f : X \to \mathbb{R}$ is any real-valued function on any set, then $d(x,y) = |f(x) - f(y)|$ is a pseudometric on X. It is a metric if and only if the function f is injective — i.e., satisfying $x \neq y \ \Rightarrow \ f(x) \neq f(y)$. A special case of this construction was given in 2.12.d.

More generally, let $\Phi = \{f_\lambda : \lambda \in \Lambda\}$ be a collection of real-valued functions on a set X. Then a gauge can be defined by $D = \{d_\lambda : \lambda \in \Lambda\}$, where $d_\lambda(x,y) = |f_\lambda(x) - f_\lambda(y)|$.

This gauge is separating if and only if the collection Φ separates points of X (in the sense of 2.6).

b. A gauge D on a set X is separating in the sense of 2.11 if and only if the collection of functions $\Phi = \{f_{x,d} : x \in X, d \in D\}$ defined by $f_{x,d}(y) = d(x,y)$ is a separating collection in the sense of 2.6.

CARDINALITY

2.16. The **cardinality** of a finite set is the number of distinct elements in that set; thus it is a nonnegative integer.

The "cardinality" of a not-necessarily-finite set is a bit harder to define; we shall postpone that concept until 6.23. However, it is much easier to define the *comparison* of cardinalities of two sets. This notion is due to Georg Cantor and is the foundation of modern set theory. We say that two sets X and Y **have the same cardinality** — written $\mathrm{card}(X) = \mathrm{card}(Y)$ — if there exists a bijection between X and Y. More generally, we write $\mathrm{card}(X) \leq \mathrm{card}(Y)$ if X has the same cardinality as some subset of Y — i.e., if there exists an injection from X into Y. Similarly, we write $\mathrm{card}(X) < \mathrm{card}(Y)$ if X and Y satisfy $\mathrm{card}(X) \leq \mathrm{card}(Y)$ but do not satisfy $\mathrm{card}(X) = \mathrm{card}(Y)$ — i.e., if there exists an injection from X into Y but there does not exist a bijection from X onto Y. (Cantor invented these ideas while investigating Fourier series; see 26.48.)

With this convention, we can now restate some of the definitions given in 1.20 and add a few more definitions. A set S is

finite if $\mathrm{card}(S) = \mathrm{card}(\{1, 2, \ldots, n\})$ for some nonnegative integer n (in which case we call n the **cardinality** of the set and write $\mathrm{card}(S) = n$);

infinite if it is not finite;

cofinite if it is being considered as a subset of some set X and its complement $X \setminus S$ is finite;

countable (or **denumerable**) if $\mathrm{card}(S) \leq \mathrm{card}(\mathbb{N})$;

uncountable if it is not countable;

countably infinite if $\mathrm{card}(S) = \mathrm{card}(\mathbb{N})$.

Caution: Some mathematicians apply the term "countable" or the term "denumerable" only to the sets that have the *same* cardinality as $\mathbb{N}$. Also, some mathematicians use a slightly different definition of "infinite" — see the remark in 6.27.

The cardinality of a set X is sometimes abbreviated $|X|$.

Much of our presentation of cardinality is based on Dalen, Doets, and Swart [1978] and Kaplansky [1977].

2.17. *Further remarks.* Throughout the mathematical literature, the letter σ (a Greek lowercase sigma) is often used to indicate countable sums or unions — e.g., in σ-ideals, σ-algebras, σ-additive measures, σ-convex sets, F_σ sets. Similarly, δ (delta) is often used to indicate countable products or intersections — e.g., in G_δ sets. We shall define these terms separately in their appropriate contexts.

2.18. *Remarks.* It is customary to use the familiar symbol $\leq$ for comparison of cardinalities. Do not assume too much on the basis of this notation, however; the comparison of cardinalities is not quite like the comparison of real numbers. Some familiar properties of real numbers are also valid for cardinalities, and some are not. For instance, it is quite easy to prove that for any sets X, Y, Z we have

$$\mathrm{card}(X) \leq \mathrm{card}(Y) \quad \text{and} \quad \mathrm{card}(Y) \leq \mathrm{card}(Z) \quad \text{imply} \quad \mathrm{card}(X) \leq \mathrm{card}(Z).$$

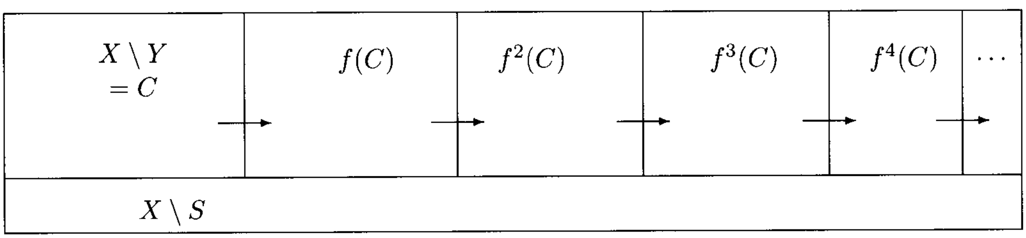
(The reader should show this now, as an exercise.) It is rather harder to prove that

$$\mathrm{card}(X) \leq \mathrm{card}(Y) \quad \text{and} \quad \mathrm{card}(Y) \leq \mathrm{card}(X) \qquad \text{imply} \qquad \mathrm{card}(X) = \mathrm{card}(Y);$$

that is the content of the Schröder-Bernstein Theorem in 2.19. Thus, comparison of cardinalities is a preordering; comparison of distinct cardinalities is a partial ordering. Still stronger properties about the comparison of cardinalities will be proved in 6.22, but the proof is deeper and also requires that we assume the Axiom of Choice.

2.19. Schröder-Bernstein Theorem. Let X and Y be sets. If there exist injections $e : Y \to X$ and $f : X \to Y$, then there exists a bijection from X onto Y. In other words, if $\mathrm{card}(Y) \leq \mathrm{card}(X)$ and $\mathrm{card}(X) \leq \mathrm{card}(Y)$, then $\mathrm{card}(X) = \mathrm{card}(Y)$.

Proof. This presentation follows Cox [1968]. We may assume that $Y \subseteq X$ and that we are given an injection $f : X \to Y$. (More precisely, since we have an injection $e : Y \to X$, by relabeling we may identify each point of Y with its image under e; see 1.10.) In the following diagram, the big box represents the set X.



Let $C = X \backslash Y$. Since f is injective and has range contained in Y, the sets C, $f(C)$, $f^2(C)$, $f^3(C)$, ... are disjoint. (Here f^n is the nth iterate of f.) Let $S = \bigcup_{n=0}^{\infty} f^n(C)$; note that $f(S) = S \setminus C \subseteq S$. (See the diagram above.) Define a function $h : X \to Y$ by taking $h(z) = f(z)$ when $z \in S$ and $h(z) = z$ when $z \in X \setminus S$. Verify that the function h takes each $f^n(C)$ bijectively to $f^{n+1}(C)$, and hence h is a bijection from X onto Y.

2.20. *Exercises and examples.*

a. $\text{card}(\varnothing) = 0$ and $\text{card}(\{\varnothing\}) = 1$.

b. $\text{card}(\varnothing) < \text{card}(\{1\}) < \text{card}(\{1,2\}) < \text{card}(\{1,2,3\}) < \cdots < \text{card}(\mathbb{N})$.

c. In $\mathbb{N}$, the subset $\{n \ : \ n > 4\}$ is cofinite.

d. The sets $\mathbb{N}$, $\mathbb{N} \cup \{0\}$, $\mathbb{Z}$, and $\{\text{even positive integers}\}$ all have the same cardinality. *Hint*: See diagram below.

$$\begin{array}{rclcccccccl}
\mathbb{N} & = & \{ & 1, & 2, & 3, & 4, & 5, & 6, & 7, & \ldots\} \\
 & & & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\
\mathbb{N} \cup \{0\} & = & \{ & 0, & 1, & 2, & 3, & 4, & 5, & 6, & \ldots\} \\
 & & & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\
\mathbb{Z} & = & \{ & 0, & 1, & -1, & 2, & -2, & 3, & -3, & \ldots\} \\
 & & & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\
\{\text{even positive integers}\} & = & \{ & 2, & 4, & 6, & 8, & 10, & 12, & 14, & \ldots\}
\end{array}$$

e. Cantor's Theorem on pairs. $\mathbb{N} \times \mathbb{N}$ is countably infinite. *Hint*: By tracing along the diagonals in the diagram below, we obtain the sequence $(1,1)$, $(2,1)$, $(1,2)$, $(3,1)$, $(2,2)$, $(1,3)$, $(4,1)$, $\ldots$, which is an enumeration of $\mathbb{N} \times \mathbb{N}$.

$$\begin{array}{ccccccccccc}
(1,1) & & (2,1) & & (3,1) & & (4,1) & & (5,1) & & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & & \\
(1,2) & & (2,2) & & (3,2) & & (4,2) & & (5,2) & & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & & \\
(1,3) & & (2,3) & & (3,3) & & (4,3) & & (5,3) & & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & & \\
(1,4) & & (2,4) & & (3,4) & & (4,4) & & (5,4) & & \cdots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & & \\
(1,5) & & (2,5) & & (3,5) & & (4,5) & & (5,5) & & \cdots \\
 & & & & & & & & & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \ddots
\end{array}$$

f. $\text{card}(\mathbb{Q}) = \text{card}(\mathbb{N})$. *Hint*: Use the preceding result, together with the Schröder-Bernstein Theorem.

Remark. Later we will show $\text{card}(\mathbb{R}) = \text{card}(\mathbb{N}^{\mathbb{N}}) = \text{card}(2^{\mathbb{N}}) > \text{card}(\mathbb{N})$. See 10.44.f.

g. If $\text{card}(X) \geq \text{card}(\mathbb{N})$ and u is any object, then $\text{card}(X \cup \{u\}) = \text{card}(X)$. *Hints*: This is trivial if $u \in X$. If $u \notin X$, use 2.20.d.

h. If X and Y are finite sets, then $\text{card}(X \times Y) = \text{card}(X)\text{card}(Y)$ and $\text{card}(X^Y) = \text{card}(X)^{\text{card}(Y)}$ (with the conventions $r^0 = 1$ for $r \geq 0$ and $0^r = 0$ for $r > 0$).

i. For any set X, we have $\text{card}(X \times X) \geq \text{card}(X)$. *Hint*: Treat separately the cases of $X = \varnothing$ and $X \neq \varnothing$.

Remarks. We have $\text{card}(X \times X) > \text{card}(X)$ when X is a finite set containing more than one element. We have $\text{card}(X \times X) = \text{card}(X)$ when X is empty or a singleton. In 6.22 we shall see that $\text{card}(X \times X) = \text{card}(X)$ when X is an infinite set; however, the proof of that result will require the Axiom of Choice.

j. If X, Y, and Z are any sets, then there is a natural bijection between $Z^{X\times Y}$ and $(Z^Y)^X$. Indeed, $f \in Z^{X\times Y}$ means that $(x,y) \mapsto f(x,y)$ is a map from $X \times Y$ into Z, while $f \in (Z^Y)^X$ means that $x \mapsto f(x,\cdot)$ is a map from X into Z^Y. It is easy to see that this correspondence between $Z^{X\times Y}$ and $(Z^Y)^X$ is no more than a change of notation.

k. The set $\{0,1\}$ is often called "2." Let X be a set; we can identify each subset $S \subseteq X$ with its characteristic function $1_S : X \to \{0,1\}$, defined in 2.2.b. Thus there is a natural bijection between the *power set* of X,

$$\mathcal{P}(X) \quad = \quad \{\text{subsets of } X\},$$

and the *Xth power of the set* 2,

$$2^X \quad = \quad \{\text{functions from } X \text{ into } 2\}.$$

These two objects are often used interchangeably.

If X is a finite set, then $\operatorname{card}(\mathcal{P}(X)) = 2^{\operatorname{card}(X)}$ is the number of subsets of X. *Exercise for beginners.* List the eight subsets of $X = \{0,1,2\}$. *Hint*: Don't forget $\varnothing$ and X.

l. Theorem (Cantor). $\operatorname{card}(\mathcal{P}(X)) > \operatorname{card}(X)$ for every set X.

Hints: Easily $\operatorname{card}(\mathcal{P}(X)) \geq \operatorname{card}(X)$. Now suppose that there exists a bijection $f : X \to \mathcal{P}(X)$. Define $R = \{x \in X : x \notin f(x)\}$, and let $r = f^{-1}(R)$. Show that $r \in R \iff r \notin R$, a contradiction. *Note*: This contradiction is not paradoxical, but it is similar to Russell's Paradox (1.43).

m. *Example.* $\operatorname{card}(2^{\mathbb{N}}) > \operatorname{card}(\mathbb{N})$.

n. *Example.* $\operatorname{card}(2^{\mathbb{N}}) = \operatorname{card}(\mathbb{N}^{\mathbb{N}})$. *Hints*:

$$\operatorname{card}(\mathbb{N}^{\mathbb{N}}) \leq \operatorname{card}((2^{\mathbb{N}})^{\mathbb{N}}) = \operatorname{card}(2^{\mathbb{N}\times\mathbb{N}}) = \operatorname{card}(2^{\mathbb{N}}) \leq \operatorname{card}(\mathbb{N}^{\mathbb{N}}).$$

2.21. *How many kinds of infinity are there?* By Cantor's Theorem (in 2.20.l),

$$\operatorname{card}(\mathbb{N}) < \operatorname{card}(\mathcal{P}(\mathbb{N})) < \operatorname{card}(\mathcal{P}(\mathcal{P}(\mathbb{N}))) < \operatorname{card}(\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))) < \cdots$$

(where $\mathcal{P}$ denotes the power set). Thus there are infinitely many different kinds of infinity. We can get still more infinities, as follows: Let S be the union of all the sets $\mathbb{N}$, $\mathcal{P}(\mathbb{N})$, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))$, ... Then S is bigger than any one of those sets. We can go further: We have

$$\operatorname{card}(S) < \operatorname{card}(\mathcal{P}(S)) < \operatorname{card}(\mathcal{P}(\mathcal{P}(S))) < \operatorname{card}(\mathcal{P}(\mathcal{P}(\mathcal{P}(S)))) < \cdots$$

and we can continue this process again and again, infinitely many times.

Are there still more infinities? Perhaps there are some even bigger than anything obtained in the "list" suggested above; or perhaps there are some lying between two consecutive elements of that list.

An **inaccessible cardinal** (also known as a *strongly inaccessible cardinal*) is, roughly, a set too big to be in the list given above; i.e., it is an uncountable set that is bigger than anything obtainable from smaller sets via power sets and unions. We shall not make this precise; refer to books on set theory and logic (e.g., Shoenfield [1967]) for details. It is not intuitively obvious whether such enormous cardinals exist. Their existence or nonexistence is taken as a hypothesis in some studies in set theory. Surprisingly, such assumptions about enormous sets lead to important conclusions about "ordinary" sets such as $\mathbb{R}$; see 14.75.

In applicable analysis one seldom has any need for infinite cardinalities other than $\mathrm{card}(\mathbb{N})$ or $\mathrm{card}(2^{\mathbb{N}}) = \mathrm{card}(\mathbb{R})$. The **Continuum Hypothesis (CH)** asserts that there are no other cardinalities *between* those two. The **Generalized Continuum Hypothesis (GCH)** asserts that for any infinite set X, there are no other cardinalities between $\mathrm{card}(X)$ and $\mathrm{card}(2^X)$. Cantor spent a large part of his last years trying to prove that CH was true or false. The question remained open for decades. Finally, Gödel and Cohen developed new methods to show that neither the truth nor the falsehood of CH can be proved from the usual axioms of set theory; thus CH is *independent* of those axioms. This is explained briefly in 14.7, 14.8, 14.53, 14.73, and 14.74.

Induction and Recursion on the Integers

2.22. We assume the reader is familiar with the basic properties of the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$. (*Caution*: Some mathematicians use the symbol $\mathbb{N}$ for the set $\{0, 1, 2, 3, \ldots\}$, but in this book $0 \notin \mathbb{N}$.)

Following are two basic principles about the natural numbers. **Induction** is a method for proving statements about objects that have already been defined; **recursion** is a method for defining new objects.

> **Principle of Countable Induction.** Suppose $1 \in T \subseteq \mathbb{N}$ and T has the property that whenever $n \in T$ then also $n + 1 \in T$. Then $T = \mathbb{N}$.

This principle can also be formulated as a method for proving that a statement $P(x)$ is true for every $x \in \mathbb{N}$ — just take $T = \{x \in \mathbb{N} : P(x)\}$. (*Note*: To logicians, this reformulation is not quite equivalent. It is usually understood that the statement $P(x)$ must be expressed using finitely many symbols from a language with only countably many symbols, so there are only countably many possible P's — but there are uncountably many sets $T \subseteq \mathbb{N}$.)

For our second principle, we shall agree that the *empty* sequence — the sequence with no components, or the sequence of length 0 — is a finite sequence and hence a member of the domain of ρ.

> **Principle of Countable Recursion.** Let T be a set, and let ρ be some mapping from {finite sequences in T} into T. Then there exists a unique sequence $(t_1, t_2, t_3, \ldots)$ in T that satisfies $t_n = \rho(t_1, t_2, \ldots, t_{n-1})$ for all n.

In other words, our definition of t_n may depend on all the preceding definitions.

Both of these principles are generalized to sets other than $\mathbb{N}$ in 1.50, 3.39.f, 3.40, and 5.51; they are then referred to as **transfinite** induction and recursion. For now, we note a few elementary applications of the countable case.

2.23. *Examples in countable induction and recursion.*

a. **Factorials** are defined recursively: $0! = 1$, and $(n+1)! = (n+1)\cdot(n!)$ for $n = 0, 1, 2, 3, \ldots$. (We read "$n!$" as "n factorial.") The first few factorials are $0! = 1$, $1! = 1$, $2! = 2$, $3! = 6$, and $4! = 24$.

b. The **binomial coefficient** $\binom{n}{k}$ (read "n choose k") can be defined directly by a formula:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad (n = 0, 1, 2, 3, \ldots;\ k = 0, 1, 2, \ldots, n)$$

or it can be defined by recursion on n: we take $\binom{n}{0} = \binom{n}{n} = 1$ for $n = 0, 1, 2, 3, \ldots$, and then

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1} \qquad (0 \le k < n).$$

Show by induction that the two methods of defining $\binom{n}{k}$ yield the same values. Also, using the second method, show that the $\binom{n}{k}$'s are the numbers in Pascal's Triangle.

Pascal's Triangle:
Each number is
the sum of the two
numbers above it.

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ & & & & \cdots & & & & \end{array}$$

By convention, we define $\binom{n}{k} = 0$ when $n \ge 0$ and $k \in \mathbb{Z} \setminus \{0, 1, 2, \ldots, n\}$.

c. By induction on n, prove the **Binomial Theorem**:

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j} \qquad (n = 1, 2, 3, \ldots).$$

An example is $(x+y)^4 = y^4 + 4xy^3 + 6x^2y^2 + 4x^3y + x^4$.

d. A **prime number** is an integer greater than 1 that is not divisible by any positive integer except itself and 1. The first few prime numbers are $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, and $p_5 = 11$. The following induction argument proves that there are infinitely many prime numbers and also gives us a crude but easy upper bound on p_n.

Assume that $p_1, p_2, \ldots, p_n$ have already been found, for some positive integer n. Then $q = p_1 p_2 \cdots p_n + 1$ is greater than p_n, and it is not divisible by any of $p_1, p_2, \ldots, p_n$. Hence either q is a new prime, or it is divisible by a new prime. In any case, $p_{n+1} \le q \le 2p_1p_2\cdots p_n$. Use induction to show that $p_n \le 2^{2^n}$.

e. *Joke.* Every positive integer has some remarkably interesting property.

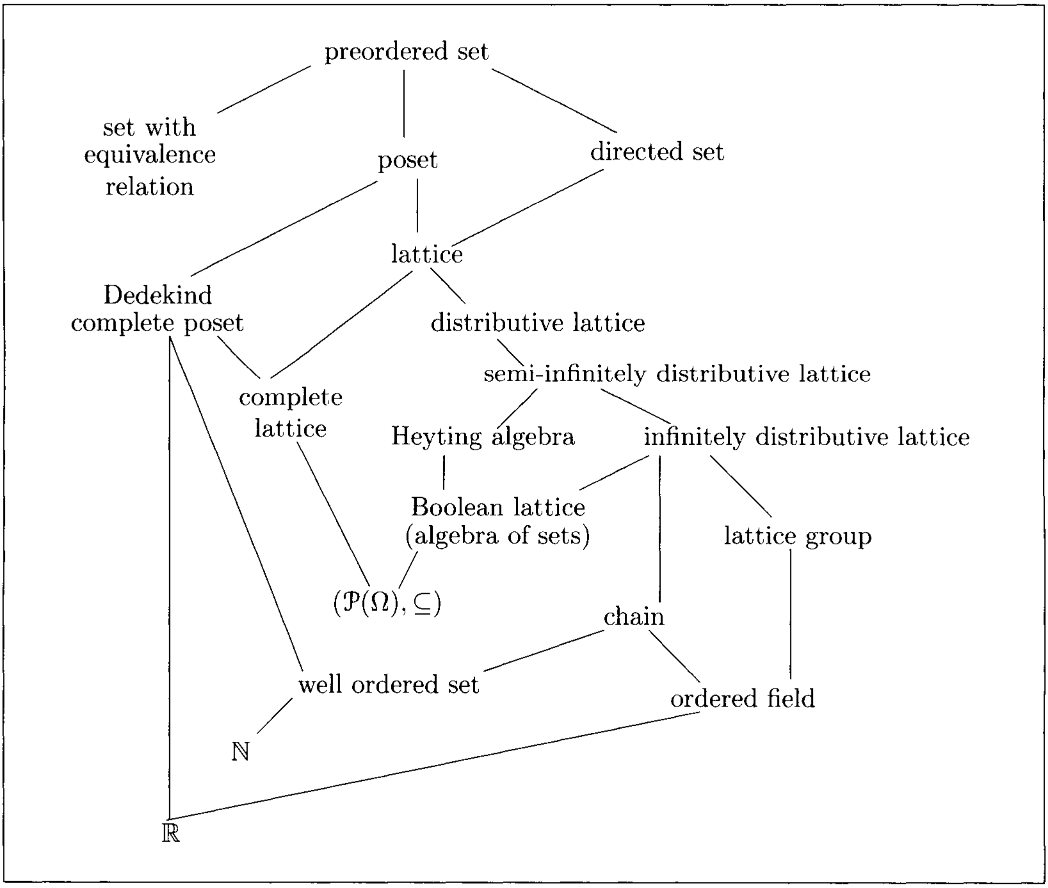
"*Proof.*" If not, let n_0 be the first uninteresting number. Then n_0 has the property that it is the *first* uninteresting number — but isn't that an interesting property?

Exercise. Carefully explain what has gone wrong here. *Hint*: See 1.11.

Chapter 3

Relations and Orderings

3.1. *Preview.* The chart below shows the connections between some kinds of preorders that we shall study in this and later chapters. Lattices and order completeness are studied in greater detail in Chapter 4; directed orderings are studied further in Chapter 7; Boolean algebras and Heyting algebras are covered in Chapter 13.



Relations

3.2. A **relation** (or *binary relation*) on a set X is simply a set $R \subseteq X \times X$, but with this change in our notation: instead of writing $(x, y) \in R$, we may write xRy. We may sometimes refer to the subset of $X \times X$ as the **graph** of the relation; we may even denote it by $\mathrm{Graph}(R)$ or $\mathrm{Gr}(R)$. Other symbols may be used in place of R. Some familiar symbols used in this fashion are $=$, $\neq$ (relations on any set X), $\leq$, $<$ (relations on $\mathbb{R}$), and $\subsetneqq$, $\subseteq$ (relations on a collection of sets).

3.3. *Examples and special kinds of relations.*

a. Equality ($=$) is a relation; its graph is the **diagonal set** $I = \{(x, x) : x \in X\}$.

b. The largest relation on a set X is

the **universal relation**: $\quad xRy$ for all $x, y \in X$.

Its graph is $X \times X$. Trivial though it may be, this relation is occasionally useful. When it is viewed as an ordering, we shall call it the **universal ordering**.

c. The smallest relation on X is the **empty relation**; its graph is the empty set.

d. The **inverse** of a relation R is the relation R^{-1} defined by $xR^{-1}y \Leftrightarrow yRx$. For instance, the inverses of $=$, $\neq$, $\subseteq$, $\subsetneqq$, $\leq$, $<$ are the relations $=$, $\neq$, $\supseteq$, $\supsetneqq$, $\geq$, $>$, respectively. Note that $(R^{-1})^{-1} = R$.

If $\preccurlyeq$ and $\succcurlyeq$ are relations that are inverses of each other, then there exists a duality between $\preccurlyeq$ and $\succcurlyeq$; each statement about either of these relations can be converted to a statement about the other relation. See 1.7.

e. The **composition** of any two relations Q and R on a set X is the relation defined by

$$Q \circ R \;=\; \{(x, y) \in X \times X \;:\; xRu \text{ and } uQy \text{ for at least one } u \in X\}.$$

This definition generalizes that in 2.3 — i.e., if Q and R are in fact functions, then the composition defined in this fashion is the same as the composition defined in 2.3. *Exercise.* Verify that the compositions of relations satisfy $(P \circ Q) \circ R = P \circ (Q \circ R)$.

f. If R is a relation on X and $Y \subseteq X$, then the **restriction** of R to Y (or **trace** of R on Y) is the relation $R\big|_Y$ defined by

$$u \; R\big|_Y \; v \qquad \text{if and only if} \qquad u, v \in Y \;\text{ and }\; uRv.$$

In other words, $\mathrm{Graph}(R\big|_Y) = \mathrm{Graph}(R) \cap (Y \times Y)$. By a slight abuse of notation, we often denote $R\big|_Y$ simply by the same symbol R — for example, a restriction of any of the relations $=$, $\neq$, $\subseteq$, $\subsetneqq$, $\leq$, $<$ is still denoted by the same symbol.

3.4. Let R be a relation on a set X, and let I be the diagonal set of X (see 3.3.a). Many relations of interest to us satisfy the condition of

transitivity: $R \circ R \subseteq R$. That is, xRy and yRz imply xRz.

Most relations of interest to us also satisfy either

reflexivity: $R \supseteq I$. That is, xRx for all $x \in X$, or
irreflexivity: $R \cap I = \varnothing$. That is, xRx for no $x \in X$.

Also, most satisfy either

symmetry: $R^{-1} = R$. That is, xRy implies yRx, or
antisymmetry: $R \cap R^{-1} \subseteq I$. That is, xRy and yRx imply $x = y$.

Some examples are given in 3.6.

3.5. *More symbols for relations.* A familiar symmetric relation is equality ($=$). For other symmetric relations, we often use the symbols $\approx$ or $\equiv$.

Inequality ($\leq$) and inclusion ($\subseteq$) are familiar relations that are not symmetric. For other relations that are not symmetric, or that are not known to be symmetric, we shall often use the symbols $\preccurlyeq$ and $\prec$. Occasionally we may also use $\sqsubseteq$ and $\sqsubset$.

Some mathematicians prefer to use the symbols $\leq$ and $<$ for any relation that is not necessarily symmetric because these symbols are more familiar and therefore easier to draw. However, beginners sometimes inadvertently attribute to those symbols some familiar properties of the ordering of the real numbers — e.g., they may implicitly assume that $\leq$ is a chain ordering (defined in 3.23). To reduce the frequency of this type of error, we will usually reserve the symbols $\leq, <$ for chain orderings, and use $\preccurlyeq, \prec$ for a more "generic" ordering. This makes it easier for beginners to disassociate themselves from the familiar properties of $\mathbb{R}$ and start over with a fresh perspective. Admittedly, $\preccurlyeq$ is difficult to draw on a blackboard; perhaps $\overset{<}{\frown}$ or $\overset{<}{\cap}$ could be used as a blackboard substitution.

In this book the symbols $\preccurlyeq$ and $\prec$ will always denote, respectively, a reflexive relation and an irreflexive relation, which are connected as follows:

$x \preccurlyeq y$	means that either $x \prec y$ or $x = y$ holds;
$x \prec y$	means that both $x \preccurlyeq y$ and $x \neq y$ hold.

In other words, the sets $I = \{(x,x) : x \in X\}$ and $\mathrm{Graph}(\prec)$ form a partition of the set $\mathrm{Graph}(\preccurlyeq)$. Because $\preccurlyeq$ and $\prec$ are connected in this fashion, we can usually state our results just in terms of $\preccurlyeq$, without explicitly mentioning corresponding results for $\prec$. A similar convention will apply to the pair $\sqsubseteq$, $\sqsubset$.

Inverses (see 3.3.d) of $\preccurlyeq$, $\prec$, $\sqsubseteq$, $\sqsubset$ will be denoted respectively by $\succcurlyeq$, $\succ$, $\sqsupseteq$, $\sqsupset$.

The symbols $\preccurlyeq$ and $\sqsubseteq$ may be read as **precedes**, is **smaller** than, is **littler** than, is **less** than, is **earlier** than. Their inverses ($\succcurlyeq$ and $\sqsupseteq$) may be read as **succeeds**, is **larger** than, is **bigger** than, is **more** than, is **later** than.

The symbols $\prec$ and $\sqsubset$ may be read as **strictly precedes**, etc., and the symbols $\succ$ and $\sqsupset$ may be read as **strictly succeeds**, etc.

3.6. *Examples and exercises.*

a. If R is a relation on X that is transitive and irreflexive, then R is also antisymmetric — but vacuously so: there cannot be x, y satisfying xRy and yRx simultaneously.

b. Six familiar relations are $=$, $\neq$, $\subseteq$, $\subsetneqq$, $\leq$, $<$. Among these examples, $=$, $\subseteq$, $\subsetneqq$, $\leq$, $<$ are transitive, $=$, $\subseteq$, $\leq$ are reflexive, $\neq$, $\subsetneqq$, $<$ are irreflexive, $=$, $\neq$ are symmetric, and $\subseteq$, $\subsetneqq$, $\leq$, $<$ are antisymmetric.

c. Show that $R = \{(1,2), (2,3), (2,2), (3,2)\}$ is a relation on the set $X = \{1,2,3\}$ that has *none* of the properties listed in 3.4 — i.e., R is not transitive, reflexive, irreflexive, symmetric, or antisymmetric.

d. Let R be a relation that is both symmetric and antisymmetric. Then (i) R is reflexive if and only if R is equality ($=$), and (ii) R is irreflexive if and only if R is the empty relation.

e. Let $\preccurlyeq$ be a reflexive relation on a set X and let $\prec$ be the corresponding irreflexive relation. Then (i) $\preccurlyeq$ is symmetric if and only if $\prec$ is symmetric, and (ii) $\preccurlyeq$ is transitive if and only if $\prec$ is transitive.

Preordered Sets

3.7. A **preorder** on a set X is a relation $\preccurlyeq$ that is both

transitive ($x \preccurlyeq y$ and $y \preccurlyeq z$ imply $x \preccurlyeq z$) and
reflexive ($x \preccurlyeq x$).

A **preordered set** is a pair $(X, \preccurlyeq)$ consisting of a set X and a preorder $\preccurlyeq$ on X; we may refer to X itself as the preordered set if $\preccurlyeq$ does not need to be mentioned explicitly. A similar syntax will be used for special kinds of preordered sets — partially ordered sets, sets with equivalence relations, directed sets, chains, well ordered sets, lattices, etc.

3.8. We note a few important special types of preorders. Let $(X, \preccurlyeq)$ be a preordered set. Then $\preccurlyeq$ is a

partial order — and $(X, \preccurlyeq)$ is a **partially ordered set**, or **poset** — if $\preccurlyeq$ is antisymmetric ($x \preccurlyeq y$ and $y \preccurlyeq x$ imply $x = y$);

equivalence relation if it is symmetric ($x \preccurlyeq x$ for all x);

directed order — and $(X, \preccurlyeq)$ is a **directed set** — if for each $x_1, x_2 \in X$ there exists some $y \in X$ satisfying $x_1 \preccurlyeq y$ and $x_2 \preccurlyeq y$.

3.9. *Basic properties and examples.*

a. In a set X equipped with a relation $\preccurlyeq$, we say that two elements x, y are **comparable** if at least one of the relations $x \preccurlyeq y$ or $y \preccurlyeq x$ holds. (This terminology is mainly used in posets.)

b. The only partial order that is also an equivalence relation is equality ($=$).

c. If $\preccurlyeq$ is an equivalence relation, then so is its inverse, $\succcurlyeq$.

d. If $\preccurlyeq$ is a partial order, then so is its inverse, $\succcurlyeq$.

e. Trivially, the empty set is directed. Any singleton $\{x\}$ is directed, when equipped with the relation $\{(x, x)\}$.

f. *Example* (from McShane [1952]). A stream or river, together with its tributaries, is directed by the relation "is upstream from." Indeed, if x, y are any two locations in the system, then there exists a third location z that is downstream (i.e., later in the water flow) from both x and y.

g. Most directed orderings of interest to us are antisymmetric. However, the universal ordering (defined in 3.3.b) is a directed ordering that we shall find useful and that is not antisymmetric. By calling the universal ordering and similar orderings "directed," we shall achieve a few simplifications in the development of the theory.

Caution: Some mathematicians make antisymmetry part of their definition of "directed set." Though we shall not follow that practice, we remark that adding antisymmetry to the definition of "directed set" does not greatly affect the ultimate applications. A not-necessarily-antisymmetric directed set can usually be replaced by a (perhaps more complicated) antisymmetric directed set, as in 7.12.

h. If $\mathcal{C}$ is any collection of subsets of a set X, then $(\mathcal{C}, \subseteq)$ is a poset. Actually, this example is as general as we could ask for: *Every* poset can be represented isomorphically in this form; see 3.16.d.

i. *Subsets of ordered sets.* Define restrictions as in 3.3.f. Verify that if the relation $\preccurlyeq$ on X has the following property, then its restriction to any set $S \subseteq X$ has the same property: reflexive, irreflexive, symmetric, antisymmetric, transitive, preorder, equivalence, or partial order.

The restriction of a directed order is *not* necessarily directed. For instance, $\mathbb{Z}^2$ with the product ordering is directed (in fact, a lattice), but its subset $\{(x, y) \in \mathbb{Z}^2 : x + y = 0\}$ is not directed.

j. For each λ in some index set Λ, let $\preccurlyeq_\lambda$ be a relation on some set X_λ. Then the **product** of the $\preccurlyeq_\lambda$'s is the relation $\preccurlyeq$ on the product of the X_λ's, defined thus:

$$f \preccurlyeq g \text{ in } \prod_{\lambda\in\Lambda} X_\lambda \qquad \text{means that} \qquad f(\lambda) \preccurlyeq_\lambda g(\lambda) \text{ for every } \lambda \in \Lambda.$$

It may also be called the *componentwise ordering* or *coordinatewise ordering* since it acts separately on each component or coordinate. We shall use the product relation on $\prod_{\lambda\in\Lambda} X_\lambda$ unless some other arrangement is specified.

Verify that if all the $\preccurlyeq_\lambda$'s have one of the following properties, then the product ordering $\preccurlyeq$ has the same property: reflexive, symmetric, antisymmetric, transitive, preorder, equivalence, directed order, or partial order.

k. There are many ways to define an ordering on a collection of functions. One commonly used method is the product ordering, defined above. Another common method is by inclusion of graphs — i.e., let $f \preccurlyeq g$ mean that $\mathrm{Gr}(f) \subseteq \mathrm{Gr}(g)$. The resulting relation $\preccurlyeq$ is a partial ordering; this is a special case of 3.9.h.

More about Equivalences

3.10. An **equivalence relation** is a relation $\approx$ that is

symmetric $(x \approx y \;\Rightarrow\; y \approx x)$,
reflexive $(x \approx x$ for all $x)$, and
transitive $(x \approx y,\, y \approx z \;\Rightarrow\; x \approx z)$.

If some equivalence relation $\approx$ has been specified, two objects x and y satisfying $x \approx y$ are said to be **equivalent**. The student is cautioned that the term "equivalent" is highly context-dependent: That one word is used for many different relations in different parts of mathematics. Our language would be more precise if we gave slightly different names to different equivalence relations — e.g., if we distinguished between "α-equivalence" and "$\mathcal{F}$-equivalence" — but unfortunately that is not customary.

Here are a few examples of ways that equivalence relations can arise:

a. On any set X, the smallest equivalence relation is equality $(=)$. The largest equivalence relation is the universal relation, defined in 3.3.b; that is, $x \approx y$ for all x and y in X.

b. Let π be a function with domain X. Define $x_1 \approx x_2$ if $\pi(x_1) = \pi(x_2)$; we easily verify that this makes $\approx$ an equivalence relation on X. Actually, *every* equivalence relation can be expressed in this form, as shown in 3.11.

c. Let $\mathcal{S} = \{S_\lambda : \lambda \in \Lambda\}$ be a partition of a set X — that is, the sets S_λ are disjoint and their union is the set X. Call two elements of X equivalent if they belong to the same S_λ. It is easy to see that this is an equivalence relation on X. The S_λ's are then called the **equivalence classes** of the relation. Actually, *every* equivalence relation can be represented in this form, as shown below.

3.11. Let $\approx$ be *any* equivalence relation on a set X.

For each $x \in X$, let $\pi(x) = \{y \in X : y \approx x\}$. We easily verify that, for any $x, x' \in X$ the sets $\pi(x)$ and $\pi(x')$ are either identical or disjoint; hence the distinct sets of the form $\pi(x)$ form a partition $\mathcal{S}$ of the given set X. Moreover, the given equivalence relation $\approx$ can be retrieved from this partition, as in 3.10.c.

The surjective mapping $\pi : X \to \mathcal{S}$ is called the **quotient map** or **quotient projection**. The given equivalence relation $\approx$ can be retrieved from this mapping, as in 3.10.b.

The collection $\mathcal{S}$ of equivalence classes is called the **quotient set**. Represented in many ways, it is most often represented by an expression of the form X/δ, where δ is any device used to define the equivalence relation. Thus, the quotient set may be represented by

$X/\approx$ where $\approx$ is the equivalence relation,
X/π if $\approx$ is determined by a mapping π as in 3.10.b,
X/S if $\approx$ is determined by some subgroup S, as in 8.14,
$X/\mathcal{F}$ if $\approx$ is determined by a filter $\mathcal{F}$, as in 9.41,
$X/\mathcal{I}$ if $\approx$ is determined by an ideal $\mathcal{I}$, as in 9.41.

3.12. Let $\approx$ be an equivalence relation on a set X; let Q be the resulting quotient set and let $\pi : X \to Q$ be the quotient mapping.

A function f defined on X is said to **respect** the equivalence $\approx$ if the value of $f(x)$ is unchanged when x is replaced by an equivalent element of X — that is, if $x_1 \approx x_2 \Rightarrow f(x_1) = f(x_2)$. Another way to say this is that each set of the form $f^{-1}(z)$ is a union of equivalence classes. Similarly, a relation R on X is said to **respect** the equivalence $\approx$ if the validity of the statement $u\,R\,v$ is unaffected when u, v are replaced by equivalent elements of X — that is, if

$$u \approx u', \qquad v \approx v', \qquad u\,R\,v \qquad \Rightarrow \qquad u'\,R\,v'.$$

Show:

a. Let $f : X \to Y$ be some function. We can define a corresponding function $\widehat{f} : Q \to Y$ by the rule $\widehat{f}(\pi(x)) = f(x)$ if and only if f respects $\approx$. We then say that the function $\widehat{f}$ is **well defined**. The hat over the f is sometimes omitted; if no confusion will result, we sometimes use the same symbol f again for the new function defined on Q.

b. Let R be some relation on X. We can define a corresponding relation $\widehat{R}$ on Q by the rule

$$\pi(u)\ \widehat{R}\ \pi(v) \qquad \Longleftrightarrow \qquad u\ R\ v$$

if and only if R respects the equivalence relation $\approx$. We then say that the relation $\widehat{R}$ is **well defined**. The hat over the R is sometimes omitted: If no confusion will result, we sometimes use the same symbol R again for the new relation defined on Q.

c. *Example.* Let $\preccurlyeq$ be a preordering on a set X, and define $x \approx y$ to mean that $x \preccurlyeq y$ and $y \preccurlyeq x$. Show that $\approx$ is an equivalence relation on X, and that $\preccurlyeq$ respects this equivalence relation. Show that the resulting relation $\widehat{\preccurlyeq}$ is a partial ordering on the quotient set Q.

d. Let (X, d) be a pseudometric space (defined in 2.11). An equivalence relation $\approx$ on X can be defined by: $x \approx y$ if and only if $d(x, y) = 0$. Then d acts as a *metric* on the quotient space $X/\approx$. More precisely, let $\pi : X \to X/\approx$ be the quotient map; then a metric D on $X/\approx$ can be defined by $D(\pi(x), \pi(y)) = d(x, y)$.

More generally, let (X, D) be a gauge space. Define an equivalence relation on X by: $x \approx y$ if and only if $d(x, y) = 0$ for all pseudometrics $d \in D$. Then D acts as a *separating gauge* on the quotient space $X/\approx$.

3.13. The term "equivalent" also has some common uses that are implicit in our mathematical language:

Two words, phrases, or definitions are *equivalent* if they have the same meaning. This is an equivalence relation on the set of all words, phrases, or definitions in our vocabulary.

Similarly, two statements are *equivalent* if each implies the other via some set of rules of inference. This is an equivalence relation on the set of all statements that can be expressed in our mathematical language. Since different rules of inference may be used, there are actually several meanings for "equivalent statements." Here are two main interpretations:

- Many mathematicians call two statements "equivalent" if each implies the other *easily* — i.e., by a fairly short and elementary proof. Of course, "elementary" is a subjective term here; what is elementary for one mathematician may not be elementary for another. Most mathematicians do not make any restriction on the use of the Axiom of Choice; it may be used freely as a "rule of inference." An example: The mathematical literature sometimes refers to Caristi's Fixed Point Theorem 19.45 and Brönsted's Maximal Principle ((DC4) in 19.51) as "equivalent" because each implies the other easily; see 19.51. Strictly speaking, the relation "each implies the other easily" is *not* really an equivalence relation, for it is not transitive: If

$$\text{(A1)} \iff \text{(A2)}, \qquad \text{(A2)} \iff \text{(A3)}, \qquad \ldots, \qquad \text{(A99)} \iff \text{(A100)}$$

by 99 easy proofs, then (A1) $\iff$ (A100) by a proof that is not necessarily easy.

- Logicians sometimes give the Axiom of Choice special status and treat it as a statement rather than as a rule of inference. When this system is followed, then the Axiom of Choice or its consequences can only be used when stated explicitly as hypotheses. This system — which will be followed in parts of this book — enables us to trace the effects of the Axiom of Choice. For emphasis, statements equivalent in this sense are sometimes called **effectively equivalent**. See 6.18. With this interpretation, Caristi's Fixed Point Theorem and Brönsted's Maximal Principle are *not* equivalent; see the discussion in 19.51.

More about Posets

3.14. *Definitions.* Recall that a **partial order** is a relation $\preccurlyeq$ that is

reflexive $(x \preccurlyeq x$ for all $x)$,
transitive $(x \preccurlyeq y,\ y \preccurlyeq z \ \Rightarrow\ x \preccurlyeq z)$, and
antisymmetric $(x \preccurlyeq y$ and $y \preccurlyeq x$ imply $x = y)$.

A set equipped with such an ordering is a **partially ordered set**, or **poset**.

Let $(X, \preccurlyeq)$ be a partially ordered set. An **order interval** in X is a subset of the form

$$[a,b] \quad = \quad \{x \in X \ :\ a \preccurlyeq x \preccurlyeq b\}$$

for some $a, b \in X$.

In $\mathbb{R}$ or $[-\infty, +\infty]$, slightly different terminology is commonly used. An **interval** is any set of one of the following types:

$$\begin{aligned}
[a,b] &= \{x \in [-\infty,+\infty] \ :\ a \le x \le b\}, \\
[a,b) &= \{x \in [-\infty,+\infty] \ :\ a \le x < b\}, \\
(a,b] &= \{x \in [-\infty,+\infty] \ :\ a < x \le b\}, \\
(a,b) &= \{x \in [-\infty,+\infty] \ :\ a < x < b\},
\end{aligned}$$

for any extended real numbers a, b. In particular, the extended real line is the interval $[-\infty, +\infty]$ (thus justifying our notation), and the real line $\mathbb{R}$ is the interval $(-\infty, +\infty)$. Two other important sets are $[0, +\infty) = \{x \in \mathbb{R} : x \geq 0\}$ and $[0, +\infty] = \{x \in \mathbb{R} : x \geq 0\} \cup \{+\infty\}$.

An interval of the form $[a, b]$ is sometimes called a **closed interval**; an interval of the form (a, b) is an **open interval**. This terminology reflects the topological structure of $\mathbb{R}$ or $[-\infty, +\infty]$, introduced in 5.15.f.

3.15. Let X be a poset. A set $S \subseteq X$ is **order bounded** if it is contained in an order interval. It is simply called "bounded" if the context is clear, but be aware that the term "bounded" has other, possibly inequivalent meanings — see 4.40, 23.1, 27.2, and 27.4. Fortunately, all the usual meanings of "bounded" coincide at least for subsets of $\mathbb{R}^n$.

Note that any subset of an order bounded set is order bounded.

Although the statement "S is bounded" does not mention the set X explicitly, boundedness of a set $S \subseteq X$ depends very much on the choice of X. For instance, $\mathbb{Z}$ is unbounded when considered as a subset of $\mathbb{R}$ (with its usual ordering), but $\mathbb{Z}$ is bounded when considered as a subset of the extended real line $[-\infty, +\infty]$ (introduced in 1.17). In fact, *every* subset of $[-\infty, +\infty]$ is bounded, since $[-\infty, +\infty]$ itself is bounded.

3.16. Let $(X, \preccurlyeq)$ be a poset. A **lower set** in X is a set $S \subseteq X$ with the property that

$$x \preccurlyeq s, \quad x \in X, \quad s \in S \qquad \Rightarrow \qquad x \in S.$$

Some older books refer to lower sets as *initial segments* or *order ideals.*

Special examples and properties.

a. Clearly, X is a lower set in itself. Any other lower set is called a **proper lower set**.

b. One lower set is the **set of predecessors** of w, defined by

$$\mathrm{Pre}(w) \quad = \quad \{x \in X \ : \ x \prec w\}.$$

It is proper. It is empty if and only if $w = \min(X)$.

c. The **principal lower set** determined by any $w \in X$ is the set $\{x \in X : x \preccurlyeq w\}$. It is sometimes denoted by $\downarrow w$. It is nonempty. It is improper if and only if $w = \max(X)$.

Exercise. A lower set is equal to the union of all the principal lower sets that it contains.

d. The mapping $w \mapsto \downarrow w$, sending each element to its principal lower set, is an order isomorphism from $(X, \preccurlyeq)$ onto a subset of the poset $(\mathcal{P}(X), \subseteq)$. Thus any poset can be represented isomorphically in the form $(\mathcal{C}, \subseteq)$ for some collection $\mathcal{C}$ of sets.

Lower sets are discussed further in 4.4.b.

3.17. Let $(X, \preccurlyeq)$ and $(Y, \sqsubseteq)$ be partially ordered sets. A mapping $p : X \to Y$ is

increasing (*isotone, order-preserving*) if $x_1 \preccurlyeq x_2 \ \Rightarrow\ p(x_1) \sqsubseteq p(x_2)$;

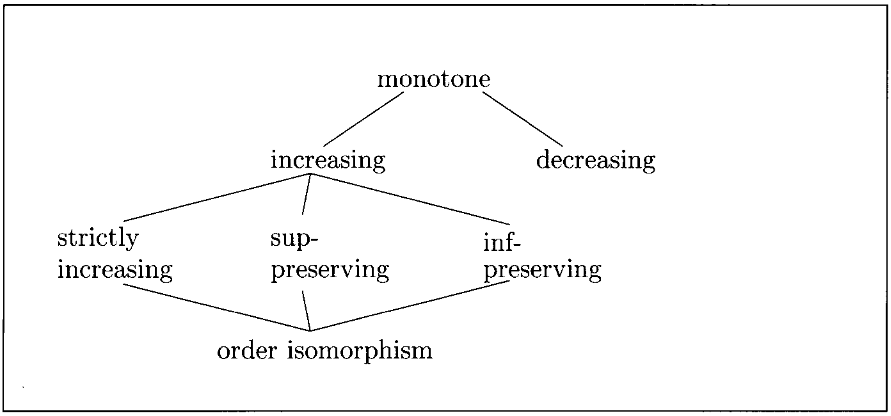
decreasing (*antitone, order-reversing*) if $x_1 \preccurlyeq x_2 \ \Rightarrow\ p(x_1) \sqsupseteq p(x_2)$;

monotone if it is increasing or decreasing;

strictly increasing or decreasing or monotone if it is injective and (respectively) increasing or decreasing or monotone;

an **order isomorphism** if it is a bijection from X onto Y such that both p and p^{-1} are increasing.

(The terms "isotone" and "antitone" are used especially if X and Y are collections of sets, ordered by inclusion.) The relationships between these kinds of mappings are explored in the next few exercises; a chart below summarizes the results. The chart also includes sup-preserving and inf-preserving, as a preview of notions that will be introduced in 3.22.



Caution: Some mathematicians use the terms *nondecreasing* or *weakly increasing* where we have used the term "increasing;" some of these mathematicians use the term "increasing" where we have used the term "strictly increasing." Analogous terminology is used for decreasing.

3.18. *Basic properties and examples.*

a. A sequence of real numbers $(r_1, r_2, r_3, \ldots)$ is increasing if $r_1 \leq r_2 \leq r_3 \leq \cdots$.

b. $S \mapsto \complement S$ is an antitone mapping from $(\mathcal{P}(X), \subseteq)$ into itself, for any set X.

c. The inverse of an increasing bijection need not be increasing. For instance, let $\leq$ be the usual ordering on $\mathbb{Z}$, and let $\preccurlyeq$ be the partial ordering on $\mathbb{Z}$ defined by

$$x \preccurlyeq y \qquad \text{if} \qquad y - x \in \{0, 5, 10, 15, 20, 25, \ldots\}.$$

Then the identity map $x \mapsto x$ is increasing from $(\mathbb{Z}, \preccurlyeq)$ into $(\mathbb{Z}, \leq)$, but not from $(\mathbb{Z}, \leq)$ into $(\mathbb{Z}, \preccurlyeq)$.

d. Let $f : X \to Y$ be any function. Then the forward image map $f : \mathcal{P}(X) \to \mathcal{P}(Y)$ and the inverse image map $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$, defined in 2.7 and 2.8, are both order-preserving — that is,

$$S_1 \subseteq S_2 \quad \Rightarrow \quad f(S_1) \subseteq f(S_2), \qquad T_1 \subseteq T_2 \quad \Rightarrow \quad f^{-1}(T_1) \subseteq f^{-1}(T_2).$$

MAX, SUP, AND OTHER SPECIAL ELEMENTS

3.19. *Definitions.* Let $(X, \preccurlyeq)$ be a partially ordered set, and let $y, z \in X$ and $S \subseteq X$.

a. We say z is an **upper bound** for S if $s \preccurlyeq z$ holds for each $s \in S$; we then say S is **bounded above**. We emphasize that z is not required to be an element of S.

Dually, z is a **lower bound** for S if $s \succcurlyeq z$ holds for each $s \in S$; we then say S is **bounded below**.

A set is order bounded (as defined in 3.15) if and only if it is bounded both above and below.

b. z is a **maximum** element of S (also known as a *greatest*, *largest*, *biggest*, *highest*, or *last* element of S) if $z \in S$ and $z \succcurlyeq s$ for all $s \in S$. Clearly, a subset of a poset has at most one maximum. If it exists, it is denoted by $\max(S)$.

Dually, z is a **minimum** element of S (also known as a *least*, *smallest*, *littlest*, *lowest*, or *first* element of S) if $z \in S$ and $z \preccurlyeq s$ for all $s \in S$. Again, a subset of a poset has at most one minimum; it may be denoted by $\min(S)$.

c. If $S \subseteq X$ is bounded above and the set of upper bounds of S has a least element, then that element is called the **least upper bound**, or **supremum** or **sup** of S. (Among algebraists, it is also known as the *join* of S.) It is denoted l.u.b.(S) or $\sup(S)$ or $\bigvee S$. If the elements of S are represented by subscripted notation, as in $S = \{x_\alpha : \alpha \in A\}$, then $\bigvee S$ may also be denoted by $\bigvee_{\alpha \in A} x_\alpha$. The sup of two elements x and y is also written as $x \vee y$. To be precise, the value $\sup(S)$ may be referred to as *the supremum of S in X*, for reasons indicated in 3.20.e.

Dually, if $S \subseteq X$ is bounded below and the set of lower bounds of S has a greatest element, then that element is called the **greatest lower bound**, or **infimum** or **inf** of S. (Among algebraists, it is also known as the *meet* of S.) It is denoted g.l.b.(S) or $\inf(S)$ or $\bigwedge S$. The infimum of a set $S = \{x_\alpha : \alpha \in A\}$ may also be denoted by $\bigwedge_{\alpha \in A} x_\alpha$. The inf of two elements x and y is also written as $x \wedge y$.

d. A **maximal** element of S is any $s_0 \in S$ with the property that no element of S is strictly greater than s_0.

Dually, a **minimal** element of S is any $s_0 \in S$ with the property that no element of S is strictly less than s_0.

3.20. *Further remarks and notational conventions.*

a. We emphasize that "max" and "min" are the abbreviations for "maximum" and "minimum," not "maximal" and "minimal."

It may be helpful to think of maximal elements and suprema as two kinds of "almost maximums" — i.e., objects with most of the properties one would find in a maximum. They can often be used in place of a maximum, in situations where a maximum is not available. (For instance, if we are trying to generalize some known theorem by modifying a known proof, we may at some point replace a maximum with a maximal element or a supremum.)

Analogously, a minimal element or an infimum is an "almost minimum."

b. We write "$\preccurlyeq$-upper bound," "$\preccurlyeq$-maximal," "$\preccurlyeq$-max(S)," "$\max_{\preccurlyeq}(S)$," etc., if we wish to emphasize or clarify which partial ordering is being used.

c. When the terms "max," "sup," etc., are applied to a collection of *sets* and no ordering is specified, then it is generally understood that $\subseteq$ is the ordering being used. Thus, for instance, a *maximal* element of a collection $\mathcal{F}$ of sets is an element of $\mathcal{F}$ that is not a subset of any other element of $\mathcal{F}$. Similarly, a *largest* element of $\mathcal{F}$ is an element of $\mathcal{F}$ that is a superset of every other element of $\mathcal{F}$. Note that a collection $\mathcal{F}$ of sets can only have a largest element if the union of all the elements of $\mathcal{F}$ is itself an element of $\mathcal{F}$ — in which case that union is the largest element. Similarly, if $\mathcal{F}$ has a smallest member, that smallest member is equal to the intersection of all the members of $\mathcal{F}$.

There are some slight similarities between our language for $(X, \preccurlyeq)$ and our language for $(\mathcal{P}(S), \subseteq)$ that may help in learning the vocabulary: $x \vee y$ is the join of two objects, while the union $A \cup B$ of two sets is obtained by "joining" them together. Also, $x \wedge y$ is the meet of two objects; two sets A and B are said to "meet" (in the sense of 1.26) if and only if their intersection $A \cap B$ is nonempty.

d. If f is a mapping from a set S into a poset, the expressions $\max f(S)$ and $\max_{s \in S} f(s)$ both mean $\max\{f(s) : s \in S\}$. Expressions for min, sup, and inf are interpreted analogously.

e. *Context dependence of the definitions.* The notation "sup(S)" does not mention X explicitly, but the value of sup(S) depends very much on the choice of the poset $(X, \preccurlyeq)$ in which S is a subset. For instance, let I denote the interval $[0, 1]$, and let

$$\begin{aligned} I^I &= \{\text{functions from } I \text{ into } I\}, \\ C(I,I) &= \{\text{continuous functions from } I \text{ into } I\}, \\ S &= \{f \in C(I,I) \ : \ f(0) = 0\}. \end{aligned}$$

Then $S \subsetneqq C(I,I) \subsetneqq I^I$. Let I^I be given the product ordering, and let $C(I,I)$ be given the restriction of that ordering. Then the supremum of S in $C(I,I)$ is the constant function 1, whereas the supremum of S in I^I is the characteristic function of the interval $(0, 1]$.

3.21. *Elementary examples and properties.* Let $(X, \preccurlyeq)$ be a partially ordered set. Let $z \in X$ and $S \subseteq X$. Then:

a. $z = \max S$ if and only if z is both an element of S and an upper bound for S.

b. $z = \min S$ if and only if z is both an element of S and a lower bound for S.

c. If max(S) exists, it is also the supremum of S and the only maximal element of S.

d. If min(S) exists, it is also the infimum of S and the only minimal element of S.

e. X itself is bounded (in its own ordering) if and only if it has both a maximum and a minimum.

f. Let $x, y \in X$. Then

$$x \preccurlyeq y \qquad \iff \qquad \max\{x,y\} = y \qquad \iff \qquad \sup\{x,y\} = y.$$

g. Suppose that $\sup(S)$ exists in X. Then

$$z \succcurlyeq \sup(S) \quad \Longleftrightarrow \quad z \succcurlyeq s \text{ for every } s \in S.$$

h. *Degenerate example.* $\varnothing$ is a bounded subset of X. Indeed, every element of X is an upper bound and a lower bound for the empty set, since the requirement involving $s \in S$ is vacuously satisfied when there are no s's.

The set $\varnothing$ has a least upper bound if and only if X has a first element, in which case those objects are the same. Similarly, $\varnothing$ has a greatest lower bound if and only if X has a last element, in which case those objects are the same.

Clearly, $\varnothing$ has neither a maximum nor minimum element, nor a maximal or minimal element, since it does not have any element.

i. A subset of a poset can have at most one maximum and at most one supremum. However, a subset of a poset may have more than one maximal element. For instance, let Ω be any set containing more than one element, and let $X = \{\text{proper subsets of } \Omega\}$ be partially ordered by inclusion. Then each complement of a singleton (i.e., each set of the form $\Omega \backslash \{\omega\}$) is a maximal element of X.

j. A subset of a poset may have an upper bound without having a maximum. For instance, let $X = \mathbb{Z}^2$ have the product ordering. Then the subset $S = \{(0,-1),(-1,0)\}$ has no maximum element, but it has $(0,0)$ as an upper bound.

k. A subset of a poset need not have any maximal elements. For instance, let X be the real line with its usual ordering. Then the set $S = \{x \in \mathbb{R} : x < 0\}$ has no maximal element, but it has 0 as a supremum. The set $\mathbb{R}$, considered as a subset of itself, has no maximum, no maximal element, and no supremum.

l. Let $(X, \preccurlyeq)$ be a poset. Then sup is an isotone map, and inf is an antitone map, from their domains into X. That is:

$$A \subseteq B \subseteq X \qquad \Rightarrow \qquad \sup A \preccurlyeq \sup B, \qquad \inf A \succcurlyeq \inf B$$

whenever those sups and infs exist.

m. *Proposition.* Suppose that $\{S_\alpha : \alpha \in A\}$ is a collection of nonempty subsets of X and $\inf(S_\alpha)$ exists for each α. Then $\inf\{\inf(S_\alpha) : \alpha \in A\}$ exists if and only if $\inf(\bigcup_{\alpha \in A} S_\alpha)$ exists, in which case they are equal. (Analogous results hold for sups.)

Hint: Show that ρ is a lower bound for $\{\inf(S_\alpha) : \alpha \in A\}$ if and only if ρ is a lower bound for $\bigcup_{\alpha \in A} S_\alpha$.

n. Let $P = \prod_{\lambda \in \Lambda} X_\lambda$ be a product of posets, with the product ordering (see 3.9.j). Let Φ be a nonempty subset of P. Verify that $\sup \Phi$ exists in P if and only if the set $\{f(\lambda) : f \in \Phi\}$ has a supremum in X_λ for each λ — in which case $\sup \Phi$ is a function defined on Λ by

$$(\sup \Phi)(\lambda) \quad = \quad \sup\{f(\lambda) : f \in \Phi\} \qquad \text{for each } \lambda \in \Lambda.$$

Thus, the supremum in P is defined coordinatewise. We shall call it the **pointwise supremum**, or sometimes simply the **supremum**, of the set Φ in P. We emphasize that $\sup \Phi$ is a member of P but not necessarily a member of Φ. Analogous notations are used for inf, max, and min. In particular, when Φ contains just two functions, we obtain

$$(x \vee y)(\lambda) = x(\lambda) \vee y(\lambda), \qquad (x \wedge y)(\lambda) = x(\lambda) \wedge y(\lambda).$$

3.22. Let $(X, \preccurlyeq)$ and $(Y, \sqsubseteq)$ be partially ordered sets. A mapping $p : X \to Y$ is

sup-preserving if, whenever S is a nonempty subset of X and $\sigma = \sup(S)$ exists in $(X, \preccurlyeq)$, then $\sup\{p(s) : s \in S\}$ exists in $(Y, \sqsubseteq)$ and is equal to $p(\sigma)$;

inf-preserving if, whenever S is a nonempty subset of X and $\iota = \inf(S)$ exists in $(X, \preccurlyeq)$, then $\inf\{p(s) : s \in S\}$ exists in $(Y, \sqsubseteq)$ and equals $p(\iota)$.

These are special kinds of *increasing* maps; see 3.17. Some basic properties follow.

a. Any order isomorphism is sup- and inf-preserving and strictly increasing.

b. Any sup- or inf-preserving map is also increasing. *Hint*: 3.21.f.

c. *Examples.* The inclusion maps $C(I, I) \xrightarrow{\subseteq} I^I$ in 3.20.e, $V \xrightarrow{\subseteq} \mathbb{R}^3$ in 4.21, and $ba(\mathcal{A}) \xrightarrow{\subseteq} \mathbb{R}^{\mathcal{A}}$ in 11.47 are order-preserving, but they are not sup-preserving or inf-preserving. The inclusion $\mathcal{T} \xrightarrow{\subseteq} \mathcal{P}(X)$ given in 5.21 is sup-preserving but not inf-preserving.

Chains

3.23. *Definition.* Let $(X, \preccurlyeq)$ be a poset. Then the following conditions are equivalent. If any, hence all, are satisfied, we say that $(X, \preccurlyeq)$ is a **chain** (or $\preccurlyeq$ is a *total order* or *linear order* or *chain order*).

(A) Any two elements of X are comparable (defined in 3.9.a).

(B) Each two-element subset of X has a first element.

(C) Each two-element subset of X has a last element.

(D) Each nonempty finite subset of X has a first element.

(E) Each nonempty finite subset of X has a last element.

(F) $(X, \preccurlyeq)$ satisfies the **Trichotomy Law**: for each $x, y \in X$, exactly one of the three conditions

$$x \prec y, \qquad y \prec x, \qquad x = y$$

holds. In other words, the sets $\mathrm{Graph}(\prec)$, $\mathrm{Graph}(\succ)$, and I form a partition of $X \times X$.

3.24. *Some important examples.* The number systems $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq [-\infty, +\infty]$ play a major role in analysis. We shall give formal introductions to $\mathbb{Q}$ and $\mathbb{R}$ in later chapters, but for now we assume that the reader is already familiar with these number systems at least informally. The reader should understand arithmetic and inequalities in $\mathbb{R}$.

All of the number systems $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are chains. Indeed, $\mathbb{R}$ is a chain, and all the inclusions $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$ are order-preserving.

3.25. *Elementary properties.*

a. Any subset of a chain is a chain.

b. If $(X, \preccurlyeq)$ is a chain, then $(X, \succcurlyeq)$ is a chain.

c. A product of chains, with the product ordering (from 3.9.j), is not necessarily a chain. For instance, $\mathbb{N}^2$ is not a chain.

Certain other orderings on a product may be chains or well orderings; see 3.44.

d. Suppose $(X, \leq)$ is a chain and $S \subseteq X$ with $\sigma = \sup(S)$. Then for each $x \in X$ with $x < \sigma$ there exists some $s \in S$ with $x < s \leq \sigma$.

3.26. A **total preorder** on a set X is a preorder $\preccurlyeq$ (i.e., a reflexive, transitive relation) that also has this property:

> Any two elements $x, y \in X$ are comparable — i.e., at least one of the relations $x \preccurlyeq y$ or $y \preccurlyeq x$ holds.

Observe that a total preorder is in fact a total order if and only if it is antisymmetric.

Let $\preccurlyeq$ be a total preorder on X; then:

a. An equivalence relation is given on X by this rule: $x \approx y$ if both $x \preccurlyeq y$ and $y \preccurlyeq x$.

b. $\preccurlyeq$ defines a total order on the equivalence classes, i.e., on the quotient set $X/\approx$.

c. $\preccurlyeq$ can be extended to a total order $\leq$ on X (so that $\mathrm{Graph}(\preccurlyeq) \supseteq \mathrm{Graph}(\leq)$) by this natural method: Define a chain ordering $\leq$ arbitrarily within each of the equivalence classes. When x and y are not equivalent, say $x \leq y$ if and only if $x \preccurlyeq y$.

3.27. The reader may be better able to appreciate transitivity and chains after considering **Condorcet's Paradox**:

> Even if we assume that each individual voter's preferences are ranked in a chain ordering, the preferences of a collection of voters (determined by majority rule) are not necessarily a chain ordering — they need not be transitive!

For instance, a recent presidential election in the United States had three main candidates: Bush, Clinton, and Perot, hereafter represented by B, C, P. (For those readers who are not interested in politics, ask which fruit is preferred: banana, cherry, or peach; the mathematics is the same.) Before the election, I took a "straw poll" and asked my students which candidate they preferred. The class preferred Clinton over Bush; the class preferred Bush over Perot; but the class preferred Perot over Clinton! How is this possible? The following chart shows the details.

Each individual voter's preferences are given by a chain ordering of the three candidates. There are six possible chain orderings of the candidates. For instance, one ordering is: Bush

	B C P	B P C	C B P	C P B	P B C	P C B	Sum	Result
3-way	1	5	6	1	2	7	22	
C > B			6	1		7	14	Clinton
C < B	1	5			2		8	beats Bush
B > P	1	5	6				12	Bush
B < P				1	2	7	10	beats Perot
P > C		5			2	7	14	Perot
P < C	1		6	1			8	beats Clinton

is first choice; Clinton is second choice; Perot is third choice. That ordering is represented by "B C P," in the first column. The row labeled "3-way" shows how many members of the class have that chain ordering; thus, that row shows the votes that would be cast in a contest between all three candidates. For instance, just one of the 22 voters had a "B C P" preference, so the number 1 appears in the "B C P" column, in the "3-way" row.

Below the "3-way" row are rows showing the results of contests between any two candidates. The totals for each contest are in the column with the heading "sum." For instance, in a contest between Bush and Clinton, 14 voters preferred Clinton over Bush, while 8 voters preferred Bush over Clinton. Thus we obtain the result "C > B."

Of course, this situation can arise with other numbers of voters and other numbers of candidates.

Exercise. The simplest case of Condorcet's Paradox involves 3 candidates and 3 voters. Work out the details.

This type of paradox was first published by Condorcet in 1785. A characterization of the combinations of numbers that yield Condorcet's paradox, and further references, were given by Weber [1993]. A generalization to infinite sets of voters (with majority rule replaced by other kinds of rule) were studied be Haddad [1989]; further considerations about finite or infinite sets of voters can also be found in Kirman and Sondermann [1972].

Van Maaren's Geometry-Free Sperner Lemma

3.28. *Discussion and preview.* The main result of this subchapter is a technical combinatorial result about preordered sets:

> **Van Maaren's Theorem.** Let $\ell : X \to P$ be some given function, where P and X are nonempty sets and P is finite. For each $p \in P$, assume $\preccurlyeq_p$ is a total preordering of X. Then there exists a function σ from some nonempty subset

of P into X, satisfying:

- $\sigma(q) \preccurlyeq_q \sigma(r)$ for all $q, r \in \mathrm{Dom}(\sigma)$.
- There is no $x \in X$ that satisfies $\sigma(q) \prec_q x$ for all $q \in \mathrm{Dom}(\sigma)$.
- $\mathrm{Dom}(\sigma) = \ell(\mathrm{Ran}(\sigma))$.

This theorem is due to Maaren [1987]; our presentation is based on the exposition given by van de Vel (see Vel [1993]). The proof will take several pages and will require several more definitions and preliminary results. We complete the proof of the theorem in 3.36, and follow it with a corollary about approximate fixed points in 3.37.

This material may be postponed. It is rather specialized and will not be used until 27.19, where we use it to prove Brouwer's Fixed Point Theorem and related results. We include van Maaren's argument this early in the book mainly in order to emphasize how elementary it is — i.e., to show that it does not depend on topology or geometry. For an abridged treatment, readers who are willing to skip some proofs may proceed directly to 3.37; the other ideas in this subchapter will not be needed elsewhere in this book.

The literature contains many different proofs of Brouwer's Theorem. Some of the proofs may appear short or elementary, but that is only because they have concealed some of the difficulty — usually by using some well-known but nontrivial theorem, about measures and Jacobian determinants or about the algebraic topology of simplicial triangulations. Those proofs, when carried out in detail, are (in this author's opinion) non-intuitive; they involve n-dimensional diagrams that are hard to visualize and that seem to have little to do with the central ideas of Brouwer's Theorem. Van Maaren's proof, though not shorter or simpler than the other proofs, avoids such drawbacks. Our presentation separates the proof of Brouwer's Theorem into two main components: a purely combinatorial result in 3.37 and a compactness argument in 27.19.

3.29. *Notations and definitions.* The cardinality of a set S will be denoted $|S|$. The symmetric difference of two sets S, T will be denoted $S \triangle T$. The domain and range of a function σ will be denoted, respectively, by $\mathrm{Dom}(\sigma)$ and $\mathrm{Ran}(\sigma)$.

Throughout this subchapter, we assume some nonempty sets P and X are given, with P finite. Also, we assume some mapping $\ell : X \to P$ is given; we shall call this function the **labeling**. An **assignment** will mean a function

$$\sigma \;:\; \mathrm{Dom}(\sigma) \to X, \qquad \text{where } \mathrm{Dom}(\sigma) \text{ is a nonempty subset of } P.$$

Note that any assignment has a finite domain, hence also a finite range. An assignment σ is **complete** (with respect to ℓ) if $\mathrm{Dom}(\sigma) = \ell(\mathrm{Ran}(\sigma))$. An assignment σ will be called **almost complete** if $|\mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))| \leq 1$. Two assignments σ_1, σ_2 will be called **neighbors** if either

$$\mathrm{Dom}(\sigma_1) = \mathrm{Dom}(\sigma_2) \text{ and } |\mathrm{Ran}(\sigma_1) \triangle \mathrm{Ran}(\sigma_2)| = 1, \text{ or}$$

$$\mathrm{Ran}(\sigma_1) = \mathrm{Ran}(\sigma_2) \text{ and } |\mathrm{Dom}(\sigma_1) \triangle \mathrm{Dom}(\sigma_2)| = 1.$$

3.30. *Observations.* For any assignment σ,

a. $|\ell(\mathrm{Ran}(\sigma))| \leq |\mathrm{Ran}(\sigma)| \leq |\mathrm{Dom}(\sigma)|$. Since $|S| - |T| = |S \setminus T| - |T \setminus S|$ for any finite sets S and T, we have

$$\begin{aligned} 0 \ \leq \ & |\mathrm{Dom}(\sigma)| - |\ell(\mathrm{Ran}(\sigma))| \\ = \ & |\mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))| - |\ell(\mathrm{Ran}(\sigma)) \setminus \mathrm{Dom}(\sigma)| \\ & \leq \ |\mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))|. \end{aligned}$$

b. If $|\mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))| = 0$ then σ is complete.

c. If σ is almost complete, then $0 \leq |\mathrm{Dom}(\sigma)| - |\ell(\mathrm{Ran}(\sigma))| \leq 1$.

d. Let us abbreviate $\widehat{\mathrm{D}} = |\mathrm{Dom}(\sigma)|$, $\widehat{\mathrm{R}} = |\mathrm{Ran}(\sigma)|$, $\widehat{\ell} = |\ell(\mathrm{Ran}(\sigma))|$. The almost complete assignments can be classified into the complete assignments and three types of noncomplete assignments:

	Complete	Type (i)	Type (ii)	Type (iii)
Is σ injective?	yes	no	yes	yes
ℓ injective on $\mathrm{Ran}(\sigma)$?	yes	yes	no	yes
$\widehat{D}, \widehat{R}, \widehat{\ell}$:	$\widehat{D} = \widehat{R}$ $= \widehat{\ell}$	$\widehat{D} - 1 =$ $\widehat{R} = \widehat{\ell}$	$\widehat{D} = \widehat{R} =$ $\widehat{\ell} + 1$	$\widehat{D} = \widehat{R}$ $= \widehat{\ell}$
$\lvert\ell(\mathrm{Ran}(\sigma)) \setminus \mathrm{Dom}(\sigma)\rvert =$	0	0	0	1

3.31. If σ is an almost complete assignment that is not complete, then $\mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))$ contains exactly one element. We shall call it the **extraneous element** for σ.

Proposition. Suppose that σ and σ' are assignments that are almost complete but not complete and they are neighbors. Then they have the same extraneous element.

Proof. Suppose that σ and σ' have extraneous elements p and p', respectively, where $p \neq p'$; we shall arrive at a contradiction.

Since $\mathrm{Dom}(\sigma)$ and $\mathrm{Dom}(\sigma')$ differ by at most one element, one must contain the other; say $\mathrm{Dom}(\sigma) \subseteq \mathrm{Dom}(\sigma')$. Since $p \in \mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))$, we have $p \in \mathrm{Dom}(\sigma')$. Since p is not the extraneous element of σ', we have $p \in \ell(\mathrm{Ran}(\sigma'))$. Since $p \notin \ell(\mathrm{Ran}(\sigma))$, the sets $\mathrm{Ran}(\sigma')$ and $\mathrm{Ran}(\sigma)$ are different, and therefore $\mathrm{Dom}(\sigma') = \mathrm{Dom}(\sigma)$. Since the sets $\mathrm{Ran}(\sigma')$ and $\mathrm{Ran}(\sigma)$ differ by at most one element, we have $\ell(\mathrm{Ran}(\sigma)) \subseteq \ell(\mathrm{Ran}(\sigma'))$. Then

$$p' \ \in \ \mathrm{Dom}(\sigma') \setminus \ell(\mathrm{Ran}(\sigma')) \ \subseteq \ \mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma)),$$

and so p' is an extraneous element of σ — a contradiction.

3.32. *More assumptions and definitions.* In the remainder of this subchapter we assume

$$P \text{ is the index set for a collection } \{\preccurlyeq_p : p \in P\} \text{ of total preorderings of } X.$$

(Recall from 3.26 that a preordering of X is *total* if it makes every two elements of X comparable. In several of the next few sections we make the additional assumption that all the $\preccurlyeq_p$'s are antisymmetric — i.e., they are chain orderings — but in 3.36 we drop that restriction.)

An assignment σ will be called a **crystal** if it satisfies these two conditions:

(CR-1) $\sigma(q) \preccurlyeq_q \sigma(r)$ for all $q, r \in \mathrm{Dom}(\sigma)$.

(CR-2) There is no $x \in X$ that satisfies $\sigma(q) \prec_q x$ for all $q \in \mathrm{Dom}(\sigma)$.

Let $\mathcal{C}$ be the set of almost complete crystals. A crystal σ with $|\mathrm{Dom}(\sigma)| = k$ will be called a ***k*-crystal**.

3.33. *Observations.* Assume all the $\preccurlyeq_p$'s are antisymmetric. Then:

a. Condition (CR-1) in 3.32 can be restated as:

$$\sigma(q) \text{ is the } \preccurlyeq_q\text{-smallest member of } \mathrm{Ran}(\sigma).$$

Thus, a crystal is uniquely determined by its domain and range.

b. Any 1-crystal is almost complete.

c. Let σ be an assignment whose domain is a singleton — i.e., $\mathrm{Dom}(\sigma) = \{p\}$ for some $p \in P$. Then σ is a crystal if and only if $\sigma(p)$ is the $\preccurlyeq_p$-largest member of X.

d. A 1-crystal is uniquely determined by its domain.

e. Suppose that $\sigma_1, \sigma_2 \in \mathcal{C}$ with $\mathrm{Ran}(\sigma_1) \subseteq \mathrm{Ran}(\sigma_2)$. Then σ_1, σ_2 agree (i.e., take the same values) on $\mathrm{Dom}(\sigma_1) \cap \sigma_2^{-1}(\mathrm{Ran}(\sigma_1))$.

Hint: Let $q \in \mathrm{Dom}(\sigma_1) \cap \sigma_2^{-1}(\mathrm{Ran}(\sigma_1))$. By (CR-1), $\sigma_j(q)$ is the $\preccurlyeq_q$-smallest member of $\mathrm{Ran}(\sigma_j)$ for $j = 1, 2$.

f. A special case of the preceding result is as follows: Suppose that $\sigma_1, \sigma_2 \in \mathcal{C}$ with $\mathrm{Ran}(\sigma_1) = \mathrm{Ran}(\sigma_2)$. Then σ_1, σ_2 agree on $\mathrm{Dom}(\sigma_1) \cap \mathrm{Dom}(\sigma_2)$.

g. Suppose that τ and τ' are neighboring almost complete crystals. Then one of τ, τ' is injective and the other is not. If τ is injective and τ' is not, then τ and τ' must be related in one of these two ways:

(a) $\mathrm{Ran}(\tau') = \mathrm{Ran}(\tau)$ and $\mathrm{Dom}(\tau') \supsetneqq \mathrm{Dom}(\tau)$. In this case τ and τ' agree on $\mathrm{Dom}(\tau)$.

(b) $\mathrm{Dom}(\tau) = \mathrm{Dom}(\tau')$ and $\mathrm{Ran}(\tau) \supsetneqq \mathrm{Ran}(\tau')$. In this case τ and τ' agree at all but one point of $\mathrm{Dom}(\tau)$.

Hints: Use 3.30.d, 3.33.e, and 3.33.f.

3.34. *Proposition.* Assume all the $\preccurlyeq_p$'s are antisymmetric. Then any noncomplete 1-crystal τ has precisely one neighbor τ' in $\mathcal{C}$.

Proof. Any 1-crystal τ is injective. Clearly τ' cannot have empty range, so 3.33.g(b) is not possible. Thus we must have $\mathrm{Ran}(\tau') = \mathrm{Ran}(\tau)$ and $\mathrm{Dom}(\tau') \supsetneqq \mathrm{Dom}(\tau)$. Say τ has graph $\{(q, b)\}$; then $\mathrm{Graph}(\tau') = \{(q, b), (q', b)\}$ for some $q' \neq q$. For τ' to be almost complete, it must satisfy $|\mathrm{Dom}(\tau') \setminus \ell(\mathrm{Ran}(\tau'))| \leq 1$; that is, $|\{q, q'\} \setminus \{\ell(b)\}| \leq 1$. Therefore at least one of q, q' must equal $\ell(b)$. By assumption (q, b) is not complete, so $\ell(b) \neq q$. Thus we must have $q' = \ell(b)$. Finally, we easily verify that $\tau' = \{(q,b), (\ell(b), b)\}$ is indeed a member of $\mathcal{C}$.

3.35. *Proposition.* Assume all the $\preccurlyeq_p$'s are antisymmetric. Then for $k \geq 2$, a noncomplete k-crystal $\sigma \in \mathcal{C}$ has precisely two neighbors in $\mathcal{C}$.

Proof. We analyze the possible values for a neighbor σ'. We consider several cases, according to the type of σ (with types as listed in 3.30.d).

σ is of Type (i). In this case σ is not injective. There is one and only one pair of elements p_1, p_2 in $\mathrm{Dom}(\sigma)$ such that $p_1 \neq p_2$ and $\sigma(p_1) = \sigma(p_2)$. We shall obtain one neighbor of σ from each of these two points.

Let p be one of p_1, p_2. We shall obtain a neighbor by either modifying or removing $\sigma(p)$ — i.e., by either changing the definition of the function at p or removing p from the domain. We do this in two cases, according to whether there does or does not exist a solution $x \in X$ to this problem:

$$(*) \qquad \sigma(q) \prec_q x \text{ for all } q \in \mathrm{Dom}(\sigma) \setminus \{p\}.$$

If $(*)$ has any solutions, let v be the $\preccurlyeq_p$-largest of those solutions. Note that $v \preccurlyeq_p \sigma(p)$, since otherwise σ and v would contradict (CR-2). Now a neighbor σ' can be defined with $\mathrm{Dom}(\sigma') = \mathrm{Dom}(\sigma)$, by taking

$$\sigma'(q) = \begin{cases} \sigma(q) & \text{when } q \neq p \\ v & \text{when } q = p. \end{cases}$$

On the other hand, if $(*)$ has no solution, then a neighbor σ' can be defined by just restricting σ to a smaller domain — i.e., taking $\mathrm{Dom}(\sigma') = \mathrm{Dom}(\sigma) \setminus \{p\}$ and taking σ' equal to σ on $\mathrm{Dom}(\sigma')$. It is tedious but straightforward to verify that the function σ' defined in either of these fashions is a neighboring almost complete crystal.

Thus we obtain one neighbor by either modifying or removing $\sigma(p_1)$ and another by either modifying or removing $\sigma(p_2)$. Now we shall show that there are no other neighbors possible besides those two.

Let σ' be a neighbor of σ in $\mathcal{C}$; what form can σ' take? By 3.33.g, σ' is injective, and there are two cases to consider:

(1) $\mathrm{Ran}(\sigma) = \mathrm{Ran}(\sigma')$ and $\mathrm{Dom}(\sigma) = \{p\} \cup \mathrm{Dom}(\sigma')$ for some $p \notin \mathrm{Dom}(\sigma')$, and σ' and σ agree on $\mathrm{Dom}(\sigma')$. Since σ' is injective, p must be one of p_1, p_2. Since σ' is a crystal, by (CR-2) we know that there is no $x \in X$ satisfying $\sigma(q) \prec_q x$ for all $q \in \mathrm{Dom}(\sigma')$. Thus there is no solution of problem $(*)$, and the function σ' can only be the one obtained by removing $\sigma(p)$.

(2) $\mathrm{Dom}(\sigma) = \mathrm{Dom}(\sigma') = D$ and $\mathrm{Ran}(\sigma') = \mathrm{Ran}(\sigma) \cup \{\sigma'(p)\}$ for some $p \in D$ with $\sigma'(p) \notin \mathrm{Ran}(\sigma)$, and σ and σ' agree on $D \setminus \{p\}$. Since $\mathrm{Ran}(\sigma) \subseteq \mathrm{Ran}(\sigma')$, we have $\sigma(p) = \sigma'(q_1)$ for some $q_1 \in D$. Since $\sigma(p)$ belongs to $\mathrm{Ran}(\sigma)$ and $\sigma'(p)$ does not, we know $\sigma'(p) \neq \sigma(p) = \sigma'(q_1)$, and therefore $p \neq q_1$. Since σ and σ' agree on $D \setminus \{p\}$, we have $\sigma(q_1) = \sigma'(q_1) = \sigma(p)$. Thus p and q_1 are distinct members of D that are mapped to the same value by σ. Therefore the set $\{p, q_1\}$ is equal to the set $\{p_1, p_2\}$. Hence p is one of p_1, p_2. Since σ' satisfies (CR-1) and σ' is injective, we have $\sigma'(q) \prec_q \sigma'(p)$ for all $q \in \mathrm{Dom}(\sigma') \setminus \{p\}$.

That is, $\sigma(q) \prec_q \sigma'(p)$ for all $q \in \mathrm{Dom}(\sigma') \setminus \{p\}$, so $\sigma'(p)$ is a solution of $(*)$. To see that $\sigma'(p)$ must be the $\preccurlyeq_p$-largest solution of $(*)$, suppose that x is a $\preccurlyeq_p$-larger solution. Then $\sigma(q) \prec_q x$ for all $q \in \mathrm{Dom}(\sigma) \setminus \{p\}$, and $\sigma'(p) \prec_p x$ as well. That is, $\sigma'(q) \prec_q x$ for all $x \in \mathrm{Dom}(\sigma')$, contradicting the fact that σ' must satisfy (CR-2). Thus, we have established that there is a solution of problem $(*)$, and the function σ' can only be the one obtained by modifying $\sigma(p)$.

σ is of Type (ii). In this case σ is injective, but ℓ is not injective on $\mathrm{Ran}(\sigma)$. Thus $|\mathrm{Dom}(\sigma)| = |\mathrm{Ran}(\sigma)| > |\ell(\mathrm{Ran}(\sigma))|$. There is a unique pair of distinct elements $w_1, w_2 \in \mathrm{Ran}(\sigma)$ that get mapped by ℓ to the same value. There are unique elements $p_1, p_2 \in \mathrm{Dom}(\sigma)$ with $\sigma(p_j) = w_j$. We shall obtain one neighbor of σ from each of these two points.

If $\mathrm{Dom}(\sigma') \supsetneqq \mathrm{Dom}(\sigma)$ and $\mathrm{Ran}(\sigma') = \mathrm{Ran}(\sigma)$, then we have $|\mathrm{Dom}(\sigma')| - 1 \geq |\mathrm{Dom}(\sigma)| \geq |\ell(\mathrm{Ran}(\sigma))| + 1 = |\ell(\mathrm{Ran}(\sigma'))| + 1$, contradicting 3.30.c. Thus 3.33.g(a) cannot hold.

Therefore $\mathrm{Dom}(\sigma) = \mathrm{Dom}(\sigma')$ and $\mathrm{Ran}(\sigma) = \mathrm{Ran}(\sigma') \cup \{\sigma(p)\}$ for some p with $\sigma(p) \notin \mathrm{Ran}(\sigma')$, and σ and σ' agree on the set $S = \mathrm{Dom}(\sigma) \setminus \{p\}$. Since $\mathrm{Ran}(\sigma') \subsetneqq \mathrm{Ran}(\sigma)$, we must have $\sigma'(p) \in \sigma(S)$.

If $p \notin \{p_1, p_2\}$, then p_1, p_2 are distinct members of S with $\ell(\sigma'(p_1)) = \ell(\sigma'(p_2))$, and $\ell(\sigma'(p)) \in \ell(\sigma'(S))$, too. It follows that $|\ell(\mathrm{Ran}(\sigma'))| \leq |\mathrm{Dom}(\sigma')| - 2$, contradicting 3.30.c. Thus we must have $p \in \{p_1, p_2\}$.

For each of those two choices of p, the value of $\sigma'(p)$ is determined uniquely: By 3.33.a, $\sigma'(p)$ must be equal to the $\preccurlyeq_p$-lowest member of $\mathrm{Ran}(\sigma')$.

We have shown that only two functions (one with $p = p_1$, the other with $p = p_2$) could possibly be almost complete crystals that neighbor σ. It is easy to verify that both of those two functions are, indeed, such crystals.

σ is of Type (iii). In this case σ is injective, and ℓ is injective on $\mathrm{Ran}(\sigma)$. We shall obtain one neighbor of σ from the unique member of $\ell(\mathrm{Ran}(\sigma)) \setminus \mathrm{Dom}(\sigma)$ and another from the unique member of $\mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))$. By 3.33.g, we can obtain a neighbor only by enlarging the domain or decreasing the range; we shall show that each of these two methods yields precisely one neighbor.

(1) Enlarging the domain: In this case $\mathrm{Dom}(\sigma') = \mathrm{Dom}(\sigma) \cup \{p\}$ for some $p \notin \mathrm{Dom}(\sigma)$, and $\mathrm{Ran}(\sigma') = \mathrm{Ran}(\sigma) = R$, and σ and σ' agree on $\mathrm{Dom}(\sigma)$. Since $p \notin \mathrm{Dom}(\sigma)$, p is not the extraneous element of σ, and therefore p is not the extraneous element of σ'. Hence $p \in \ell(\mathrm{Ran}(\sigma')) = \ell(R) = \ell(\mathrm{Ran}(\sigma))$. Thus p is the unique member of $\ell(R) \setminus \mathrm{Dom}(\sigma)$. By 3.33.a, $\sigma'(p)$ must be the $\preccurlyeq_p$-least member of R. Thus we have specified σ' uniquely. It is easy to verify that the function σ' defined in this fashion is indeed an almost complete neighboring crystal.

(2) Decreasing the range: In this case $\mathrm{Dom}(\sigma) = \mathrm{Dom}(\sigma') = D$ and $\mathrm{Ran}(\sigma') = \mathrm{Ran}(\sigma) \setminus \{\sigma(p)\}$ for some $p \in D$, and σ and σ' agree on $D \setminus \{p\}$. Since ℓ is injective on the range of σ, we have $\ell(\sigma(p)) \in \ell(\mathrm{Ran}(\sigma)) \setminus \ell(\mathrm{Ran}(\sigma'))$. Since $\ell(\sigma(p)) \in \ell(\mathrm{Ran}(\sigma))$, we have $\ell(\sigma(p)) \notin \mathrm{Dom}(\sigma) \setminus \ell(\mathrm{Ran}(\sigma))$. That is, $\ell(\sigma(p)) \notin \mathrm{Dom}(\sigma') \setminus \ell(\mathrm{Ran}(\sigma'))$, since σ and σ' have the same extraneous point. But $\ell(\sigma(p)) \notin \ell(\mathrm{Ran}(\sigma'))$, so we conclude $\ell(\sigma(p)) \notin \mathrm{Dom}(\sigma') = \mathrm{Dom}(\sigma)$. Thus we have identified $\ell(\sigma(p))$ uniquely: it is the unique member of $\ell(\mathrm{Ran}(\sigma)) \setminus \mathrm{Dom}(\sigma)$. Since ℓ and σ are injective, we have determined p uniquely: It is the unique member of $\mathrm{Dom}(\sigma)$ that satisfies $\ell(\sigma(p)) \notin \mathrm{Dom}(\sigma)$. The functions σ and σ' agree on $D \setminus \{p\}$, and the value of $\sigma'(p)$ is determined uniquely by 3.33.a. Thus we have defined σ' uniquely. It is a

tedious but straightforward exercise to verify that the function σ' defined in this fashion is indeed an almost complete crystal that neighbors σ.

3.36. Van Maaren's Theorem. Suppose that P and X are finite sets and for each $p \in P$ we are given a total preorder $\preccurlyeq_p$ on X (not necessarily antisymmetric). Let any labeling $\ell : X \to P$ be given. Then (X, P) has at least one complete crystal with respect to ℓ.

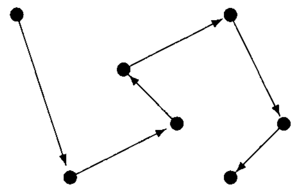


Diagram for proof of 3.36

Proof. A preliminary first step is this: We can replace each total preorder $\preccurlyeq_p$ with a total order, hereafter denoted $\preccurlyeq_p$, by arbitrarily choosing a total ordering on each of the equivalence classes of $\preccurlyeq_p$. This replacement results in fewer crystals. Thus, it suffices to prove the theorem under the additional assumption that each preordering $\preccurlyeq_p$ is a total ordering.

Since X is a finite set, $\mathcal{C}$ is finite also. By 3.33.c, there exists a 1-crystal σ_0, with a singleton domain $\mathrm{Dom}(\sigma_0) = \{p_0\}$. In $\mathcal{C}$, each incomplete 1-crystal has exactly one neighbor, and each incomplete 2-crystal has exactly two neighbors. Follow a path, starting at σ_0, going from each crystal to its neighbor. If we do not encounter any complete crystals along the path, then our route is uniquely determined; it must begin and end at distinct 1-crystals (see the preceding diagram). However, at each step the extraneous point is preserved, by 3.31; thus the beginning and ending 1-crystals must have the same extraneous point — contradicting the fact that they are distinct. This proves that the path must include at least one complete crystal. (Incidentally, we have given a constructive *algorithm* for finding a complete crystal: just follow the path until one is encountered.)

3.37. The first theorem below is included only for motivation; we give references for it in lieu of a proof. The second theorem, though more complicated to state, is easier to prove, and we shall do so below. It will be used to prove Brouwer's Theorem in 27.19. For both theorems, let $\mathbb{R}^n$ be metrized by $d(x, y) = \max\{|x_j - y_j| : 1 \le j \le n\}$.

> **First Approximate Fixed Point Theorem.** Let n be a positive integer, let $f = (f_1, f_2, \ldots, f_n) : [0,1]^n \to [0,1]^n$ be a function, and let any number $\varepsilon > 0$ be given. Then there exists a set $S \subseteq [0,1]^n$ with diameter less than ε, with the following property: For each $j \in \{1, 2, \ldots, n\}$ there exist some points $x = (x_1, x_2, \ldots, x_n)$ and $x' = (x'_1, x'_2, \ldots, x'_n)$ in K such that $x_j \le f_j(x)$ and $x'_j \ge f_j(x')$.
>
> **Second Approximate Fixed Point Theorem.** Let n be a positive integer.

Let Δ be the *standard n-simplex*; that is, the set

$$\Delta \quad = \quad \left\{ u \in \mathbb{R}^n \ : \ u_1, u_2, \ldots, u_n \geq 0 \text{ and } \sum_{j=1}^{n} u_j \leq 1 \right\}.$$

Let any function $f : \Delta \to \Delta$ and any number $\varepsilon > 0$ be given. Then there exists a set $S \subseteq \Delta$ that acts as an approximate fixed point of f, in the following sense:

- $\mathrm{diam}(S) \leq \varepsilon$
- For each $i = 1, 2, \ldots, n$, there is some $u \in S$ such that $u_i - \varepsilon \leq f(u)_i$.
- There exists some point $v \in S$ satisfying $\sum_{i=1}^{n} v_i + \varepsilon \geq \sum_{i=1}^{n} f(v)_i$.

Remarks. We emphasize that f is not assumed to be continuous or even measurable. Aside from the domain and codomain, we make no assumption at all about f. Thus, these theorems are not really "about" f; they are theorems about the combinatorial structure of $\mathbb{R}^n$. An analogous theorem about infinite dimensional vector spaces will be given in 27.19.

A similar argument in two dimensions, more geometrical and elementary in presentation, was given by Shashkin [1991]. Theorem 2 of Baillon and Simons [1992] is also very similar.

The First Approximate Fixed Point Theorem can be proved by methods similar to those below, using Wolsey's [1977] Cubical Sperner Lemma instead of our 3.36. It would be interesting to know if the First Approximate Fixed Point Theorem can also be proved by some short argument using 3.36; no such argument is presently known to this author.

Proof of the Second Approximate Fixed Point Theorem. By writing $u_{n+1} = 1 - \sum_{j=1}^{n} u_j$, we may rewrite

$$\Delta \quad = \quad \left\{ u \in \mathbb{R}^{n+1} \ : \ u_1, u_2, \ldots, u_n, u_{n+1} \geq 0 \text{ and } \sum_{j=1}^{n+1} u_j = 1 \right\}.$$

The $(n+1)$st coordinate will be treated just like the other coordinates in the following argument.

Let M be an integer large enough so that $2(n+1)/M < \varepsilon$. Let X consist of the collection of all points $u \in \Delta$ for which all of $Mu_1, Mu_2, \ldots, Mu_n, Mu_{n+1}$ are integers. Let $P = \{1, 2, 3, \ldots, n, n+1\}$. For $1 \leq j \leq n+1$ define a preordering of X by taking $u \preccurlyeq_j v$ when $u_j \leq v_j$.

Let σ be a crystal, and let S be its range. If

$$\sum_{i \in \mathrm{Dom}(\sigma)} \sigma(i)_i \quad \leq \quad 1 - \frac{n+1}{M},$$

then there exists a member $x \in X$ that satisfies $\sigma(i)_i < x_i$ for all $i \in \mathrm{Dom}(\sigma)$, contradicting (CR-2) in 3.32. Thus the inequality above does not hold.

For any $i, j \in \mathrm{Dom}(\sigma)$, we have $\sigma(i)_i \leq \sigma(j)_i$, and therefore

$$1 - \frac{n+1}{M} \quad < \quad \sum_{i \in \mathrm{Dom}(\sigma)} \sigma(i)_i \quad \leq \quad \sum_{i \in \mathrm{Dom}(\sigma)} \sigma(j)_i \quad \leq \quad 1$$

from which it follows that $0 \leq \sigma(j)_i - \sigma(i)_i \leq \frac{n+1}{M}$ for each $i, j \in \mathrm{Dom}(\sigma)$. Therefore $|\sigma(j)_i - \sigma(k)_i| \leq 2(n+1)/M$ whenever $i, j, k \in \mathrm{Dom}(\sigma)$. Hence $|u_i - v_i| \leq \varepsilon$ for all $u, v \in S$ and $i \in \mathrm{Dom}(\sigma)$.

On the other hand, since $\sum_{i=1}^{n+1} \sigma(j)_i = 1$, we must have $\sum_{i \notin \mathrm{Dom}(\sigma)} \sigma(j)_i \leq \frac{n+1}{M}$ for $j \in \mathrm{Dom}(\sigma)$ and, in particular, $\sigma(j)_i \leq \frac{n+1}{M}$ for each $i \notin \mathrm{Dom}(\sigma)$. Thus $0 \leq u_i \leq \varepsilon/2$ for all $u \in S$ and $i \notin \mathrm{Dom}(\sigma)$. Therefore $\mathrm{diam}(S) \leq \varepsilon$.

Define a labeling $\ell : \Delta \to \{1, 2, \ldots, n, n+1\}$ as follows: let $\ell(u) = i$ if i is the first coordinate that satisfies $u_i \leq f(u)_i$. By 3.36 there exists a complete crystal σ with respect to that labeling. When $i \in \mathrm{Dom}(\sigma) = \ell(\mathrm{Ran}(\sigma))$, then $i = \ell(u)$ for some $u \in S$, so $u_i \leq f(u)_i$. On the other hand, we noted earlier in this proof that when $i \notin \mathrm{Dom}(\sigma)$ and $u \in S$, then $u_i \leq \varepsilon$; hence $f(u)_i \geq u_i - \varepsilon$. This completes the proof.

Well Ordered Sets

3.38. *Definition.* Let $(X, \preccurlyeq)$ be a poset. We say $\preccurlyeq$ is a **well ordering** if each nonempty subset of X has a first element. Then X is a **well ordered set**, or a **woset**.

Examples. The set $\mathbb{N}$ is well ordered. Also see 3.43 and 5.44.

Remark. Well ordered sets are only used infrequently in analysis. This subchapter may be postponed or omitted if the reader is concerned only with the usual topics of analysis.

3.39. *Basic properties of wosets.*

a. Any woset is a chain.

b. Any subset of a woset is a woset.

c. Let S be a subset of a woset X. Then S is a proper lower set in X if and only if $S = \mathrm{Pre}(b)$ for some $b \in X$, with notation as in 3.16.b.

Hint for the "only if" part: Let b be the first element of $X \setminus S$.

d. Let X be a woset. Then the lower sets of X form a woset $\widehat{X}$, when ordered by $\subseteq$. The last element of $\widehat{X}$ is X. If X is not empty, then the first element of $\widehat{X}$ is $\varnothing = \mathrm{Pre}(\min(X))$, where $\min(X)$ is the first member of X.

e. Any woset X is a proper lower set of some larger woset Y. Indeed, one way to form such a larger set is by adjoining some new element — call it $\square$ — that is not already present in X and defining $\square$ to be larger than all the elements of X.

f. **Induction on Wosets.** Let $(X, \preccurlyeq)$ be a woset, and let S be a subset of X with the property that $\mathrm{Pre}(b) \subseteq S \Rightarrow b \in S$. Then in fact $S = X$.

Hint: If not, let b be the first element of $X \setminus S$.

3.40. *Notation.* For the result below, if $(X, \preccurlyeq)$ is a well ordered set and T is a nonempty set, then an *X-based sequence in T* will mean a function whose domain is some proper lower set of X and whose range is contained in T. As a degenerate case, we may view the empty function (with graph equal to the empty set) as an X-based sequence in T.

Theorem of Recursion on Wosets. Let $(X, \preccurlyeq)$ be a woset and let T be a nonempty set. Let any function

$$\rho \;:\; \{X\text{-based sequences in } T\} \;\to\; T$$

be given. Then there exists a unique function $F : X \to T$ satisfying

$$F(x) \;=\; \rho\Big(F\big|_{\mathrm{Pre}(x)}\Big) \qquad \text{for each } x \in X.$$

Here $F\big|_{\mathrm{Pre}(x)}$ denotes the restriction of F to the set $\mathrm{Pre}(x) = \{w \in X : w \prec x\}$. Thus, the value of F at any x is determined, via the rule ρ, by the values of F at all the predecessors of x.

Remark. Compare this result with 2.22.

Proof of theorem. First we prove uniqueness. Suppose F_1, F_2 are two such functions, and $F_1 \neq F_2$. Let x be the first member of X that satisfies $F_1(x) \neq F_2(x)$. Then $F_1(w) = F_2(w)$ for all $w \in \mathrm{Pre}(x)$ — that is, the restrictions $F_1\big|_{\mathrm{Pre}(x)}$ and $F_2\big|_{\mathrm{Pre}(x)}$ are the same function φ. But then $F_1(x) = \rho(\varphi) = F_2(x)$, a contradiction. This proves uniqueness.

We now turn to the existence proof. It will be convenient to replace X with a slightly larger set. Let $Y = X \cup \{\natural\}$, where $\natural$ is some object not belonging to X. Extend the ordering of X to an ordering on Y by setting $x \prec \natural$ for all $x \in X$; then Y is also a woset. Note that for each $y \in Y$, the set $\mathrm{Pre}(y)$ is a lower set in X; in particular, $\mathrm{Pre}(\natural) = X$.

We shall prove that for each $y \in Y$ there is a function $F_y : \mathrm{Pre}(y) \to T$ satisfying

$$F_y(x) \;=\; \rho\Big(F_y\big|_{\mathrm{Pre}(x)}\Big) \qquad \text{for each } x \in \mathrm{Pre}(y).$$

(Once this is established, we simply take $F = F_\natural$ to prove the theorem.) First note that for each y, we may unambiguously use the notation "F_y" because there is *at most* one such function F_y; that is clear by a uniqueness argument similar to the one at the beginning of the proof of the theorem.

The proof of the existence of F_y's is by induction on y. Assume, then, that some $\eta \in Y$ is given and that F_y's exist for all $y \prec \eta$; we are to demonstrate the existence of F_η. We demonstrate that in two different ways, according to the nature of η:

First, suppose η has an immediate predecessor ξ — that is, suppose η is the first member of Y after some member ξ. Then $\mathrm{Pre}(\eta) = \{\xi\} \cup \mathrm{Pre}(\xi)$, and $F_\xi : \mathrm{Pre}(\xi) \to T$ is a function of the sort described above. Define a function $F_\eta : \mathrm{Pre}(\eta) \to T$ by

$$F_\eta(x) \;=\; \begin{cases} F_\xi(x) & \text{when } x \in \mathrm{Pre}(\xi) \\ \rho\,(F_\xi) & \text{when } x = \xi. \end{cases}$$

It is easy to verify that F_η has the required properties.

On the other hand, suppose η has no immediate predecessor in Y. Then $\mathrm{Pre}(\eta) = \bigcup_{y \prec \eta} \mathrm{Pre}(y)$. Also, $\mathrm{Graph}(F_y)$ is an increasing function of y — that is,

$$y \prec y' \prec \eta \qquad \Rightarrow \qquad \mathrm{Graph}(F_y) \subseteq \mathrm{Graph}(F_{y'}).$$

Verify that $\mathrm{Graph}(F_\eta) = \bigcup_{y \prec \eta} \mathrm{Graph}(F_y)$ defines F_η with the required properties.

3.41. Comparability Theorem. If $(X, \leq)$ and $(Y, \preccurlyeq)$ are wosets, then exactly one of these three conditions holds:

- There exists an order isomorphism between X and Y.
- There exists an order isomorphism from X onto a lower set of Y.
- There exists an order isomorphism from Y onto a lower set of X.

Furthermore, in each case the isomorphism is uniquely determined.

Proof. For each proper lower set $L \subseteq X$ and each function $\varphi : L \to Y$, define

$$\rho(\varphi) \quad = \quad \begin{cases} \min(Y \setminus \mathrm{Range}(\varphi)) & \text{if } \mathrm{Range}(\varphi) \neq Y \\ \min(Y) & \text{if } \mathrm{Range}(\varphi) = Y. \end{cases}$$

Now define $F : X \to Y$ by recursion, as in 3.40. Then $F(\mathrm{Pre}(x))$ is an increasing function of x, so $X_0 = \{x \in X : F(\mathrm{Pre}(x)) \neq Y\}$ is a lower set in X. Show that $F(X_0)$ is a lower set in Y and F gives an order isomorphism from X_0 onto $F(X_0)$. If $F(X_0) \neq Y$, then $X_0 = X$. This establishes the existence of at least one isomorphism.

If f and g were distinct order isomorphisms from X onto a lower set of Y, then we could take x_0 to be the first element of X satisfying $f(x_0) \neq g(x_0)$; show that this leads to a contradiction. This proves uniqueness in either direction.

Suppose f is an order isomorphism from X onto a lower set of Y and g is an order isomorphism from Y onto a lower set of X. Then $g \circ f$ is an order isomorphism from X onto a lower set of X — but by the uniqueness result of the previous paragraph, $g \circ f$ must then be the identity map of X.

3.42. *Corollaries.*

a. If X and Y are wosets, then $\mathrm{card}(X) \leq \mathrm{card}(Y)$ or $\mathrm{card}(Y) \leq \mathrm{card}(X)$.

b. If X is an infinite woset, then $\mathrm{card}(X) \geq \mathrm{card}(\mathbb{N})$, and for any $\xi \notin X$ we have $\mathrm{card}(X) = \mathrm{card}(X \cup \{\xi\})$.

3.43. Canonical Well Ordering Theorem. Let X be a set, and let some mapping

$$\rho \ : \ \{\text{proper subsets of } X\} \ \to \ X$$

be given that satisfies $\rho(S) \in X \setminus S$ for each S. Then there exists a unique well ordering $\preccurlyeq$ on X with the following property:

$$x \ = \ \rho\Big(\{u \in X \ : \ u \prec x\}\Big) \qquad \text{for each } x \in X.$$

(In other words, to find the next term in the ordering, just apply ρ to the set of all the terms that have already been ordered. Contrast this with (AC4) in 6.20.)

Proof (modified from Malitz [1979]). Consider well orderings of subsets of X. When $\preccurlyeq$ is such a well ordering, let $S_{\preccurlyeq}$ be the subset of X that it well orders, and let its sets of predecessors be denoted by

$$\mathrm{Pre}_{\preccurlyeq}(x) \quad = \quad \{s \in S_{\preccurlyeq} \ : \ s \preccurlyeq x, \ s \neq x\}$$

for points $x \in S_{\preccurlyeq}$. Say $\preccurlyeq$ is a *ρ-well ordering* if it also satisfies

$$x \quad = \quad \rho\Big(\mathrm{Pre}_{\preccurlyeq}(x)\Big) \qquad\qquad \text{for each } x \in S_{\preccurlyeq}.$$

Let $\mathcal{K}$ be the set of all ρ-well orderings. It is clear that $\mathcal{K}$ is nonempty, since the empty relation is a ρ-well ordering of the empty set. We are to show that $\mathcal{K}$ has a unique member $\preccurlyeq$ satisfying $S_{\preccurlyeq} = X$. As a preliminary step, we shall show that

> (∗∗) Whenever $\preccurlyeq_1$ and $\preccurlyeq_2$ are ρ-well orderings, then one of the wosets $(S_{\preccurlyeq_1}, \preccurlyeq_1)$, $(S_{\preccurlyeq_2}, \preccurlyeq_2)$ is a lower set of the other, whose ordering is just the restriction of the other's ordering.

Indeed, by 3.41, we know that there exists an order isomorphism between one of the wosets $(S_{\preccurlyeq_1}, \preccurlyeq_1)$, $(S_{\preccurlyeq_2}, \preccurlyeq_2)$ and a lower set of the other. Say $p : S_{\preccurlyeq_1} \to S_{\preccurlyeq_2}$ is an order isomorphism from $(S_{\preccurlyeq_1}, \preccurlyeq_1)$ onto a lower set of $(S_{\preccurlyeq_2}, \preccurlyeq_2)$. Then $p\Big(\mathrm{Pre}_{\preccurlyeq_1}(u)\Big) = \mathrm{Pre}_{\preccurlyeq_2}\Big(p(u)\Big)$ for any $u \in S_{\preccurlyeq_1}$. To prove (∗∗), it suffices to show that this isomorphism is actually an inclusion map — i.e., that $p(x) = x$ in X for all $x \in S_{\preccurlyeq_1}$. If $\beta = p(\alpha) \neq \alpha$ for some $\alpha \in S_{\preccurlyeq_1}$, choose the $\preccurlyeq_1$-first such α in $S_{\preccurlyeq_1}$ and the corresponding β. Then p acts as the identity map on $\mathrm{Pre}_{\preccurlyeq_1}(\alpha)$. Thus

$$\mathrm{Pre}_{\preccurlyeq_1}(\alpha) \quad = \quad p\Big(\mathrm{Pre}_{\preccurlyeq_1}(\alpha)\Big) \quad = \quad \mathrm{Pre}_{\preccurlyeq_2}\Big(p(\alpha)\Big) \quad = \quad \mathrm{Pre}_{\preccurlyeq_2}(\beta),$$

and therefore $\alpha = \rho\Big(\mathrm{Pre}_{\preccurlyeq_1}(\alpha)\Big) = \rho\Big(\mathrm{Pre}_{\preccurlyeq_2}(\beta)\Big) = \beta$, a contradiction. This proves the claim (∗∗).

Each member of $\mathcal{K}$ is a relation on X, which may be viewed as a subset of $X \times X$. From (∗∗) it follows easily that the union of the elements of $\mathcal{K}$ is itself a member of $\mathcal{K}$. Hence it is the largest member of $\mathcal{K}$; let us denote it by $\sqsubseteq$. If $S_{\sqsubseteq}$ is not equal to X, then $\sqsubseteq$ extends to a strictly larger ρ-well ordering on $S_{\sqsubseteq} \cup \{\rho(S_{\sqsubseteq})\}$, by defining $\rho(S_{\sqsubseteq})$ to be larger than all the members of $S_{\sqsubseteq}$ — contradicting the maximality of $\sqsubseteq$. Thus $S_{\sqsubseteq} = X$. Uniqueness follows from (∗∗).

3.44. *Products of wosets.* A product of wosets, with the product ordering, is not necessarily a woset; an example is given by $\mathbb{N}^2$. Other orderings on a product may be well ordered, however:

a. Let $(\Lambda, \leq)$ be a woset, and for each $\lambda \in \Lambda$ let $(X_\lambda, \leq)$ be a chain. (The $\leq$'s may represent different orderings.) Let $P = \prod_{\lambda \in \Lambda} X_\lambda$. The **lexicographical order** (or *dictionary order*) on P is defined as follows: $p < q$ in P if $p(\nu) < q(\nu)$ where $\nu \in \Lambda$ is the first component in which p and q differ. Show:

(i) The lexicographical ordering is a chain ordering on P.

(ii) If each $(X_\lambda, \leq)$ is well ordered and Λ is a finite set, then the lexicographical ordering is a well ordering on P.

(iii) In general, the lexicographical ordering on an infinite product is not a well ordering. Indeed, if Λ is an infinite woset with no last element and each X_λ is a woset containing at least two elements, then P is not well ordered. To see this, let ξ be the function whose value at λ is the smallest member of X_λ; show that $P \setminus \{\xi\}$ has no smallest element. For a more concrete special case, show that $\{x \in \{0,1\}^{\mathbb{N}} : x \neq (0,0,0,\ldots)\}$ has no first element in $\{0,1\}^{\mathbb{N}}$.

b. (This construction will be used in 3.45.) Let $(X, \leq)$ be a well ordered set. Define an ordering $\sqsubseteq$ on $X \times X$, as follows: For (x,y) and (u,v) in $X \times X$, say $(u,v) \sqsubset (x,y)$ means that

- $\max\{u,v\} < \max\{x,y\}$, or
- $\max\{u,v\} = \max\{x,y\}$ and $u < x$, or
- $\max\{u,v\} = \max\{x,y\}$ and $u = x$ and $v < y$.

Verify that this is a well ordering on $X \times X$.

3.45. Theorem on card(X^2). Let X be an infinite set, and suppose that X can be well ordered. Then $\mathrm{card}(X \times X) = \mathrm{card}(X)$.

Remarks. The present result does not require the Axiom of Choice, which tells us that *every* set can be well ordered; see 6.20 and 6.22.

Proof of theorem. Let $|\ \ |$ denote cardinality. Clearly, $|X| \leq |X \times X|$. Suppose $|X| < |X \times X|$ for some infinite woset $(X, \leq)$; we shall obtain a contradiction. Clearly, we can replace X by any other woset with the same cardinality; by 3.39.d we may replace X with the *first* lower set in X that is infinite and satisfies $|X| < |X \times X|$. Observe that if K is any proper lower set in X, then either K is finite or $|K| = |K \times K|$; hence $|K| \neq |X|$, hence (since $K \subseteq X$) we have $|K| < |X|$. In particular, X does not have a last element — for, if X were an infinite woset with last element ξ, then $X \setminus \{\xi\}$ would be a proper lower set with the same cardinality as X, a contradiction.

Define a well ordering $\sqsubseteq$ on $X \times X$ as in 3.44.b. Since X and $X \times X$ are well ordered, one of these sets is uniquely order isomorphic to a lower set of the other. Since $|X| < |X \times X|$, the order isomorphism must be from X onto a set L that is a proper lower set of $X \times X$. Then $|L| = |X|$ is an infinite cardinal.

Let (u_0, v_0) be the $\sqsubseteq$-first member of $(X \times X) \setminus L$. Let w_0 be the maximum of u_0 and v_0 in $(X, \leq)$. Let $M = \{x \in X : x \leq w_0\}$. Then M is a lower set in X. Since X does not have a last element, M is a proper lower set in X, and therefore $|M| < |X|$. Observe that

$$(u,v) \in L \quad \Rightarrow \quad (u,v) \sqsubset (u_0, v_0) \quad \Rightarrow \quad \max\{u,v\} \leq w_0 \quad \Rightarrow \quad u, v \in M,$$

and thus $L \subseteq M \times M$. Hence $|L| \leq |M \times M|$. Since L is an infinite set, M must be infinite, too. By our choice of X, then, $|M \times M| = |M|$. Now $|X| = |L| \leq |M \times M| = |M| < |X|$, a contradiction. This completes the proof.

3.46. *Definition.* A collection $\mathcal{F}$ of subsets of a set X is said to have **finite character** if for each set $S \subseteq X$,

$$S \text{ is a member of } \mathcal{F} \text{ if and only if each finite subset of } S \text{ is a member of } \mathcal{F}.$$

Example. If $(X, \preccurlyeq)$ is a poset, then $\{S \subseteq X \ : \ S \text{ is chain ordered by } \preccurlyeq\}$ has finite character. Other examples will be given in 5.7.e, 11.10, 12.17.f, and 14.31.

We shall now prove the following theorem:

> **Finite Character Theorem (canonical choice version).** Let X be a set that can be well ordered and let $\mathcal{F}$ be a collection of subsets of X that has finite character. Then $\mathcal{F}$ has a $\subseteq$-maximal element.

Remark. Contrast this with (AC5) in 6.20. The present theorem does not require the Axiom of Choice, which would tell us that *every* set can be well ordered.

Proof of theorem. Let $\preccurlyeq$ be a well ordering of X. We shall determine a maximal set $M \in \mathcal{F}$ by defining its characteristic function $1_M : X \to \{0, 1\}$, by transfinite recursion. At the αth step, we have defined 1_M on all of $\mathrm{Pre}(\alpha)$, and thus we have determined which elements of $\mathrm{Pre}(\alpha)$ should be members of M — i.e., we have determined the set $\mathrm{Pre}(\alpha) \cap M$. Now define $1_M(\alpha)$, by taking it to be 1 if the set $\big(\mathrm{Pre}(\alpha) \cap M\big) \cup \{\alpha\}$ is a member of $\mathcal{F}$ and 0 otherwise. Verify that the resulting set M is a $\subseteq$-maximal member of $\mathcal{F}$.

Chapter 4

More about Sups and Infs

4.1. *Chapter overview.* Sups and infs were introduced briefly in Chapter 3. This chapter investigates sups and infs further and introduces the related notions of Moore closures and order completeness. A Moore collection is a collection of sets that is closed under arbitrary intersection — i.e., under arbitrary infimum with respect to the ordering given by inclusion. Order completeness of a poset refers to the existence of sups and infs in that poset. Order completions can be constructed most easily using *polars*, a special type of Moore closure.

Moore Collections and Moore Closures

4.2. The term "closure" has several different meanings in mathematics. Most of the meanings of "closure" are specializations of the *Moore closure*, defined below. (An exception: the "pretopological convergence closures" introduced in 15.4 need not be Moore closures; see the example in 15.6.)

Many mathematicians write the closure of a set S as $\overline{S}$. However, that notation has certain disadvantages: (i) It is used for other purposes (e.g., complex conjugation). (ii) It becomes awkward if one wishes to work simultaneously with two or more closures (e.g., from two different topologies). In this book we shall write a closure of a set S as $\mathrm{cl}(S)$. We shall use subscripted notation, such as $\mathrm{cl}_{\mathcal{T}}(S)$ and $\mathrm{cl}_{\mathcal{U}}(S)$, if we need to distinguish between several different closures.

4.3. Let X be a set, and let $\mathcal{C}$ be a collection of subsets of X. We shall say $\mathcal{C}$ is a **Moore collection of sets** if:

(i) $X \in \mathcal{C}$, and

(ii) $\mathcal{C}$ is closed under arbitrary intersection — i.e., if $\{S_\lambda : \lambda \in \Lambda\} \subseteq \mathcal{C}$, then $\bigcap_{\lambda\in\Lambda} S_\lambda \in \mathcal{C}$.

(If we adopt the convention that the intersection of no subsets of X is just X, then condition (i) can be omitted; it follows from (ii) by taking $\Lambda = \varnothing$.) In the present context, members of $\mathcal{C}$ will be called **Moore closed** sets, or just **closed** sets if the context is understood.

Now, let any set $S \subseteq X$ (not necessarily a member of $\mathcal{C}$) be given. Then there exist closed sets that are supersets of S — for instance, X itself is such a set. Among all the closed supersets of S, there is a *smallest* — namely, the intersection of all the closed supersets of S. We shall call it the **Moore closure** (or more simply, the **closure**) of S, relative to the collection $\mathcal{C}$, and we shall denote it by $\mathrm{cl}(S)$. (It is easy to see that a set $T \subseteq X$ is closed if and only if $\mathrm{cl}(T) = T$.)

In this fashion we define a mapping $\mathrm{cl} : \mathcal{P}(X) \to \mathcal{C}$, called the **Moore closure operator** associated with the collection $\mathcal{C}$. In some cases (i.e., for some choices of X and $\mathcal{C}$) we can also give some other, equivalent description of $\mathrm{cl}(S)$ that may be more convenient — e.g., a characterization directly in terms of S, which does not mention the collection of all closed supersets of S.

The Moore closure of S is also known as the member of $\mathcal{C}$ that is **generated** by S; we may call S a **generating set** for $\mathrm{cl}(S)$. This terminology is particularly common when the elements of X and S are themselves subsets of some set Ω — i.e., when S is a collection of sets, and $\mathrm{cl}(S)$ is the collection of sets generated by S. (Of course, from the viewpoint of axiomatic set theory, *all* sets are sets of sets, but most mathematicians do not share that viewpoint.)

Though Moore closures appear in many parts of mathematics, the terminology varies greatly. In most applications of the concept, the "closed" / "closure" terminology is commonly used, or the "generated" / "generating" terminology is commonly used, but not both. Another term sometimes used for a Moore closed set is **saturated set**; other terms sometimes used for a Moore closure are **hull**, **saturation**, or **saturated hull**. The following table previews a few kinds of Moore closures that will be studied later. We shall introduce appropriate terminology separately in each context. The top part of the table deals with sets of points; the bottom part of the table deals with sets of sets.

Moore closed set	Moore closure
sublattice	sublattice generated
full (order convex) set	full hull
upper set	up closure
max closed gauge	max closure
saturated set	saturation
topologically closed set	topological closure
ideal (in a variety)	ideal generated by a set
convex set	convex hull
balanced set	balanced hull
linear subspace	linear span
filter of sets	filter generated
ideal of sets	ideal generated
(σ-)algebra	(σ-)algebra generated
topology	topology generated
monotone class	monotone class generated
Moore collection	Moore collection generated

Further remarks about the terminology. Most Moore closures of interest are either *algebraic* (introduced in 4.8) or *topological* (introduced in 5.16.b and 5.19). Very few closures of interest are *both* algebraic and topological; that is clear from 16.8.b.

The term "closure" by itself is commonly used by algebraists to refer to any Moore closure (as defined above), but the term "closure" usually is used by analysts to refer only to topological closures. We shall follow the analysts' convention in some parts of this book since this book is largely devoted to the foundations of analysis.

The basic property given in 4.5.a is due to Moore [1910], although the notation certainly has changed since then. Most properties of closures in this chapter are taken from Cohn [1965], Evers and Maaren [1985], McKenzie, McNulty, and Taylor [1987], and Tsinakis [1993].

4.4. *A few examples of Moore closures.*

a. Let $(X, \preccurlyeq)$ be a preordered set. For $a, b \in X$ let $[a,b] = \{x \in X : a \preccurlyeq x \preccurlyeq b\}$. A set $S \subseteq X$ is called **full** (or **order convex**) if $a, b \in S \Rightarrow [a,b] \subseteq S$. (For example, in $[-\infty, +\infty]$, any interval $[a,b]$ or $[a,b)$ or $(a,b]$ or (a,b) is full.) Show that the full subsets of X form a Moore collection. Hence any set $T \subseteq X$ is contained in a smallest full superset, called the **full hull** of T. Show that the full hull of T is equal to $\bigcup_{a,b \in T}[a,b]$.

For later applications we note some further properties of full subsets of chains. Let $(X, \leq)$ be a chain. Show:

(i) If S_1, S_2 are full sets that are not disjoint, then $S_1 \cup S_2$ is full.

(ii) For any $T \subseteq X$ and $p \in T$, let $C(p,T)$ be the union of all the full sets S that satisfy $p \in S \subseteq T$. Show that $C(p,T)$ is a full subset of T and that $C(p,T)$ is maximal for that property — i.e., $C(p,T)$ is not a proper subset of some other full subset of T. Show that the sets $C(p,T)$ form a partition of T — that is, any two sets $C(p_1,T), C(p_2,T)$ are either identical or disjoint. We may refer to the $C(p,T)$'s as the **full components** of T. This will be used in 15.35.c and 17.24.

b. Let $(X, \preccurlyeq)$ be a partially ordered set, and let $S \subseteq X$. We say that S is

up-closed, or an **upper set**, if $x \succcurlyeq y,\ y \in S \Rightarrow x \in S$;

down-closed, or a **lower set**, if $x \preccurlyeq y,\ y \in S \Rightarrow x \in S$;

sup-closed if, whenever A is a nonempty subset of S and $\sigma = \sup(A)$ exists in X, then σ is a member of S;

inf-closed if, whenever A is a nonempty subset of S and $\iota = \inf(A)$ exists in X, then ι is a member of S.

(Lower sets were defined in 3.16.) Show that the collections of such sets are Moore collections, with resulting Moore closures as follows:

$$\begin{aligned} \text{up-cl}(S) &= \{x \in X : x \succcurlyeq s \text{ for some } s \in S\}, \\ \text{down-cl}(S) &= \{x \in X : x \preccurlyeq s \text{ for some } s \in S\}, \end{aligned}$$

$$\begin{array}{rcl} \text{sup-cl}(S) & = & \{x \in X \ : \ x = \sup(A) \text{ for some nonempty } A \subseteq S\}, \\ \text{inf-cl}(S) & = & \{x \in X \ : \ x = \inf(A) \text{ for some nonempty } A \subseteq S\}. \end{array}$$

Show also that

(i) Any up-closed set is sup-closed; any down-closed set is inf-closed.

(ii) A set S is up-closed if and only if $X \setminus S$ is down-closed.

(iii) Any *union* of up-closed sets is up-closed; any *union* of down-closed sets is down-closed. (This property is not shared by most Moore collections.)

(iv) $\varnothing$ is up-closed and down-closed.

c. If $P = \{d_1, d_2, \ldots, d_n\}$ is a finite collection of pseudometrics on a set X (defined in 2.11), then another pseudometric, $\vee P$, can be defined by

$$(\vee P)(x,y) \quad = \quad \max\{d_1(x,y),\, d_2(x,y),\, \ldots,\, d_n(x,y)\}\,.$$

Clearly, $\vee P$ is the supremum of P in the family of all pseudometrics — i.e., it is the smallest pseudometric that is larger than or equal to all the d_j's. We may also denote it by $d_1 \vee d_2 \vee \cdots \vee d_n$. When P contains just a single pseudometric, then $\vee P$ equals that pseudometric.

Let D be a gauge on X — that is, a collection of pseudometrics. We shall say that D is **max-closed** if $d_1, d_2 \in D \ \Rightarrow \ d_1 \vee d_2 \in D$ or, equivalently, if $d_1, d_2, \ldots, d_n \in D \ \Rightarrow \ d_1 \vee d_2 \vee \cdots \vee d_n \in D$. (In the wider literature, another name for max-closed is **saturated**.) Clearly, this determines a type of Moore closure on the collection of all pseudometrics on X; the **max closure** of a gauge D is the gauge

$$\text{max-cl}(D) \quad = \quad \{\vee P \ : \ P \text{ is a finite subset of } D\}\,.$$

Similarly, a gauge D is *closed under addition*, or *sum-closed*, if $d_1, d_2 \in D \ \Rightarrow \ d_1 + d_2 \in D$. This also determines a Moore closure, which we shall call the *sum closure.*

A gauge D is **directed** if for each finite set $P \subseteq D$, there exists some pseudometric $d \in D$ such that $\vee P \leq d$. Note that if D is max-closed or closed under addition, then D is directed.

Preview. In 5.15.h we shall see that any gauge D is topologically equivalent to its max closure and its sum closure; in 18.13 we shall see that any gauge D is uniformly equivalent to its max closure and its sum closure. Hence, for many purposes, D may be replaced with its max closure or sum closure — i.e., D may be replaced by a directed gauge. For some theorems, this replacement will not affect the hypotheses, but may simplify the proofs.

d. If $\mathcal{M}_\alpha, \mathcal{M}_\beta, \mathcal{M}_\gamma, \ldots$ are collections of subsets of X, each of which is closed under finite union, then $\bigcap_{\lambda \in \{\alpha,\beta,\gamma,\ldots\}} \mathcal{M}_\lambda$ is also closed under finite union. Thus, the collection $\{\mathcal{M} \subseteq \mathcal{P}(X) : \mathcal{M} \text{ is closed under finite union}\}$ is a Moore collection of subsets of $\mathcal{P}(X)$. The resulting Moore closure may be described as follows: If $\mathcal{S}$ is a collection of subsets of X, then the smallest collection that is closed under finite union and contains $\mathcal{S}$ is

$$\cup\text{-cl}(\mathcal{S}) \ = \ \{M \subseteq X \ : \ M \text{ is the union of finitely many members of } \mathcal{S}\}.$$

Similarly, the other "closures" defined in 1.30 are also Moore closures.

In particular, the closure under arbitrary intersections is a Moore closure. Thus the collection of all Moore collections of subsets of X is a Moore collection of subsets of $\mathcal{P}(X)$.

e. Let $\approx$ be an equivalence relation on a set X; let Y be the quotient set; let $\pi : X \to Y$ be the quotient map. Say that a set $S \subseteq X$ is **π-saturated**, or **$\approx$-saturated**, if it is closed under this equivalence — i.e., if

$$x_1 \in S, \quad x_1 \approx x_2 \qquad \Rightarrow \qquad x_2 \in S$$

— that is, if S is a union of equivalence classes.

The collection of saturated subsets of X is a Moore collection — i.e., it is closed under intersection. Hence we can define the corresponding Moore closure: The **π-saturation**, or **π-saturated hull**, of a set $A \subseteq X$ is the smallest π-saturated set that contains A; it is the intersection of the π-saturated sets that contain A. Show that

$$\pi^{-1}(\pi(A)) \quad = \quad \bigcup_{a \in A} \{x \in X : x \approx a\} \quad = \quad \text{the } \pi\text{-saturation of } A.$$

The forward image mapping $S \mapsto \pi(S) = \{\pi(x) : x \in S\}$ is usually defined as a mapping from $\mathcal{P}(X)$ into $\mathcal{P}(Y)$ (see 2.7); but if we restrict it to a smaller domain, we get a bijection from $\{\pi\text{-saturated subsets of } X\}$ onto $\mathcal{P}(Y)$, whose inverse is given by the inverse image mapping $T \mapsto \pi^{-1}(T) = \{x \in X : \pi(x) \in T\}$. This bijection preserves the basic set operations: complementations, intersections, and unions. That is,

$$\pi(X \backslash S) = Y \backslash \pi(S), \qquad \pi\left(\bigcap_{\alpha \in A} S_\alpha\right) = \bigcap_{\alpha \in A} \pi(S_\alpha), \qquad \pi\left(\bigcup_{\alpha \in A} S_\alpha\right) = \bigcup_{\alpha \in A} \pi(S_\alpha)$$

for any π-saturated sets S and S_α.

4.5. *Some basic properties of Moore closures.*

a. Axioms for a Moore Closure Operator. Let X be a set, and suppose we are given a function $\mathrm{cl} : \mathcal{P}(X) \to \mathcal{P}(X)$. Show that "cl" is the Moore closure operator for some Moore collection of subsets of X (as defined in 4.3) if and only if "cl" satisfies these three rules:

$$S \subseteq \mathrm{cl}(S) \qquad \text{(extensive)}$$

$$\mathrm{cl}(\mathrm{cl}(S)) = \mathrm{cl}(S) \qquad \text{(idempotent)}$$

$$S \subseteq T \quad \Rightarrow \quad \mathrm{cl}(S) \subseteq \mathrm{cl}(T) \qquad \text{(isotone)}$$

for all sets $S, T \subseteq X$. Of course, if "cl" does satisfy these axioms, then the corresponding Moore collection $\mathcal{C} \subseteq \mathcal{P}(X)$ is uniquely determined by "cl:" it consists of those sets $S \subseteq X$ that satisfy $\mathrm{cl}(S) = S$.

b. Note that any Moore closure on X is an isotone mapping from $\mathcal{P}(X)$ into $\mathcal{P}(X)$; hence it satisfies the conclusions of 4.29.c. It also satisfies

$$\mathrm{cl}\left(\bigcup_{\alpha\in A} S_\alpha\right) \quad = \quad \mathrm{cl}\left(\bigcup_{\alpha\in A} \mathrm{cl}(S_\alpha)\right)$$

for any sets $S_\alpha \subseteq X$ $(\alpha \in A)$.

c. Let $\mathcal{C}$ be a Moore collection of subsets of X. Then each subset of $\mathcal{C}$ has a sup and an inf in $\mathcal{C}$. (Thus $(\mathcal{C}, \subseteq)$ is a *complete lattice*; see the definition in 4.13.)

Indeed, let any collection $\mathcal{S} = \{S_\lambda : \lambda \in \Lambda\} \subseteq \mathcal{C}$ be given. Then $A = \bigcap_{\lambda\in\Lambda} S_\lambda$ is the largest member of $\mathcal{C}$ that is contained in all the S_λ's; thus it is the **infimum** of the S_λ's. Also, $B = \mathrm{cl}(\bigcup_{\lambda\in\Lambda} S_\lambda)$ is the smallest member of $\mathcal{C}$ that contains all the S_λ's; thus it is the **supremum** of the S_λ's.

Some Special Types of Moore Closures

4.6. Many Moore closures used in applications are formulated as closures with respect to some sort of operations. Let X be a set, and let Ψ be a collection of Λ-ary operations on X (defined as in 1.41). Different members of Ψ may have different Λ's, and we permit the Λ's to be empty or nonempty, finite or infinite. We shall say that a set $E \subseteq X$ is **closed under the operations** Ψ if it has this property:

> Whenever ψ is a Λ-ary operation in Ψ, with index set $\Lambda = \{\alpha, \beta, \gamma, \ldots\}$, and $e_\alpha, e_\beta, e_\gamma, \ldots$ are members of E, then $\psi(e_\alpha, e_\beta, e_\gamma, \ldots) \in E$ also.

Here the notation is as in 1.32; it is not intended to imply that the index set $\Lambda = \{\alpha, \beta, \gamma, \ldots\}$ is a countable or ordered.

It is easy to see that the sets that are "closed" in this sense satisfy Moore's axioms 4.3(i) and (ii). Hence they are the closed sets for a Moore closure operator, cl, defined as in 4.3. The closure obtained in this fashion is called the **closure with respect to the operations** Ψ. Here is an elementary example: A collection $\mathcal{S}$ of subsets of a set X is closed under finite union if and only if $\mathcal{S}$ is closed under the binary operation $\cup : \mathcal{P}(X) \times \mathcal{P}(X) \to \mathcal{P}(X)$.

Observation. The empty set is closed under the operations Ψ if and only if none of the operations in Ψ is nullary.

4.7. *Exercise (optional).* Actually, *every* Moore closure can be represented as a closure under operations (although such a representation is not necessarily helpful).

Hints: Let cl be a Moore closure on a set X — that is, let $\mathrm{cl} : \mathcal{P}(X) \to \mathcal{P}(X)$ be a given mapping satisfying the axioms in 4.5.a. For each set $\Lambda \subseteq X$ and each point $z \in \mathrm{cl}(\Lambda)$, define a Λ-ary operation $\psi_{\Lambda,z} : X^\Lambda \to X$ by taking

$$\psi_{\Lambda,z}(x_\alpha, x_\beta, x_\gamma, \ldots) \;=\; \begin{cases} z & \text{if } x_\lambda = \lambda \text{ for every } \lambda \in \Lambda \\ \text{some member of } \{x_\lambda : \lambda \in \Lambda\} & \text{otherwise.} \end{cases}$$

(In particular, when Λ is the empty set, then we form a nullary operation $\psi_{\varnothing,z}() = z$ for each z in $\mathrm{cl}(\varnothing)$, if there are any such z's.) Let Ψ be the collection of all operations formed in this fashion; verify that closure under the operations Ψ is the same as the given Moore closure.

4.8. *Theorem and definition.* Let X be a set, let $\mathcal{K}$ be a Moore collection of subsets of X, and let $\mathrm{cl} : \mathcal{P}(X) \to \mathcal{K}$ be the resulting Moore closure operator. Then the following conditions are equivalent. If one, hence all, of these conditions are satisfied, we say that $\mathcal{K}$ is an **algebraic closure system** and cl is an **algebraic closure operator**. (For some important examples, see 9.21.d and 12.3. See also the related exercise in 16.8.b.)

(A) cl is the closure with respect to some collection Ψ of finitary operations on X — that is, Λ-ary operations, where the Λ's are finite sets.

(B) Whenever $\mathcal{C}$ is a subset of $\mathcal{K}$ that is directed by inclusion, then the union of all the members of $\mathcal{C}$ is a closed set. In other words, if $\mathcal{C} \subseteq \mathcal{K}$ and the union of any two members of $\mathcal{C}$ is a subset of a member of $\mathcal{C}$, then the union of all the members of $\mathcal{C}$ is a member of $\mathcal{K}$.

(C) Whenever $\mathcal{D}$ is a collection of subsets of X that is directed by inclusion, then $\mathrm{cl}\left(\bigcup_{D\in\mathcal{D}} D\right) \subseteq \bigcup_{D\in\mathcal{D}} \mathrm{cl}(D)$.

(D) For set $S \subseteq X$ we have $\mathrm{cl}(S) = \bigcup\{\mathrm{cl}(F) : F \text{ is a finite subset of } S\}$.

Proof. (A) $\Rightarrow$ (B) is an easy exercise. For (B) $\Rightarrow$ (C), let $\mathcal{C} = \{\mathrm{cl}(D) : D \in \mathcal{D}\}$. For (C) $\Rightarrow$ (D), take $\mathcal{D} = \{F \subseteq S : F \text{ is finite}\}$.

It remains only to prove (D) $\Rightarrow$ (A). Let Ψ be the collection of all finitary operations $f : X^n \to X$ (for nonnegative integers n) that satisfy

$$f(\underbrace{K \times K \times \cdots \times K}_{n \text{ times}}) \quad \subseteq \quad K \qquad\qquad \text{for each } K \in \mathcal{K}.$$

Let $\mathcal{L}$ be the collection of subsets of X that are closed under the operations Ψ. We shall show that $\mathcal{K} = \mathcal{L}$. It is easy to see that $\mathcal{K} \subseteq \mathcal{L}$.

Now let any $S \in \mathcal{L}$ be given; we wish to show that $S \in \mathcal{K}$. Thus, it suffices to show that $\mathrm{cl}(S) \subseteq S$. We may assume $\mathrm{cl}(S)$ is nonempty and let any $\sigma \in \mathrm{cl}(S)$ be given; it suffices to show that $\sigma \in S$. By (D), there is some finite set $F \subseteq S$ such that $\sigma \in \mathrm{cl}(F)$.

We may write $F = \{u_1, u_2, \ldots, u_n\}$ for some nonnegative integer n (which is 0 if F is the empty set). Now define $f : X^n \to X$ by

$$f(x_1, x_2, \ldots, x_n) \quad = \quad \begin{cases} \sigma & \text{if } n = 0 \text{ or } (x_1, x_2, \ldots, x_n) = (u_1, u_2, \ldots, u_n) \\ x_1 & \text{otherwise.} \end{cases}$$

Here it is understood that if $n = 0$, then f is a nullary operation — i.e., a constant function. In that case the list of arguments $x_1, x_2, \ldots, x_n$ is empty; that is, $f(x_1, x_2, \ldots, x_n) = f()$.

We claim that $f \in \Psi$. Indeed, let any $K \in \mathcal{K}$ be given and any $x_1, x_2, \ldots, x_n \in K$; we are to show that $f(x_1, x_2, \ldots, x_n) \in K$. This is clear in the case where $f(x_1, x_2, \ldots, x_n) = x_1$. In the remaining case, we have $F \subseteq K$ and, therefore,

$$f(x_1, x_2, \ldots, x_n) \quad = \quad \sigma \quad \in \quad \mathrm{cl}(F) \quad \subseteq \quad \mathrm{cl}(K) \quad = \quad K.$$

Thus $f \in \Psi$. By assumption $S \in \mathcal{L}$, and so the set S is closed under the operation f. Hence $\sigma = f(u_1, u_2, \ldots, u_n)$ is a member of S.

4.9. *Definitions.* Let X and Y be sets. A **polar** from X to Y is a mapping $p : \mathcal{P}(X) \to \mathcal{P}(Y)$ that satisfies

$$p\left(\bigcup_{i \in I} A_i\right) = \bigcap_{i \in I} p(A_i)$$

for any collection $\{A_i : i \in I\}$ of subsets of X. The **dual** of p is the mapping $q : \mathcal{P}(Y) \to \mathcal{P}(X)$ defined by

$$q(B) = \bigcup\{S \subseteq X \ : \ B \subseteq p(S)\} \qquad \text{or, equivalently,} \qquad q(B) = \{x \in X : B \subseteq p(\{x\})\}.$$

We shall also write $p(A) = A^{\triangleleft}$ and $q(B) = B^{\triangleright}$. (In many books, one symbol is used for both $\triangleleft$ and $\triangleright$. Typically it is $\circ$ or $\perp$ or $\top$.)

Exercises/basic properties:

a. Any polar is antitone — i.e., if $A_1 \subseteq A_2$ in X, then $p(A_1) \supseteq p(A_2)$ in Y.

b. The dual of a polar is also a polar. Moreover, if q is the dual of p, then p is the dual of q. Thus we may speak of p, q as a **polar pair** between X and Y.

Examples of polars will be given in studying Dedekind cuts (see 4.34) and topological vector spaces (see 28.25); several other examples are also mentioned in 4.12.

4.10. Let X and Y be sets, and let $p : \mathcal{P}(X) \to \mathcal{P}(Y)$ and $q : \mathcal{P}(Y) \to \mathcal{P}(X)$ be some given functions. Then the following four conditions are equivalent.

(A) p and q are a polar pair.

(B) $p(A) = \{y \in Y : A \subseteq q(\{y\})\}$ and $q(B) = \{x \in X : B \subseteq p(\{x\})\}$ for all $A \subseteq X$ and $B \subseteq Y$.

(C) p and q are antitone (as in 4.9.a), and $q \circ p$ and $p \circ q$ are *extensive*:

$$A \subseteq q(p(A)), \qquad B \subseteq p(q(B)) \qquad \text{for all } A \subseteq X,\ B \subseteq Y.$$

(D) There exists a set $\Gamma \subseteq X \times Y$ such that, for all $A \subseteq X$ and $B \subseteq Y$,

$$p(A) = \{y \in Y \ : \ A \times \{y\} \subseteq \Gamma\}, \qquad q(B) = \{x \in X \ : \ \{x\} \times B \subseteq \Gamma\}.$$

This condition can be restated in the triangle notation as:

$$A^{\triangleleft} = \{y \in Y \ : \ A \times \{y\} \subseteq \Gamma\}, \qquad B^{\triangleright} = \{x \in X \ : \ \{x\} \times B \subseteq \Gamma\}.$$

4.11. *Further properties of polars.* Suppose $p : \mathcal{P}(X) \to \mathcal{P}(Y)$ and $q : \mathcal{P}(Y) \to \mathcal{P}(X)$ are a polar pair. Then:

a. $p \circ q \circ p = p$ and $q \circ p \circ q = q$. *Hint.* This follows easily from 4.10(C).

b. $q \circ p$ and $p \circ q$ are Moore closure operators in X and Y, respectively. The resulting closed subsets of X or Y are the sets A or B that satisfy $A = q(p(A))$ or $B = p(q(B))$, respectively.

c. Applied to just the collections of closed sets, the polar maps $A \mapsto p(A)$ and $B \mapsto q(B)$ are inverses of each other; they give a bijection between the closed subsets of A and the closed subsets of B.

d. Let $\mathrm{cl} = q \circ p$. Then $\mathrm{cl}\left(\bigcap_{i \in I} A_i\right) = \bigcap_{i \in I} \mathrm{cl}(A_i)$ for any collection $\{A_i : i \in I\}$ of subsets of X — that is, the closure of an intersection equals the intersection of the closures. An analogous result also holds in Y for $p \circ q$.

This result does *not* generalize to all Moore closures. For a simple counterexample, let cl be the usual topological closure when $\mathbb{R}$ is equipped with its usual topology. Let $A_1 = \{\text{rational numbers}\}$ and $A_2 = \{\text{irrational numbers}\}$. Then $A_1 \cap A_2 = \varnothing$, but $\mathrm{cl}(A_1) = \mathrm{cl}(A_2) = \mathbb{R}$. Consequently, $\mathrm{cl}(A_1 \cap A_2) = \varnothing$ but $\mathrm{cl}(A_1) \cap \mathrm{cl}(A_2) = \mathbb{R}$.

4.12. *Generalized orthogonality.* We now describe a special type of polar pair. Assume X is a set, 0 is some special member of X, and $\perp$ is a relation on X with these properties:

(i) $\perp$ is symmetric; that is, $x \perp y \iff y \perp x$.

(ii) $0 \perp x$ for all $x \in X$.

(iii) $x \perp x \iff x = 0$.

If $x \perp y$, usually we say that x and y are **orthogonal** (or **perpendicular**).

We now apply the results of the preceding sections, using 4.10(D) with $\Gamma = \{(x, y) \in X \times X \ : \ x \perp y\}$. Then the polars $p(S) = S^{\lhd}$ and $q(S) = S^{\rhd}$ are the same; we shall denote them both by $S^{\perp}$. Thus, $S^{\perp} = \{y \in X : x \perp y \text{ for all } x \in S\}$. This set is called the **orthogonal complement** of S. Let us first restate, in the present notation, some of the conclusions already reached in the preceding sections:

$S \subseteq S^{\perp\perp}$ and $S^{\perp} = S^{\perp\perp\perp}$.

$S \subseteq T \Rightarrow S^{\perp} \supseteq T^{\perp}$. (Thus the mapping $S \mapsto S^{\perp}$ is antitone.)

$S \mapsto S^{\perp\perp}$ is a Moore closure operator on X. We shall denote it by cl, at least for the moment. *Caution*: This operator is *not* called a "closure" in most specialized contexts where it is applied. Instead it is given other names, such as "closed linear span."

$\left(\bigcup_{\lambda \in \Lambda} S_\lambda\right)^{\perp} = \bigcap_{\lambda \in \Lambda} \left(S_\lambda^{\perp}\right)$ and $\left(\bigcap_{\lambda \in \Lambda} S_\lambda\right)^{\perp} = \mathrm{cl}\left(\bigcup_{\lambda \in \Lambda} \left(S_\lambda^{\perp}\right)\right)$.

$\mathrm{cl}(\bigcap_{\lambda \in \Lambda} S_\lambda) = \bigcap_{\lambda \in \Lambda} \mathrm{cl}(S_\lambda)$.

We also have a few new conclusions, which do not apply to all polar pairs. Show:

$\varnothing^{\perp} = \{0\}^{\perp} = X$ and $X^{\perp} = \{0\}$. Hence $\mathrm{cl}(\varnothing) = \{0\}$. Thus the empty set is not a closed set. (Hence this Moore closure is not a topological closure; see 5.19.)

Examples.

a. Let $X = \mathbb{R}^\Omega$ for some set Ω, and let $x \perp y$ mean that $x(\omega)y(\omega) = 0$ for all $\omega \in \Omega$. Show that a set $S \subseteq X$ is closed (in the sense of the Moore closure) if and only if it is of the form $C_M = \{x \in X : x\big|_M = 0\}$ for some set $M \subseteq \Omega$. Show also that $(C_M)^\perp = C_{\Omega \setminus M}$.

b. (This example requires some familiarity with analytic geometry from college calculus.) Let $X = \mathbb{R}^2$ and define $x \perp y$ to mean $x \cdot y = 0$ — that is, $x_1y_1 + x_2y_2 = 0$. Represent X by points in the plane. When x and y are both nonzero, show that $x \cdot y$ means that the line segment from the origin to x is perpendicular (in the usual geometric sense) to the line segment from the origin to y. Show that a subset of X is closed (in the sense of this Moore closure) if and only if it is either $\{0\}$, X, or a straight line through 0. For an example of a set S that is not closed, let S be the line segment from the point $(1,0)$ to the point $(2,0)$; show that the closure of S is the entire line $\{(x,y) : y = 0\}$.

More examples will be given later in this book, in Riesz spaces (see 11.59) and in Hilbert spaces (see 22.50). In 13.4.f we shall see that the collection of closed sets obtained in the fashion indicated above is a complete Boolean lattice.

Lattices and Completeness

4.13. *Definitions.* Let $(X, \preccurlyeq)$ be a poset. We say that

> $(X, \preccurlyeq)$ is a **Dedekind complete** poset if either of the following equivalent conditions holds. (*Exercise.* Prove the equivalence.) (A) Whenever $S \subseteq X$ is nonempty and bounded above, then S has a least upper bound in X. (B) Whenever $S \subseteq X$ is nonempty and bounded below, then S has a greatest lower bound in X.
>
> $(X, \preccurlyeq)$ is a **lattice** if every two-element subset of X has a sup and an inf in X.
>
> $(X, \preccurlyeq)$ is **complete** (or **order complete**, or a **complete lattice**) if every subset of X has a sup and an inf in X.

Examples are given later in this chapter.

4.14. *Remarks.* The term "complete" generally means "not missing any parts," or "not having any holes or gaps," but this has several different meanings in different parts of mathematics. We also caution that the term "complete poset" has a more specialized and technical meaning among some algebraists — e.g., in domain theory.

Uniform completions will be studied in later chapters. There are some strong analogies between the theories of order completions and uniform completions. It is possible to develop those analogies into a unifying theory,[1] but that theory is rather technical and complicated and not recommended for beginners. Readers of this book are urged to instead view order completeness and uniform completeness as two entirely unrelated concepts that, just by coincidence, use some of the same words and have slightly analogous meanings.

Formal logic also uses the term "complete" to mean "without holes," but the precise meaning is not closely related to order or uniform completeness. See 14.58.

4.15. *Relations between types of posets.*

- **a.** If $\preccurlyeq$ has any of the following properties, then $\succcurlyeq$ has the same property: lattice ordering; Dedekind complete; order complete.
- **b.** Any well ordered set is Dedekind complete.
- **c.** Every chain is a lattice.
- **d.** Any lattice is both a poset and a directed set.
- **e.** A poset $(X, \preccurlyeq)$ is order complete if and only if it is order bounded and Dedekind complete.
- **f.** If $(X, \preccurlyeq)$ is a Dedekind complete poset and both $\preccurlyeq$ and $\succcurlyeq$ are directed, then $(X, \preccurlyeq)$ is a lattice.

4.16. *Observations on products.* With the product ordering 3.9.j, a product of lattices is a lattice; a product of complete lattices is a complete lattice; a product of Dedekind complete posets is a Dedekind complete poset. In each case the supremum or infimum in the product is defined pointwise; see 3.21.n.

See also the corollary in 4.28.

More about Lattices

4.17. If $(X, \preccurlyeq)$ is a lattice, then the binary operation $\wedge : X \times X \to X$ is both

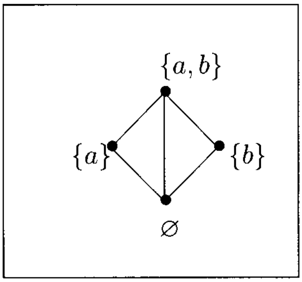
$$\text{commutative:} \quad x_1 \wedge x_2 = x_2 \wedge x_1 \quad \text{and}$$
$$\text{associative:} \quad (x_1 \wedge x_2) \wedge x_3 = x_1 \wedge (x_2 \wedge x_3).$$

It follows that the operations in $x_1 \wedge x_2 \wedge x_3 \wedge \cdots \wedge x_n$ can be evaluated in any order — left to right, right to left, center to outside, etc. — and thus parentheses are not necessary. The value of the expression is the same as $\inf\{x_1, x_2, \ldots, x_n\}$. (*Hint*: 3.21.m.) Analogous conclusions apply for $\vee$'s and sups.

An equivalent definition of lattice is: A poset in which every finite nonempty subset has a sup and an inf.

[1]See "reflective subcategories," in books on category theory, for other examples besides order completions and uniform completions.

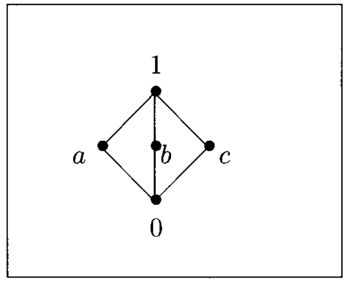
4.18. Lattices, particularly finite ones, can be illustrated with **lattice diagrams**. Elements of the lattice are indicated by vertices — i.e., dots. In these diagrams, we have $x \succcurlyeq y$ if there is a downward path from x to y. Two examples are given below.



2^2

Lattice diagrams

M_3



The first diagram shows the inclusion relation between the subsets of a two-element set. This lattice is known (among some lattice theorists) as 2^2.

The second diagram shows a lattice containing five members; 0 is the smallest member and 1 is largest. This lattice is sometimes known as M_3.

4.19. *Miscellaneous properties.*

a. In a lattice, the union of two order bounded sets is order bounded. (Hence, in a lattice, the order bounded sets form an *ideal* of sets, in the sense of 5.2.)

b. Not every subset of a lattice is a lattice; not every subset of a directed set is directed. For instance, $\mathbb{Z}^2$ with the product ordering is a lattice but its subset $\{(x, y) \in \mathbb{Z}^2 : x + y = 0\}$ is not directed.

4.20. Meet-join characterization of lattices. We have defined "lattice" in terms of its ordering $\preccurlyeq$, but we shall now show that "lattice" can be defined instead in terms of the binary operations $\wedge$ and $\vee$. Show that these laws are satisfied, for all x, y, z in a lattice:

L1 (commutative): $x \wedge y = y \wedge x$ and $x \vee y = y \vee x$,

L2 (associative): $x \wedge (y \wedge z) = (x \wedge y) \wedge z$ and $x \vee (y \vee z) = (x \vee y) \vee z$,

L3 (absorption): $x \wedge (x \vee y) = x$ and $x \vee (x \wedge y) = x$,

and

$$(*) \quad x \preccurlyeq y \iff x \wedge y = x \iff x \vee y = y.$$

Conversely, suppose X is a set equipped with two binary operations $\wedge$, $\vee$ that satisfy L1–L3. Show that $x \wedge y = x \iff x \vee y = y$. Define $\preccurlyeq$ by $(*)$; then show that $(X, \preccurlyeq)$ is a lattice. (*Hint*: First use L3 to prove that $x \vee x = x \wedge x = x$.)

4.21. Let $(X, \preccurlyeq)$ be a lattice. Then a **sublattice** of X is a subset S that is closed under the lattice operations $\vee, \wedge$ — i.e., that satisfies

$$s_1, s_2 \in S \quad \Rightarrow \quad s_1 \vee s_2,\ s_1 \wedge s_2 \in S.$$

It then follows that S is also a lattice, when equipped with the restrictions of $\vee$, $\wedge$.

If $(X, \preccurlyeq)$ is a lattice, S is a subset, and S is a lattice when equipped with the restriction of the ordering $\preccurlyeq$, it does *not* follow that S is necessarily a sublattice of X.

The collection of all sublattices of a lattice X is a Moore collection of subsets of X. The closure of any set $S \subseteq X$ is the smallest sublattice containing S; it is called the **sublattice generated by** S.

Example. Let $V = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \;:\; x_1 + x_2 = x_3\right\}$. Then V is a subset (in fact, a linear subspace) of $\mathbb{R}^3$. We shall order V by the restriction of the product ordering; that is,

$$(x_1, x_2, x_3) \preccurlyeq (y_1, y_2, y_3) \qquad \text{means} \qquad x_j \leq y_j \quad \text{for all } j.$$

Then V is a lattice (in fact, a vector lattice), but the lattice operations $\vee, \wedge$ determined on V by the ordering $\preccurlyeq$ are *not* simply the restrictions of the lattice operations on $\mathbb{R}^3$. Rather, the reader should verify that

$$\begin{aligned}(x \vee y)_1 &= \max\{x_1, y_1\},\\ (x \vee y)_2 &= \max\{x_2, y_2\},\\ (x \vee y)_3 &= \Big[(x \vee y)_1 + (x \vee y)_2\Big]\end{aligned}$$

and $\wedge$ is computed analogously with minima.

For instance, let $x = (1, 2, 3)$ and $y = (3, -1, 2)$. Then the lattice operations of $\mathbb{R}^3$ yield $x \vee y = (3, 2, 3)$, which is not a member of V; the lattice operations of V (defined by the formulas above) yield $x \vee y = (3, 2, 5)$.

This example may seem somewhat contrived, but it is actually quite typical of the behavior one sees in lattices of measures, which are discussed in later chapters.

4.22. *Example.* The set $\mathbb{N} = \{\text{positive integers}\}$ is a lattice when ordered by this rule: $x \preccurlyeq y$ if x is a divisor of y — that is, if $xu = y$ for some $u \in \mathbb{N}$. With this ordering, $u \vee v$ and $u \wedge v$ are the least common multiple and greatest common divisor of u and v, respectively.

A sublattice of $\mathbb{N}$ is given by $\{\text{divisors of } m\}$, for any positive integer m.

4.23. *Definition.* For a lattice $(X, \preccurlyeq)$, the following two conditions are equivalent:

(A) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ for all $x, y, z \in X$.

(B) $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ for all $x, y, z \in X$.

If these conditions are satisfied, we say $(X, \preccurlyeq)$ is a **distributive** lattice.

4.24. *Definition.* We shall say that a lattice $(X, \preccurlyeq)$ is **semi-infinitely distributive** if it satisfies either of the following conditions:

(A′) $x \wedge \sup(S) = \sup_{s \in S}(x \wedge s)$,

(B′) $x \vee \inf(S) = \inf_{s \in S}(x \vee s)$;

where the equations are to be interpreted in this sense: If the left side of the equation exists, then so does the right side, and they are equal. If both of these two conditions are satisfied, the lattice X is **infinitely distributive**. It is clear that any semi-infinitely distributive lattice is distributive. In 5.21 we shall give an example of a semi-infinitely distributive lattice that is not infinitely distributive; thus the two laws (A′) and (B′) are not equivalent to each other.

Exercise. Let X be a lattice. Show that conditions (A′) and (B′) are respectively equivalent to the following two conditions:

$$\text{(A}''\text{)} \quad \sup(R) \wedge \sup(S) = \sup\{r \wedge s \ : \ r \in R, \ s \in S\},$$

$$\text{(B}''\text{)} \quad \inf(R) \vee \inf(S) = \inf\{r \vee s \ : \ r \in R, \ s \in S\},$$

for any nonempty sets $R, S \subseteq X$. Again, each equation is to be interpreted in this fashion: If the left side of the equation exists, then so does the right side, and they are equal. *Hint*: See 3.21.m.

4.25. *Examples.*

a. If Ω is any set, then $(\mathcal{P}(\Omega), \subseteq)$ is an infinitely distributive lattice. (See 1.29.b.)

b. The five-element lattice M_3 is not distributive. (See 4.18.)

c. Every chain is an infinitely distributive lattice.

Further examples of infinitely distributive lattices will be given in 8.43.

4.26. A **lattice homomorphism** is a mapping $f : X \to Y$, from one lattice into another, that satisfies $f(x_1 \vee x_2) = f(x_1) \vee f(x_2)$ and $f(x_1 \wedge x_2) = f(x_1) \wedge f(x_2)$ for all x_1, x_2 in X. Lattice homomorphisms will be studied further in 8.48 and thereafter.

More about Complete Lattices

4.27. *Some important examples.* In Chapter 10, in our formal development of $\mathbb{R}$, we shall show that $\mathbb{R}$ is Dedekind complete. (More precisely, we shall prove that there exists a *unique* Dedekind complete ordered field and then *define* $\mathbb{R}$ to be that field.) For now, however, we shall "borrow" that result from Chapter 10: We shall accept the fact that $\mathbb{R}$ is Dedekind complete and use that fact in some examples below.

The extended real line, $[-\infty, +\infty]$, was introduced in 1.17. Recall that it is obtained by adjoining two new objects, $-\infty$ and $+\infty$, to the real number system and defining $-\infty < r < +\infty$ for all real numbers r. It follows that $[-\infty, +\infty]$ is a chain that is order complete.

4.28. *Observation.* Let Λ be any nonempty set. Then $\mathbb{R}^\Lambda = \{\text{functions from } \Lambda \text{ into } \mathbb{R}\}$ is Dedekind complete, and $[-\infty, +\infty]^\Lambda = \{\text{functions from } \Lambda \text{ into } [-\infty, +\infty]\}$ is a complete lattice, when these products are equipped with the product ordering.

4.29. *Miscellaneous properties.*

a. For any set X, the ordering $\subseteq$ makes $\mathcal{P}(X)$ into a complete lattice; hence it is both a directed ordering and a partial ordering. It is not a chain ordering if X contains more than one element.

b. Not every subset of a complete lattice is a complete lattice; not every subset of a Dedekind complete poset is Dedekind complete. For instance, $[0,1]$ (with its usual ordering) is order complete, but $\mathbb{Q} \cap [0,1]$ is not Dedekind complete.

c. Let X and Y be complete lattices, and suppose $f : X \to Y$ is an order-preserving function — that is, $x_1 \preccurlyeq x_2 \Rightarrow f(x_1) \preccurlyeq f(x_2)$. Then

$$\bigvee_{\lambda \in \Lambda} f(x_\lambda) \quad \preccurlyeq \quad f\left(\bigvee_{\lambda \in \Lambda} x_\lambda\right) \qquad \text{and} \qquad f\left(\bigwedge_{\lambda \in \Lambda} x_\lambda\right) \quad \preccurlyeq \quad \bigwedge_{\lambda \in \Lambda} f(x_\lambda)$$

for any set $\{x_\lambda : \lambda \in \Lambda\} \subseteq X$. Neither of these $\preccurlyeq$'s is necessarily equality; that is evident from an example in 15.11.

4.30. Tarski's Fixed Point Theorem. Let $(L, \preccurlyeq)$ be a complete lattice, and suppose $f : L \to L$ is isotone. Then f has at least one fixed point — i.e., there exists at least one point $p \in L$ such that $f(p) = p$.

Furthermore, among the fixed points there is a largest one. In fact, that largest fixed point is also the largest member of the set $S = \{x \in L : x \preccurlyeq f(x)\}$.

Hints: S is nonempty, since it includes the first member of L. Let $p = \sup(S)$; show that $p \in S$; show that $f(p) = p$.

Order Completions

4.31. *Definitions.* A set $D \subseteq X$ is **sup-dense** in X if X is the sup closure of D — i.e., if each point ξ in X is the sup (in X) of some nonempty subset of D. It is easy to see that in this case

$$L_\xi \;=\; \{d \in D \,:\, d \preccurlyeq \xi\} \quad \text{is nonempty, and} \quad \xi = \sup_X(L_\xi).$$

Dually, a set $D \subseteq X$ is **inf-dense** in X if X is the inf closure of D — i.e., if point ξ in X is the inf (in X) of some nonempty subset of D. It is easy to see that in this case

$$U_\xi \;=\; \{d \in D \,:\, d \succcurlyeq \xi\} \quad \text{is nonempty, and} \quad \xi = \inf_X(U_\xi).$$

4.32. *Proposition.* Let $(X, \preccurlyeq)$ be a poset, and let $D \subseteq X$. Let D be ordered by the restriction of the ordering $\preccurlyeq$; let this restriction be denoted again by $\preccurlyeq$. Then:

(i) If D is inf-dense in $(X, \preccurlyeq)$, then the inclusion map $D \xrightarrow{\subseteq} X$ is sup-preserving from $(D, \preccurlyeq)$ to $(X, \preccurlyeq)$.

(ii) If D is sup-dense in $(X, \preccurlyeq)$, then the inclusion map $D \xrightarrow{\subseteq} X$ is inf-preserving from $(D, \preccurlyeq)$ to $(X, \preccurlyeq)$.

Outline of proof. We shall only prove (i); then (ii) follows since it is dual to (i).

We shall use "$\sup_D$" and "$\sup_X$" to denote the supremum in D or in X, respectively; denote an infimum analogously.

Let $S \subseteq D$ be nonempty, and assume that $\sigma = \sup_D(S)$ exists. Then σ is also an upper bound for S in X; we wish to show that it is the *least* upper bound in X. Let β be any upper bound for S in X; we wish to show that $\sigma \preccurlyeq \beta$.

Define L_ξ, U_ξ as in 4.31. Consider any $d \in U_\beta$. Then $d \succcurlyeq \beta \succcurlyeq s$ for every $s \in S$. Hence d is an upper bound for S, and d lies in D. Since σ is the *least* of all the upper bounds for S that lie in D, it follows that $\sigma \preccurlyeq d$. Thus $d \in U_\sigma$ — i.e., we have shown that $U_\beta \subseteq U_\sigma$. Since the infimum operation is antitone (see 3.21.1), it follows that $\inf_X(U_\beta) \succcurlyeq \inf_X(U_\sigma)$ — that is, $\beta \succcurlyeq \sigma$.

4.33. The literature contains many different kinds of order completions. (A survey of different kinds of completions applicable to lattice groups was given by Ball [1989].) The following notion of completion seems to be best suited for the purposes of this book.

Definition. Let D and X be posets, with partial orderings both denoted by $\preccurlyeq$. We shall say that X is a **Dedekind completion** of D if

(i) $D \subseteq X$, and the ordering of D is the restriction to D of the ordering of X;

(ii) D is both sup-dense and inf-dense in X; and

(iii) $(X, \preccurlyeq)$ is Dedekind complete.

Note that the inclusion $D \xrightarrow{\subseteq} X$ is then sup-preserving and inf-preserving, by 4.32.

This type of completion might be more precisely named a "generalized Dedekind completion," since the term "Dedekind completion" usually refers to chains. See also 4.36.c.

4.34. Existence Theorem. Every poset has a Dedekind completion.

Proof. Let $(D, \preccurlyeq)$ be the given poset. Define $\Gamma = \{(u, v) \in D \times D \ : \ u \preccurlyeq v\}$, and define a polar pair between D and itself as in 4.10(D). By a **cut** we shall mean an ordered pair (A, B) such that

(1) A and B are nonempty subsets of D, and

(2) $A = B^{\triangleright}$ and $B = A^{\triangleleft}$ in the sense of 4.9.

Note that condition (2) can be restated as:

(2a) A is the set of lower bounds of B, and

(2b) B is the set of upper bounds of A.

It follows that A is down-closed and B is up-closed.

Let X be the set of all cuts. Show that any cuts (A_1, B_1) and (A_2, B_2) satisfy $A_1 \subseteq A_2 \iff B_1 \supseteq B_2$; hence we may define a partial ordering $\sqsubseteq$ on X by:

$$(A_1, B_1) \sqsubseteq (A_2, B_2) \qquad \iff \qquad A_1 \subseteq A_2 \qquad \iff \qquad B_1 \supseteq B_2.$$

For each $d \in D$ define

$$j(d) \quad = \quad \Big(\{e \in D : e \preccurlyeq d\},\ \{e \in D : e \succcurlyeq d\} \Big).$$

It is easy to verify that $j(d)$ is a cut and that the mapping $j : D \to X$ is injective; hence we may view D as a subset of X by identifying each $d \in D$ with its image $j(d)$. Verify that $d_1 \preccurlyeq d_2 \iff j(d_1) \sqsubseteq j(d_2)$; thus the ordering of D is the restriction of the ordering of X.

To show that $(X, \sqsubseteq)$ is Dedekind complete, let $S = \{(A_\lambda, B_\lambda) : \lambda \in \Lambda\}$ be a nonempty subset of X. Verify that

> if S has a $\sqsubseteq$-lower bound, then the $\sqsubseteq$-inf of S is the pair (A_-, B_-), where $A_- = \bigcap_{\lambda \in \Lambda} A_\lambda$ and $B_- = p(A_-)$; and
>
> if S has a $\sqsubseteq$-upper bound, then the $\sqsubseteq$-sup of S is the pair (A_+, B_+), where $B_+ = \bigcap_{\lambda \in \Lambda} B_\lambda$ and $A_+ = q(B_+)$.

Use those two facts to show that D is inf-dense and sup-dense in X; verify that any cut (A, B) is the $\sqsubseteq$-infimum of $\{j(b) : b \in B\}$ and the $\sqsubseteq$-supremum of $\{j(a) : a \in A\}$.

4.35. *Example.* Let $\mathbb{Q} = \{\text{rational numbers}\}$ have its usual ordering. Let

$$L \;=\; \{q \in \mathbb{Q} \;:\; q < 0 \text{ or } q^2 < 2\}, \qquad U \;=\; \{q \in \mathbb{Q} \;:\; q > 0 \text{ and } q^2 > 2\}.$$

Then the pair (L, U) is a cut in $\mathbb{Q}$, in the sense of the proof in 4.34. We shall see in Chapter 10 that the order completion of $\mathbb{Q}$ is the real number system $\mathbb{R}$; the cut (L, U) described above corresponds to the number $\sqrt{2}$ in $\mathbb{R}$.

4.36. *Further properties of the Dedekind completion.* Let X be a Dedekind completion of a poset D. Then:

a. X has a first element if and only if D has a first element, in which case they are the same. Similarly for last elements.

b. If D is a lattice, then X is a lattice. *Hint*: 4.15.f.

c. Let $(D, \preccurlyeq)$ be any poset. Then there exists a complete lattice $(X, \preccurlyeq)$ with $D \xrightarrow{\subseteq} X$, such that the inclusion is both sup- and inf-preserving. (Such a complete lattice X is sometimes called a **MacNeille completion** of D.)

Hint: Adjoin a lowest element and a highest element, and then take the Dedekind completion. (Or take the Dedekind completion first if you prefer; the result will be the same.)

4.37. *Remarks.* In the theorem below we shall consider Dedekind completions only for chains. The theorem can be extended to a more general setting, but it then becomes more complicated; we shall not need the greater generality later in this book. We are mainly concerned with completing $\mathbb{Q}$ to obtain $\mathbb{R}$. Most other Dedekind complete structures used in analysis can be obtained by putting together copies of $\mathbb{R}$ in various ways.

4.38. Theorem on completions of chains.

(i) **(Linear ordering.)** If X is a Dedekind completion of a chain D, then X is also a chain.

(ii) **(Extension of mappings.)** If X is a Dedekind completion of a chain D, and Q is another chain that is Dedekind complete, and $f : D \to Q$ is a sup-preserving mapping, then f extends uniquely to a sup-preserving mapping $F : X \to Q$. In fact, F must be defined by this formula:

$$F(\xi) \quad = \quad \sup f(L_\xi) \quad = \quad \sup\{f(d) : d \in L_\xi\}, \qquad (**)$$

where $L_\xi = \{d \in D : d \leq \xi\}$.

(iii) **(Uniqueness of completions.)** The Dedekind completion of a chain D is *unique up to isomorphism over* D, in the following sense: If X_1, X_2 are two Dedekind completions of D, then there exists an order isomorphism from X_1 onto X_2 that maps each member of D to itself.

Proof of linear ordering. Define L_ξ, U_ξ as in 4.31. Suppose X is not a chain. Then there exist two distinct elements $x, y \in X$ that are not comparable — i.e., such that neither $x \preccurlyeq y$ nor $y \preccurlyeq x$ is valid. Consider any $a \in L_x$ and $b \in U_y$. Then we cannot have $a \geq b$, since that would imply $x \succcurlyeq a \succcurlyeq b \succcurlyeq y$. Since D is a chain, it follows that $a < b$. Hold a fixed, and let b vary over all of U_y; thus a is a lower bound for U_y, so $a \in L_y$. This reasoning is applicable for every $a \in L_x$, so $L_x \subseteq L_y$. Since $x = \sup(L_x)$ and $y = \sup(L_y)$, it follows that $x \preccurlyeq y$.

Proof of extension of mappings. Define L_ξ, U_ξ as in 4.31. For each $x \in X$, the set L_x is nonempty and is bounded above in D by any $d_1 \in U_x$. Therefore $f(d_1)$ is an upper bound for the set $f(L_x) = \{f(d) : d \in L_x\}$. Since Q is Dedekind complete, $\sup f(L_x)$ exists in Q. Hence a function $F : X \to Q$ can be defined by $(**)$.

It is easy to see that this function is increasing and is an extension of f, since f is sup-preserving on D. If f has a sup-preserving extension $F : X \to Q$, that extension must satisfy $(**)$, since $x = \sup(L_x)$.

It suffices to show the function F defined by $(**)$ is indeed sup-preserving. Let S be a nonempty subset of X, and suppose $\sigma = \sup(S)$ in X; we are to show that $q = \sup\{F(s) : s \in S\}$ exists in Q and equals $F(\sigma)$.

For simplicity of notation, we may replace S with the set $\{x \in X : x \leq s$ for some $s \in S\}$; this does not affect our hypotheses or desired conclusion. Thus we may assume S is down-closed in X. Hence $S \cap D = \bigcup_{s \in S} L_s$. For each $s \in S$ we have $s = \sup(L_s)$, and therefore

$$\sigma \; = \; \sup(S) \; = \; \sup\{\sup(L_s) : s \in S\} \; = \; \sup\left(\bigcup_{s \in S} L_s\right) \; = \; \sup(S \cap D)$$

by 3.21.m. Also, from 3.25.d we see that $L_\sigma \subseteq S \cup \{\sigma\}$.

For each $s \in S$ we have $s \leq \sigma$ and hence $F(s) \leq F(\sigma)$; thus the set $\{F(s) : s \in S\}$ is bounded above by $F(\sigma)$. Since Q is Dedekind complete, it follows that $q = \sup\{F(s) : s \in S\}$ exists in Q and that $q \leq F(\sigma)$. It remains to show the reverse of this inequality. If $\sigma \notin D$, then $L_\sigma \subseteq S$, and so $F(\sigma) = \sup\{f(d) : d \in L_\sigma\} \leq \sup\{F(s) : s \in S\} = q$. On the other hand, if $\sigma \in D$, then (since f is sup-preserving on D)

$$f(\sigma) \;=\; f(\sup(S \cap D)) \;=\; \sup(f(S \cap D)) \;\leq\; \sup(F(S)) \;=\; q.$$

Proof of uniqueness of completions. For $k = 1, 2$, let $f_k : D \xrightarrow{\subseteq} X_k$ be the inclusion map. Using the Extension Property with $X = X_j$ (for $j = 1, 2$) and $Q = X_k$, we see that f_k extends uniquely to a sup-preserving mapping $F_{jk} : X_j \to X_k$. Thus F_{jk} is the only sup-preserving mapping from X_j into X_k that leaves elements of D fixed.

Since the identity map of X_1 is a sup-preserving map that leaves elements of D fixed, it follows that F_{11} is the identity map on X_1 and that this is the only sup-preserving map from X_1 into itself that leaves elements of D fixed. Analogous statements are valid for F_{22} and X_2.

The compositions $F_{21} \circ F_{12} : X_1 \to X_1$ and $F_{12} \circ F_{21} : X_2 \to X_2$ are sup-preserving maps that leave elements of D fixed. Hence these maps are the identity maps on X_1 and X_2, respectively. Therefore $F_{12} : X_1 \to X_2$ is an order isomorphism.

4.39. (*Optional remarks.*) Although the "Dedekind completion" defined in 4.33 is probably the simplest for the purposes of this book, some mathematicians may prefer a different sort of completion.

Let $(X, \preccurlyeq)$ be a poset. Let $S \subseteq X$ be partially ordered by the restriction of $\preccurlyeq$. We shall say that $(X, \preccurlyeq)$ is a **sup completion** of $(S, \preccurlyeq)$ if these further properties are satisfied:

(i) $(X, \preccurlyeq)$ is Dedekind complete.

(ii) The inclusion map $S \xrightarrow{\subseteq} X$ is sup-preserving, from $(S, \preccurlyeq)$ to $(X, \preccurlyeq)$.

(iii) If $(Q, \sqsubseteq)$ is any Dedekind complete poset and $f : (S, \preccurlyeq) \to (Q, \sqsubseteq)$ is any sup-preserving function, then f extends uniquely to a sup-preserving function $F : (X, \preccurlyeq) \to (Q, \sqsubseteq)$.

This definition is slightly more complicated than the one in 4.33. However, it has the following advantages: Every poset S has a sup completion X that is *unique*, in the sense that any two sup completions X_1, X_2 are order isomorphic via a map that acts as the identity on members of S. Moreover, S is sup-dense in its sup completion X; X is bounded if and only if S is bounded, in which case the two posets have the same maximum and same minimum; X is a chain if and only if S is a chain, in which case the sup completion agrees with the Dedekind completion (defined in 4.33). The Dedekind complete posets form a reflective subcategory of the category of posets, if we use sup-preserving maps for morphisms; this notion is developed in books on category theory. We shall not prove these results, but for the ambitious reader we provide a hint: To prove existence of a sup completion of S, let $X = \{C \subseteq S : C \text{ is nonempty, bounded above, down-closed, and sup-closed}\}$; then partially order X by $\subseteq$.

SUPS AND INFS IN METRIC SPACES

4.40. Let (X, d) be a pseudometric space (defined in 2.11), let $S \subseteq X$ be a nonempty subset, and let $x \in X$. Then the **distance** from x to S and the **diameter** of the set S are the numbers

$$\operatorname{dist}_d(x, S) \;=\; \inf_{s \in S} d(x, s), \qquad\qquad \operatorname{diam}_d(S) \;=\; \sup_{u,v \in S} d(u, v)$$

in $[0, +\infty]$. We may omit the subscript d if no confusion is likely. By convention, we define $\operatorname{diam}(\varnothing) = 0$.

(The existence of these infs and sups follows from the fact that $[0, +\infty]$ is order complete — a fact that will not be established rigorously until we investigate the real numbers carefully in Chapter 10. We shall "borrow" that result now, to give some important examples of sups and infs; we promise not to engage in any circular reasoning.)

A set S is **bounded** (or, more specifically, *metrically bounded*) if it has finite diameter. A function is sometimes called **bounded** if its range is a metrically bounded set. *Caution*: The term "bounded" has several other meanings (see 3.19.a, 23.1, 27.2, and 27.4). Fortunately, most meanings of "bounded" coincide, at least when applied to subsets of $\mathbb{R}$.

4.41. *Basic properties and examples.* Let (X, d) be a pseudometric space. Then:

a. $d(x, y) = \operatorname{dist}(x, \{y\})$.

b. $|\operatorname{dist}(x, S) - \operatorname{dist}(y, S)| \leq d(x, y)$ for any nonempty set $S \subseteq X$.

c. Any subset of a bounded set is bounded, and the union of finitely many bounded sets is bounded. Thus, the bounded subsets of X form an *ideal of sets*, in the sense of 5.2.

d. In a metric space, a set with diameter 0 contains at most one point.

e. Show that $d(x, y) = |x - y|$ and $e(x, y) = \min\{1, |x - y|\}$ are metrics on $\mathbb{R}$ that yield *different* collections of bounded sets. In most other respects, however, these metrics are equivalent — they yield the same topological structure and the same uniform structure (see 18.14).

f. Let Λ be any set. Then $\rho(f, g) = \sup\{|f(\lambda) - g(\lambda)| : \lambda \in \Lambda\}$ is a metric on $B(\Lambda) =$ {bounded functions from Λ into $\mathbb{R}$}.

Suppose d is a metric on the given set Λ. Then we may embed the metric space (Λ, d) in the metric space $(B(\Lambda), \rho)$, as follows: Fix any point in Λ; we shall denote it by "0" (although we do not assume any additive structure here). For each $\mu \in \Lambda$ define a function $f_\mu \in B(\Lambda)$ by

$$f_\mu(\lambda) \;=\; d(\lambda, \mu) \;-\; d(\lambda, 0) \qquad\qquad (\lambda \in \Lambda).$$

Verify that $\rho(f_\mu, f_\nu) = d(\mu, \nu)$. Thus $\mu \mapsto f_\mu$ is a distance-preserving map from Λ into $B(\Lambda)$, and so we may view Λ as a subset of $B(\Lambda)$.

The space $B(\Lambda)$ has certain special properties that will be of interest later: It is a Banach space. Thus the example above shows that *every metric space can be embedded isometrically in a Banach space.* See 19.11.f and 22.14.

4.42. Let X be a set, and let $f : X \times X \to [0, +\infty)$ be some function satisfying $f(x, y) = f(y, x)$. Then we can define a pseudometric d on X by

$$d(x,y) \quad = \quad \inf\left\{\sum_{j=1}^{m} f(a_{j-1}, a_j) \quad : \quad m \geq 0,\ a_0 = x,\ a_m = y\right\};$$

here the infimum is over all nonnegative integers m and all finite sequences $(a_j)_{j=0}^m$ in X that go from x to y. The existence of the infimum follows from the fact that $[0, +\infty]$ is order complete. We permit $m = 0$ in the case when $x = y$; then the sum is interpreted to be 0. This construction can be summarized informally as "the distance between two points is the shortest route connecting them." It is not hard to show that $d(x, y) \leq f(x, y)$ and that in fact d is the *largest* pseudometric that is less than or equal to f.

4.43. More generally, the formula above defines a pseudometric d if the function f is merely defined on a subset $D \subseteq X \times X$, provided that subset is large enough that for each pair $(x, y) \in X \times X$ there exists at least one finite sequence $(a_j)_{j=0}^m$ from x to y satisfying

$$(a_0, a_1),\ (a_1, a_2),\ (a_2, a_3),\ \ldots,\ (a_{m-1}, a_m) \ \in\ D.$$

We then define d to be the infimum of the sums over all such sequences. Again, we permit $m = 0$ when $x = y$. This more general construction will be used in 19.48.

4.44. Weil's Pseudometrization Lemma. Let $V_0, V_1, V_2, V_3, \ldots$ be a sequence of reflexive symmetric relations on a set X, satisfying $V_{n-1} \supseteq V_n \circ V_n \circ V_n$ for all $n \in \mathbb{N}$, with $V_0 = X \times X$. Then there exists a pseudometric d on X that satisfies

$$\{(x,y) \in X^2 : d(x,y) < 2^{-n}\} \ \subseteq\ V_n \ \subseteq\ \{(x,y) \in X^2 : d(x,y) \leq 2^{-n}\} \qquad (***)$$

for $n = 0, 1, 2, \ldots$. In fact, d may be selected as follows: Define

$$f(x,y) \quad = \quad \inf\{2^{-n} : (x,y) \in V_n\} \quad = \quad \begin{cases} 2^{-n} & \text{if } (x,y) \in V_n \setminus V_{n+1} \\ 0 & \text{if } (x,y) \in \bigcap_{n=0}^{\infty} V_n, \end{cases}$$

and then define d as in 4.42.

Remark. The literature contains several variants of this lemma. Some other formulations may be simpler, but the present formulation — which follows Murdeshwar [1983] — has the advantage that it can be applied directly in the settings of uniform spaces, topological groups, topological vector spaces, and locally solid vector lattices; see 16.16, 26.29, and 26.57.

Outline of proof. We begin by observing that

$$f(u_1, u_4) \leq 2\max\{f(u_1, u_2), f(u_2, u_3), f(u_3, u_4)\} \quad \text{for all } u_1, u_2, u_3, u_4 \in X. \qquad (1)$$

Indeed, if $2^{-n} = \max\{f(u_1, u_2), f(u_2, u_3), f(u_3, u_4)\}$, then (u_1, u_2), (u_2, u_3), (u_3, u_4) are all in V_n, so $(u_1, u_4) \in V_n \circ V_n \circ V_n \subseteq V_{n-1}$.

Next, by induction on m, we shall show that

$$f(x_0, x_m) \;\le\; 2\sum_{i=1}^{m} f(x_{i-1}, x_i) \qquad \text{for any } m \in \mathbb{N} \text{ and } x_0, x_1, \ldots, x_m \in X. \tag{2}$$

To see this, let $b = \sum_{i=1}^{m} f(x_{i-1}, x_i)$; we are to show $f(x_0, x_m) \le 2b$. If $b = 0$, then $(x_{i-1}, x_i) \in V_n$ for all i and n, hence $(x_0, x_m) \in V_n$ for all n, and we are done. Thus we may assume $b > 0$. Choose j as large as possible satisfying $\sum_{i=1}^{j} f(x_{i-1}, x_i) \le b/2$. Then $j < m$ and $\sum_{i=1}^{j+1} f(x_{i-1}, x_i) > b/2$; hence $\sum_{i=j+2}^{m} f(x_{i-1}, x_i) < b/2$. By two uses of the induction hypothesis we have

$$f(x_0, x_j) \le b \qquad \text{and} \qquad f(x_{j+1}, x_m) < b.$$

Also $f(x_j, x_{j+1}) \le b$ by our definition of b. By (1) we have $f(x_0, x_m) \le 2b$, completing the induction proof of (2).

Now define d as in 4.42. Then d is a pseudometric and $d \le f$. From (2) we have $f \le 2d$. The second inclusion in $(***)$ is obvious, since $d \le f$. For the first inclusion in $(***)$, suppose $d(x, y) < 2^{-n}$. By the definition of d, then, there exists a finite sequence $a_0, a_1, \ldots, a_m \in X$, with $a_0 = x$ and $a_m = y$ and $\sum_{j=1}^{m} f(a_{j-1}, a_j) < 2^{-n}$. By (2), then, $f(x, y) < 2^{-n+1}$. Since f takes on only the values 2^0, 2^{-1}, 2^{-2}, $\ldots$, and 0, we must have $f(x, y) \le 2^{-n}$, and hence $(x, y) \in V_n$.

Chapter 5

Filters, Topologies, and Other Sets of Sets

Filters and Ideals

5.1. Let $\mathcal{F}$ be a nonempty collection of subsets of a set X. We say $\mathcal{F}$ is a **filter** on X if

(i) $S \in \mathcal{F}$ and $S \subseteq T \subseteq X$ imply $T \in \mathcal{F}$, and

(ii) $S, T \in \mathcal{F} \Rightarrow S \cap T \in \mathcal{F}$.

(For clarification or emphasis we may sometimes call $\mathcal{F}$ a **filter of sets**.) Note that any such collection $\mathcal{F}$ necessarily satisfies $X \in \mathcal{F}$. Clearly, $\mathcal{P}(X)$ is a filter on X; we shall refer to it as the **improper filter**. Any other filter on X will be called a **proper filter**. It is easy to see that a filter $\mathcal{F}$ is proper if and only if

(iii) $\varnothing \notin \mathcal{F}$.

Our terminology here follows that of algebraists. However, we remark that many mathematicians — particularly topologists — use the term "filter" to refer only to collections satisfying all of (i), (ii), and (iii). We prefer the algebraists' terminology because in later chapters we shall use the duality between filters and ideals.

Elementary examples of filters and ideals are given in 5.5, and further examples (particularly of interest to analysts) are previewed in 5.6. An intuitive discussion is given in 5.3.

5.2. A nonempty collection $\mathcal{J}$ of subsets of a set X is an **ideal** on X if

(i) $S \in \mathcal{J}$ and $S \supseteq T$ imply $T \in \mathcal{J}$, and

(ii) $S, T \in \mathcal{J} \Rightarrow S \cup T \in \mathcal{J}$.

If the context is not clear, we might say that $\mathcal{J}$ is an **ideal of sets**, to distinguish it from the "ideal in an algebra" introduced in 9.25. We can also avoid ambiguity by referring to $\mathcal{J}$ as an **ideal in the Boolean algebra $\mathcal{P}(X)$** because in that setting the two notions of "ideal" coincide (see 13.17.a).

For any ideal $\mathcal{J}$, we have $\varnothing \in \mathcal{J}$, by (i). Clearly, $\mathcal{P}(X) = \{\text{subsets of } X\}$ is an ideal on X; we shall call it the **improper ideal**. Any other ideal on X is called a **proper ideal**. It is easy to see that an ideal $\mathcal{I}$ is proper if and only if

(iii) $X \notin \mathcal{I}$.

A **σ-ideal** is an ideal $\mathcal{I}$ that is closed under *countable* unions — i.e., such that $S_1, S_2, S_3, \ldots \in \mathcal{I} \Rightarrow S_1 \cup S_2 \cup S_3 \cup \cdots \in \mathcal{I}$.

There is a simple correspondence between filters and ideals. Let $\mathcal{F}$ be a collection of subsets of X, and let $\mathcal{I} = \{\complement S : S \in \mathcal{F}\}$, where $\complement$ denotes complementation in X; then $\mathcal{F}$ is a filter (proper filter, improper filter) if and only if $\mathcal{I}$ is an ideal (proper ideal, improper ideal respectively). We say that $\mathcal{F}$ and $\mathcal{I}$ are **dual** to each other. Any statement about $\mathcal{F}$ can be translated into a statement about $\mathcal{I}$, and vice versa, but some concepts can be expressed more simply in terms of filters or in terms of ideals.

Caution: The dual ideal $\{\complement S : S \in \mathcal{F}\}$ should not be confused with the other complementary set, $\mathcal{P}(X) \setminus \mathcal{F} = \{T \subseteq X : T \notin \mathcal{F}\}$. In general, $\mathcal{P}(X) \setminus \mathcal{F}$ is neither a filter nor an ideal. However, under special circumstances $\{\complement S : S \in \mathcal{F}\}$ is *equal* to the ideal $\mathcal{P}(X) \setminus \mathcal{F}$; see 5.8.

5.3. To better understand the definitions of filter and ideal, suppose $\mathcal{I}$ is a nonempty collection of subsets of a set X, and let $\mathcal{F}$ be the dual collection $\{X \setminus S : S \in \mathcal{I}\}$. Say that a set $S \subseteq X$ is "*small*" if $S \in \mathcal{I}$, or "*large*" if $S \in \mathcal{F}$ (i.e., if $X \setminus S$ is small).

Then $\mathcal{I}$ is an ideal and $\mathcal{F}$ is a filter if and only if

(i) any subset of a small set is small, and

(ii) the union of two small sets (or finitely many small sets) is small.

The ideal and filter are proper if and only if also

(iii) not every set is small.

If this third condition is satisfied, then a set $S \subseteq X$ cannot be both small and large. Can a set be neither small nor large? That depends on what $\mathcal{I}$ and $\mathcal{F}$ are; see 5.5.d and 5.8(B).

Our three rules (i), (ii), and (iii) are compatible with common nonmathematical usage of the words "small" and "large." However, different rules would also be compatible with the nonmathematical usage of those words, since nonmathematical usage deals only with finite sets; the mathematical usage also covers infinite sets. We have drawn this connection, not so much to explain small and large, but rather to explain ideal and filter.

Different ideals give us different collections of small sets. A set may be small with respect to one ideal while large with respect to another ideal; see the example in 24.39. Other words sometimes used in place of small are **negligible** and **null** (although the latter term also sometimes refers to the empty set).

Other words sometimes used in place of large are **residual** or **generic** — especially in the context of directed sets or in the context of Baire category theory. Also, a large subset of X is **almost all** of X. We might also say that a condition K on points $x \in X$ is satisfied **almost everywhere** or **almost always**, or is **$\mathcal{F}$-true** or **almost true**, if the set $\{x \in X \ : \ K \text{ is satisfied at } x\}$ is a member of $\mathcal{F}$.

This interpretation of "true" preserves some, but not all, of the usual features of that word — for instance, the conjunction of finitely many $\mathcal{F}$-true statements is $\mathcal{F}$-true (as with ordinary truth), but the conjunction of infinitely many $\mathcal{F}$-true statements is not necessarily

$\mathcal{F}$-true (unlike ordinary truth). This slightly unusual interpretation of "truth" is occasionally useful to logicians.

5.4. We have noted that $\mathcal{P}(X)$ is a filter on X, and we can easily verify that the intersection of any collection of filters on X is another filter on X. Thus, the collection of all filters on X is a Moore collection of subsets of $\mathcal{P}(X)$. Similarly, the collection of all ideals on X is a Moore collection.

Hence, given any $\mathcal{G} \subseteq \mathcal{P}(X)$, there exists a smallest filter (or ideal) that contains $\mathcal{G}$ — namely, the intersection of all the filters (or ideals) that contain $\mathcal{G}$. We call it the filter (respectively, the ideal) **generated** by $\mathcal{G}$; we say that $\mathcal{G}$ is a **generating set** for it. (This is a special case of the Moore closure, introduced in 4.3, but the terms "closed" and "closure" are generally not used for filters and ideals.)

We shall say that $\mathcal{F}$ is a **superfilter** of $\mathcal{G}$ whenever $\mathcal{F}$ is a filter and $\mathcal{F} \supseteq \mathcal{G}$. With this terminology, the filter generated by $\mathcal{G}$ is simply the smallest superfilter of $\mathcal{G}$.

Show that the filter generated by a collection $\mathcal{G} \subseteq \mathcal{P}(X)$ is

$$\mathcal{F} = \Big\{F \subseteq X \ : \ F \supseteq G_1 \cap G_2 \cap \cdots \cap G_n \quad \text{for some finite set } \{G_1, G_2, \ldots, G_n\} \subseteq \mathcal{G}\Big\}.$$

Dually, the ideal generated by a collection $\mathcal{G} \subseteq \mathcal{P}(X)$ is

$$\mathcal{I} = \Big\{I \subseteq X \ : \ I \subseteq G_1 \cup G_2 \cup \cdots \cup G_n \quad \text{for some finite set } \{G_1, G_2, \ldots, G_n\} \subseteq \mathcal{G}\Big\}.$$

We permit $n = 0$ in both formulas, with the conventions that the intersection of no subsets of X is just X and the union of no subsets of X is just $\varnothing$.

5.5. *Examples and elementary properties.* Let X be a set.

a. Degenerate examples: The singleton $\{X\}$ is the smallest filter; the singleton $\{\varnothing\}$ is the smallest ideal. These are dual to each other; they both are generated by the empty set.

b. Let A be a nonempty subset of X, and let $\mathcal{F}$ be a filter on X. Then the following conditions are equivalent:

(A) $\mathcal{F}$ is fixed (i.e., has nonempty intersection — see 1.26) and the intersection of its members is A.

(B) $\mathcal{F}$ is the filter generated by the singleton $\{A\}$.

(C) $\mathcal{F} = \{S \subseteq X : S \supseteq A\}$.

Assume that those conditions are satisfied. Then the filter $\mathcal{F}$ is dual to the ideal $\mathcal{P}(X \backslash A)$.

c. We note an important special case of the preceding example: Let $p \in X$, and take A to be the singleton $\{p\}$. Thus we obtain the filter

$$\mathcal{F}_p \quad = \quad \{S \subseteq X \,:\, S \supseteq \{p\}\} \quad = \quad \{S \subseteq X \,:\, p \in S\}.$$

It is the fixed filter generated by the singleton $\{\{p\}\}$. It is actually an *ultrafilter* (defined in 5.8); hence it is called the **ultrafilter fixed at** $\boldsymbol{p}$.

d. The **ideal of finite sets** is $\{S \subseteq X : S \text{ is finite}\}$. It is generated by the collection of all the singletons in X. The filter that is dual to this is the **cofinite filter**, $\{S \subseteq X : \complement S$ is finite$\}$, also known as the **Fréchet filter**. This ideal and filter are proper if and only if the set X is infinite.

Example. If subsets of $\mathbb{N}$ are classified as small or large as in 5.3, using the ideal of finite sets and the cofinite filter, then the set $\{1, 3, 5, 7, \ldots\}$ is neither small nor large.

e. Let $\mathcal{F}$ and $\mathcal{I}$ be a filter and ideal on X, dual to each other. Show that the following are equivalent:

(A) $\mathcal{F}$ is free (i.e., has empty intersection).

(B) $\mathcal{F}$ contains the cofinite filter.

(C) $\mathcal{I}$ is a cover of X.

(D) $\mathcal{I}$ contains the ideal of finite sets.

f. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be filters on X. We shall say that $\mathcal{G}_1$ **meets** $\mathcal{G}_2$ if every member of $\mathcal{G}_1$ meets (i.e., has nonempty intersection with) every member of $\mathcal{G}_2$. Show that there exists a proper filter $\mathcal{H} \supseteq \mathcal{G}_1 \cup \mathcal{G}_2$ if and only if $\mathcal{G}_1$ meets $\mathcal{G}_2$.

g. Let $\mathcal{F}$ be a filter on X, and suppose $Y \subseteq X$. Then $\mathcal{F}_Y = \{Y \cap F : F \in \mathcal{F}\}$ is a filter on Y, sometimes called the **trace** of $\mathcal{F}$ on Y. It is a proper filter if and only if $X \setminus Y \notin \mathcal{F}$.

h. Kowalsky's filter. (Optional; this will be used in 15.10.) Let I and X be sets, let $\mathcal{G}$ be a filter on I, and for each $i \in I$ let $\mathcal{F}_i$ be a filter on X. Then $\bigcup_{G \in \mathcal{G}} \bigcap_{i \in G} \mathcal{F}_i$ is a filter on X.

i. How to enlarge a filter. Let $\mathcal{F}$ be a proper filter on X, and let $K \subseteq X$; suppose that neither K nor $\complement K$ is an element of $\mathcal{F}$. Then $\{F \cap L : F \in \mathcal{F} \text{ and } K \subseteq L \subseteq X\}$ is a proper filter that contains $\{K\} \cup \mathcal{F}$.

5.6. *Preview of more examples.* Other important filters studied later are the collection of

- absorbing subsets of a vector space (see 12.8);
- neighborhoods of a point in a topological space (see 5.16.a);
- eventual sets of a net (see 7.9).

Other important ideals studied later are the collection of

- relatively compact subsets of a Hausdorff topological space (see 17.7.c);
- equicontinuous sets of maps between two uniform spaces (see 18.30.e);

- bounded subsets of a lattice (4.19.a), a metric space (4.41.c), or a topological vector space (27.3.b);
- precompact or totally bounded subsets of a uniform space (see 19.15.f);
- nowhere-dense subsets of a topological space, and meager subsets of a topological space — the latter is in fact a σ-ideal (see Chapter 20);
- null sets with respect to a positive charge — i.e., a finitely additive, positive set function; this is in fact a σ-ideal if the measure is countably additive (see 21.15).
- shy sets of a Banach space; this is in fact a σ-ideal — see 21.21.

5.7. We consider several useful generalizations of the notion of a proper filter. Let $\mathcal{G}$ be a nonempty collection of subsets of X; then $\mathcal{G}$ may or may not satisfy the following conditions:

(i) $\mathcal{G}$ is a proper filter on X.

(ii) Every member of $\mathcal{G}$ is nonempty, and $\mathcal{G}$ is closed under finite intersection.

(iii) $\mathcal{G}$ is a **filterbase** on X — that is, each member of $\mathcal{G}$ is nonempty, and for each pair of sets $A, B \in \mathcal{G}$ there exists some $C \in \mathcal{G}$ with $C \subseteq A \cap B$.

(iv) $\mathcal{G}$ is a **filter subbase** on X, or $\mathcal{G}$ has the **finite intersection property** — that is, the intersection of finitely many elements of $\mathcal{G}$ is always nonempty.

Clearly, (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). Show the following further results:

a. Let $\mathcal{G}$ be a collection of subsets of X. Show that the filter generated by $\mathcal{G}$ is a *proper* filter if and only if $\mathcal{G}$ is a filter subbase.

b. If $\mathcal{F}$ is a filter subbase on X and $S_1, S_2, \ldots, S_n$ are disjoint subsets of X, then at most one of the S_i's is an element of $\mathcal{F}$.

c. If $\mathcal{G}$ is a filterbase, then the filter generated by $\mathcal{G}$ is the proper filter

$$\mathcal{F} \quad = \quad \{S \subseteq X \quad : \quad S \supseteq G \text{ for some } G \in \mathcal{G}\}.$$

We then say that $\mathcal{G}$ is a **base** for the filter $\mathcal{F}$.

d. If $\mathcal{S}$ is a filter subbase, then $\mathcal{B} = \{$intersections of finitely many members of $\mathcal{S}\}$ satisfies condition (ii) above, and it generates the same filter as $\mathcal{S}$ does.

e. Let $X = \mathcal{P}(\Omega)$ for some set Ω. Then the set of all filter subbases on Ω is a collection of subsets of X that has finite character (defined in 3.46).

5.8. *Definition and exercise.* Let $\mathcal{F}$ be a nonempty collection of subsets of X. Show that the following conditions on $\mathcal{F}$ are equivalent. If any (hence all) of them are satisfied, we say $\mathcal{F}$ is an **ultrafilter**. *Hint*: For 5.8(C) $\Rightarrow$ 5.8(B), use 5.5.i.

(A) $\mathcal{F}$ is a proper filter, and the complementary set $\mathcal{P}(X) \setminus \mathcal{F} = \{S \subseteq X : S \notin \mathcal{F}\}$ is an ideal. (This is the special circumstance mentioned in the cautionary remark at the end of 5.2.)

(B) $\mathcal{F}$ is a proper filter on X that also satisfies: for each set $K \subseteq X$, either $K \in \mathcal{F}$ or $\complement K \in \mathcal{F}$. (In the terminology of 5.3, every subset of X is either large or small. Thus, the dual ideal $\{\complement S : S \in \mathcal{F}\}$ is equal to $\mathcal{P}(X) \setminus \mathcal{F}$.)

(C) $\mathcal{F}$ is a *maximal filter* on X (or more precisely, a maximal proper filter). That is, $\mathcal{F}$ is a proper filter on X, and no other proper filter on X contains $\mathcal{F}$.

(D) $\mathcal{F}$ is a maximal filter subbase on X — i.e., $\mathcal{F}$ is a filter subbase on X, and no other filter subbase contains $\mathcal{F}$.

(E) $\mathcal{F}$ is a proper filter on X, and whenever $S_1 \cup S_2 \cup \cdots \cup S_n \in \mathcal{F}$, then at least one of the S_i's is an element of $\mathcal{F}$.

(F) $\mathcal{F}$ is a proper filter on X, and whenever $S_1 \cup S_2 \in \mathcal{F}$, then at least one of S_1, S_2 is an element of $\mathcal{F}$.

5.9. Let $\mathcal{U}$ be an ultrafilter on X. Then one of the following two cases must hold:

(1) $\mathcal{U}$ is a **fixed ultrafilter** (i.e., having nonempty intersection — see 1.26 and 5.5.b). In this case $\mathcal{U}$ is also known as a **principal ultrafilter**. Show that in this case $\mathcal{U}$ is the ultrafilter $\mathcal{U}_p$ fixed at some point $p \in X$ (defined in 5.5.c).

(2) $\mathcal{U}$ is a **free ultrafilter** (i.e., having empty intersection — see 1.26). In this case $\mathcal{U}$ is also called a **nonprincipal ultrafilter**. Show that in this case $\mathcal{U}$ is a superset of the cofinite filter and no element of $\mathcal{U}$ is a finite set. In particular, this case cannot occur if X itself is a finite set.

5.10. *Remarks.* Free ultrafilters will play an important role in some later parts of this book. A free ultrafilter on a set X can be described as a classification of subsets of X into small sets and large sets, satisfying conditions 5.3 (i), (ii), and (iii) and also satisfying

(iv) every set is either small or the complement of a small set, and

(v) every finite set is small.

Our description of fixed ultrafilters in 5.5.c is quite constructive: It tells us explicitly how to form such objects. In contrast, our description of a free ultrafilter is indirect, and we find it difficult to visualize such an object. Before continuing to the next sentence, the reader is urged to try to give a completely explicit example of a free ultrafilter.

Surprisingly, free ultrafilters do exist, but explicit *examples* of free ultrafilters do *not* exist! Thus, free ultrafilters are our first *intangibles*. This bizarre situation will be explained in 14.76 and 14.77. Basically, it arises because the customary criteria for "explicit examples" are somewhat stricter than the customary criteria for existence proofs.

5.11. *Remarks.* Ultrafilters will be studied further in the last part of Chapter 6 and thereafter. For purposes of convergences, filters can be used interchangeably with nets; this concept is developed in 7.9 and 7.14, and used extensively thereafter.

Topologies

5.12. *Definition.* A **topology** on a set X is a collection $\mathcal{T}$ of subsets of X satisfying these three axioms:

(i) $\varnothing, X \in \mathcal{T}$,

(ii) $S_1, S_2 \in \mathcal{T} \Rightarrow S_1 \cap S_2 \in \mathcal{T}$, and

(iii) $\{S_\lambda : \lambda \in \Lambda\} \subseteq \mathcal{T} \Rightarrow \bigcup_{\lambda \in \Lambda} S_\lambda \in \mathcal{T}$.

That is, $\mathcal{T}$ contains $\varnothing$ and X, and $\mathcal{T}$ is closed under finite intersections and arbitrary unions. A **topological space** is a pair $(X, \mathcal{T})$ consisting of a set X and a topology $\mathcal{T}$ on X; we may refer to X itself as the topological space if $\mathcal{T}$ does not need to be mentioned explicitly. The members of $\mathcal{T}$ are called the **open** subsets of X.

A point x is called **isolated** if it is the only member of some open set; the topological space X is **disconnected** if it can be partitioned into two disjoint nonempty open sets. If no such partition exists, the space X is **connected**.

Topological spaces can be described in other ways — in terms of closed sets (5.13), convergences (15.8.b and 15.10), closure or interior operators (5.19, 5.20, and 15.7), neighborhood systems (5.22 and 15.8.a), bases (15.36.b), or distances (5.15.i).

Topological spaces will be studied briefly in the next few sections and in Chapter 9, and then in much greater detail in Chapter 15 and thereafter.

5.13. *More definitions.* Let $(X, \mathcal{T})$ be a topological space. The complements of the open sets are the **closed** subsets of X.

Closed sets are dual to open sets (in the sense of 1.7). Although it is customary to define a topological space in terms of its open sets, we could as easily define it in terms of the closed sets, as follows. Let X be a set, and let $\mathcal{K}$ be a collection of subsets of X; then $\mathcal{K}$ is the collection of closed sets for a topology on X if and only if $\mathcal{K}$ satisfies these conditions:

(i) $\varnothing, X \in \mathcal{K}$,

(ii) $S_1, S_2 \in \mathcal{K} \Rightarrow S_1 \cup S_2 \in \mathcal{K}$, and

(iii) $\{S_\lambda : \lambda \in \Lambda\} \subseteq \mathcal{K} \Rightarrow \bigcap_{\lambda \in \Lambda} S_\lambda \in \mathcal{K}$.

That is, $\mathcal{K}$ contains $\varnothing$ and X, and $\mathcal{K}$ is closed under finite unions and arbitrary intersections.

5.14. *Remarks and more definitions.* In common nonmathematical English, "open" and "closed" are opposites. This could lead beginners to expect that every subset of a topological space X must be either open or closed, but that expectation is incorrect. Some sets may be neither open nor closed. In fact, in topological spaces commonly used, *most* subsets are neither open nor closed. We shall demonstrate this in 15.37.c in the case where X is the real line, a space typical of the topological spaces used by analysts.

Also, some sets may be both open and closed. Such sets are called **clopen**. Indeed, in any topological space X, both $\varnothing$ and X are clopen. *Exercise.* The space X is *connected* (as defined in 5.12) if and only if it has no other clopen subsets besides $\varnothing$ and X.

5.15. *Elementary examples of topologies.* Let X be any set. Then:

a. The **indiscrete topology** (also called *chaotic topology*) is $\{\varnothing, X\}$; it is the smallest topology on X.

b. The **discrete topology** is $\mathcal{P}(X) = \{\text{subsets of } X\}$; it is the largest topology on X. It is the only topology that makes every subset of X clopen. It is also the only topology that makes every point of X isolated.

Finite sets are usually equipped with the discrete topology. The set $2 = \{0, 1\}$ will be used in many different contexts; we shall understand it to be equipped with the discrete topology unless some other arrangement is specified.

$\mathbb{Z} = \{\text{integers}\}$ is also usually equipped with the discrete topology. (That topology can be described another way; see 5.15.f.)

c. The **cofinite topology** is $\{S \subseteq X \ : \ \text{either } S \text{ is empty or } \complement S \text{ is finite}\}$. The cofinite topology coincides with the discrete topology when X is a finite set, but the two topologies are different when X is an infinite set.

The cofinite topology is the smallest topology on X that makes every singleton $\{x\}$ in X a closed set. That is, a topology on X makes every singleton a closed set if and only if that topology contains the cofinite topology.

d. Lower set topology (optional). The set $\mathbb{N}$ is most often equipped with its discrete topology. However, another interesting topology on $\mathbb{N}$ is given by

$$\mathcal{U} \quad = \quad \Big\{\varnothing,\ \{1\},\ \{1,2\},\ \{1,2,3\},\ \{1,2,3,4\},\ \ldots,\ \mathbb{N}\Big\}.$$

More generally, let $(X, \preccurlyeq)$ be a preordered set, and let $\mathcal{T} = \{\text{lower sets of } X\}$. Show that

(i) $\mathcal{T}$ is a topology on X. We shall call it the **lower set topology** on X.

(ii) The preorder $\preccurlyeq$ can be recovered from the topology $\mathcal{T}$. Thus, the mapping $\preccurlyeq\ \mapsto\ \mathcal{T}$ is injective (i.e., different preorders determine different lower set topologies), so we may view $\{\text{preordered spaces}\}$ as a subclass of $\{\text{topological spaces}\}$.

(iii) Analogous properties are easily verified for the **upper set topology**, which is defined to be $\{\text{upper sets of } X\}$. (Indeed, if $\preccurlyeq$ is a preorder on X, then $\succcurlyeq$ is another preorder.)

Example. The upper set topology on $\mathbb{N}$ is

$$\mathcal{V} \quad = \quad \Big\{\{1,2,3,\ldots\},\ \{2,3,4,\ldots\},\ \{3,4,5,\ldots\},\ \ldots,\ \varnothing\Big\}.$$

e. Let X be a subset of a topological space $(Y, \mathcal{T})$. Verify that $\{X \cap T : T \in \mathcal{T}\}$ is a topology on X. It is called the **relative topology**, or **subspace topology**, induced on X by Y. Any subset of a topological space will be understood to be equipped with its relative topology, unless some other arrangement is specified. Show that

(i) A subset of X is open in the relative topology if and only if it is of the form $X \cap G$ for some set $G \subseteq Y$ that is $\mathcal{T}$-open.

(ii) A subset of X is closed in the relative topology if and only if it is of the form $X \cap F$ for some set $F \subseteq Y$ that is $\mathcal{T}$-closed.

(iii) Suppose X itself is $\mathcal{T}$-open. Then a subset of X is open in the relative topology if and only if it is $\mathcal{T}$-open.

(iv) Suppose X itself is $\mathcal{T}$-closed. Then a subset of X is closed in the relative topology if and only if it is $\mathcal{T}$-closed.

(v) Suppose $W \subseteq X \subseteq Y$. Then the relative topology induced on W by Y is the same as the relative topology induced on W by the relative topology induced on X by Y.

f. Let $(X, \leq)$ be a chain. Let $\mathcal{T}$ be the collection of all sets $T \subseteq X$ that satisfy the following condition:

For each $p \in T$, there exists some set J of the form $\{x \in X : a < x\}$ or $\{x \in X : a < x < b\}$ or $\{x \in X : x < b\}$ such that $p \in J \subseteq T$.

Then $\mathcal{T}$ is a topology on X, called the **order interval topology**.

The usual topologies on $\mathbb{R}$ and $[-\infty, +\infty]$ are their order interval topologies. These sets will always be understood to be equipped with these topologies, unless some other arrangement is specified. The topology of $\mathbb{R}$ is in many ways typical of topologies used in analysis. In fact, most topological spaces used in analysis are built from copies of $\mathbb{R}$, in one way or another.

Any subset of $\mathbb{R}$ is a chain, but such sets are *not* always equipped with their order interval topologies. Rather, they are usually equipped with the relative topology induced by $\mathbb{R}$ (as defined in 5.15.e). That topology does not always agree with the order interval topology; we shall compare the two topologies in 15.46.

g. *Definitions.* Let (X, d) be a pseudometric space (defined as in 2.11). For any $z \in X$ and $r > 0$, we define the **open ball** centered at z with radius r to be the set

$$B_d(z, r) \quad = \quad \{x \in X \ : \ d(x, z) < r\}.$$

(We may omit the subscript d when no confusion will result.) A set $T \subseteq X$ is said to be **open** if

for each $z \in T$, there exists some $r > 0$ such that $B_d(z, r) \subseteq T$.

The reader should verify that the collection of all such sets T is a topology $\mathcal{T}_d$ on X; we call it the **pseudometric topology** (or the **metric topology**, if d is known to be a metric). Any pseudometric space will be understood to be equipped with this topology, unless some other is specified.

The reader should verify that $B_d(z, r)$ is an open set in the topological space $(X, \mathcal{T}_d)$, thus justifying our calling it the open ball. We also define the **closed ball** with center z and radius r to be the set

$$K_d(z, r) \quad = \quad \{x \in X : d(z, x) \leq r\}.$$

The reader should verify that this is a $\mathcal{T}_d$-closed set.

The **usual metric on** $\mathbb{R}$ is that given by the absolute value function — that is, $d(x,y) = |x-y|$. The set $\mathbb{R}$ is always understood to be equipped with this metric, unless some other arrangement is specified.

Exercise. Show that the resulting metric topology on $\mathbb{R}$ is the same as the order interval topology on $\mathbb{R}$. (This result will be easier to prove later; see 15.43.)

Two of the usual metrics on the extended real line $[-\infty, +\infty]$ are

$$d(x,y) \;=\; |\arctan(x) - \arctan(y)| \qquad \text{and} \qquad d(x,y) = \left| \frac{x}{1+|x|} - \frac{y}{1+|y|} \right|.$$

(It follows easily from 2.15.a that these are both metrics.) In fact, there are many usual metrics on $[-\infty, +\infty]$, all of them slightly more complicated than one might wish. Fortunately, they are interchangeable for most purposes: They all yield the same topology, and in fact we shall see in 18.24 that they all yield the same uniformity.

Exercise. Show that the two metrics given above both yield the order interval topology on $[-\infty, +\infty]$. (This exercise may be postponed; it will be easier after 15.43.)

Further examples of pseudometric topologies are given in 5.34 and elsewhere.

h. For many applications we shall need a generalization of pseudometric topologies:

Let D be a gauge (i.e., a collection of pseudometrics) on a set X. For each $d \in D$ let B_d be the corresponding open ball, as in 5.15.g. Let $\mathcal{T}_D$ be the collection of all sets $T \subseteq X$ having the property that

> for each $x \in T$, there is some finite set $D_0 \subseteq D$ and some number $r > 0$ such that $\bigcap_{d \in D_0} B_d(x,r) \subseteq T$.

Then (*exercise*) $\mathcal{T}_D$ is a topology on X. We may call it the **gauge topology** determined by D. Any gauge space (X, D) will be understood to be equipped with this topology unless some other arrangement is specified. We may write $\mathcal{T}$, omitting the subscript D, if the choice of D is clear or does not need to be mentioned.

Exercises. If D is a gauge and E is its max closure or its sum closure (as defined in 4.4.c), then D and E determine the same topology. If D is a directed gauge (as defined in 4.4.c), then we can always choose D_0 to be a singleton in the definition of $\mathcal{T}_D$ given above.

Remarks continued. Whenever convenient, we shall treat pseudometric spaces as a special case of gauge spaces, with gauge D consisting of a singleton $\{d\}$. When no confusion will result, we may write d and $\{d\}$ interchangeably and consider d itself as a gauge. Conversely, in 5.23.c we shall see that any gauge topology $\mathcal{T}_D$ can be analyzed in terms of the simpler pseudometric topologies $\{\mathcal{T}_d : d \in D\}$.

Different gauges on a set X may determine the same topology or different topologies. Two gauges D and E are called **equivalent** (or **topologically equivalent**) if they determine the same topology. This terminology is discussed further in 9.4.

A topological space $(X, \mathcal{T})$ is *metrizable* (or *pseudometrizable*) if there exists at least one metric d on X (respectively, at least one pseudometric d on X) for which $\mathcal{T} = \mathcal{T}_d$; the topological space is *gaugeable* if there exists at least one gauge D on X for which $\mathcal{T} = \mathcal{T}_D$. This (pseudo)metric or gauge is not necessarily unique. When we state that a topological space is (pseudo)metrizable or gaugeable, we do not necessarily have some particular (pseudo)metric or gauge in mind. We say that a topology $\mathcal{T}$ and a gauge D are **compatible** if $\mathcal{T} = \mathcal{T}_D$; this term also applies to topologies and pseudometrics.

Most topologies used in analysis are gaugeable. In 16.18 we present some examples of topologies that are not gaugeable, but these examples are admittedly somewhat contrived.

Actually, the term "gaugeable" is seldom used in practice. We shall see in 16.16 that a topology is gaugeable if and only if it is uniformizable and if and only if it is completely regular; the terms "uniformizable" and "completely regular" are commonly used in the literature.

Exercise. If (X, D) is a gauge space and $S \subseteq X$, then the relative topology on S is also gaugeable; it can be given by the restriction of D to S.

i. (*Optional.*) We can generalize the notion of pseudometric topologies still further. Let D be a **quasigauge** on X — i.e., a collection of quasipseudometrics on X (which are not necessarily symmetric; see 2.11). We can use D to define a **quasigauge topology** $\mathcal{T}_D$ in a fashion analogous to that in 5.15.h. That is, $\mathcal{T}_D$ is the collection of all sets $T \subseteq X$ that have the property that

> for each $x \in T$, there is some finite set $D_0 \subseteq D$ and some number $r > 0$ such that $\{u \in X : \max_{d \in D_0} d(x, u) < r\} \subseteq T$.

(This is the supremum of the topologies $\mathcal{T}_d$ determined by the individual quasipseudometrics $d \in D$; see 5.23.c.)

Reilly's Representation. Actually, *every* topology $\mathcal{T}$ on a set X can be determined by a quasigauge D. Show this with $D = \{d_G : G \in \mathcal{T}\}$, where

$$d_G(x, x') \quad = \quad \begin{cases} 1 & \text{if } x \in G \text{ and } x' \notin G \\ 0 & \text{otherwise.} \end{cases}$$

Consequently, many of the ideas that we commonly associate with gauge spaces — uniform continuity, equicontinuity, completeness, etc. — can be extended (in a weaker and more complicated form) to arbitrary topological spaces.

This presentation follows Reilly [1973]. Similar ideas have been discovered independently in other forms; for instance, see Kopperman [1988] and Pervin [1962].

5.16. *Definitions.* Let $(X, \mathcal{T})$ be a topological space, and let $S \subseteq X$.

a. We shall say that S is a **neighborhood** of a point z if $z \in G \subseteq S$ for some open set G. Then $\mathcal{N}(z) = \{\text{neighborhoods of } z\}$ is a proper filter on X, which we shall call the **neighborhood filter at z** or the **filter of neighborhoods of z**.

Caution: Some mathematicians define neighborhood as we have done, but other mathematicians also require the set S to be open, as part of their definition of a neighborhood of a point. With the latter approach, the neighborhoods of a point generally do not form a filter. The two definitions yield similar results for the main theorems of general topology, but the open-neighborhoods-only approach is not compatible with the pedagogical style with which general topology is developed in this book: We shall use filters frequently.

b. There exist some closed sets that contain S (for instance, X itself), and among all such sets there is a smallest (namely, the intersection of all the closed supersets of S). The smallest closed set containing S is called the **topological closure** of S; we shall denote it by $\mathrm{cl}(S)$. It is a special case of the Moore closure. It is probably the type of closure that is most often used by analysts. It may be called simply the **closure** of S, if the context is clear.

c. There exist some open sets that are contained in S (for instance, $\varnothing$), and among all such sets there is a largest (namely, the union of all the open subsets of S). The largest open set contained in S is called the **interior** of S; we shall denote it by $\mathrm{int}(S)$.

d. Let X and Y be sets, and suppose some element of Y is designated "0" — e.g., if Y is a vector space, or if $Y \subseteq [-\infty, +\infty]$. Let $f : X \to Y$ be some function. If X is a topological space, then the **support** of f means the set

$$\mathrm{supp}(f) \quad = \quad \mathrm{cl}\big(\,\{x \in X \;:\; f(x) \neq 0\}\,\big).$$

If X is not equipped with a topology, then the support of f usually means the set $\{x \in X : f(x) \neq 0\}$. Note that these two definitions agree if X has the discrete topology.

5.17. *Elementary properties.*

a. $\mathrm{int}(S) \subseteq S \subseteq \mathrm{cl}(S)$. A set S is open if and only if $S = \mathrm{int}(S)$, and a set S is closed if and only if $S = \mathrm{cl}(S)$.

b. The notions of closure and interior are dual to each other, in the sense of 1.7. Show that

$$\complement\,\mathrm{cl}(S) \;=\; \mathrm{int}(\complement\, S), \qquad\qquad \complement\,\mathrm{int}(S) \;=\; \mathrm{cl}(\complement\, S),$$

where $\complement\, A = X \setminus A$.

c. A set $S \subseteq X$ is open if and only if S is a neighborhood of each of its points.

d. $z \in \mathrm{cl}(S)$ if and only if S meets every neighborhood of z.

e. If G is open and $\mathrm{cl}(S) \cap G$ is nonempty, then $S \cap G$ is nonempty.

5.18. *Closures and distances.* Let (X, d) be a pseudometric space. The diameter of a set and the distance from a point to a set were defined in 4.40. Let S be a nonempty subset of X, and let $z \in X$. Then:

a. $\mathrm{dist}(z, S) = \mathrm{dist}(z, \mathrm{cl}(S))$, and $\mathrm{dist}(z, S) = 0 \iff z \in \mathrm{cl}(S)$.

b. $\mathrm{diam}(\mathrm{cl}(S)) = \mathrm{diam}(S)$.

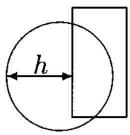
c. $\mathrm{cl}(B(z,r)) \subseteq K(z,r)$, where B and K are the open and closed balls, defined as in 5.15.g.

Show that $\mathrm{cl}(B(z,r)) \subsetneqq K(z,r)$ may sometimes occur, by taking $X = \mathbb{R}$ and $d(x,y) = \min\{1, |x-y|\}$. (See 26.4.a for a further related result.)

d. (*Optional.*) Assume (X, d) is a metric space. Let $\mathcal{K} = \{$nonempty, closed, metrically bounded subsets of $X\}$. For $S, T \in \mathcal{K}$, let

$$h(S,T) \quad = \quad \max\left\{\sup_{s\in S} \operatorname{dist}(s,T),\ \sup_{t\in T} \operatorname{dist}(t,S)\right\}.$$

(See example in the figure below.) Show that h is a metric on $\mathcal{K}$; it is called the **Hausdorff metric**.



Example. Hausdorff distance h between a circle and a rectangle

5.19. Kuratowski's Closure Axioms. Let X be a set, and let $\operatorname{cl} : \mathcal{P}(X) \to \mathcal{P}(X)$ be some mapping. Show that cl is the closure operator for a topology on X if and only if cl satisfies these four conditions:

$$\operatorname{cl}(\varnothing) = \varnothing, \qquad S \subseteq \operatorname{cl}(S), \qquad \operatorname{cl}(\operatorname{cl}(S)) = \operatorname{cl}(S), \qquad \operatorname{cl}(S \cup T) = \operatorname{cl}(S) \cup \operatorname{cl}(T)$$

for all sets $S, T \subseteq X$. Of course, if these conditions are satisfied, then the corresponding topology is uniquely determined by cl; its closed sets are those sets $S \subseteq X$ that satisfy $S = \operatorname{cl}(S)$.

5.20. The dual of Kuratowski's axioms follows easily; we include it here for convenient later reference.

Let X be a set, and let $\operatorname{int} : \mathcal{P}(X) \to \mathcal{P}(X)$ be some mapping. Then int is the interior operator for a topology on X if and only if int satisfies these four conditions:

$$\operatorname{int}(X) = X, \qquad S \supseteq \operatorname{int}(S), \qquad \operatorname{int}(\operatorname{int}(S)) = \operatorname{int}(S), \qquad \operatorname{int}(S \cap T) = \operatorname{int}(S) \cap \operatorname{int}(T)$$

for all sets $S, T \subseteq X$. Of course, if these conditions are satisfied, then the corresponding topology is uniquely determined by int; its open sets are those sets $S \subseteq X$ that satisfy $S = \operatorname{int}(S)$.

5.21. *The lattice of open sets (optional).* Let $(X, \mathcal{T})$ be a topological space. Then $(\mathcal{T}, \subseteq)$ is a complete lattice. Indeed, for any open sets $G_\lambda \in \mathcal{T}$ $(\lambda \in \Lambda)$, the smallest open set containing all the G_λ's is

$$\bigvee_{\lambda\in\Lambda} G_\lambda \quad = \quad \bigcup_{\lambda\in\Lambda} G_\lambda,$$

while the largest open set contained in all the G_λ's is

$$\bigwedge_{\lambda\in\Lambda} G_\lambda \quad = \quad \operatorname{int}\left(\bigcap_{\lambda\in\Lambda} G_\lambda\right).$$

Note that when the index set Λ is finite, then $\cap_{\lambda\in\Lambda}G_\lambda$ is an open set, and so $\bigwedge_{\lambda\in\Lambda} G_\lambda$ generally is equal to $\bigcap_{\lambda\in\Lambda} G_\lambda$. Hence the inclusion $\mathcal{T} \xrightarrow{\subseteq} \mathcal{P}(X)$ preserves finite sups and infs; thus it is a lattice homomorphism. However, when the index set Λ is infinite, $\bigwedge_{\lambda\in\Lambda} G_\lambda$ generally is *not* equal to $\bigcap_{\lambda\in\Lambda} G_\lambda$. Thus the inclusion $\mathcal{T} \xrightarrow{\subseteq} \mathcal{P}(X)$ is sup-preserving, but it generally is not inf-preserving.

It is easy to verify that $(\mathcal{T}, \subseteq)$ satisfies one of the infinite distributive laws:

$$H \wedge \bigvee_{\lambda\in\Lambda} G_\lambda \quad = \quad \bigvee_{\lambda\in\Lambda} (H \wedge G_\lambda). \tag{1}$$

(See also the related results in 13.28.a.) However, $(\mathcal{T}, \subseteq)$ does not necessarily satisfy the other infinite distributive law,

$$H \vee \bigwedge_{\lambda\in\Lambda} G_\lambda \quad = \quad \bigwedge_{\lambda\in\Lambda} (H \vee G_\lambda). \tag{2}$$

For instance, that law is not satisfied in the following example, taken from Vulikh [1967]:

Let $\mathcal{T}$ be the usual topology on the real line. Let $H = (0,1)$, $\Lambda = \mathbb{N}$, and $G_n = (1 - \frac{1}{n}, 2)$ for $n = 1, 2, 3, \ldots$. Verify that $\bigwedge_{n\in\mathbb{N}} G_n = (1,2)$, hence the left side of equation (2) is $(0,1) \cup (1,2)$. On the other hand, $H \vee G_n = (0,2)$, hence the right side of equation (2) is $(0,2)$.

It is possible to study at least some properties of a topological space purely in terms of its lattice of open sets; one can disregard the individual points that make up those sets. (See the related result in 16.5.d and the related comments in 13.3.) An introduction to this "pointless topology" was given by Johnstone [1983]. However, this pointless topology is seldom useful in applied analysis, which is greatly concerned with points.

5.22. Neighborhood Axioms. Let X be a set. For each $x \in X$, suppose $\mathcal{N}(x)$ is some filter on X, such that x is a member of every member of $\mathcal{N}(x)$. Let

$$\mathcal{T} \quad = \quad \{G \subseteq X \ : \ G \in \mathcal{N}(z) \text{ for every } z \in G\}.$$

(In particular, $\varnothing \in \mathcal{T}$, since there is no $z \in G$ in that case.) Then the following three conditions are equivalent:

(A) There exists a topology on X for which $\{\mathcal{N}(z) : z \in X\}$ is the system of neighborhood filters.

(B) For each $z \in X$, the collection of sets $\mathcal{T} \cap \mathcal{N}(z)$ is a base for the filter $\mathcal{N}(z)$. That is, every member of $\mathcal{N}(z)$ contains some member of $\mathcal{N}(z) \cap \mathcal{T}$.

(C) For each $z \in X$ and each $S \in \mathcal{N}(z)$, there is some $G \in \mathcal{N}(z)$ with the property that $u \in G \Rightarrow S \in \mathcal{N}(u)$.

Moreover, if (A), (B), (C) are satisfied, then the topology in (A) must be $\mathcal{T}$.

Hints: Let us first restate (B) as follows:

(B′) For every $S \in \mathcal{N}(z)$, there is some $G \in \mathcal{N}(z) \cap \mathcal{T}$ such that $G \subseteq S$.

For (A) $\Rightarrow$ (B′), let int be the interior operator of the given topology; show that $G = \text{int}(S)$ is a member of the collection of sets $\mathcal{T}$ described above. For (B′) $\Rightarrow$ (C), note that $u \in G \Rightarrow G \in \mathcal{N}(u) \Rightarrow S \in \mathcal{N}(u)$ since $S \supseteq G$. For (C) $\Rightarrow$ (A), define an operator $\text{int} : \mathcal{P}(X) \to \mathcal{P}(X)$ by $\text{int}(S) = \{z \in X : S \in \mathcal{N}(z)\}$; then verify that this operator int satisfies the conditions of 5.20.

5.23. Here are a few more ways to make topologies:

a. If $\{\mathcal{T}_\lambda : \lambda \in \Lambda\}$ is a collection of topologies on a set X, then

$$\bigcap_{\lambda \in \Lambda} \mathcal{T}_\lambda \quad = \quad \{S \subseteq X \,:\, S \in \mathcal{T}_\lambda \text{ for every } \lambda\}$$

is also a topology on X. It is sometimes called the **infimum** of the $\mathcal{T}_\lambda$'s, since it is their greatest lower bound — i.e., it is the largest topology that is contained in all the $\mathcal{T}_\lambda$'s.

b. From the preceding result we see that the collection of all topologies on X is a Moore collection of subsets of $\mathcal{P}(X)$. Thus, if $\mathcal{G}$ is any collection of subsets of X, then there exists a smallest topology $\mathcal{T}$ containing $\mathcal{G}$ — namely, the intersection of all the topologies that contain $\mathcal{G}$. The topology $\mathcal{T}$ obtained in this fashion is the **topology generated by** $\mathcal{G}$; the generating set $\mathcal{G}$ is also called a **subbase** for the topology $\mathcal{T}$. (The topology generated is a special case of the Moore closure, but the terms "closed" and "closure" generally are not used in this context.)

Example. The order interval topology on a chain X (defined in 5.15.f) is the topology generated by the sets that can be expressed in either of the forms

$$S_a = \{x \in X \,:\, a < x\} \qquad \text{or} \qquad S^b = \{x \in X \,:\, x < b\}$$

for points $a, b \in X$.

Exercise. Let $\mathcal{G}$ be a collection of subsets of a set X. A set $T \subseteq X$ is a neighborhood of a point $x \in X$ with respect to the topology generated by $\mathcal{G}$ if and only if T has the following property:

There is some finite set $\{G_1, G_2, \ldots, G_n\} \subseteq \mathcal{G}$ such that $x \in \bigcap_{j=1}^n G_j \subseteq T$.

(We permit $n = 0$, with the convention that the intersection of no subsets of X is all of X.)

c. If $\{\mathcal{T}_\lambda : \lambda \in \Lambda\}$ is a collection of topologies on a set X, then the topology generated by

$$\mathcal{G} \quad = \quad \bigcup_{\lambda \in \Lambda} \mathcal{T}_\lambda \quad = \quad \{S \subseteq X \,:\, S \in \mathcal{T}_\lambda \text{ for some } \lambda\}$$

is called the **supremum** of the $\mathcal{T}_\lambda$'s, since it is their least upper bound — i.e., it is the smallest topology that contains all the $\mathcal{T}_\lambda$'s.

The collection of all topologies on X is a complete lattice when ordered by $\subseteq$ since each subcollection has an inf (see 5.23.a) and a sup.

Important example. On any gauge space (X, D), the gauge topology $\mathcal{T}_D$ is the supremum of the pseudometric topologies $\{\mathcal{T}_d : d \in D\}$ (defined as in 5.15.g).

5.24. *Remarks.* The theory of topological spaces will be developed a little further in Chapter 9. It will be continued in much greater detail in Chapter 15 and thereafter.

ALGEBRAS AND SIGMA-ALGEBRAS

5.25. Let X be a set, and let $\complement$ denote complementation in X. An **algebra** (or *field*) of subsets of X is a collection $\mathcal{S} \subseteq \mathcal{P}(X)$ with the following properties:

(i) $X \in \mathcal{S}$,

(ii) $S \in \mathcal{S} \Rightarrow \complement S \in \mathcal{S}$, and

(iii) $S, T \in \mathcal{S} \Rightarrow S \cup T \in \mathcal{S}$.

In the terminology of 1.30, that says: $\mathcal{S}$ contains X itself and $\mathcal{S}$ is closed under complementation and finite union. It follows that $\varnothing \in \mathcal{S}$ and that $\mathcal{S}$ is closed under finite intersection and relative complementation: $S, T \in \mathcal{S}$ implies $S \cap T$, $S \backslash T \in \mathcal{S}$.

Caution: The term "algebra" has many different meanings in mathematics; several meanings will be given in 8.47 and one more in 11.3. When we need to distinguish the algebra defined above from other kinds of algebras, the algebra defined in the preceding paragraph will be called an **algebra of sets**.

A **σ-algebra** (or *σ-field*) of subsets of X is an algebra that is closed under countable union:

(iii′) $S_1, S_2, S_3, \ldots \in \mathcal{S} \Rightarrow \bigcup_{j=1}^{\infty} S_j \in \mathcal{S}$.

It follows immediately that any σ-algebra $\mathcal{S}$ is also closed under countable intersection: $S_1, S_2, S_3, \ldots \in \mathcal{S} \Rightarrow \bigcap_{j=1}^{\infty} S_j \in \mathcal{S}$.

A **measurable space** is a pair $(X, \mathcal{S})$, where X is a set and $\mathcal{S}$ is a σ-algebra of subsets of X. The elements of $\mathcal{S}$ are referred to as the **measurable sets** in X. We may refer to X itself as a measurable space if $\mathcal{S}$ does not need to be mentioned explicitly. Measurable spaces $(X, \mathcal{S})$ should not be confused with *measure spaces* $(X, \mathcal{S}, \mu)$, introduced in 21.9, or with *spaces of measures* $\{\mu_\alpha\}$, introduced in 11.47, 11.48, and 29.29.f. Somewhat imprecisely, we may say that a measure is device for measuring how big sets are; a measurable space is a space that is capable of being equipped with any of several different measures; a measure space is a space that has been equipped with a particular measure; and a space of measures is a collection of measures that is equipped with some additional structure (linear, topological, etc.) that leads us to call it a "space."

5.26. *Examples of* (σ-) *algebras.* In the following examples, any statement involving σ- in parentheses should be read once with the σ- omitted and once with it included. Let X be any set; then:

- **a.** $\{\varnothing, X\}$ is the smallest (σ-)algebra on X; we shall call this the **indiscrete (σ-)algebra**.
- **b.** $\mathcal{P}(X) = \{\text{subsets of } X\}$ is the largest (σ-)algebra on X. We shall call it the **discrete (σ-)algebra**.

c. Let $J \subseteq \mathbb{R}$ be an interval (possibly all of $\mathbb{R}$). Let $\mathcal{A}$ be the collection of all unions of finitely many subintervals of J (where a singleton is considered to be an interval and, by convention, $\varnothing \in \mathcal{A}$ also). Show that $\mathcal{A}$ is an algebra of subsets of J.

d. Let $\{\mathcal{S}_\alpha : \alpha \in A\}$ be a collection of (σ-)algebras on X. Then

$$\bigcap_{\alpha \in A} \mathcal{S}_\alpha \quad = \quad \{T \subseteq X \quad : \quad T \in \mathcal{S}_\alpha \text{ for every } \alpha \in A\}$$

is also a (σ-)algebra on X.

e. In view of the preceding exercise, the collection of all (σ-)algebras on X is a Moore collection of subsets of $\mathcal{P}(X)$. Hence, given any collection $\mathcal{G}$ of subsets of X, there exists a smallest (σ-)algebra that contains $\mathcal{G}$ — namely, the intersection of all the (σ-)algebras that contain $\mathcal{G}$. We call it the **(σ-)algebra generated** by $\mathcal{G}$; we say that $\mathcal{G}$ is a **generating set** for that (σ-)algebra. (The (σ-)algebra thus generated is a special case of the Moore closure, introduced in 4.3. However, the terms "closed" and "closure" generally are not used in this context.) The σ-algebra generated by $\mathcal{G}$ is sometimes denoted by $\sigma(\mathcal{G})$.

f. $\{S \subseteq X : S \text{ or } \complement S \text{ is finite}\}$ is the algebra generated by the singletons of X.

g. $\{S \subseteq X : S \text{ or } \complement S \text{ is countable}\}$ is the σ-algebra generated by the singletons of X. (The proof of this result assumes some familiarity with the most basic properties of countable sets; see particularly 6.26.)

h. Let $\mathcal{G}$ be a collection of subsets of X. Then:

(i) The algebra generated by $\mathcal{G}$ is equal to the union of the algebras generated by finite subsets of $\mathcal{G}$.

(ii) The σ-algebra generated by $\mathcal{G}$ is equal to the union of the σ-algebras generated by countable subcollections of $\mathcal{G}$.

i. Some of the most important σ-algebras are determined in one way or another by topologies.

Let $(X, \mathcal{T})$ be a topological space. The **Borel σ-algebra** is the σ-algebra generated by $\mathcal{T}$ — that is, the smallest σ-algebra containing all the open sets. Its members are called the **Borel sets**. (*Remark.* In 15.37.e we shall see that when X is any subinterval of the real line, equipped with its usual topology, then the Borel σ-algebra is generated by the algebra in 5.26.c.)

Some other σ-algebras based on topologies are

- the *almost open sets*, also known as the *sets with the Baire property*, studied in 20.20 and thereafter;
- the *Baire sets*, mentioned in 20.34; and
- in $\mathbb{R}^n$, the *Lebesgue measurable sets*, studied in Chapters 21 and 24.

Caution: The "Baire sets" are not the same as the "sets with the Baire property," and the "Lebesgue measurable sets" are not the same as the "Lebesgue sets" (introduced in 25.16).

j. The clopen subsets of a topological space X form an algebra of subsets of X.

5.27. *More definitions (optional).* Let Ω be a set. A **ring** of subsets of Ω (also known as a **clan**) is a collection $\mathcal{R}$ of subsets of Ω that satisfies $\varnothing \in \mathcal{R}$ and also

$$A, B \in \mathcal{R} \qquad \Rightarrow \qquad A \cup B,\ A \setminus B \in \mathcal{R}.$$

A ***σ*-ring** (also known as a **tribe**) is a ring $\mathcal{R}$ that also satisfies

$$A_1, A_2, A_3, \ldots \in \mathcal{R} \qquad \Rightarrow \qquad A_1 \cup A_2 \cup A_3 \cup \cdots \in \mathcal{R}.$$

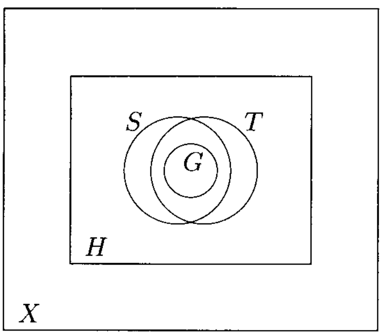
Clearly, a collection $\mathcal{R} \subseteq \mathcal{P}(\Omega)$ is an algebra (or σ-algebra) if and only if it is a ring (or σ-ring) in which Ω is a member.

Most treatments of measure theory use either σ-algebras or σ-rings. The σ-algebra approach has the advantage that it is more algebraic — i.e., the algebraic structure of a σ-algebra is simpler than that of a σ-ring. On the other hand, σ-rings are more general, and more useful in the study of regular measures on locally compact Hausdorff spaces — see the remarks in 20.35 — but we shall not study such measures in this book.

5.28. *Exercise.* If $\mathcal{A}$ is a (σ-)algebra on X and $\mathcal{I}$ is a (σ-)ideal, then

$$\mathcal{A} \triangle \mathcal{I} \quad = \quad \{A \triangle I \ : \ A \in \mathcal{A} \text{ and } I \in \mathcal{I}\}$$

is a (σ-)algebra on X. It is the smallest (σ-)algebra of sets that contains both $\mathcal{A}$ and $\mathcal{I}$. (See hint in diagram below. This result will be used in 20.21 and 21.16.)



Hint for exercise on $\mathcal{A} \triangle \mathcal{I}$: Show that a set $S \subseteq X$ is an element of $\mathcal{A} \triangle \mathcal{I}$ if and only if there exist sets G, H, and T such that $G \subseteq S \subseteq H$, $G \subseteq T \subseteq H$, $T \in \mathcal{A}$, and $H \setminus G \in \mathcal{I}$.

5.29. Let Ω be a set. A **monotone class** of subsets of Ω is a collection $\mathcal{M}$ of subsets of Ω with both of these properties:

(i) If (A_j) is a sequence in $\mathcal{M}$ and $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$.

(ii) If (A_j) is a sequence in $\mathcal{M}$ and $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$ then $\bigcap_{n=1}^{\infty} A_n \in \mathcal{M}$.

It is easy to see that the monotone classes form a Moore collection — i.e., any intersection of monotone classes is a monotone class. Hence, given any collection $\mathcal{A}$ of subsets of Ω, there exists a smallest monotone class containing $\mathcal{A}$; we shall call it the **monotone class generated by** $\mathcal{A}$.

Monotone Class Theorem. Let $\mathcal{A}$ be an algebra of subsets of Ω. Then the monotone class generated by $\mathcal{A}$ is equal to the σ-algebra generated by $\mathcal{A}$.

Proof. Let $\mathcal{M}$ and $\mathcal{S}$ be, respectively, the monotone class and the σ-algebra generated by $\mathcal{A}$. Thus $\mathcal{A} \subseteq \mathcal{M} \cap \mathcal{S}$. Temporarily fix any $M \in \mathcal{M}$, and let

$$\mathcal{N}_M \quad = \quad \{N \in \mathcal{M} \,:\, M \cap N,\ \complement M \cap N,\ M \cap \complement N \text{ all belong to } \mathcal{M}\}.$$

Verify that

a. $\mathcal{N}_M$ is a monotone class.

b. If $A \in \mathcal{A}$, then $\mathcal{A} \subseteq \mathcal{N}_A$; hence $\mathcal{N}_A = \mathcal{M}$ (by the minimality of $\mathcal{M}$ among monotone classes that contain $\mathcal{A}$).

c. If $A \in \mathcal{A}$ and $M \in \mathcal{M}$, then $M \in \mathcal{N}_A$; hence $A \in \mathcal{N}_M$. Thus $\mathcal{A} \subseteq \mathcal{N}_M$. Therefore $\mathcal{N}_M = \mathcal{M}$ (by the minimality of $\mathcal{M}$ again).

d. $\mathcal{M}$ is an algebra of sets. (Indeed, if $L, M \in \mathcal{M}$, use the fact that $L \in \mathcal{M} = \mathcal{N}_M$.)

e. $\mathcal{M}$ is a σ-algebra.

f. $\mathcal{M} \supseteq \mathcal{S}$, by minimality of $\mathcal{S}$ among σ-algebras containing $\mathcal{A}$.

g. $\mathcal{S}$ is a monotone class containing $\mathcal{A}$; hence $\mathcal{S} \supseteq \mathcal{M}$ by minimality of $\mathcal{M}$.

5.30. *Remarks.* Measurable spaces will be studied a little more in Chapter 9. Algebras of sets will be related to Boolean algebras in Chapter 13. Algebras of sets and σ-algebras will be used in measure theory in several later chapters.

UNIFORMITIES

5.31. *Definitions.* Let X be a set; its diagonal is the set $I = \{(x,x) : x \in X\}$. Define inverses and compositions as in 3.2. Then a **preuniformity** on X is a collection $\mathcal{U}$ of subsets of $X \times X$ that satisfies:

(i) $U \in \mathcal{U} \Rightarrow U \supseteq I$ (i.e., each $U \in \mathcal{U}$ is reflexive);

(ii) for each $U \in \mathcal{U}$, there is some $V \in \mathcal{U}$ with $V \subseteq U^{-1}$; and

(iii) for each $U \in \mathcal{U}$, there is some $V \in \mathcal{U}$ with $V \circ V \subseteq U$.

We say $\mathcal{U}$ is a **uniformity** on X if, in addition,

(iv) $\mathcal{U}$ is a filter on $X \times X$.

(It is necessarily a *proper* filter since each $U \in \mathcal{U}$ contains I and is therefore nonempty.)

An element of $\mathcal{U}$ is called an **entourage** (or a *vicinity*). A **uniform space** is a pair $(X, \mathcal{U})$ consisting of a set X and a uniformity $\mathcal{U} \subseteq \mathcal{P}(X \times X)$. We may refer to X itself as a uniform space if $\mathcal{U}$ does not need to be mentioned explicitly.

Caution: The definitions of "uniformity" and "uniform space" vary slightly in the literature. (See 16.17.) Also, the "preuniformity" defined above is related to, but *not* the same as, the "subbase for a uniformity" defined in some books.

5.32. *Uniformities constructed from distances.* Let d be a pseudometric on a set X (defined as in 2.11). For each number $r > 0$, let

$$U_r \quad = \quad \{(x, x') \in X \times X \ : \ d(x, x') < r\}.$$

Then let $\mathcal{U}_d = \{V \subseteq X \times X \ : \ V \supseteq U_r \text{ for some } r > 0\}$. Then (*exercise*) $\mathcal{U}_d$ is a uniformity on X. We shall call it the **uniformity on X determined by d**. A uniformity that can be determined by a pseudometric (or a metric) in this fashion is called a **pseudometrizable uniformity** (or a **metrizable uniformity**). Some uniformities are not pseudometrizable; an example of this is given in 18.20.

More generally, let D be a gauge on a set X. Let $\mathcal{U}_D$ be the collection of all sets $U \subseteq X \times X$ that have this property:

> There is some number $r > 0$ and some finite set $F \subseteq D$ such that $\{(x, x') \in X \times X : \max_{d \in F} d(x, x') < r\} \subseteq U$.

(The set F can be taken to be a singleton, if D is directed, as in 4.4.c.) Then (*exercise*) $\mathcal{U}_D$ is a uniformity on X. We shall call it the **uniformity for X determined by D**. Any gauge space (X, D) will be understood to be equipped with this uniformity unless some other arrangement is specified.

Two different gauges D and E on a set X may determine different uniformities $\mathcal{U}_D, \mathcal{U}_E$ or the same uniformity. We say D and E are **uniformly equivalent** if they determine the same uniformity. This is an equivalence relation on the set of all gauges on X; the terminology is discussed further in 9.4. We say that a uniformity $\mathcal{U}$ and a gauge D are **compatible** if $\mathcal{U} = \mathcal{U}_D$; this term also applies to uniformities and pseudometrics.

Exercise. If D is a gauge and E is its max closure or its sum closure (as defined in 4.4.c), then D and E determine the same uniformity.

It would be natural to say that a uniformity $\mathcal{U}$ is "*gaugeable*" if $\mathcal{U} = \mathcal{U}_D$ for some gauge D. However, it turns out that *every* uniformity is gaugeable; see 16.16.

5.33. *Topologies constructed from uniformities.* Let $(X, \mathcal{U})$ be a uniform space, and let $x \in X$. Then

$$\begin{array}{lcll} U[x] & = & \{y \ : \ (x, y) \in U\} & \text{is a subset of } X, \text{ for each } U \in \mathcal{U}, \quad \text{and} \\ \mathcal{U}[x] & = & \{U[x] \ : \ U \in \mathcal{U}\} & \text{is a filter on } X. \end{array}$$

It is an easy exercise to verify that the system of filters $\{\mathcal{U}[x] : x \in X\}$ satisfies condition 5.22(C), and hence it is the system of neighborhood filters for a topology $\mathcal{T}$ on X. That topology is called the **uniform topology** determined by $\mathcal{U}$. We may sometimes denote it by $\mathcal{T}_{\mathcal{U}}$.

A topology that can be represented in this fashion is said to be **uniformizable**. Most topologies in applications are uniformizable. Some examples of nonuniformizable topologies — admittedly, rather artificial — are given in 16.18.

Different uniformities may determine different topologies or the same topology; we do not necessarily have a particular uniformity in mind when we say that a topology $\mathcal{T}$ is uniformizable. We say that a topology $\mathcal{T}$ and a uniformity $\mathcal{U}$ are **compatible** if $\mathcal{T} = \mathcal{T}_{\mathcal{U}}$.

Exercise. Let D be a gauge on X. Then the gauge topology $\mathcal{T}_D$ determined by D as in 5.15.h is the same as the uniform topology $\mathcal{T}_{\mathcal{U}}$ determined as above from the uniformity $\mathcal{U} = \mathcal{U}_D$ defined from the gauge D as in 5.32.

5.34. *Examples.*

a. $\{U \subseteq X \times X \ : \ U \supseteq I\}$ is the largest uniformity on X; we shall call it the **discrete uniformity**. The topology that it determines is the discrete topology. The discrete uniformity is determined by some discrete metrics — for instance, by the Kronecker metric $1 - \delta$ (see 2.12.b).

Any discrete metric (defined in 2.12.b) yields the discrete topology (defined in 5.15.b). However, *not every discrete metric yields the discrete uniformity*; an example is given in 19.11.e.

b. Let $d : X \times X \to \mathbb{R}$ be the constant function 0. Then d is a pseudometric. The uniformity it determines is the singleton $\{X \times X\}$. This is the smallest uniformity on X; we shall call it the **indiscrete uniformity**. It determines the indiscrete topology (defined in 5.15.a).

c. Let X be a nonempty set, and let ξ be some particular specified element of X; we may refer to ξ as the "knob" of X. Let $S = X \setminus \{\xi\}$. Define a pseudometric on X by

$$d(u,v) \quad = \quad \begin{cases} 0 & \text{if both or neither of } u,v \text{ are equal to } \xi \\ 1 & \text{if exactly one of } u,v \text{ is equal to } \xi \end{cases}$$

for $u, v \in X$. A set X equipped with this pseudometric will be called a **knob space**. The resulting uniformity is

$$\mathcal{U} \quad = \quad \{U \subseteq X \times X \ : \ U \supseteq S \times S \text{ and } (\xi, \xi) \in U\}$$

and the resulting topology is $\{\varnothing, \ \{\xi\}, \ S, \ X\}$. Knob spaces will be important in certain arguments concerning the Axiom of Choice, in 17.20 and 19.13, using "Kelley's choice" (see 6.24).

5.35. *Some basic properties of uniformities.* Let $(X, \mathcal{U})$ be a uniform space. Show that

a. $U \in \mathcal{U} \Rightarrow U^{-1} \in \mathcal{U}$. (This property is not always satisfied by a preuniformity.)

b. If U is an entourage, then $V = U \cap U^{-1}$ is a *symmetric* entourage (i.e., an entourage satisfying $V = V^{-1}$) contained in U.

c. If U is an entourage and k is a positive integer, then there exists an entourage V satisfying

$$V^k \quad = \quad \underbrace{V \circ V \circ V \circ \cdots \circ V}_{k \text{ of the } V\text{'s}} \quad \subseteq \quad U.$$

Moreover, we may choose V symmetric.

In particular, if U is an entourage then there exists a symmetric entourage V such that $V^3 = V \circ V \circ V \subseteq U$. We shall use that fact in our proof of 16.16.

5.36. *Pathological example.* An intersection of uniformities is not necessarily a uniformity. (In this respect, uniformities are not like σ-algebras or topologies.)

For instance, take $X = \mathbb{R} \times \mathbb{R}$. Let $\pi_1, \pi_2 : X \to \mathbb{R}$ be the coordinate projections — that is, $\pi_1(x_1, x_2) = x_1$ and $\pi_2(x_1, x_2) = x_2$. Define pseudometrics d_1, d_2 on X by

$$d_1(x,y) \quad = \quad |\pi_1(x) - \pi_1(y)|, \qquad\qquad d_2(x,y) \quad = \quad |\pi_2(x) - \pi_2(y)|,$$

for x and y in X. Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be the resulting pseudometric uniformities on X, and let $\mathcal{W} = \mathcal{U}_1 \cap \mathcal{U}_2$. Show that $\mathcal{W}$ is not a uniformity. Use that fact to show also that, although there do exist uniformities that contain $\mathcal{W}$, there does not exist a smallest uniformity containing $\mathcal{W}$.

5.37. It will sometimes be useful to "generate" a uniformity $\mathcal{U}$ from a smaller collection $\mathcal{S}$ of sets. However, we saw in 5.36 that an intersection of uniformities is not necessarily a uniformity. Given a collection $\mathcal{S}$ of sets, we cannot simply look for the "smallest uniformity $\mathcal{U}$ that contains $\mathcal{S}$;" such a uniformity $\mathcal{U}$ need not exist. The smaller, "generating" collection $\mathcal{S}$ cannot be chosen arbitrarily, but preuniformities (defined in 5.31) will serve quite well for this purpose.

Suppose that $\mathcal{S}$ is a preuniformity on X. Then $\mathcal{S}$ is a filter subbase on $X \times X$, by 5.31(i). Hence $\mathcal{S}$ generates a proper filter on $X \times X$; that filter is

$$\mathcal{U} \quad = \quad \Big\{U \subseteq X \times X \quad : \quad U \supseteq S_1 \cap S_2 \cap \cdots \cap S_n \quad \text{for some finite set } \{S_1, S_2, \ldots, S_n\} \subseteq \mathcal{S}\Big\}.$$

It is easy to show that $\mathcal{U}$ is a uniformity on X, and in fact $\mathcal{U}$ is the smallest uniformity containing $\mathcal{S}$. We shall call $\mathcal{U}$ the uniformity **generated by** $\mathcal{S}$. *Caution*: Despite the similar language, this is *not* a special case of a Moore closure.

5.38. Here are some further noteworthy properties of a preuniformity $\mathcal{S}$ on X and the uniformity $\mathcal{U}$ it generates:

- **a.** The union of any family of preuniformities on X is a preuniformity on X.
- **b.** In many cases of interest, the preuniformity $\mathcal{S}$ has the property that it is closed under finite intersection — i.e., $S_1, S_2 \in \mathcal{S} \Rightarrow S_1 \cap S_2 \in \mathcal{S}$. In this case, we can simplify our formula for the uniformity $\mathcal{U}$ generated by $\mathcal{S}$; the formula becomes

  $$\mathcal{U} \quad = \quad \{U \subseteq X \times X \quad : \quad U \supseteq S \text{ for some } S \in \mathcal{S}\}.$$

  Moreover, in this case let us denote

  $$S[x] \; = \; \{y \, : \, (x,y) \in S\} \qquad \text{and} \qquad \mathcal{S}[x] \; = \; \{S[x] \, : \, S \in \mathcal{S}\};$$

  then $\mathcal{S}[x]$ is a neighborhood base at x for the uniform topology.

Most of these ideas are from Kelley [1955/1975], although that book does not use the term "preuniformity."

5.39. *Remarks.* The theory of uniform spaces will be developed a little further in Chapter 9 and in 16.16. It will be continued in much greater detail in Chapter 18 and thereafter.

Images and Preimages of Sets of Sets

5.40. Let $g : X \to Y$ be some function. Show that

a. *Forward image of a filter subbase.* If $\mathcal{S}$ is a filter subbase or a filterbase on X, then

$$g(\mathcal{S}) \quad = \quad \{g(S) \ : \ S \in \mathcal{S}\}$$

is a filter subbase or a filterbase, respectively, on Y.

b. *Sets with suitable inverse images.* Let $\mathcal{S}$ be a collection of subsets of X, and let

$$\mathcal{T} \quad = \quad \{T \subseteq Y \ : \ g^{-1}(T) \in \mathcal{S}\}.$$

If $\mathcal{S}$ is closed under complementation in X, under finite or countable or arbitrary union, or under finite or countable or arbitrary intersection, then $\mathcal{T}$ is closed under the same operation in Y. If $\mathcal{S}$ is a filter subbase, a filter, an ultrafilter, a fixed ultrafilter, or a collection of nonempty subsets of X, a topology, or a σ-algebra on X, then $\mathcal{T}$ has the same property on Y.

If $\mathcal{S}$ is a filter generated by a filterbase $\mathcal{B}$ on X, then $g(\mathcal{S})$ and $g(\mathcal{B})$ (defined as in 5.40.a) are filterbases on Y, both of which generate the filter $\mathcal{T}$ (defined as above).

c. *Inverse image of a collection.* Let $\mathcal{T}$ be a collection of subsets of Y, and let

$$g^{-1}(\mathcal{T}) \quad = \quad \{g^{-1}(T) \ : \ T \in \mathcal{T}\}.$$

If $\mathcal{T}$ is closed under complementation in Y, under finite or countable or arbitrary union, or under finite or countable or arbitrary intersection, then $g^{-1}(\mathcal{T})$ is closed under the same operation in X. If $\mathcal{T}$ is a topology or a (σ-)algebra, then $g^{-1}(\mathcal{T})$ is, too.

If g is *surjective* and $\mathcal{T}$ is a filter subbase or filterbase on Y, then $g^{-1}(\mathcal{T})$ has the same property on X.

d. Define $g \times g : X^2 \to Y^2$ by $(g \times g)(x_1, x_2) = (g(x_1), g(x_2))$. If $\mathcal{U} \subseteq Y \times Y$ is a preuniformity on Y, then

$$(g \times g)^{-1}(\mathcal{U}) \ = \ \left\{(g \times g)^{-1}(U) \ : \ U \in \mathcal{U}\right\}$$

is a preuniformity on X (regardless of whether g is surjective).

Transitive Sets and Ordinals

5.41. *Remark.* The ideas in the remainder of this chapter are mainly needed for set theory and logic; they can be skipped if one is only concerned with the traditional topics of analysis.

5.42. *Definitions and remarks.* Let X be a set. Then the following conditions are equivalent. If one (hence all) are satisfied, we say X is **transitive** (or more precisely, **$\in$-transitive**).

(A) Whenever $A \in S$ and $S \in X$, then $A \in X$.

(B) Each member of X is also a subset of X — that is, $A \in X \Rightarrow A \subseteq X$. That can also be restated as $X \subseteq \mathcal{P}(X)$.

(C) $\mathrm{Un}(X) \subseteq X$, where $\mathrm{Un}(X)$ is the union of the members of X, as defined with the Axiom of Unions in 1.47.

Remarks. As Doets [1983] points out, the notion of transitive sets is slightly alien to most mathematicians — i.e., those not actively involved in set theory. After all, for most mathematicians, it is enough to consider sets of sets and occasionally sets of sets of sets. But if X is transitive, then each member of X is a subset of X, and so each member of each member of X is a subset of X, and so on. For some purposes in set theory, this process must be continued to an infinite depth; see 5.44.

5.43. *Basic properties of transitive sets.*

a. *Examples.* The sets $\varnothing$ and $\{\varnothing\}$ and $\{\varnothing, \{\varnothing\}\}$ are transitive; the set $\{\{\varnothing\}\}$ is not.

b. The intersection of any nonempty collection of transitive sets is a transitive set.

c. If S is any set, then S is a *subset* of some transitive set. For instance,

$$\mathrm{cl}(S) \quad = \quad S \;\cup\; \mathrm{Un}(S) \;\cup\; \mathrm{Un}(\mathrm{Un}(S)) \;\cup\; \mathrm{Un}(\mathrm{Un}(\mathrm{Un}(S))) \;\cup\; \cdots$$

is a transitive set with $\mathrm{cl}(S) \supseteq S$. In fact, $\mathrm{cl}(S)$ is the smallest transitive superset of S; it is the intersection of all the transitive supersets of S. We shall call it the **transitive closure** of S. It is a special case of Moore closures (discussed in 4.3), except that in this case the domain of cl is a proper class, not a set.

d. If S is any set, then S is a *member* of some transitive set. For instance, one such set is the transitive closure of the singleton $\{S\}$.

5.44. *Preview/examples.* Because the definition of ordinals is somewhat abstract and complicated, we shall precede that definition with a few examples. Of course, the assertions that we now make about these examples cannot be proved until a few pages later.

The first few ordinals are the *finite ordinals.* Set theorists find it convenient to attach the labels "0," "1," "2," etc. to these sets. (See the related discussions in 1.16 and 1.46.) The nonnegative integers are thus defined to be the sets

$$\begin{aligned} 0 &= \varnothing, \\ 1 &= \{\varnothing\}, \\ 2 &= \Big\{\varnothing, \{\varnothing\}\Big\}, \\ 3 &= \left\{\varnothing, \{\varnothing\}, \Big\{\varnothing, \{\varnothing\}\Big\}\right\}, \end{aligned}$$

and so on. Thus the set $n = \{0, 1, 2, \ldots, n-1\}$ contains exactly n elements, and $n + 1 = n \cup \{n\}$ is the successor of n.

After the finite ordinals come the countably infinite ordinals. The first few of these are

$$\begin{array}{rcl} \omega & = & \{0,1,2,3,\ldots\}, \\ \omega+1 & = & \{0,1,2,3,\ldots,\omega\}, \\ \omega+2 & = & \{0,1,2,3,\ldots,\omega,\omega+1\}, \\ & & \cdots \quad \cdots \\ 2\omega & = & \{0,1,2,3,\ldots,\omega,\omega+1,\omega+2,\ldots\}, \\ 2\omega+1 & = & \{0,1,2,3,\ldots,\omega,\omega+1,\omega+2,\ldots,2\omega\} \\ 2\omega+2 & = & \{0,1,2,3,\ldots,\omega,\omega+1,\omega+2,\ldots,2\omega,2\omega+1\} \\ & & \cdots \quad \cdots \\ 3\omega & = & \{0,1,2,3,\ldots,\omega,\omega+1,\omega+2,\ldots,2\omega,2\omega+1,2\omega+2,\ldots\} \end{array}$$

and so on. Again, note that the successor of any ordinal S is the ordinal $S \cup \{S\}$. The ordinal ω is an ordered version of the unordered set $\mathbb{N} \cup \{0\}$.

After the countable ordinals come the uncountable ordinals. They are a bit harder to visualize, but we can easily sketch a proof of their existence. Consider any uncountable set — for instance, $2^{\mathbb{N}}$ — and give it a well ordering (see (AC4) in 6.20). Then find the ordinal that is order isomorphic to it (see 5.46.f). Among the uncountable ordinals that are isomorphic to subsets of $2^{\mathbb{N}}$, there is a first one, by 5.46.g. The first uncountable ordinal is equal to the set of all countable ordinals.

5.45. Before proceeding further, the reader may find it helpful to briefly review the theory of well ordered sets developed in Chapter 3.

Definitions. In the discussion below, the expression $A \overset{\in}{=} B$ will mean "$A \in B$ or $A = B$."

Let X be a set. Then the following conditions are equivalent. If one (hence all) are satisfied, we say X is an **ordinal**.

(A) X is a transitive set, and the relation $\overset{\in}{=}$ is a chain ordering of X.

(B) X is a transitive set, and the relation $\overset{\in}{=}$ is a well ordering of X.

(C) X is a transitive set, and all the members of X are transitive sets.

Proof of equivalence of conditions. The implication (B) $\Rightarrow$ (A) is trivial, and the implication (A) $\Rightarrow$ (C) is a fairly easy exercise. For (A) $\Rightarrow$ (B) use the Axiom of Regularity in 1.47: If S is a nonempty subset of X, it is easy to show that any $\in$-minimal element of S is also a $\overset{\in}{=}$-minimum element of S.

It remains to prove (C) $\Rightarrow$ (A). Suppose (A) is false; thus there exist sets $A, B \in X$ that are not $\overset{\in}{=}$-comparable. Using the Axiom of Regularity, let A_0 be a $\in$-minimal element of the nonempty set

$$S_0 \quad = \quad \{A \in X \ : \ \text{some } B \in X \text{ exists such that } A, B \text{ are not } \overset{\in}{=}\text{-comparable}\}.$$

Then let B_0 be some $\in$-minimal element of the nonempty set

$$T_0 \quad = \quad \{B \in X \ : \ A_0 \text{ and } B \text{ are not } \overset{\in}{=}\text{-comparable}\}.$$

Then A_0 and B_0 are not $\stackrel{\in}{=}$-comparable. Both A_0 and B_0 are members of X, hence they are transitive sets.

We first show that $B_0 \subseteq A_0$. Indeed, let $C_0 \in B_0$; we shall show $C_0 \in A_0$. Since B_0 is an $\in$-minimal element of T_0, it follows that C_0 is not a member of T_0. Since $C_0 \in B_0 \in X$ and X is transitive, we have $C_0 \in X$. If $A_0 \stackrel{\in}{=} C_0 \in B_0$, it would contradict the fact that A_0 and B_0 are not $\stackrel{\in}{=}$-comparable; thus we do not have $A_0 \stackrel{\in}{=} C_0$. Since $C_0 \in X \setminus T_0$, it follows that A_0 and C_0 are $\stackrel{\in}{=}$-comparable; thus we must have $C_0 \in A_0$. This proves $B_0 \subseteq A_0$.

Since A_0 and B_0 are not $\stackrel{\in}{=}$-comparable, they are not equal; thus $A_0 \setminus B_0$ is nonempty. Let D_0 be some member of $A_0 \setminus B_0$. Since A_0 is a $\in$-minimal element of S_0 and $D_0 \in A_0$, it follows that $D_0 \notin S_0$. Thus D_0 is $\stackrel{\in}{=}$-comparable with every member of X, and in particular D_0 is $\stackrel{\in}{=}$-comparable with B_0. By our choice of D_0 we know that $D_0 \notin B_0$; thus we must have $B_0 \stackrel{\in}{=} D_0$. But then $B_0 \stackrel{\in}{=} D_0 \in A_0$, contradicting the fact that A_0 and B_0 are not $\stackrel{\in}{=}$-comparable. (This argument follows Shoenfield [1967].)

Remarks. In recent years the definitions above (due to von Neumann) have become standard. However, some of the earlier literature used slightly different definitions of "ordinal." To some mathematicians an ordinal meant any well ordered set. To others it meant any equivalence class of well ordered sets, where two well ordered sets are considered to be equivalent if there exists an order isomorphism between them. The latter definition may cause some difficulties, since that equivalence class is a proper class, not a set. The von Neumann definition removes these difficulties by specifying a natural representative from that equivalence class, as we shall see in 5.46.f.

5.46. *Basic properties of ordinals.*

- **a.** If X is an ordinal, then we understand X to be equipped with the ordering given by $\stackrel{\in}{=}$, which makes X a well ordered set. Note that with this ordering, if $\alpha \in X$, then the set of predecessors of α is $\text{Pre}(\alpha) = \{x \in X : x \in \alpha\} = \alpha$.
- **b.** If X is an ordinal, then $X = \{\text{proper lower sets of } X\}$. *Hint*: 3.39.c.
- **c.** All the members of an ordinal are ordinals.
- **d.** If X and Y are ordinals, then $Y \subsetneqq X \iff Y \in X$.

  Hints: $Y \in X \Rightarrow Y \subsetneqq X$ since X is transitive. Conversely, suppose $Y \subsetneqq X$. Let α be the first member of $X \setminus Y$. Show that $\text{Pre}(\alpha) = Y$.
- **e.** The only order isomorphism from an ordinal onto an ordinal is the identity map from an ordinal to itself. *Hint*: Induction on lower sets.
- **f.** If $(X, \preccurlyeq)$ is a well ordered set, then there is one and only one mapping that is an order isomorphism from X onto an ordinal. (*Hint*: Induction on lower sets.) That ordinal is sometimes referred to as the **ordinal type** of X.
- **g.** If $\mathcal{C}$ is any nonempty subclass of the class of ordinals, then the intersection of all the members of $\mathcal{C}$ is an ordinal, and furthermore that ordinal is a member of $\mathcal{C}$ — in fact, it is the smallest member of $\mathcal{C}$.

  If X and Y are ordinals, then $X \in Y$ or $X = Y$ or $Y \in X$. Thus, $\stackrel{\in}{=}$ is a chain ordering on the class of all ordinals.

In fact, it is a well ordering: If $\mathcal{C}$ is any nonempty subclass of the class of all ordinals, then $\mathcal{C}$ has a smallest member — namely, the intersection of all the members of $\mathcal{C}$.

(Later we shall show that the class of all ordinals is a proper class — i.e., it is not a set.)

h. The union of any set of ordinals is an ordinal.

i. If X is an ordinal, then so is $X \cup \{X\}$. It is called the **successor** of X; it is sometimes written X^+ or $X + 1$. It is the smallest ordinal greater than X — i.e., it is the first ordinal after X. Any ordinal that can be written in the form X^+ for some X is called a **successor ordinal**.

Note that any successor ordinal X^+ has a largest element — namely, X. Conversely, show that any ordinal with a largest element is a successor ordinal. Thus, we may equivalently define a *successor ordinal* to be an ordinal that has a largest element.

j. A **limit ordinal** is an ordinal that does not have a largest element.

Show that if X is a limit ordinal, then $\bigcup_{S \in X} S = X$.

Examples. Refer to 5.44. The ordinals $\omega, 2\omega, 3\omega, \ldots$ and the first uncountable ordinal are limit ordinals. The ordinals 1, 2, 3, ... and $\omega + 1,\ \omega + 2,\ \omega + 3,\ \ldots$ and $2\omega + 1,\ 2\omega + 2,\ 2\omega + 3,\ \ldots$, etc. are successor ordinals. By our definition, the empty set is a limit ordinal — but some mathematicians use a slightly different definition for limit ordinal that excludes the empty set.

5.47. Note that different ordinals may have the same cardinality. For instance, it is easy (*exercise*) to give bijections between the ordinals ω and $\omega + 1$ and 2ω. (Of course, such bijections cannot be order preserving.)

An **initial ordinal**, also known as a **cardinal** or a **cardinal number**, is an ordinal X with the property that no earlier ordinal has the same cardinality as X. (*Note*: Some mathematicians add the restriction that the set be infinite as part of the definition of initial ordinal, but we shall not impose that restriction.)

It follows from 3.42.b that any infinite cardinal must be a limit ordinal.

Examples. Refer to 5.44. All the finite ordinals are cardinals; ω is a cardinal; the first uncountable ordinal is a cardinal. The ordinals $\omega + 1$ and 2ω are not cardinals, since they have the same cardinality as ω.

Preview. It will follow from (AC4) in 6.20 that any set S can be well ordered, and hence we can assign a cardinal number to each set. See further remarks in 6.23.

5.48. (*Optional.*) The infinite cardinals are also called **alephs**; they are written

$$\aleph_0,\ \aleph_1,\ \aleph_2,\ \aleph_3,\ \ldots,\ \aleph_\omega,\ \aleph_{\omega+1},\ \aleph_{\omega+2},\ \ldots,\ \aleph_{2\omega},\ \ldots,\ \text{etc.}$$

"Aleph" is the name of $\aleph$, the first letter of the Hebrew alphabet.

The first infinite ordinal is $\omega = \aleph_0$; it is countable. The first uncountable ordinal is $\aleph_1$. The Continuum Hypothesis is the statement that $2^{\aleph_0} = \aleph_1$.

We have these inclusions:

$$\{\text{alephs}\} \subsetneqq \{\text{ordinals}\} \subsetneqq \{\text{sets}\}.$$

We saw in 1.45 that {sets} is a proper class, not a set; likewise we shall show in 5.50 that {ordinals} is a proper class. The class of all alephs is also a proper class, but we shall not prove that; a proof is given by Krivine [1971].

The Class of Ordinals

5.49. *Definition.* For any set X, the **Hartogs number** of X is defined to be the smallest ordinal α that satisfies $\text{card}(\alpha) \not\leq \text{card}(X)$ — i.e., the smallest ordinal α that does not satisfy $\text{card}(\alpha) \leq \text{card}(X)$. (It is a cardinal number.)

We do *not* require that $\text{card}(\alpha) > \text{card}(X)$. That slightly stronger statement will follow as a consequence only if we assume the Axiom of Choice — see 6.22. The Hartogs number is mainly useful if one wishes to avoid using the Axiom of Choice — e.g., to study its alternatives, as we shall do briefly in this book.

Exercise. Prove that the definition of the Hartogs number makes sense — i.e., prove that there do exist ordinals α satisfying $\text{card}(\alpha) \not\leq \text{card}(X)$, and among such ordinals there is a smallest.

Hint: Let $\alpha = \{\beta : \beta \text{ is an ordinal with } \text{card}(\beta) \leq \text{card}(X)\}$. The only hard part is showing that α is actually a set. (After all, the collection of *all* ordinals is *not* a set; we shall prove this below.) First use the Axiom of Comprehension in 1.47 to show that

$$\mathcal{A} \quad = \quad \Big\{(S, R) \quad : \quad S \subseteq X,\ R \subseteq S \times S, \text{ and } R \text{ is a well ordering of } S\Big\}$$

is a set. Then use 3.41, 5.46.f, and the Axiom of Replacement to prove α is a set. Finally, show that α is an ordinal, and α is the Hartogs number of X.

5.50. Theorem. Let $\mathcal{O} = \{\text{ordinals}\}$. Then $\mathcal{O}$ is a proper class, not a set.

We offer two slightly different proofs, because both are interesting. The first proof is more elementary in that it does not use the Hartogs number; the second proof may be preferable to some readers because it does not use the Axiom of Regularity.

First proof. Suppose that $\mathcal{O}$ is a set. Show that $\mathcal{O}$ is then an ordinal and hence a member of itself, contradicting 1.49. (This is the *Burali-Forti Paradox.*)

Second proof. Suppose $\mathcal{O}$ is a set. Let β be the Hartogs number of $\mathcal{O}$. Then β is an ordinal, hence $\beta \subseteq \mathcal{O}$, hence $\text{card}(\beta) \leq \text{card}(\mathcal{O})$, contradicting the definition of the Hartogs number.

5.51. The collection of all ordinals is a proper class — i.e., a "very big" collection, much like the collection of all sets. Nevertheless, the ordinals have some interesting structure; they are well ordered by $\overset{\in}{=}$, as we noted in 5.46.g. Consequently we have the following two principles:

Induction on the Ordinals. Suppose $\mathcal{C}$ is a class of ordinals, such that whenever X is an ordinal whose members all belong to $\mathcal{C}$, then X also belongs to $\mathcal{C}$. Then $\mathcal{C}$ contains all the ordinals.

Recursion on the Ordinals. By an *ordinal-based map* we shall mean a function from some ordinal into some set. Let $\mathcal{M}$ be the class of all ordinal-based maps. Let ρ be some function of classes, from $\mathcal{M}$ into {sets}. Then there exists a unique function F : {ordinals} $\to$ {sets} that satisfies $F(X) = \rho\left(F\big|_X\right)$ for each ordinal X.

In the last line above, $F\big|_X$ is the function F restricted to X. The members of its domain are the *members* of X — i.e., the preceding ordinals — and not X itself. Note that any ordinal X is equal to its own set of predecessors.

Proofs. If the induction principle does not hold, then there is some ordinal $M \notin \mathcal{C}$. Let X be the first member of M^+ that does not belong to $\mathcal{C}$; then X is the first ordinal that does not belong to $\mathcal{C}$, a contradiction. The recursion principle can be proved by an argument similar to 3.40; we omit the details.

5.52. Zermelo's Fixed Point Theorem. Let $(X, \preccurlyeq)$ be a nonempty poset, with the property that each $\preccurlyeq$-chain in X has a $\preccurlyeq$-supremum in X. Suppose $f : X \to X$ is a function that satisfies $f(x) \succcurlyeq x$ for all x. Then f has at least one fixed point.

Proof. Suppose not — i.e. suppose $f(x) \succ x$ for all $x \in X$. Then $u(C) = f(\sup C)$ defines a function

$$u : \{\text{chains in } X\} \to X, \qquad \text{satisfying} \qquad u(C) \succ c \text{ for all } c \in C.$$

We shall show that such a function yields a contradiction. Let $\mathcal{H}$ be either the Hartogs number of X or the class of all ordinals, according to the reader's taste — the rest of the argument will work with either $\mathcal{H}$. Recursively define a strictly increasing mapping $\psi : \mathcal{H} \to X$ by

$$\psi(S) \quad = \quad \varphi\Big(u\{\psi(T) \ : \ T \in A\} \Big).$$

To see that this definition makes sense, note that if ψ is strictly increasing on some ordinal $S \in \mathcal{H}$, then $C(S) = \{\psi(T) : T \in S\}$ is a chain, and so $u(C(S))$ exists and is an upper bound for $C(S)$. Hence $f(u(C(S)))$ exists and is strictly greater than every member of $C(S)$ — hence ψ is defined and strictly increasing on $S^+ = S \cup \{S\}$. This completes the definition of ψ. However, since $\psi : \mathcal{H} \to X$ is strictly increasing, it is injective, and therefore $\text{card}(\mathcal{H}) \leq \text{card}(X)$, a contradiction.

Remarks. The proof above is from Howard [1992]. Slightly longer proofs that avoid the use of ordinals are given by Fuchssteiner [1986] and Mańka [1988]. *None of these proofs requires the Axiom of Choice or any of its consequences.* Thus Zermelo's Fixed Point Theorem is occasionally useful in the study of set theory without the Axiom of Choice, as in 19.45. If we permit the use of the Axiom of Choice and its equivalents, then Zermelo's Fixed Point Theorem is a trivial corollary of Zorn's Lemma, which is (AC7) in 6.20.

5.53. The class of all sets specified by the ZF axioms is often denoted by **V**, because an interesting and useful description of it was given by von Neumann. This scheme is also known as the **cumulative hierarchy**. Using recursion on the ordinals (in 5.51), we define a function of classes

$$\text{Stage} \quad : \quad \{\text{ordinals}\} \to \{\text{sets}\}$$

by this rule:

$$\text{Stage}(\alpha) \quad = \quad \bigcup_{\beta \in \alpha} \mathcal{P}\Big(\text{Stage}(\beta)\Big).$$

In other words, the αth stage is the collection of all subsets of all sets that have *already* been formed in previous stages. (The literature contains minor variants on this definition. Some mathematicians prefer to define Stage(α) by two slightly different formulas when α is a limit ordinal or when α is a successor ordinal. However, the ultimate effect is the same. Also, some mathematicians use the term "rank" instead of "stage.") Von Neumann's universe is the class

$$V \quad = \quad \bigcup_{\alpha \in \{\text{ordinals}\}} \text{Stage}(\alpha)\,.$$

Although each Stage(α) is a set, V is a proper class.

In 1.44 we stated, somewhat imprecisely and intuitively, that a set is a collection of "already fixed" sets. We have not used that statement in our formal development of ZF set theory; instead we have simply assumed that the class of all sets is some collection of objects that satisfies the ZF axioms. However, we are now ready to use the ZF axioms to prove a precise version of that earlier intuitive statement.

Theorem. In ZF set theory, every set is in some stage — that is, the von Neumann class V is the class of all sets.

Proof (following Shoenfield [1967]). Let X be any given set; we wish to show $X \in V$. By 5.43.d, let T be a transitive set with $X \in T$. If $T \subseteq V$, then we are done. Assume, then, that $T \setminus V$ is nonempty. Note that the class $T \setminus V$ is actually a set, by the Axiom of Comprehension. By the Axiom of Regularity, let M be a $\in$-minimal member of the set $T \setminus V$.

Let A be any member of the set M. Then $A \in T$ by transitivity of T, but $A \notin (T \setminus V)$ by minimality of M. Thus $A \in V$. Therefore $A \in \text{Stage}(\alpha_A)$ for some ordinal α.

Now, $\{\alpha_A : A \in M\}$ is a *set* of ordinals. Its union is an ordinal β, which has some successor β^+. For every $A \in M$, we have $\alpha_A \overset{\in}{=} \beta$; hence $A \in \text{Stage}(\beta)$. Thus $M \subseteq \text{Stage}(\beta)$, so $M \in \text{Stage}(\beta^+)$ — contradicting our choice of M as a set that does not belong to V.

5.54. By a slight modification of von Neumann's cumulative construction, we shall obtain *Gödel's constructible universe*, L. This will appear briefly in our discussions in Chapter 14. The idea is that instead of taking *arbitrary* subsets of V_α to get $V_{\alpha+1}$, we shall use *describable* subsets.

Define ordered pairs and ordered triples in terms of sets, as in 1.46; define the product of two sets as a set of ordered pairs. Then the **Gödel operations** are defined as follows. First, for any sets X and Y, let

$$\mathcal{F}_1(X,Y) = \{X,Y\},$$

$$\mathcal{F}_2(X,Y) = X \backslash Y,$$

$$\mathcal{F}_3(X,Y) = X \times Y = \{(x,y) : x \in X, y \in Y\}.$$

Also, for any set X, let

$$\mathcal{F}_4(X) = \mathrm{Dom}(X) = \{u : (u,v) \in X \text{ for some } v\},$$

$$\mathcal{F}_5(X) = \{(u,v) \in X \times X : u \in v\},$$

$$\mathcal{F}_6(X) = \{(u,v,w) : (v,w,u) \in X\},$$

$$\mathcal{F}_7(X) = \{(u,v,w) : (w,v,u) \in X\},$$

$$\mathcal{F}_8(X) = \{(u,v,w) : (v,u,w) \in X\}.$$

Now, for any set X, define

$$\mathcal{G}(X) \;=\; X \cup \{\mathcal{F}_i(u,v) : u,v \in X,\ 1 \le i \le 3\} \cup \{\mathcal{F}_i(u) : u \in X,\ 4 \le i \le 8\}.$$

Let $\mathcal{G}^2(X) = \mathcal{G}(\mathcal{G}(X))$, etc., and define

$$\mathrm{cl}(X) \;=\; X \cup \mathcal{G}(X) \cup \mathcal{G}^2(X) \cup \mathcal{G}^3(X) \cup \cdots .$$

Then $\mathrm{cl}(X)$ is the smallest set that contains X and is closed under the Gödel operations. We now recursively define

(i) $L_0 = \varnothing$,

(ii) $L_{\alpha+1} = \mathcal{P}(L_\alpha) \cap \mathrm{cl}(L_\alpha \cup \{L_\alpha\})$ for ordinals α,

(iii) $L_\alpha = \bigcup_{\beta<\alpha} L_\beta$ when α is a limit ordinal,

and finally $L = \bigcup_{\alpha \in \text{ordinals}} L_\alpha$. Of course, L is a proper class, since Ord is not a set. The members of L are said to be **Gödel constructible**, or **constructible relative to the ordinals**. For further discussion we refer to Jech [1973], Manin [1977], or other books on logic and set theory.

Is every set in von Neumann's universe "constructed" at some stage of Gödel's hierarchy? Or are there some other sets in V that cannot be so "constructed?" In other words, is L equal to V, or are these classes different? This question cannot be answered either way except by making additional assumptions beyond those of conventional set theory. The statement $V = L$ is called the **Axiom of Constructibility**; it is discussed further in 14.7.

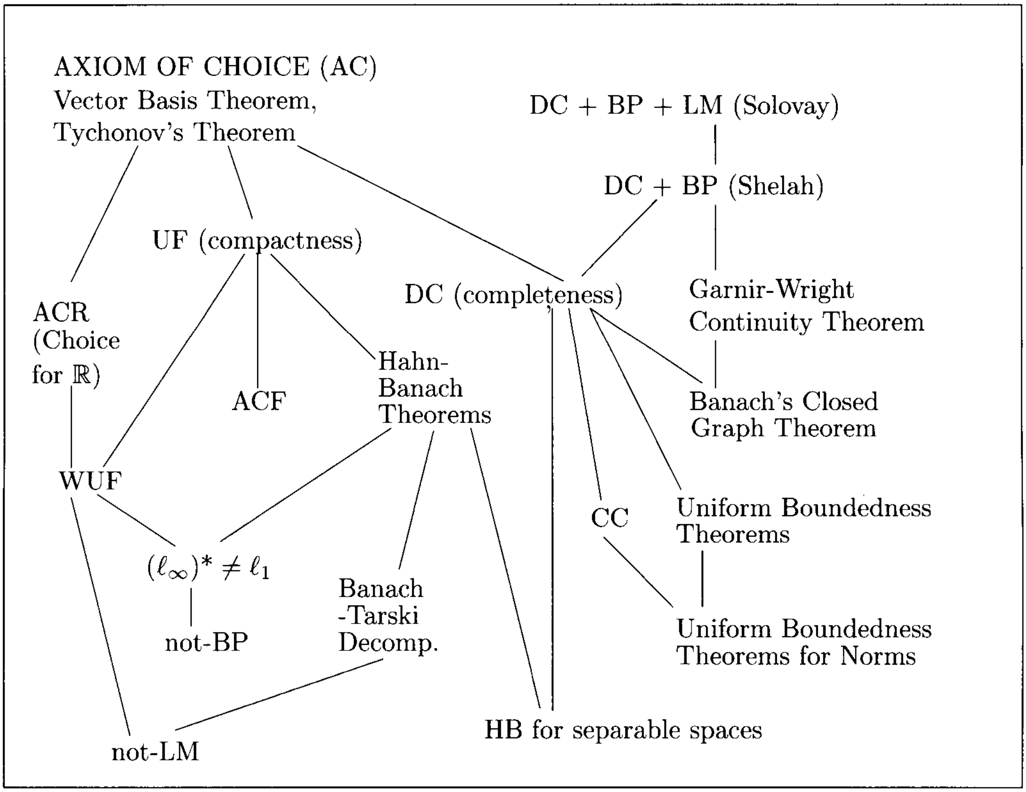
Gödel's constructions may take uncountably many steps. They are quite different from, and should not be confused with, Bishop's constructions, introduced in 6.2, which permit only countably many steps.

Chapter 6

Constructivism and Choice

6.1. *Preview.* Conventional set theory is ZF + AC; that is, Zermelo-Fraenkel set theory plus the Axiom of Choice. ZF was introduced in 1.47; for the most part, it is just a formalization of our intuition about sets.

This chapter introduces the Axiom of Choice (AC) and a few weakened forms of Choice. Some relations between these principles are summarized in the chart below, which is based



partly on a chart of Pincus [1974]. All assertions in the chart are understood to be in conjunction with ZF. Implications in the chart are downward, and we shall prove most of these implications. For instance, in this chapter we include proofs of AC $\Rightarrow$ DC $\Rightarrow$ CC and AC $\Rightarrow$ UF $\Rightarrow$ ACF. That DC + BP implies the Garnir-Wright Theorem is proved in 27.45; a proof of $(\ell_\infty)^* \neq \ell_1 \Rightarrow$ not-BP is given in 29.38. The proof of WUF $\Rightarrow$ not-LM is somewhat complicated and is not included in this book; it was given by Sierpiński [1938].

Most of the implications are known to be irreversible — for instance, it is known that HB $\not\Leftarrow$ UF and UF $\not\Leftarrow$ AC — but proofs of these irreversibility results are beyond the scope of this book. Most of them can be found in Jech [1973] or Pincus [1974] or in references cited therein. An enormous survey of the weak forms of Choice, including implications and irreversibility results, is given by [Howard and Rubin, in preparation].

For our purposes, the most important principles are AC (the Axiom of Choice), DC (Dependent Choice), UF (the Ultrafilter Principle), and HB (the Hahn-Banach Theorem). These four principles appear in many equivalent forms in later chapters.

Most interesting consequences of AC actually follow from either DC or UF. One might almost think of those principles as the "constructive component" and the "nonconstructive component" of AC. However, that description would be slightly misleading, for it is known that ZF + DC + UF does not imply AC; see Pincus [1977].

Particular attention will also be devoted to the principles

BP = "every subset of $\mathbb{R}$ has the Baire property," and

not-BP = "there exists a subset of $\mathbb{R}$ that lacks the Baire property."

The topological meaning of the Baire property will be discussed in Chapter 20, but its foundational significance must be mentioned now. The statement not-BP is the weakest nonconstructive consequence of AC that we shall consider as such in this book; thus BP is our strongest negation of the Axiom of Choice. **Shelah's Theorem,**

$$\mathrm{Con}(\mathrm{ZF}) \Rightarrow \mathrm{Con}(\mathrm{ZF} + \mathrm{DC} + \mathrm{BP}),$$

is a remarkable accomplishment; it gives us a unified method of proving the intangibility of many of the pathological objects that arise in analysis — i.e., of proving that those objects have no explicitly constructible examples. This is discussed in greater detail in 14.76 and 14.77.

Examples of Nonconstructive Mathematics

6.2. Most of this book follows mainstream mathematics, which is *not* constructivist. However, a brief discussion of constructivism will be helpful.

The terms "construct" and "construction" are used loosely by most mathematicians; these terms may be applied to any argument that builds something complicated from seemingly simpler things. However, the terms "constructive" and "constructivist" are used more narrowly. An existence proof is **constructive** if the proof actually *finds* the object in question by a procedure involving just finitely many steps — or, in some cases, if the proof *approximates* the object arbitrarily closely by a procedure involving just countably many steps. **Constructivists** are mathematicians who study such proofs and/or who prefer such proofs; **constructivism** is the study of such proofs. To be more specific, we might call this *constructivism in the sense of Errett Bishop.* Actually, the literature now contains many schools of constructivism that differ slightly from Bishop's view. The survey given in the next few pages is too superficial to distinguish between these different schools; Bridges and Richman [1987] give a much more detailed survey. (Bishop's constructibility should not be confused with *Gödel constructibility,* indicated in 5.54, which permits uncountably many steps and is very different in nature.)

Most mathematicians are part-time informal constructivists, in this respect: In teaching or learning mathematics, we try to follow any abstract idea with one or more concrete examples. Indeed, new teachers of mathematics probably hear that pedagogical practice recommended more often than any other. We follow that pedagogical practice — i.e., of giving examples — whenever possible, but we must depart from that practice at times, for some mathematical ideas are inherently nonconstructive. Indeed, some of the objects studied in this book are **intangible**: We shall see that the objects "exist," but that explicitly constructible examples of these objects do *not* exist. By this we do not mean merely that no examples have been found yet; rather, we mean that *it can be proven that no explicit examples can ever be given.* We shall see that this peculiar status is shared by free ultrafilters, nontrivial universal nets, subsets of $\mathbb{R}$ that lack the Baire property, well orderings of $\mathbb{R}$, finitely additive probabilities that are not countably additive, certain kinds of linear maps, and diverse other objects. Although intangibles can be avoided in applied mathematics, they are conceptually useful in pure mathematics and appear frequently in the literature (usually without much explanation). The lack of examples may be disconcerting to students. We shall give some explanation of the lack of examples, here and later; see especially 14.77.

6.3. *Some examples of nonconstructive mathematics.* Two of the axioms of conventional set theory are nonconstructive. The Axioms of Regularity and Choice *postulate* the existence of certain sets without giving any indication of how to find those sets. The Axiom of Regularity (introduced in 1.47) is largely a formality, included in set theory for convenience; it has little effect on mathematics outside of set theory. Indeed, for most purposes it can be replaced by the *Principle of* $\in$*-Induction* (in 1.50), which is in some sense constructive — or, perhaps more precisely, it is not nonconstructive. In contrast, the Axiom of Choice (introduced later in this chapter) has enormous effects on many branches of mathematics, and cannot be replaced so easily with a constructive variant.

Different mathematicians have different interpretations for the term "constructive," and attach different degrees of importance to that notion as well. It is ironic that Baire, Borel, and Lebesgue, three of the founders of this century's analysis, were philosophically opposed to any uses of arbitrary choices, and yet Countable Choice — a mildly nonconstructive principle involving a sequence of arbitrary choices (introduced in 6.25) — was crucial to their work. They used it without noticing it; only later was this use pointed out explicitly by Sierpiński. See Moore [1983].

6.4. The Axiom of Choice is a highly visible form of nonconstructive reasoning. Some mathematicians are not aware of other kinds of nonconstructive reasoning, and consequently they use the term "constructive" simply to mean "not using the Axiom of Choice." However, that is an erroneous usage. There are other kinds of nonconstructive existence proofs, one of which we shall now describe.

Proof by contradiction was introduced in 1.9; it can be stated as $\neg\neg P \Rightarrow P$, where $\neg$ means "not." In some constructive frameworks (e.g., in intuitionist logic — see 14.35), the principle of proof by contradiction is equivalent to the **Law of the Excluded Middle**:

For every proposition P, either P holds or not-P holds

(or more briefly, $P \vee \neg P$). For example, we don't yet know whether **Goldbach's Conjecture**[1] is true or false, but most mathematicians would agree that surely it is one or the other. Thus, most mathematicians are in wholehearted agreement with the Law of the Excluded Middle and might have trouble seeing how the constructivists could reject it.

However, formal constructivists use language a bit differently from mainstream mathematicians. For constructivists, Goldbach's Conjecture is not yet true or false, although someday it may attain one of those states. Interpreted in the language of constructivists, the expression "P or Q" means "we have a constructive proof of P, or we have a constructive proof of Q, or both." With this convention, the Law of the Excluded Middle becomes:

> For every proposition P, either P holds or not-P holds, *and we can determine which one.*

Of course, with this interpretation, the Law of the Excluded Middle is blatantly false; constructivists and mainstream mathematicians agree on that. But why do constructivists use language in this fashion? The example below may help us to understand why; another explanation will be given in 14.36.

6.5. *An example with irrationals.* The following example is taken from Troelstra and Dalen [1988]. We shall prove the following proposition.

(P) There exist positive, irrational numbers a and b such that a^b is rational.

A quick, easy, nonconstructive proof is as follows: Either

(i) ${\sqrt{2}}^{\sqrt{2}}$ is rational — then take $a = b = \sqrt{2}$; or

(ii) ${\sqrt{2}}^{\sqrt{2}}$ is irrational — then take $a = {\sqrt{2}}^{\sqrt{2}}$ and $b = \sqrt{2}$.

---

[1] In 1742, Goldbach conjectured that every even integer greater than 2 can be written as the sum of two prime numbers. This is one of the most famous unsolved problems of mathematics: As of the time of this writing, no one has yet proved or disproved Goldbach's Conjecture, though many mathematicians have spent much time trying and have proved slightly weakened versions of the conjecture. Goldbach's Conjecture was part of Problem 8 in Hilbert's famous list of 23 problems for the twentieth century. See Yuan [1984] for a survey of Goldbach's Conjecture.

For the purposes of this book, Goldbach's Conjecture is of interest not because of what it would tell us about prime numbers, but rather because it is a simple example of an unsolved problem that could be solved if we could carry out a countable infinity of steps. Any other unsolved problem that can be solved in that fashion will do as well for the discussions in this section and in 10.46 and 15.48. If Goldbach's Conjecture gets proved or disproved between the time this book is written and the time this book is read, simply replace it with some other such problem. (An earlier draft of this book used **Fermat's Last Theorem**, a more famous problem that went unsolved for 300 years. However, shortly before this book was finished, a proof of Fermat's Last Theorem was finally completed by Taylor and Wiles [1995].)

However, this proof does not tell us *which* of the two possibilities (i) or (ii) is valid, so we have not found a particular explicit example of a pair (a, b) satisfying (P). This proof could not be used as a subroutine in a numerical computer program: It yields not one answer, but two possible answers with no method of choosing between them.

Actually, (ii) is true and can be proved constructively, by a much longer argument. By a theorem of Gelfond and Schneider, if a and b are positive algebraic numbers, $a \neq 1$, and b is irrational, then a^b is transcendental. See page 106 of Gelfond [1960]; related results are surveyed by Tijdeman [1976].

Further Comments on Constructivism

6.6. *Making mathematics constructive.* Nonconstructive arguments often can be replaced by constructive ones. Sometimes this is only with difficulty, as in the preceding example with $\sqrt{2}^{\sqrt{2}}$. Sometimes it is easier. For instance, the **Trichotomy Law for Real Numbers**:

$$\text{for all real numbers } x \text{ and } y, \text{ either } x < y \text{ or } x = y \text{ or } x > y$$

is not constructively provable; there is no algorithm that takes constructive descriptions of x and y and yields the assertion of one of those three relations. We shall illustrate and demonstrate this unprovability in two ways in 14.9 and 10.46. However, Bishop [1973/1985] points out that in most applications, the Trichotomy Law is not needed in its full strength; it can be replaced by the following weaker law.

> **Comparison Law.** For any real numbers u, v, and y, if $u < v$ then at least one of $u < y$ or $y < v$ must hold.

This law *is* constructively provable.

The alterations one makes while translating classical mathematics to constructive mathematics generally have little or no effect on the ultimate applications. For instance, one of the fundamental theorems of classical functional analysis is the *Hahn-Banach Theorem*; we shall study several versions of this theorem in later chapters. Some versions assert the existence of a certain type of linear functional on a normed space X. The theorem is inherently nonconstructive, but a constructive proof can be given for a variant involving normed spaces X that are *separable* — i.e., normed spaces that have a countable dense subset; see Bridges [1979]. Little is lost in restricting one's attention to separable spaces, for in applied math most or all normed spaces of interest are separable. The constructive version of the Hahn-Banach Theorem is more complicated, but it has the advantage that it actually *finds* the linear functional in question.

6.7. Constructivists and mainstream mathematicians use the same words in different ways; in fact, different schools of constructivists use the same words in different ways.

A basic example is in the meaning of "real number." Mainstream mathematicians have several different equivalent definitions of real numbers (see Chapter 10). One way to define a real number is as an equivalence class of Cauchy sequences of rational numbers (see 19.33.c). But constructivists prefer to indicate a real number by a Cauchy sequence that is *accompanied by some estimate of the rate of convergence* — e.g., a sequence (r_n) of rational numbers that satisfies $|r_m - r_n| \leq \max\{\frac{1}{m}, \frac{1}{n}\}$. Of course, in mainstream mathematics, every real number can be represented as the limit of such a sequence, but such sequences are not essential to our way of thinking about real numbers. In constructivist mathematics, all computations about real numbers are expressed, either directly or indirectly, in terms of such sequences. (Constructivist "real numbers" are discussed further in 10.46.)

Here is a more complicated example of the differences in language:

In constructive analysis, the continuous functions that are of chief interest are the uniformly continuous ones. Indeed, it is hard to constructively establish that a function is continuous except by giving a modulus of uniform continuity — and thus establishing that the function is indeed uniformly continuous. Of course, in mainstream mathematics, any continuous function on a compact interval is uniformly continuous, but that fact is not provable in constructive mathematics.

In the terminology of Bishop and Bridges [1985], a function on a compact interval is *continuous* if it has a modulus of uniform continuity — i.e., that book's definition of "continuity" is the usual definition of uniform continuity, but the context is one where the two notions are classically equivalent anyway.

Among some constructivists, the only functions that can really be called "functions" are the *representable* ones. Moreover, it is a theorem (in certain axiom systems of constructive mathematics) that every representable function is continuous. Thus, under certain uses of the language, the following is true:

Ceitin's Theorem. Every function is continuous.

A proof of this startling result can be found on page 69 of Bridges and Richman [1987]. The result is slightly less startling when we consider that, even in mainstream mathematics, any function with certain "good" properties is continuous; theorems to this effect are given in 24.42, 27.28.c, and 27.45.

The introduction to constructivism given by Bridges and Mines [1984] also discusses the importance of language.

6.8. *Constructivism versus mainstream mathematics.* This book, which is intended to introduce the reader to the literature, is frequently nonconstructive, since much of the literature is nonconstructive. Indeed, the constructivist viewpoint is foreign to most mathematicians today; we are so used to nonconstructive proofs that we tend to believe one cannot do much interesting mathematics constructively. And, until a few decades ago, we would have been right. Brouwer's intuitionism was more a matter of philosophy than mathematics, and Heyting extended the matter from philosophy to formal logic. But then, finally, Bishop [1967] showed how to develop a large portion of analysis constructively. (See also the revised version, Bishop and Bridges [1985].) Since then, several other mathematicians have extended Bishop's style of reasoning and written constructive versions of many other parts of mathematics. In particular, the reader may refer to Bridges [1979] for functional analysis, to Beeson [1985] for foundations (i.e., logic and set theory), and to Bridges and Richman [1987] for a recent survey of the several different schools of constructivism.

Despite its growing literature, constructivism remains separated from the mainstream of mathematics. This may be largely because constructivism's finer distinctions necessitate a use of language quite different from, and more complicated than, that of the mainstream mathematician. For instance, among some constructive analysts, $x \neq y$ simply means the negation of $x = y$, while $x \# y$ means[2] the slightly stronger condition of **apartness**: We can find a positive lower bound for the distance between approximations to x and y. Thus, constructivists distinguish between notions that the classical mathematician is accustomed to viewing as identical. Consequently, a mainstream mathematician can only learn constructivism by relearning his or her entire language — a sizable undertaking.

Some philosophical questions deserve at least a brief mention here, although we shall not address them in any depth. Bishop [1973/1985] suggested that mainstream mathematicians, in pursuit of form, have lost track of content; Bishop exhorted mathematicians to return to a more meaningful mathematics. Perhaps the contentless mathematics that he condemned would include the intangibles studied elsewhere in this book (free ultrafilters, etc.), which lack examples and do not seem to be a direct reflection of anything in the "real world." However, an argument can be made for the conceptual usefulness of such objects. For instance, free ultrafilters provide a basis for nonstandard analysis, which yields new insights into calculus and other limit arguments. Moreover, we may be surprised by just what kinds of mathematics can reflect the real world; for instance, Augenstein [1994] suggests that the Banach-Tarski Decomposition may be a useful model of some interactions of subatomic particles.

Both constructive and nonconstructive thinking have their advantages. A constructive proof may be more informative (e.g., it tells us that $\sqrt{2}^{\sqrt{2}}$ is irrational — see 6.5), but a nonconstructive proof is often quicker and simpler. Extending a metaphor of Urabe: To feed one's family, it is not enough to prove that a certain pond contains a fish; ultimately one must *catch* the fish. On the other hand, it would be helpful to have an inexpensive device that quickly and easily determines which ponds contain fish.

6.9. Much of this book is concerned with nonconstructive mathematics. Moreover, to better understand some of the nonconstructible objects studied in this book, we shall sometimes find it helpful to vary the amount and kind of nonconstructiveness that we are willing to accept. In particular, we may compare results requiring the Axiom of Choice with results that only require a weakened form of the Axiom of Choice. At first glance, that looks like a rather strange notion; after all, either we can find a certain mathematical object, or we can't. How we can say that one object is *harder* to find than another object, when in fact we can't find *either* of them?

The metaphor of "oracles" was introduced in recursion theory by Turing [1939] (see the discussion by Enderton [1977 recursion theory]); a similar metaphor may be helpful in the present context. Imagine we have access to an **oracle**, who has frequent conversations with some deity. We present the oracle with various questions that we have been unable to answer by merely mortal, human methods. The oracle is able and willing to answer *some*, but not all, of these questions. For instance, the oracle might tell us whether Goldbach's

[2]Caution: Some constructive analysts use $x \neq y$ to denote apartness and use $\neg(x = y)$ to denote inequality.

conjecture is true, but refuse to comment on the Riemann Hypothesis. In some of the literature, such an oracle is referred to as a "limited principle of omniscience."

Now, in some cases, if the oracle gives us an answer to question A, we may use that information to deduce an answer to question B — even if the oracle has not given us an answer to B. Thus, one answer may be *stronger* than another. Similarly, two answers may be considered *equivalent* to each other if each is stronger than the other — i.e., if either answer would enable us to deduce the other.

It must be emphasized that when we use the oracle's answer to A to deduce an answer to B, then we are using human, mortal reasoning — i.e., the oracle is not helping in such deductions. Thus, our *relation* — of one answer being stronger than another — is determined without the aid of the oracle; this relation does not depend on our actually having answers to *either* of the questions A or B. It is *these relations between the answers*, not the actual answers themselves, that will concern us later, when we compare different levels of nonconstructiveness. Since the oracle is not actually used to determine and compare those different levels, we may now dispense with the oracle altogether.

6.10. *Proposition.* The Axiom of Regularity implies the Law of the Excluded Middle, if interpreted in the language of constructivism. (Hence the Axiom of Regularity is nonconstructive.)

Proof. The following proof is modified from Beeson [1985]. Since most readers of this book probably are not familiar with constructivist language, we shall restate the proof in terms of the oracle metaphor of 6.9.

Interpreted in constructivist terms, the Axiom of Regularity says that we have an oracle of the following type:

> We may describe to the oracle some nonempty set S, in terms that do not necessarily give a clear understanding of the set but that do at least uniquely determine the set. Then the oracle will specify to us some element $x \in S$ such that $x \cap S = \varnothing$.

Let P be a proposition (such as Goldbach's conjecture) that we can state precisely, but that we do not necessarily know to be true or false. Now define

$$S = \begin{cases} \{\varnothing, \{\varnothing\}\} & \text{if } P \text{ is true} \\ \{\{\varnothing\}\} & \text{if } P \text{ is false.} \end{cases}$$

Then S is nonempty, since $\{\varnothing\} \in S$. The oracle will tell us either "$\varnothing$ is a member of S that does not meet S" — in which case P is obviously true — or "$\{\varnothing\}$ is a member of S that does not meet S" — in which case we can deduce that P is false. Thus, the oracle can be used to deduce the truth or falsehood of *any* proposition P.

Remark. The Axiom of Choice, if interpreted in constructivist terms, can also be shown to imply the Law of the Excluded Middle. The proof of this implication, though short, depends on a deeper understanding of constructivist language; it does not translate readily into the language of mainstream mathematicians. We omit it here; it is given by Beeson [1985] and Bridges and Richman [1987].

6.11. Constructivism (in the sense of Errett Bishop) will be discussed further in 6.13, 10.46, and 15.48.

Logicians have another notion that is similar to constructibility. An object x_0 is said to be *definable* if there exists a proposition $P(x)$ in first-order logic for which $x = x_0$ is the unique element for which $P(x)$ is true. See Lévy [1965].

Constructibility in the sense of Bishop, constructibility in the sense of Gödel, and definability in the sense of Lévy are far outside the mainstream of thinking of most analysts. In 14.76 we shall introduce "quasiconstructibility," which is (in this author's opinion) closer to the way that most analysts think.

THE MEANING OF CHOICE

6.12. Conventional set theory is Zermelo-Fraenkel set theory plus the Axiom of Choice, abbreviated ZF + AC. We described Zermelo-Fraenkel set theory in 1.47.

The **Axiom of Choice** has many equivalent forms; we shall study several in this and later chapters. (A much longer list of equivalents is given by Rubin and Rubin [1985].) We shall denote our equivalents of Choice by (AC1), (AC2), (AC3), (AC4), etc.; collectively we shall refer to them as AC. Most of these equivalents are discussed in next few pages. A few more equivalents are the *Vector Basis Theorem* in 11.29, and *Tychonov's Theorem* and similar results on product topologies in 15.29, 17.16, and 19.13.

Here are three of the simplest forms of Choice:

> **(AC1) Choice Function for Subsets.** Let X be a nonempty set. Then for each nonempty subset $S \subseteq X$ it is possible to *choose* some element $s \in S$. That is, there exists a function f that assigns to each nonempty set $S \subseteq X$ some representative element $f(S) \in S$.
>
> **(AC2) Set of Representatives.** Let $\{X_\lambda : \lambda \in \Lambda\}$ be a nonempty set of nonempty sets that are pairwise disjoint. Then there exists a set C containing exactly one element from each X_λ.
>
> **(AC3) Nonempty Products.** If $\{X_\lambda : \lambda \in \Lambda\}$ is a nonempty set of nonempty sets, then the Cartesian product $\prod_{\lambda \in \Lambda} X_\lambda$ is nonempty. That is, there exists a function $f : \Lambda \to \bigcup_{\lambda \in \Lambda} X_\lambda$ satisfying $f(\lambda) \in X_\lambda$ for each λ.

A function f that specifies choices, in this or similar contexts, is called a **choice function**. We postpone until 6.19 the proof of equivalence of these three principles.

The Axiom of Choice is "obviously true," in that it agrees with the intuition of most mathematicians. For instance, consider (AC1). Each nonempty set $S \subseteq X$ certainly contains *some* element s, and thus to define $f(S)$ it suffices to "just pick any such s." It requires only a small stretch of the imagination to make all such choices simultaneously and thus to define the function f.

In fact, AC is so much a part of the way of thinking of most mathematicians that it can easily sneak into a proof unnoticed; then we say AC is used *unconsciously* or *implicitly*.

Cantor and other mathematicians used Choice implicitly in their early work in set theory in the late 19th century; only in 1908 did Zermelo become aware of this assumption in their work and explicitly formulate it as an axiom. (See Moore [1982].)

AC is so very "obviously true" that the reader may wonder why it is considered to be an axiom, rather than just a consequence of definitions. To see why, let us consider the choice function f in (AC1). If a choice function can be given *explicitly*, by some completely describable procedure or rule, then the choices are said to be **canonical**; otherwise the choices are said to be **arbitrary**. Canonical choice functions are sometimes available. For instance, if $X = \mathbb{N} = \{1, 2, 3, \ldots\}$, then we can satisfy (AC1) by taking $f(S)$ to be the smallest (i.e., first) element of S.

However, in other cases we *cannot* find f explicitly. For instance, one specialization of AC is this principle:

> **(ACR) Axiom of Choice for the Reals.** There exists a function f that assigns to each nonempty set $S \subseteq \mathbb{R}$ some element $f(S) \in S$.

Such a function is, for analysts, perhaps the simplest **intangible** — i.e., the simplest instance of an object that exists but that we cannot illustrate with a specific example. The reader is urged to try for a moment to think of an explicit choice function f for $\mathbb{R}$. Some partial solutions suggest themselves — e.g., when S is a bounded nonempty interval, then let $f(S)$ be the midpoint of that interval. More complicated answers will choose points from larger collections of sets; for instance, let (r_n) be an enumeration of the rationals; then every set S with nonempty interior contains some rational number and so we may take $f(S)$ to be the *first* rational number in S. But what about a choice function that works for *all* nonempty subsets of $\mathbb{R}$? No explicit choice function has ever been found[3] for $\mathbb{R}$, and, in fact, it can be proved that no explicit choice function ever *will* be found for $\mathbb{R}$ (see 14.77 and 6.34). Hence we need to assume a principle such as AC or ACR to tell us that such a function f exists.

The Axiom of Choice makes selections for us that we do not know how to make for ourselves; Sah [1990] calls it a "mathematicians' Maxwell demon." Bertrand Russell illustrated it with this example:

> To select one sock from each of infinitely many pairs of socks requires the Axiom of Choice; but for shoes the Axiom is not needed.

For instance, one way to choose shoes *canonically* would be to take all the left shoes. (Actually, socks do not require the full strength of the Axiom of Choice; we can choose the socks using a slightly weakened form of Choice discussed in 6.15.)

In situations where we need the Axiom of Choice, usually there are *infinitely many* choices available. However, we cannot establish the existence of infinitely many choices, or even the existence of one choice, except by giving an example or applying some nonconstructive principle such as AC. This is discussed further in 14.77.

---

[3] Logicians and set theorists may have a slightly different view of this matter, for they are more comfortable with the Axiom of Constructibility $V = L$ (introduced in 5.54). Without assuming $V = L$ or the Axiom of Choice or any of its relatives such as DC, CC, UF, etc., we can write down a formula φ that has the following property: When we assume $V = L$, then φ becomes a well ordering of $\mathbb{R}$. Thus, φ provides an "explicit example" of a well ordering of $\mathbb{R}$ and hence an "explicit example" of a choice function for $\mathbb{R}$.

6.13. *A defective "proof" of Choice.* The reader may find it instructive to consider the following "proof" of (AC3). If each X_λ contains exactly one element x_λ, then we can conclude $\prod_{\lambda\in\Lambda} X_\lambda$ is nonempty without using any existential axiom: We know that $\prod_{\lambda\in\Lambda} X_\lambda$ contains the function x that assigns to each coordinate λ the value x_λ. If we make all the X_λ's larger, then this could only make $\prod_{\lambda\in\Lambda} X_\lambda$ larger, and so it would still be nonempty; thus the Axiom of Choice is "proved."

The flaw in this reasoning is a subtle one: By what method do we "make all the X_λ's larger?" If the enlarged set X_λ still contains the original member x_λ, and if we still know which of the elements of X_λ is that original member x_λ, then indeed the original function x would still be available to us; we can make a canonical choice. But if we lose track of the x_λ's when we enlarge the X_λ's and all we know about the enlarged sets is that they are nonempty, then we no longer have an explicit formula or rule for choosing one element from each X_λ. We have no way to "construct" the function x except via some additional assumption such as (AC3).

Of course, intuitively it is obvious that $\prod_{\lambda\in\Lambda} X_\lambda$ is nonempty — in fact, in most cases of interest, $\prod_{\lambda\in\Lambda} X_\lambda$ contains infinitely many elements. But we may be unable to find any particular element of $\prod_{\lambda\in\Lambda} X_\lambda$.

Variants and Consequences of Choice

6.14. To understand better what the Axiom of Choice is really assuming, let us contrast AC with several related principles, some of which are much weaker. The weakest of these is:

> **Finite "Axiom" of Choice.** If n is a positive integer and S_1, S_2, ..., S_n are nonempty sets, then $S_1 \times S_2 \times \ldots \times S_n$ is nonempty.

Although this principle is sometimes called an axiom, it really is not an axiom, for it follows "for free" from conventional logic and the axioms of ZF set theory without any additional assumptions. Indeed, ordinary mathematical logic permits us to apply an operation finitely many times. Each S_i is nonempty, hence it contains (and we can choose from it) some element s_i. Repeat this operation n times; then $(s_1, s_2, \ldots, s_n)$ is a member of the product.

We only need the Axiom of Choice, or some form of the Axiom of Choice, when we need to make *infinitely many* arbitrary choices.

6.15. Following are two variants of Choice that do not follow "for free" from logic and ZF set theory:

> **(ACF) Axiom of Choice for Finite Sets.** Let $\mathcal{C}$ be a set whose members are nonempty *finite* sets. Then it is possible to choose some member s from each set $S \in \mathcal{C}$.

> **(MC) Multiple Choice Axiom.** Let $\mathcal{C}$ be a set whose members are nonempty

sets. Then it is possible to choose some nonempty finite subset F from each set $S \in \mathcal{C}$.

The Axiom of Choice for Finite Sets must not be confused with the *Finite Axiom of Choice*, although the names are similar.

It can be shown that ACF implies the Law of the Excluded Middle (introduced in 6.4). A short proof of this implication is given by Goodman and Myhill [1978], but we shall not reproduce it here because it does not translate readily into the language of mainstream mathematics (see 6.8).

ACF, though weaker than the Axiom of Choice, is still strong enough to act as a mathematical "Maxwell's demon," making choices for us that we cannot make for ourselves. For instance, ACF is strong enough to choose Bertrand Russell's socks (see 6.12).

Obviously AC is equivalent to ACF + MC. Actually, it can be proved that the Multiple Choice Axiom by itself is equivalent to the Axiom of Choice. However, the proof is too long to include here; it can be found in Jech [1973] or in Rubin and Rubin [1985]. The proof requires the Axiom of Regularity, unlike most other proofs mentioned in this book.

6.16. *Pathological consequences of AC.* The Axiom of Choice is a nonconstructive assertion of existence: It postulates the existence of certain objects without giving any indication of how to find those objects. There may not even *be* a way to find those objects. We shall see in 14.77 that many of the objects generated by AC are intangibles — i.e., objects for which no constructive examples can *ever* be given.

Moreover, some of the intangible consequences of the Axiom of Choice are **pathological** — i.e., they are so very different from familiar, constructible examples that they are contrary to our intuition. Perhaps the most dramatic of these pathologies is:

Banach-Tarski Decomposition. The closed unit ball in three dimensions,

$$B \;=\; \{(x,y,z) \in \mathbb{R}^3 \;:\; x^2 + y^2 + z^2 \le 1\},$$

can be partitioned into finitely many pieces, which can be rearranged by rigid motions (i.e., rotations and translations) and recombined to form *two* closed unit balls, each identical to the original ball B.

At first glance, the Banach-Tarski Decomposition seems preposterous. It blatantly contradicts our intuition about the conservation of mass or volume. In fact, the theorem above is often called the "Banach-Tarski Paradox." In mathematics, the term "paradox" usually refers to an impossibility.

However, the Banach-Tarski Decomposition only appears impossible at first; its paradoxical appearance can be explained away. The ordinary "volume" of a subset of $\mathbb{R}^n$ is its n-dimensional Lebesgue measure. That number is defined if the set is Lebesgue measurable, but — as we shall see later in this book — not all subsets of $\mathbb{R}^n$ are Lebesgue measurable. In particular, the pieces in the Banach-Tarski Decomposition are not Lebesgue measurable.

Thus, the Banach-Tarski Decomposition does not actually violate any rules concerning volume. It simply tells us that, if we accept the Axiom of Choice, then the rules of volume are more complicated than we might like. The intuition about volumes that we have obtained from our experience with everyday macroscopic objects in the real, physical world is only applicable to some, not all, subsets of the mathematical world $\mathbb{R}^3$. (In fact, that intuition is not even applicable to submicroscopic objects in the real world; Augenstein [1994] suggests that the Banach-Tarski Decomposition is a possible model for some kinds of interactions of subatomic particles.) Most mathematicians have learned to live with such pathological consequences of the Axiom of Choice, feeling that the pathological consequences are outweighed by the advantages of AC.

We shall not prove the Banach-Tarski Decomposition Theorem in this book — a proof and much related material are given by Wagon [1985] — but in 21.22 we shall give Vitali's classical short proof of the existence of a Lebesgue nonmeasurable set. The existence of Lebesgue nonmeasurable sets is the chief reason that measure theory is generally developed for an algebra or σ-algebra $\mathcal{S}$ of subsets of a set X, rather than in the simpler setting of $\mathcal{P}(X)$.

Actually, the Banach-Tarski Decomposition does not require the full strength of AC. A recent proof of Pawlikowski [1991] shows that the Banach-Tarski Theorem is implied by the *Hahn-Banach Theorem*, a weakened form of Choice that will be studied extensively later in this book. We shall not give Pawlikowski's proof, but a crucial ingredient of that proof is Luxemberg's Boolean reformulation of the Hahn-Banach Theorem, which we shall prove in 23.19.

6.17. *What makes AC true or false?* When we accept the Axiom of Choice, we declare that we intend to treat certain mathematical objects as if they exist, regardless of whether we can find examples of those objects. This implies a particular interpretation of some words, such as "choose," "exist," "set," and "function."

Constructivist mathematics and classical (mainstream) mathematics give us two intuitive interpretations of language, which make the Axiom of Choice either false or true. Axiomatic set theory takes a more rigorous approach that does not rely on intuitive interpretations. Axiomatic set theory divests symbols and words such as $\in$, $\subseteq$, and "set" of their usual meanings and investigates how certain relations between the meaningless symbols and words imply certain other relations. In axiomatic set theory, one is not concerned with "true" or "false" (because ultimately these things are unknowable), but only with "implies." With this viewpoint, AC is simply another axiom that we may accept or reject.

Alternative axiom systems are also possible, and some of them are just as consistent as conventional set theory. Although we shall use AC freely throughout most of this book, in a few brief discussions we shall also consider some of its alternatives, which have important consequences in functional analysis. Even if the reader is a "firm believer" in the Axiom of Choice, nevertheless there are strong reasons for considering its alternatives: Such considerations will improve our understanding of the consequences of Choice. Some of the alternatives to Choice, though not compatible with AC itself, are at least compatible with weakened forms of AC; thus we shall also study such weakened forms.

Some Equivalents of Choice

6.18. To clarify the role played by Choice, we shall keep track of its uses in some parts of this book. A proof is **effective** if it does not use AC or consequences of AC except as explicitly stated hypotheses. Two statements are **equivalent** (or *effectively equivalent*) if each can be proved effectively from the other. The oracle metaphor in 6.9 may be helpful in understanding these "effective" proofs.

6.19. The preceding discussions may have made clear just what the rules are that govern proofs of equivalence. The reader may now try to prove the equivalence of (AC1), (AC2), and (AC3).

Hint for (AC2) $\Rightarrow$ (AC1): See 1.10. Relabel copies of the subsets of X so that they are all disjoint. For instance, $S \times \{S\}$ is a bijective copy of S, since $\{S\}$ is a singleton and the sets $S \times \{S\}$ are disjoint subsets of $X \times \mathcal{P}(X)$.

6.20. *Maximal principles.* The following statements are equivalent to the Axiom of Choice.

> **(AC4) Well Ordering Principle (Zermelo).** Every set can be well ordered.
>
> **(AC5) Finite Character Principle (Tukey, Teichmuller).** Let X be a set, and let $\mathcal{F}$ be a collection of subsets of X; suppose that $\mathcal{F}$ has finite character (as defined in 3.46). Then any member of $\mathcal{F}$ is a subset of some $\subseteq$-maximal member of $\mathcal{F}$.
>
> **(AC6) Maximal Chain Principle (Hausdorff).** Let $(X, \preccurlyeq)$ be a poset. Then any $\preccurlyeq$-chain in X is included in a $\subseteq$-maximal $\preccurlyeq$-chain.
>
> **(AC7) Zorn's Lemma (Hausdorff, Kuratowski, Zorn, others).** Let $(X, \preccurlyeq)$ be a poset. Assume every $\preccurlyeq$-chain in X has a $\preccurlyeq$-upper bound in X. Then X has a $\preccurlyeq$-maximal element.
>
> **(AC8) Weakened Zorn Lemma.** Let $(X, \preccurlyeq)$ be a poset. Assume every subset of X that is directed by $\preccurlyeq$ has a $\preccurlyeq$-upper bound in X. Then X has a $\preccurlyeq$-maximal element.

Hint for (AC1) $\Rightarrow$ (AC4): Use 3.43, with $\gamma(S) = f(X \backslash S)$.

Hint for (AC4) $\Rightarrow$ (AC5): Use the theorem in 3.46.

Hint for (AC5) $\Rightarrow$ (AC6): The $\preccurlyeq$-chains form a collection of finite character.

Hint for (AC6) $\Rightarrow$ (AC7): Use the upper bound of the maximal chain.

Hint for (AC7) $\Rightarrow$ (AC8): Any chain is a directed set.

Proof of (AC8) $\Rightarrow$ (AC1). Let Ω be a nonempty set. By a "partial choice function" for Ω we shall mean a function f whose domain is some collection of nonempty subsets of

Ω, satisfying $f(S) \in S$ for each $S \in \mathrm{Dom}(f)$. Let X be the collection of partial choice functions; partially order X by taking $f \preccurlyeq g$ if $\mathrm{Graph}(f) \subseteq \mathrm{Graph}(g)$. It follows from the Finite Axiom of Choice (see 6.14) that X is nonempty and that any maximal element of X must be a function f with domain $\mathcal{P}(\Omega) \setminus \{\varnothing\}$; thus it suffices to show that X has a maximal element. Verify that the hypotheses of (AC8) are satisfied.

6.21. We ask again: Is the Axiom of Choice "true?" According to Bona [1977],

> The Axiom of Choice is obviously true; the Well Ordering Principle is obviously false; and who can tell about Zorn's Lemma?

The joke is that the three principles are *equivalent*, as we have just seen. Still, there is a point to the jest: Our intuition isn't reliable here.

In fact, Bona's aphorism does agree with most mathematicians' intuition. The Axiom of Choice seems true, because the Axiom of Choice is worded in such a way that the *simultaneity* of the choices has little psychological impact. Indeed, as we noted in 6.12, (AC1) seems true because we can "just pick any $s \in S$." In contrast, the Well Ordering Principle seems false, since the simultaneity of choices is built into the well ordering. Well orderings are quite difficult to find. Indeed, a partial ordering chosen "at random" generally is not a well ordering, and we are altogether unable to find an explicit well ordering for $\mathbb{R}$. Finally, Zorn's Lemma is too complicated to seem "obviously true" or "obviously false" to most mathematicians, although of course those who use it repeatedly become accustomed to it and begin to think of it as "true."

6.22. *Choice and cardinality.* For this section, let $|A|$ denote the cardinality of a set A.

The Axiom of Choice and its equivalents deal with infinite sets. We understand finite sets fairly well, but it is difficult to extrapolate from finite sets and describe how infinite sets should behave; several different descriptions seem equally plausible. Our first few results about cardinality — the Schröder-Bernstein Theorem, Cantor's Theorem, etc. — did not depend on the Axiom of Choice and may have given the impression that every set has some definite "size," definable in some absolute way. Thus it may be surprising that some basic properties of cardinality are actually equivalent to the Axiom of Choice.

> **(AC9) Well Ordering of Cardinals.** Comparison of cardinalities is a well ordering. That is, if $\mathcal{S}$ is a set whose elements are sets, then there is some $S_0 \in \mathcal{S}$ that satisfies $|S_0| \leq |T|$ for all $T \in \mathcal{S}$.

> **(AC10) Trichotomy of Cardinals.** Comparison of cardinalities is a chain ordering. That is, for any two sets S and T, precisely one of these three conditions holds: $|S| < |T|$; $|S| = |T|$; $|S| > |T|$.

> **(AC11) Comparability of the Hartogs Number.** If $H(S)$ is the Hartogs number of a set S, then $|H(S)|$ and $|S|$ are comparable — i.e., one is bigger than or equal to the other (and hence $|H(S)| > |S|$).

> **(AC12) Squaring of Cardinals.** If X is an infinite set, then $|X \times X| = |X|$.

(AC13) Multiplication of Cardinals. If X is an infinite set, Y is a nonempty set, and $|X| \geq |Y|$, then $|X \times Y| = |X|$.

(AC14) If X and Y are disjoint sets, $|X| \geq |\mathbb{N}|$, and Y is nonempty, then $|X \cup Y| = |X \times Y|$.

(AC15) If X is an infinite set, then the cardinality of X is equal to the cardinality of $\bigcup_{n=1}^{\infty} X^n = \{\text{finite sequences in } X\}$.

Proofs. We shall first prove that (AC4), (AC9), (AC10), and (AC11) are equivalent. For a proof of (AC4) $\Rightarrow$ (AC9), use 5.46.g. The implication (AC9) $\Rightarrow$ (AC10) is obvious. The proof of (AC10) $\Rightarrow$ (AC11) is immediate from the definition of the Hartogs number. Finally, for a proof of (AC11) $\Rightarrow$ (AC4), note that we have an injection from any set X into a well ordered set $H(X)$; use 3.39.b.

Next we shall prove that (AC12), (AC13), and (AC15) are equivalent. For a proof of (AC12) $\Rightarrow$ (AC13), use relabeling; thus we may assume X and Y are disjoint. Choose any $y_0 \in Y$. Then

$$|X| = |X \times \{y_0\}| \leq |X \times Y| \leq |X \times X| = |X|.$$

For a proof of (AC13) $\Rightarrow$ (AC15), show by induction that $|X^n| = |X|$ for all positive integers n. Hence $\bigcup_{n=1}^{\infty} X^n$ has the same cardinality as the union of $\mathbb{N}$ disjoint copies of X — i.e., the same cardinality as $X \times \mathbb{N}$. Finally, (AC15) $\Rightarrow$ (AC12) is obvious.

A proof of (AC4) $\Rightarrow$ (AC12) is immediate from 3.45.

To prove (AC13) $\Rightarrow$ (AC14), pick any $y_0 \in Y$ and any object v that is *not* in X. Then $|X| = |X \cup \{v\}|$ by 2.20.g. Hence

$$\begin{array}{rcl} |X \cup Y| & = & |(X \times \{y_0\}) \cup (\{v\} \times Y)| \\ & \leq & |(X \cup \{v\}) \times Y| \quad = \quad |X \times Y| \quad \leq \quad |(X \cup Y) \times Y| \quad = \quad |X \cup Y|, \end{array}$$

where the last equation follows from (AC13) since $|X \cup Y| \geq |Y|$.

Finally, we shall show that (AC14) $\Rightarrow$ (AC11); this proof takes a bit longer but it will complete our cycle of equivalences. Let any set S be given; let $H = H(S)$ be its Hartogs number; we wish to show that $|H|$ and $|S|$ are comparable. We may assume S is not finite. It follows easily that H is infinite also. Since H is an infinite ordinal, we have $|H| \geq |\mathbb{N}|$ by 3.42.b. By relabeling, let Y be a copy of S (i.e., a set with the same cardinality as that of S) that is disjoint from H. By (AC14), we have $|H \cup Y| = |H \times Y|$. Therefore there exists a bijection $\beta : H \cup Y \to H \times Y$. Thus $H_1 = \beta(H)$ and $Y_1 = \beta(Y)$ are disjoint sets whose union equals $H \times Y$ and such that $|H_1| = |H|$ and $|Y_1| = |Y|$.

We now consider two cases. In the first case, there exists some $y \in Y$ such that $H \times \{y\} \subseteq Y_1$. In that case, the mapping $h \mapsto (h, y)$ is an injection from H into Y_1, proving that $|H| \leq |Y_1| = |Y| = |S|$.

In the second case, there is no such y. Hence for every $y \in Y$ there exists at least one $h \in H$ such that $(h, y) \notin Y_1$ — i.e., such that $(h, y) \in H_1$. Since H is well ordered, let $\alpha(y)$ be the *first* such h. Thus $y \mapsto (\alpha(y), y)$ is an injection from Y into H_1, and therefore $|S| = |Y| \leq |H_1| = |H|$.

6.23. In 2.16, we defined the cardinal number of a finite set, but for infinite sets we merely indicated how to *compare* cardinalities. How can we define the "cardinal number" of an infinite set? We would like to define an object "card(S)" separately for every set S in such a way that our definition of "card(S) $\leq$ card(T)" in 2.16 remains valid.

One naive approach would be to observe that equality of cardinality is an equivalence relation; thus it would seem that we can define the "cardinal number" of a set to be the equivalence class to which that set belongs. However, this approach involves an equivalence relation on the class of all sets, which is a very large proper class — perhaps too large for some purposes.

If we assume the Axiom of Choice, then we can make a *canonical* selection from each equivalence class, as follows: Every set can be well ordered, and hence has the same cardinality as some ordinal; see 5.46.f. There may be many such ordinals — for instance, ω, $\omega+1$, $\omega+2$, ... all have the same cardinality — but we can choose canonically among such ordinals, by taking the *first* such ordinal. It is a *cardinal number*, as defined in 5.47. Thus, for any set S, let card(S) be the first ordinal that has the same cardinality as S. With this definition, "card" is a function of classes, from the collection of all sets (a proper class) to the collection of all ordinals (another proper class). Note in particular that if S is a cardinal number (i.e., an initial ordinal), then card(S) $= S$.

The preceding definition uses the Axiom of Choice but not the Axiom of Regularity. Some mathematicians may prefer the following alternate definition, which uses Regularity but not Choice; it follows Enderton [1977 set theory]. Given any set S, there is some stage in which S occurs, as defined in 5.53. Say $S \in \mathrm{Stage}(\alpha)$. Recall that each stage is a set, not a proper class. Let β be the first ordinal with the property that some set $T \in \mathrm{Stage}(\beta)$ satisfies card(T) = card(S) — i.e., the first ordinal with the property that there exists a bijection between S and some member of Stage(β). Now let

$$\mathrm{kard}(S) \quad = \quad \{T \in \mathrm{Stage}(\beta)\ :\ \mathrm{card}(T) = \mathrm{card}(S)\}.$$

Then kard(S) is a set (not a proper class), uniquely determined by S, and two sets have the same "kardinality" if and only if they have the same cardinality. However, when S is an initial ordinal, it is not equal to kard(S).

6.24. Kelley's Choice. In later chapters, the Axiom of Choice will be used to prove certain important topological principles. Some of these principles also *imply* AC and thus are equivalent to it. We shall now sketch a general argument, which will be used several times in later chapters to prove that certain topological principles imply (AC3). This type of argument apparently was first used by Kelley [1950], in a proof we present in 17.16.

Let $\{S_\lambda : \lambda \in \Lambda\}$ be a nonempty set of nonempty sets; we wish to show that $\prod_{\lambda\in\Lambda} S_\lambda$ is nonempty.

For each λ, let ξ_λ be some object that is not an element of S_λ. (For instance, we could take $\xi_\lambda = S_\lambda$, since $S_\lambda \notin S_\lambda$. Thus, the ξ_λ's can be selected without making any arbitrary choices. However, it is probably better not to think of ξ_λ as being equal to S_λ, for such an assignment is an irrelevant distraction. It does not really matter what we choose for ξ_λ, so long as it is not a member of S_λ.)

Let $Y_\lambda = S_\lambda \cup \{\xi_\lambda\}$, let $X = \prod_{\lambda\in\Lambda} Y_\lambda$, and let $\pi_\lambda : X \to Y_\lambda$ be the λth coordinate projection. Obviously the function ξ defined by $\pi_\lambda(\xi) = \xi(\lambda) = \xi_\lambda$ is an element of $X =$

$\prod_{\lambda\in\Lambda} Y_\lambda$, which is therefore nonempty.

Now let M be any finite subset of Λ. By the Finite Axiom of Choice (see 6.14), $\prod_{\lambda\in M} S_\lambda$ is nonempty. Then the set

$$T_M \quad = \quad \bigcap_{\lambda\in M} \pi_\lambda^{-1}(S_\lambda) \quad = \quad \left(\prod_{\lambda\in M} S_\lambda\right) \times \left(\prod_{\lambda\in\Lambda\setminus M} Y_\lambda\right)$$

is a nonempty subset of X, as it includes $\left(\prod_{\lambda\in M} S_\lambda\right) \times \left(\prod_{\lambda\in\Lambda\setminus M}\{\xi_\lambda\}\right)$. Observe that $T_M \cap T_N = T_{M\cup N}$; hence the collection of sets $\mathcal{F} = \{T_M \ : \ M \text{ is a finite subset of } \Lambda\}$ is a filterbase on X.

The remainder of the argument is topological and takes a different form for different topological principles. The details cannot be given here, but will be given in 15.29, 19.13, and 17.16. In brief: We equip each Y_λ with some simple topology — e.g., the discrete topology, the indiscrete topology, or the "knob" topology — and let $X = \prod_{\lambda\in\Lambda} Y_\lambda$ be equipped with the product topology. Some assumed topological principle is then used to prove that $\prod_{\lambda\in\Lambda} S_\lambda = \bigcap_M T_M$ is a nonempty subset of X.

Countable Choice

6.25. An important weakened form of Choice is:

> **(CC) Axiom of Countable Choice.** We can choose representative elements from a sequence of nonempty sets. In other words, if $S_1, S_2, S_3, \ldots$ is a sequence of nonempty sets, then $\prod_{n=1}^{\infty} S_n$ is nonempty — i.e., we can choose a sequence $(x_1, x_2, x_3, \ldots)$ with $x_n \in S_n$ for each n.

Countable Choice is strong enough for many applications — for instance, Garnir, De Wilde, and Schmets [1968] develop a sizable portion of functional analysis using this axiom rather than the Axiom of Choice. However, Countable Choice is strictly weaker than the Axiom of Choice (see Jech [1973]). In fact, CC is so weak that the reader may again ask why this is an axiom, rather than just an "obviously true" statement or a consequence of definitions. To answer that question, we shall contrast CC with countable recursion (see examples in 2.23). Using either CC or recursion, one constructs a sequence (x_n). A recursive definition only allows one possible value for each x_n, and so no choices need to be made; the resulting sequence (x_n) is uniquely determined. In contrast, the sequence described in CC is not uniquely determined (unless all the S_n's are singletons), and so some arbitrary choices must be made. Of course, if the S_n's have some sort of known structure — e.g., if each S_n is a nonempty subset of $\mathbb{N}$ — then it may be possible to make canonical choices. But when we do not know of any structure, the axiom of Countable Choice still permits us to make an infinite sequence of arbitrary choices. (Contrast this also with AC, which permits *arbitrarily many* arbitrary choices.)

Countable Choice will now be used to prove two very basic properties of cardinality; our presentation follows that of Jech [1973].

6.26. (Assume CC. Then:) The union of countably many countable sets is countable.

Discussion and hints. Let $\{S_\lambda : \lambda \in \Lambda\}$ be a countable collection of countable sets; we are to prove that $S = \bigcup_{\lambda\in\Lambda} S_\lambda$ is countable. Since Λ is countable, by relabeling we may assume that $\Lambda = \mathbb{N}$ or $\Lambda = \{1, 2, \ldots, N\}$ for some positive integer N. To reflect this relabeling, let us replace the λ's with n's. We wish to show that $S = \bigcup_{n\in\Lambda} S_n$ is countable.

For each $n \in \Lambda$, the set S_n is countable, and so there exists at least one injection from S_n into $\mathbb{N}$. For each n, let us *choose* some injection $\iota_n : S_n \to \mathbb{N}$. Since there are countably many n's, we are making countably many choices; it is at this step that we need the Axiom of Countable Choice. (If the S_n's were given to us with some sort of listing already provided, so that we could choose the ι_n's canonically, then CC would not be needed.)

Now, for each $x \in S$, let $n(x)$ be the first integer n that satisfies $x \in S_n$, and let $p(x) = \iota_{n(x)}(x)$. Then the mapping $x \mapsto (n(x), p(x))$ is an injection from S into $\mathbb{N} \times \mathbb{N}$, which is countable by 2.20.e.

6.27. (Assume CC. Then:) A set X is infinite (i.e., not finite) if and only if it contains a countably infinite set — i.e., if and only if $\mathrm{card}(X) \geq \mathrm{card}(\mathbb{N})$.

Proof. It is clear that $\mathrm{card}(X) \geq \mathrm{card}(\mathbb{N})$ implies X is not finite; that implication does not require the Axiom of Countable Choice.

Conversely, assume X is an infinite set. We claim that

> ($*$) for each nonnegative integer j, the set X contains a subset A_j having exactly j elements.

Indeed, take $A_0 = \varnothing$. If ($*$) is false for some $j > 0$, consider the smallest such j. Then A_{j-1} is finite and X is not, so $A_{j-1} \subsetneqq X$. Thus $X \setminus A_{j-1}$ is nonempty; let x be any element of $X \setminus A_{j-1}$. Now take $A_j = A_{j-1} \cup \{x\}$. This contradiction proves ($*$). Then $\bigcup_{j=0}^{\infty} A_j$ is a countably infinite subset of X.

Remarks. A set is **Dedekind infinite** if it has the same cardinality as some proper subset of itself; otherwise it is **Dedekind finite**. Some mathematicians take Dedekind infiniteness, or the condition $\mathrm{card}(X) \geq \mathrm{card}(\mathbb{N})$, as a definition of X being infinite. If we assume Countable Choice, then the three notions of "infinite" coincide.

Dependent Choice

6.28. Between the Axiom of Choice and Countable Choice lies an important but more complicated principle, the **Principle of Dependent Choice (DC)**. We shall give two versions of this principle now and a few more versions in Chapters 19 and 20.

> **(DC1) Dependent Choice (version without history).** Let any nonempty set S and any function $f : S \to \{\text{nonempty subsets of } S\}$ be given. Then there exists a sequence (x_n) in S such that $x_{n+1} \in f(x_n)$ for each n.

> **(DC2) Dependent Choice (version with history).** Let S_1, S_2, S_3, ... be nonempty sets. For each $n \geq 1$, let f_n be a mapping from $S_1 \times S_2 \times \cdots \times S_n$ into {nonempty subsets of S_{n+1}}. Then there exists a sequence $(x_1, x_2, x_3, \ldots)$ such that $x_{n+1} \in f_n(x_1, x_2, \ldots, x_n)$ for each n.

In either formulation, the idea is that we can choose some x_1; with it in mind we can then choose some x_2; with x_1 and x_2 (or just x_2) in mind we can then choose some x_3, etc.

Remarks. It is known that AC is strictly stronger than DC, and DC is strictly stronger than CC. Recently, Howard and Rubin [1996] have shown that UF + CC does not imply DC. (The Ultrafilter Principle, UF, is discussed in 6.32 and thereafter.)

The Axiom of Dependent Choice (DC) should not be confused with the Axiom of Determinacy (AD). Though the two names sound similar, the two axioms are entirely different. We shall not study AD in this book; a good introduction to it is given by Dalen, Doets, and Swart [1978].

6.29. *Exercise.* (DC1) $\iff$ (DC2), and (AC4) $\Rightarrow$ (DC2) $\Rightarrow$ CC.

Hint for (DC1) $\Rightarrow$ (DC2): Let

$$S \;=\; \bigcup_{n=1}^{\infty} (S_1 \times S_2 \times \cdots \times S_n)$$

and $f(x_1, x_2, \ldots, x_n) = \{(x_1, x_2, \ldots, x_n)\} \times f_n(x_1, x_2, \ldots, x_n)$.

6.30. *Optional exercise.* Assume Dependent Choice. Let $(X, \leq)$ be a chain ordered set. Show that $\leq$ is a well ordering of X if and only if there does not exist an infinite sequence $x_1 > x_2 > x_3 > \cdots$ in X.

6.31. *Optional exercise* (from Johnstone [1987]). If we assume Dependent Choice, then the Axiom of Regularity is equivalent to the following principle.

> **No Infinite Regress.** There does not exist an infinite sequence of sets S_0, S_1, S_2, S_3, ... that satisfies $\cdots \in S_3 \in S_2 \in S_1 \in S_0$.

Proof. In 1.49 we proved that the Axiom of Regularity implies No Infinite Regress. Conversely, suppose that the Axiom of Regularity is false. Say S_0 is a nonempty set that meets each of its elements. If S_n meets S_0, then there is some $S_{n+1} \in S_n \cap S_0$. Use DC to form a sequence $(S_0, S_1, S_2, \ldots)$.

The Ultrafilter Principle

6.32. (We assume some familiarity with filters, which were introduced in Chapter 5.) Midway between the Axiom of Choice (AC) and the Axiom of Choice for Finite Sets (ACF) is a more complicated but very important principle, the *Ultrafilter Principle* (also known as the *Ultrafilter Theorem*). Like the Axiom of Choice, the Ultrafilter Principle has many important equivalents in many branches of mathematics. The equivalents discussed in this book will be denoted (UF1), (UF2), (UF3), etc.; collectively we shall refer to them as UF. They can be found in the paragraph below, and in 6.35, 7.24, 9.54, 13.22, 14.57, 14.59, 14.61, 17.4, 17.22, 17.42.f, 19.17, and 28.29.

The version of greatest use for purposes of this book is

> **(UF1) Ultrafilter Principle (Cartan).** Any proper filter is included in an ultrafilter. That is, if $\mathcal{F}$ is a proper filter on a set X, then there exists an ultrafilter $\mathcal{U}$ on X with $\mathcal{U} \supseteq \mathcal{F}$.

This result follows easily from (AC5) or (AC7), in 6.20. The Ultrafilter Principle is strictly weaker than the Axiom of Choice; that was proved by Halpern and Lévy [1971], but the proof is beyond the scope of this book.

This book contains nearly two dozen equivalents of the Ultrafilter Principle; the wider literature contains many more. The equivalents have not all been collected into one source. A few more equivalents are given by Jech [1973], Morillon [1986], Rav [1977], and Rubin and Rubin [1985]. Mathematicians who wish to study equivalents of UF are urged to search not only under "ultrafilter" but also under "Compactness Principle," "Boolean Prime Ideal Theorem," and "Stone Representation Theorem;" those equivalents of UF (considered later in this book) are less essential to analysts but are more famous among logicians and algebraists.

6.33. *Existence of free ultrafilters.* Recall that an ultrafilter on an infinite set is free if and only if it contains the cofinite filter (see 5.5.d). Hence, using (UF1) to extend the cofinite filter is one method of "constructing" free ultrafilters. Thus we obtain this corollary of (UF1):

$$\text{On every infinite set there exists a free ultrafilter.} \qquad (**)$$

A special case of this result is important enough to have its own name:

> **(WUF) Weak Ultrafilter Theorem.** A free ultrafilter exists on $\mathbb{N}$.

Actually, WUF plus CC implies statement (**) above. (*Proof.* CC tells us that X contains a countably infinite set X_0; see 6.27. Show that if $\mathcal{F}$ is a free ultrafilter on X_0, then $\{S \subseteq X : S$ contains some member of $\mathcal{F}\}$ is a free ultrafilter on X.)

Neither of the implications AC $\Rightarrow$ UF $\Rightarrow$ WUF is reversible; this is proved in Jech [1973]. Thus, UF is strictly weaker than the Axiom of Choice, and WUF is weaker still.

Free ultrafilters — on $\mathbb{N}$ or on any infinite set — are **intangibles**, in the sense of 14.76 and 14.77: They exist in conventional set theory, but we cannot prove their existence using just ZF + DC. This result was proved by Pincus and Solovay [1977]; see the discussion in 14.74. It also follows from Shelah's result Con(ZF + DC + BP), via an argument of WUF $\Rightarrow$ not-BP given later in this book.

Although the free ultrafilters are harder to illustrate or imagine than the fixed ones, they are also far more numerous. A theorem of Tarski states that if X is an infinite set, then card{free ultrafilters on X} = card$\{\mathcal{P}(\mathcal{P}(X))\}$. This is in contrast with card{fixed ultrafilters on $X\} = \mathrm{card}(X)$. The proof, which uses the Axiom of Choice, is too long to be included here. Proofs can be found in Tarski [1939], Bell and Slomson [1974], and Gähler [1977].

6.34. Show that ACR (in 6.12) implies WUF.

Hints: We may replace $\mathbb{R}$ with $\mathcal{P}(\mathbb{N})$, by 10.44.f. By 3.43, the subsets of $\mathbb{N}$ can be well ordered. The collection $\mathcal{C}$ of filter subbases on $\mathbb{N}$ is a collection with finite character; hence by the theorem in 3.46 the cofinite filter can be extended to a maximal member of $\mathcal{C}$.

6.35. (UF2) Cowen-Engeler Lemma. Let Λ and X be sets. Let Φ be a collection of functions from subsets of Λ, into X. Assume that

(i) $\Phi(\lambda) = \{f(\lambda) : f \in \Phi \text{ with } \lambda \in \mathrm{Dom}(f)\}$ is a finite subset of X, for each $\lambda \in \Lambda$;

(ii) each finite set $S \subseteq \Lambda$ is the domain of at least one element of Φ; and

(iii) Φ has *finite character*; i.e., a function f from some subset of Λ into X is a member of Φ if and only if each restriction of f to a finite subset of Domain(f) is a member of Φ.

Then Λ is the domain of at least one element of Φ.

Remarks. The proof of (UF2) $\Rightarrow$ (UF1) will be given via several other propositions in 13.22. Actually, (UF2) remains equivalent if we make the further stipulation that $X = \{0, 1\}$; that will be evident from the argument in 13.22.

The principle (UF2) is very similar to several principles that are known as **Rado's Selection Lemma**; the reader is cautioned that those principles are not all known to be equivalent to each another. For a few results on Rado's Lemma(s) see Howard [1984] and [1993], Jech [1977], Rav [1977], and Thomassen [1983].

The Cowen-Engeler Lemma, particularly with $X = \{0, 1\}$, is in many respects similar to the Compactness Principle of Propositional Logic, which is (UF16) in 14.61. In fact, as Rav [1977] points out, the Cowen-Engeler Lemma is a sort of combinatorial, non-logicians' version of (UF16); the Cowen-Engeler Lemma can often be used in place of (UF16) but does not require any knowledge of formal logic.

Proof of (UF1) $\Rightarrow$ (UF2). This proof is modified from arguments of Rav [1977] and Luxemburg [1962]. Let $\mathrm{Fin}(\Lambda) = \{\text{finite subsets of } \Lambda\}$. For each $S \in \mathrm{Fin}(\Lambda)$, let $\Gamma_S = \{f \in \Phi : \mathrm{Dom}(f) \supseteq S\}$. Then Γ_S is nonempty, by hypothesis (ii). Since $\Gamma_S \cap \Gamma_T = \Gamma_{S \cup T}$, the collection of sets $\{\Gamma_S : S \in \mathrm{Fin}(\Lambda)\}$ has the finite intersection property. By (UF1), there exists a (not necessarily unique) ultrafilter $\mathcal{U}$ on Φ that includes $\{\Gamma_S : S \in \mathrm{Fin}(\Lambda)\}$.

To define $\varphi : \Lambda \to X$, temporarily fix any $\lambda \in \Lambda$. Note that $\Phi(\lambda) = \{f(\lambda) : f \in \Gamma_{\{\lambda\}}\}$. The sets $\{f \in \Gamma_{\{\lambda\}} : f(\lambda) = x\}$ (for $x \in \Phi(\lambda)$) are disjoint and their union is $\Gamma_{\{\lambda\}}$, which is a member of the ultrafilter $\mathcal{U}$. By 5.7.b and 5.8(E), precisely one of the x's in $\Phi(\lambda)$ satisfies $\{f \in \Gamma_{\{\lambda\}} : f(\lambda) = x\} \in \mathcal{U}$. Let that x be denoted by $\varphi(\lambda)$.

Thus, we define a function $\varphi : \Lambda \to X$, satisfying

$$\{f \in \Gamma_{\{\lambda\}} : f(\lambda) = \varphi(\lambda)\} \in \mathcal{U} \qquad \text{for all } \lambda \in \Lambda.$$

It suffices to show that $\varphi \in \Phi$. Let any $S \in \mathrm{Fin}(\Lambda)$ be given. Since Φ has finite character, it suffices to show that φ agrees on S with some $f \in \Phi$. The set

$$\Psi \;=\; \bigcap_{\lambda \in S} \{f \in \Gamma_{\{\lambda\}} : f(\lambda) = \varphi(\lambda)\}$$

is also an element of $\mathfrak{U}$, hence a nonempty subset of Φ. Now any $f \in \Psi$ will do.

6.36. *Exercise.* Show that (UF2) implies the Axiom of Choice for Finite Sets, which was stated in 6.15 as (ACF).

Hint: Use the Finite Axiom of Choice (6.14).

6.37. Marriage Theorems. Let $\{S_\gamma : \gamma \in \Gamma\}$ be a collection of sets. Assume either

(i) Γ is finite (for **P. Hall's Theorem**), or

(ii) each S_γ is finite (for **M. Hall's Theorem**).

Then the following two conditions are equivalent:

(A) there exists an injective function $x \in \prod_{\gamma \in \Gamma} S_\gamma$.

(B) $\mathrm{card}(\bigcup_{\gamma \in F} S_\gamma) \geq \mathrm{card}(F)$ for each finite set $F \subseteq \Gamma$.

Remarks. In both cases, the implication (A) $\Rightarrow$ (B) is obvious; it is (B) $\Rightarrow$ (A) that we must prove. We cannot omit hypotheses (i) and (ii). For instance, the sets $\{1, 2, 3, \ldots\}$, $\{1\}$, $\{2\}$, $\{3\}$, $\ldots$ satisfy (B) but not (A).

The theorem above gives necessary and sufficient conditions for the solvability of the "marriage problem" of combinatorics: Let Γ be a collection of heterosexual people of one gender, and let S_γ be the set of suitors (of the other gender) of person γ; then condition (A) says that all the elements of Γ can be married simultaneously to suitors.

We have attributed the two theorems above to Phillip Hall and Marshall Hall, respectively, because they were apparently the earliest publishers of those theorems — see Hall [1935] and Hall [1948] — but both theorems have been subsequently rediscovered many times. P. Hall's Theorem is also equivalent to several other important combinatorial matching theorems, including theorems of König and Menger in graph theory, Dilworth's Theorem on partially ordered sets, and the Ford-Fulkerson Max-flow Min-cut Theorem of network theory. (By "equivalent" we mean in this instance that each theorem implies the others easily.) Surveys of related material are given by Mirsky [1971] and Reichmeider [1984].

Our proof of M. Hall's Theorem will use (UF2); later we shall use M. Hall's Theorem to prove Löwig's Theorem in 11.31. It is not yet known whether M. Hall's Theorem or Löwig's Theorem is equivalent to UF.

Proof of P. Hall's Theorem — i.e., assuming (i). This proof follows Halmos and Vaughan [1950]. Let $\Gamma = \{1, 2, 3, \ldots, n\}$; we proceed by induction on n. For $n = 1$ the result is trivial. For larger n we consider two cases:

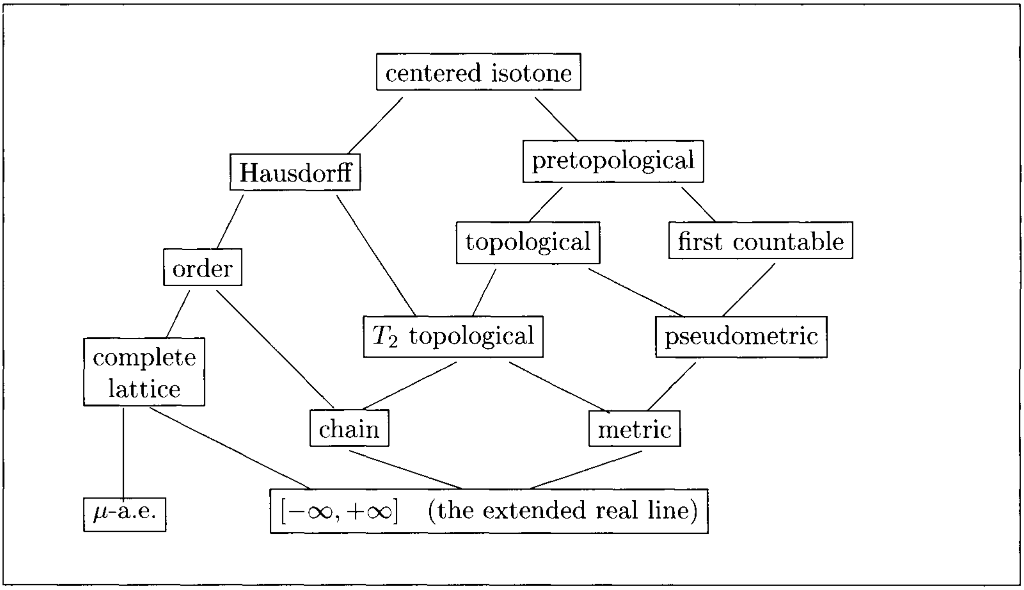
- First, suppose that each union of k S_i's $(1 \leq k \leq n-1)$ contains at least $k+1$ elements. In this case we may choose any $x_n \in S_n$ and then apply the induction hypothesis to the $n-1$ sets $S_1 \backslash \{x_n\}$, $S_2 \backslash \{x_n\}$, $\ldots$, $S_{n-1} \backslash \{x_n\}$.

- On the other hand, suppose that some union of k of the S_i's contains exactly k elements, for some k with $1 \leq k \leq n-1$. By relabeling we may assume that these are the sets S_1, S_2, ..., S_k. Let their union be T. Clearly the inductive hypothesis can be applied to the k sets S_1, S_2, ..., S_k. It suffices for us to show that the inductive hypothesis also can also be applied to the $n-k$ sets $S_{k+1}\backslash T$, $S_{k+2}\backslash T$, ..., $S_n\backslash T$. To see that, note that for $1 \leq r \leq n-k$, the union of any r of these $n-k$ sets contains at least r elements, since the union of those r sets together with T is the union of $k+r$ of the original S_i's, and hence contains at least $k+r$ elements.

Proof of M. Hall's Theorem — i.e., assuming (ii). This proof is modified from Mirsky [1971]. Let Φ be the collection of all injective functions f defined on subsets of Γ that satisfy $f(\gamma) \in S_\gamma$ for each $\gamma \in \text{Domain}(f)$. It is easy to see that Φ has finite character, in the sense of (UF2)(iii). Also, each set $\Phi(\gamma) = \{f(\gamma) : f \in \Phi, \gamma \in \text{Domain}(f)\}$ is finite, since it is contained in the finite set S_γ. By P. Hall's Theorem, each finite subset of Γ is the domain of at least one member of Φ. Thus, (UF2) is applicable, and Γ is the domain of at least one member of Φ.

Chapter 7

Nets and Convergences



7.1. *An elementary special case.* A sequence (x_n) in a metric space (X, d) is said to **converge** to a limit $x \in X$ if

> for each number $\varepsilon > 0$, there exists an integer $N = N(\varepsilon)$ such that $n \geq N \Rightarrow d(x, x_n) < \varepsilon$.

We then write $x_n \to x$ or $x = \lim_{n\to\infty} x_n$.

7.2. *Chapter overview.* Much of analysis can be formulated in terms of convergence of sequences in metric spaces, but occasionally we need greater generality.

Nets are a generalization of sequences. A sequence is a function whose domain is $\mathbb{N}$; a net (or "generalized sequence") is a function whose domain is any directed set D. Most of this chapter can be postponed; it will not be needed until much later in this book.

A **convergence space** is a set X equipped with some rule that specifies which nets

— or equivalently,[1] which filters — converge to which "limits" in X. Analysts who are already familiar with convergent sequences in metric spaces should have little difficulty with convergent nets, for — as we shall see in this chapter — nets and convergence spaces are natural generalizations of sequences in metric spaces.

The chart at the beginning of this chapter shows the relations between some of the main types of convergences we shall consider in this book. In later chapters we shall be primarily concerned with topological convergences — and, to a much smaller degree, order convergences. The other kinds of convergences — Hausdorff, pretopological, first countable, etc. — are introduced here mainly to give a clearer understanding of the basic properties of topological and order convergences.

Nets are particularly helpful for understanding topologies that are known to be nonmetrizable — e.g., the weak topology of an infinite-dimensional normed vector space — or understanding topologies that are not known to be metrizable. But nets are also occasionally useful in metric spaces; two examples of this are the proof of Caristi's Theorem given in 19.45 and the explanation of Riemann integrals given in 24.7.

One very important order convergence that is *not* topological is the convergence almost everywhere of $[-\infty,+\infty]$-valued random variables over a positive measure; this topic is considered briefly in 21.43. Other nontopological order convergences are important in the study of vector lattices, but that subject is not studied in great depth in this book. We are more concerned with order convergences that *are* topological. For instance, the order convergence and the topological convergence in $\mathbb{R}$ are identical, but the order viewpoint and the topological viewpoint yield different kinds of information about that convergence.

Nets are an aid to the intuition and to the process of discovery, but they are not always essential; many proofs involving nets can be rewritten so that nets are not mentioned. Some researchers prefer to rewrite their proofs in that fashion: The original insight may thereby be obscured, but the result becomes readable by a wider audience since familiarity with nets is no longer required.

Although nets are used mainly for convergences, it is conceptually simpler to first study nets without regard to convergences — i.e., as devices for a modified sort of "counting," without any regard to limits. That is the subject of the first half of this chapter.

7.3. *Review of directed sets.* Before reading this chapter, it may be helpful to briefly review the introduction to filters in Sections 5.1 through 5.11.

Also, recall from 3.8 the definition of directed set: It is a set X equipped with a relation $\preccurlyeq$ that is reflexive ($x \preccurlyeq x$ for all x) and transitive (i.e., if $x \preccurlyeq y$ and $y \preccurlyeq z$, then $x \preccurlyeq z$) and that also satisfies this condition:

$$\text{for each } x, y \in X, \text{ there exists } u \in X \text{ such that } x, y \preccurlyeq u.$$

We review a few basic properties of directed sets from Chapter 3:

- The universal ordering ($x \preccurlyeq x$ for all $x \in X$) is a directed ordering that is not antisymmetric.

[1] To make the convergence of nets equivalent to the convergence of filters simplifies our theory substantially, but it imposes a mild restriction on the kinds of net convergences that we shall consider. This is discussed further in 7.31.

- Any product of directed sets, with the product ordering, is directed.
- A subset of a directed set (when equipped with the restriction ordering) is not necessarily directed.

Directed sets will be used as generalizations of $\mathbb{N}$ or $\mathbb{R}$, so this book will often denote directed sets by $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$, $\mathbb{D}$, etc. (The reader must determine from context whether $\mathbb{C}$ means a directed set or the complex numbers.) In accordance with much of the literature, we shall usually denote elements of a directed set by lowercase Greek letters (α, β, γ, ...).

7.4. Let $\mathcal{B}$ be a nonempty collection of nonempty subsets of a set X. Then $(\mathcal{B}, \supseteq)$ is a directed set if and only if $\mathcal{B}$ is a filterbase on X. We refer to $\supseteq$ as the **ordering by reverse inclusion**.

Since every filter is also a filterbase, we see in particular that reverse inclusion is a directed ordering on any filter. Reverse inclusion is the most common directed ordering used on filters. It will be understood to be in use whenever we use a filter as a directed set, except when some other ordering is specified. Note that, unfortunately, larger sets are "smaller" in this ordering, and smaller sets are "larger" — i.e., $S \subseteq T \iff S \succcurlyeq T$.

Ordinary inclusion ($\subseteq$) is also a directed ordering on any filter, but that ordering is seldom useful.

Nets

7.5. Before turning to generalized sequences, let us first review the notation of sequences. Recall that a *sequence* in a set X is a mapping from $\mathbb{N}$ into X, where $\mathbb{N} = \{1, 2, 3, \ldots\}$ has its usual ordering. A sequence can be viewed as a function, with values $x(1), x(2), x(3), \ldots$, and when it is helpful we shall adopt that notation. However, it is more common to view a sequence as a set parametrized by $\mathbb{N}$; then a sequence is written as $(x_1, x_2, x_3, \ldots)$ or $(x_n : n \in \mathbb{N})$ or (x_n). Subscripts i, j, k, m, n generally will mean elements of $\mathbb{N}$ if no other index set is indicated. If we disregard the ordering of (x_n), we obtain the countable set $\{x_n\} = \{x_1, x_2, x_3, \ldots\}$, which is the *range* of the sequence; note the use of braces instead of parentheses.

7.6. Nets are a generalization of sequences — in fact, they are also sometimes called **generalized sequences** or **Moore-Smith sequences**.

A **net** in a set X is any function from a nonempty directed set into X. Thus, in the notation of Chapter 1, a net in X is any function $x : \mathbb{D} \to X$, where $\mathbb{D}$ is any nonempty directed set; the values of a net may sometimes be written as $x(\alpha), x(\beta), \ldots$. However, it is more common to view a net as a set parametrized by a directed set $\mathbb{D}$; thus we usually represent such a net by the expression $(x_\delta : \delta \in \mathbb{D})$. We may abbreviate this as (x_δ) if $(\mathbb{D}, \preccurlyeq)$ does not need to be mentioned explicitly, but the set $\mathbb{D}$ and its ordering $\preccurlyeq$ are still understood to be part of the structure of the net. If we disregard the order on (x_δ), we obtain the *set* $\{x_\delta : \delta \in \mathbb{D}\}$, which is the range of the net (x_δ); again note the use of braces instead of parentheses.

Sequences are a special case of nets, and so most statements we shall make about nets will apply to sequences as well. Of course, sequences are conceptually simpler than nets, and so whenever possible we prefer to use sequences. However, for some purposes (e.g., the study of convergence in nonmetrizable topological spaces) nets are a more natural tool.

Remark. The word "net" is perhaps unfortunate — it does not have any intuitive justification, as far as this author knows. The alternate term "stream" was suggested by McShane [1952], for reasons indicated in 3.9.f, but "net" is the standard word.

7.7. Let $x : \mathbb{D} \to X$ be a net, and let $S \subseteq X$. We shall say that

S is a **tail set** of the net if S is of the form $\{x_\delta : \delta \succcurlyeq \delta_0\}$ for some $\delta_0 \in \mathbb{D}$;

S is an **eventual** (or *residual*) set of the net if S contains some tail set — i.e., if there is some $\delta_0 \in \mathbb{D}$ such that $\{x_\delta : \delta \succcurlyeq \delta_0\} \subseteq S$. In this case we say that $x_\delta \in S$ happens **eventually**, or that $x_\delta \in S$ happens **for all δ sufficiently large**.

S is a **frequent** (or *cofinal*) set of the net if S meets every tail set — i.e., if for each $\delta_0 \in \mathbb{D}$ there is some $\delta \succcurlyeq \delta_0$ such that $x_\delta \in S$. In this case we say that $x_\delta \in S$ happens **frequently**, or that $x_\delta \in S$ happens **for arbitrarily large values of δ**.

S is **infrequent** if it is not frequent.

Of course, these definitions all depend on the directed set $(\mathbb{D}, \preccurlyeq)$, the codomain X, and the net $(x_\delta : \delta \in \mathbb{D})$.

These terms can also be applied to subsets of a directed set $\mathbb{D}$, by viewing the identity map $i : \mathbb{D} \to \mathbb{D}$ as a $\mathbb{D}$-valued net (with $i_\delta = \delta$). Thus, a subset $\mathbb{S} \subseteq \mathbb{D}$ is a *tail set* if $\mathbb{S}$ is of the form $\{\delta \in \mathbb{D} : \delta \succcurlyeq \delta_0\}$, an *eventual set* if $\mathbb{S}$ contains some set of the form $\{\delta \in \mathbb{D} : \delta \succcurlyeq \delta_0\}$, or a *frequent set* if $\mathbb{S}$ meets every set of the form $\{\delta \in \mathbb{D} : \delta \succcurlyeq \delta_0\}$.

Caution: The term "tail set" has another, unrelated meaning; see 20.31.

7.8. *Examples and basic properties.*

a. A set is eventual if and only if its complement is infrequent.

b. Let $\mathbb{N}$ have its usual ordering, and consider the identity map $i : \mathbb{N} \to \mathbb{N}$ as a net. Then i_δ is eventually greater than 5, and i_δ is frequently a multiple of 17. A subset of $\mathbb{N}$ is eventual if and only if it is cofinite (i.e., has finite complement), and frequent if and only if it is an infinite set.

c. Let $\mathbb{N}$ be partially ordered by $m \preccurlyeq n$ if m is a factor of n. Then $(\mathbb{N}, \preccurlyeq)$ is a directed set. Let $i : \mathbb{N} \to \mathbb{N}$ be the identity map; then i_δ is eventually a multiple of 17.

7.9. *Correspondence between nets and filters.* Let any net $(x_\delta : \delta \in \mathbb{D})$ in a set X be given. Then $\mathcal{B} = \{\text{tail sets of the net } (x_\delta)\}$ is a filterbase on X; the proper filter that it generates is $\mathcal{F} = \{\text{eventual sets of the net } (x_\delta)\}$. We shall call these the **filterbase of tails** and the

eventuality filter of (x_δ), respectively. (Some mathematicians call $\mathcal{F}$ the *filter of tails* of (x_δ).)

The proper ideal that is dual to the filter $\mathcal{F}$ is the collection of all infrequent subsets of X . In other words, a set $S \subseteq X$ is eventual if and only if $X \setminus S$ is infrequent. Using this duality, we can convert statements about frequent sets to statements about eventual sets, and vice versa. Referring to the discussion in 5.3, the reader may find it helpful to think of "eventual" as meaning "large," "infrequent" as meaning "small," and "frequent" as meaning "not small."

7.10. *Further properties and examples.*

a. The constant net at z — i.e., the net satisfying $x_\alpha = z$ for all α — has eventuality filter equal to the ultrafilter fixed at z.

b. If X is a set directed by the universal ordering — that is, $x \preccurlyeq y$ for all $x, y \in X$ — then the identity map $i : X \to X$ is a net whose eventuality filter is the singleton $\{X\}$.

c. Any eventual set is frequent.

d. Any superset of a frequent set is frequent; any superset of an eventual set is eventual.

e. If n is a positive integer and $S_1 \cup S_2 \cup \cdots \cup S_n$ is frequent, then at least one of the S_i's is frequent.

f. Let $(\mathbb{D}, \preccurlyeq)$ be a directed set, and suppose $\mathbb{S} \subseteq \mathbb{D}$ is frequent. Then $\mathbb{S}$ is itself a directed set, when ordered by the restriction of the given ordering.

g. Let $p : X \to Y$ be a mapping from one set into another. Let $(x_\delta : \delta \in \mathbb{D})$ be a net in a set X; then (x_δ) has tail filterbase equal to $\mathcal{B} = \{\{x_\delta : \delta \succcurlyeq \gamma\} : \gamma \in \mathbb{D}\}$ and eventuality filter $\mathcal{F} = \{S \subseteq X : S \supseteq B \text{ for some } B \in \mathcal{B}\}$. Show that

(i) $p(\mathcal{B}) = \{p(B) : B \in \mathcal{B}\}$ is the tail filterbase for the net $(p(x_\delta) : \delta \in \mathbb{D})$ in Y.

(ii) $p(\mathcal{F}) = \{p(F) : F \in \mathcal{F}\}$ is also a filterbase on Y.

(iii) $\mathcal{G} = \{S \subseteq Y : p^{-1}(S) \in \mathcal{F}\}$ is the eventuality filter of the net $(p(x_\delta) : \delta \in \mathbb{D})$ in Y.

(iv) $p(\mathcal{B}) \subseteq p(\mathcal{F}) \subseteq \mathcal{G}$, and $\mathcal{G}$ is the filter generated by both $p(\mathcal{B})$ and $p(\mathcal{F})$. (Refer to 5.40.b.)

7.11. In 7.9 we saw how each net determines a proper filter, which we call the eventuality filter. Conversely, now let $\mathcal{B}$ be a proper filter on a set X; we wish to construct a net (x_δ) in X whose eventuality filter is $\mathcal{B}$. Many such nets are available, but we shall describe one that is canonical — i.e., one that can be constructed from $\mathcal{B}$ by a straightforward algorithm without any arbitrary choices. This construction is taken from Bruns and Schmidt [1955]; it was independently rediscovered by Wilansky [1970].

Let $\mathcal{B}$ be any proper filter — or more generally, any filterbase — on a set X. Let X have the universal ordering (as in 3.9.g), let $\mathcal{B}$ be ordered by reverse inclusion (as in 7.4), and let $X \times \mathcal{B}$ have the product ordering. Show that $\mathbb{D} = \{(x, S) \in X \times \mathcal{B} : x \in S\}$ is a frequent subset of $X \times \mathcal{B}$, and hence is a directed set by 7.10.f. Then show that the map $(x, S) \mapsto x$, from $\mathbb{D}$ into X, is a net whose filterbase of tails is $\mathcal{B}$; its eventuality filter is $\mathcal{B}$ if that filterbase is a filter.

We shall call this net the **canonical net** of $\mathcal{B}$.

This canonical construction is admittedly a bit complicated. We shall use it occasionally, but more often we shall merely need to use the fact that some canonical construction *exists* — i.e., that there is *some* canonical way to construct a net with a given eventuality filter; the specific details of the construction will not enter into most applications.

7.12. (*Optional.*) If $\mathcal{B}$ is any filterbase on a set X, then there also exists a net $(x_\alpha : \alpha \in \mathbb{A})$ whose filterbase of tails is $\mathcal{B}$ and such that the directed set $\mathbb{A}$ is *antisymmetric* — i.e., it is also a poset. (Consequently, our applications would not be greatly affected if we made antisymmetry a part of our definition of directed set.)

Construction: Let

$$\mathbb{A} \quad = \quad \{(u, n, B) \in X \times \mathbb{N} \times \mathcal{B} \quad : \quad u \in B\}$$

Order it as follows: $(u, n, B) \prec (u', n', B')$ if and only if either (i) $B \supsetneqq B'$ or (ii) $B = B'$ and $n < n'$. Define $x_{(u,n,B)} = u$. Verify that $(\mathbb{A}, \preccurlyeq)$ is a directed poset and that $\{x_\alpha : \alpha \succcurlyeq (u_0, n_0, B_0)\} = B_0$ for each $(u_0, n_0, B_0) \in \mathbb{A}$ — hence $\mathcal{B}$ is the filterbase of tails for the net. This construction is also from Bruns and Schmidt [1955].

7.13. *Remarks: nets versus filters.* Filters have many other uses — in set theory, logic, algebra, etc. — but filters can also be used to study convergences. In fact, nets and filters yield essentially the same results about convergences. Some mathematicians prefer nets or prefer filters, and use only one system or the other. It is this author's opinion that the ideas of nets and filters complement each other; they should not be viewed as two separate systems of ideas. In 7.9 we showed how to switch back and forth between nets and filters, so that each system can be used to its best advantage. That interchangeability is strengthened by the ideas of Aarnes and Andenæs; see especially 7.15(C). This book will make frequent use of nets and filters and of their interchangeability.

For any proper filter $\mathcal{F}$, the eventuality filter of the canonical net of $\mathcal{F}$ is $\mathcal{F}$. This gives us a bijection between the proper filters on X and a certain collection of nets in X. However, this bijection is not onto the class of *all* nets in X. "Most" nets are not canonical nets. In fact, the class of all nets in X is a proper class (see 1.44) — it is far too big to be a set, since we make no restriction on the choice of the underlying directed set. In contrast, the collection of all filters on X is a set of ordinary size — clearly, $\{\text{filters on } X\} \subseteq \mathcal{P}(\mathcal{P}(X))$.

Nevertheless, the correspondence between filters and nets is quite good, and so we may use the two tools interchangeably, thereby gaining the advantages of each. Nets are a natural generalization of sequences, so they may be intuitively appealing to analysts, who are already familiar with sequences. On the other hand, many proofs are easier in terms of filters, since filters are always "canonical." For instance, the filter $\mathcal{N}(x)$ of all neighborhoods of a point — studied in 15.2 and 15.7 and thereafter — plays a useful special role in many proofs, since it is the smallest filter that converges to the point x. It can be replaced by a net, as in 7.11, but that replacement is somewhat complicated and artificial, and removes much of the available intuition; we may prefer to work with $\mathcal{N}(x)$.

SUBNETS

7.14. *Preview and historical remarks.* In the following pages we shall compare several types of subnets. In order of increasing generality, they are

$$\{\text{subsequences}\} \subseteq \left\{\begin{array}{c}\text{frequent}\\ \text{subnets}\end{array}\right\} \subseteq \left\{\begin{array}{c}\text{Willard}\\ \text{subnets}\end{array}\right\} \subseteq \left\{\begin{array}{c}\text{Kelley}\\ \text{subnets}\end{array}\right\} \subseteq \left\{\begin{array}{c}\text{AA}\\ \text{subnets}\end{array}\right\}.$$

The last three types — Willard, Kelley, and AA — are our main types of subnets. Any one of these by itself would make a good definition of "subnet," and has been used as such elsewhere in the literature. Although the three definitions require slightly different proofs of theorems, they yield essentially the same statements of theorems; their near-interchangeability will follow from results in 7.19 and 15.38. The Kelley definition is oldest and is most widely used in the literature, but the other two definitions are simpler. The AA definition is the most general and yields the simplest proofs. For those reasons and other reasons indicated below, *this book will use the term "subnet" to mean "AA subnet" except where noted explicitly.* For an abridged treatment, the reader may skip over Willard and Kelley subnets.

Frequent subnets (introduced in 7.16.c) are important enough to deserve mention, but they are much more specialized. In general, they cannot be used interchangeably with the other three kinds of subnets; this will be shown in 17.29.

Subnets are a generalization of subsequences. Recall that $(y_p : p \in \mathbb{N})$ is a **subsequence** of $(x_n : n \in \mathbb{N})$ if we can write $y_p = x_{\varphi(p)}$ or $y_p = x_{\varphi_p}$ for some positive integers $\varphi(1) < \varphi(2) < \varphi(3) < \cdots$. Analogous ideas for nets were gradually developed by Moore, Smith, Birkhoff, Tukey, and Kelley. The theory was popularized by Kelley's textbook [1955/1975]. We shall say that $(y_\beta : \beta \in \mathbb{B})$ is a *Kelley subnet* of $(x_\alpha : \alpha \in \mathbb{A})$ if we can write $y_\beta = x_{\varphi(\beta)}$ or $y_\beta = x_{\varphi_\beta}$ for a function $\varphi : \mathbb{B} \to \mathbb{A}$ satisfying certain technical conditions discussed in 7.15.b below. A slight variant on Kelley's definition was given by Willard [1970]; we present it in 7.15.c.

While Kelley et al. were investigating nets, several other mathematicians — notably Cartan and Bourbaki — were developing an analogous theory of filters. Soon it became clear that the two systems of ideas yielded the same kinds of conclusions about uniform convergence, compactness, weak topologies, etc. Each system offered certain advantages: Nets look more like sequences and thus appeal more to the intuition of analysts; filters are amenable to arguments involving elementary set-theoretic operations and the Ultrafilter Principle. However, the two systems were not easily interchangeable; there was some awkwardness in the translation. Most mathematicians in convergence theory ended up using either nets or filters, but not both.

The difficulty is removed by a more general approach to subnets that has been suggested independently by several mathematicians (Smiley [1957], Aarnes and Andenæs [1972], Murdeshwar [1983], and perhaps others) but which, nevertheless, seems not to be widely known yet. We shall name this approach after Aarnes and Andenæs, because they investigated it in greatest depth. The Aarnes and Andenæs (AA) approach moved further away from the original notion of subsequence, and dispensed altogether with the connecting function $\varphi : \mathbb{B} \to \mathbb{A}$. Kelley's definition related two nets $x : \mathbb{A} \to X$ and $y : \mathbb{B} \to X$ by their behavior in the domains $\mathbb{A}$ and $\mathbb{B}$, but the AA approach relates the nets by their behavior in the codomain X. This approach makes nets and filters easily interchangeable, thus offering mathematicians the advantages of both systems.

7.15. *Definitions.* Let $(x_\alpha : \alpha \in \mathbb{A})$ and $(y_\beta : \beta \in \mathbb{B})$ be nets in a set X, with eventuality filters $\mathcal{F}$ and $\mathcal{G}$, respectively. Then:

a. The following conditions are equivalent. If any (hence all) of them are satisfied, we shall say that (y_β) is a **subnet** of (x_α) (or more precisely, an **AA subnet**, or a **subnet in the sense of Aarnes and Andenæs**).

(A) Every (y_β)-frequent subset of X is also (x_α)-frequent. That is, if $y_\beta \in S$ for arbitrarily large values of β, then $x_\alpha \in S$ for arbitrarily large values of α.

(B) Every (x_α)-eventual subset of X is also (y_β)-eventual. That is, if $x_\alpha \in S$ for all sufficiently large values of α, then $y_\beta \in S$ for all sufficiently large β.

(C) $\mathcal{G} \supseteq \mathcal{F}$. (In other words, *an AA subnet corresponds to a superfilter.*)

(D) Each (x_α)-tail set contains some (y_β)-tail set. In other words, for each $\alpha_0 \in \mathbb{A}$ there is some $\beta_0 \in \mathbb{B}$ such that $\{y_\beta : \beta \succcurlyeq \beta_0\} \subseteq \{x_\alpha : \alpha \succcurlyeq \alpha_0\}$.

(E) For each eventual set $\mathbb{S} \subseteq \mathbb{A}$, the set $y^{-1}(x(\mathbb{S}))$ is eventual in $\mathbb{B}$.

b. We shall say (y_β) is a **Kelley subnet** of (x_α) if there exists a function $\varphi : \mathbb{B} \to \mathbb{A}$ such that

(i) $y = x \circ \varphi$ — that is, $y_\beta = x_{\varphi(\beta)}$ for all $\beta \in \mathbb{B}$ and

(ii) for each eventual set $\mathbb{S} \subseteq \mathbb{A}$, the set $\varphi^{-1}(\mathbb{S})$ is eventual in $\mathbb{B}$.

Condition (ii) can be restated in either of these equivalent forms:

(ii′) For each $\alpha_0 \in \mathbb{A}$, there is some $\beta_0 \in \mathbb{B}$ such that $\beta \succcurlyeq \beta_0 \Rightarrow \varphi(\beta) \succcurlyeq \alpha_0$.

(ii″) The $\mathbb{A}$-valued net $\varphi : \mathbb{B} \to \mathbb{A}$ is an Aarnes-Andenæs subnet of the identity map $i_{\mathbb{A}} : \mathbb{A} \to \mathbb{A}$.

c. Willard [1970] modified Kelley's definition slightly, adding a requirement of monotonicity; this may make the definition more palatable to many readers. We shall say (y_β) is a **Willard subnet** of (x_α) if there exists a function $\varphi : \mathbb{B} \to \mathbb{A}$ such that

(i) $y = x \circ \varphi$ — that is, $y_\beta = x_{\varphi(\beta)}$ for all $\beta \in \mathbb{B}$;

(ii) φ is monotone; that is, $\beta_1 \preccurlyeq \beta_2 \Rightarrow \varphi(\beta_1) \preccurlyeq \varphi(\beta_2)$; and

(iii) for each $\alpha_0 \in \mathbb{A}$ there is some $\beta_0 \in \mathbb{B}$ such that $\varphi(\beta_0) \succcurlyeq \alpha_0$.

7.16. *Comparison of the definitions.*

a. Show that any Kelley subnet is also an Aarnes-Andenæs subnet.

The converse is not valid. For instance, each of the sequences $(0, 5, 6, 7, 8, \ldots)$ and $(1, 5, 6, 7, 8, \ldots)$ is an AA subnet of the other, but neither is a Kelley subnet of the other.

b. Show that any Willard subnet is also a Kelley subnet.

The converse is not valid. For instance, each of the sequences $(2, 1, 4, 3, 6, 5, \ldots)$ and $(1, 2, 3, 4, 5, 6, \ldots)$ is a Kelley subnet of the other, but neither is a Willard subnet of the other.

c. Suppose $(x_\alpha : \alpha \in \mathbb{A})$ is a net in a set X and $\mathbb{F}$ is a frequent subset of the directed set $\mathbb{A}$. Then $\mathbb{F}$ is a directed set (see 7.10.f), and so $(x_\alpha : \alpha \in \mathbb{F})$ is a net. We shall say that it is a **frequent subnet** of the net $(x_\alpha : \alpha \in \mathbb{A})$. (In some of the literature, this is called a **cofinal subnet**.)

Show that any frequent subnet is a Willard subnet (by using the inclusion map $i : \mathbb{F} \xrightarrow{\subseteq} \mathbb{A}$ for the map φ in definition 7.15.c).

The converse is not valid. For instance, show that $(1, 1, 2, 2, 3, 3, \ldots)$ is a Willard subnet, but not a frequent subnet, of the sequence $(1, 2, 3, \ldots)$.

Frequent subnets cannot be used interchangeably with Willard, Kelley, or AA subnets; see 17.29.

d. Frequent subnets are a generalization of subsequences.

Let $(x_m : m \in \mathbb{N})$ and $(y_n : n \in \mathbb{N})$ be two sequences. Show that (y_n) is a subsequence of (x_m) if and only if (y_n) is a frequent subnet of (y_n).

7.17. *Further elementary properties.*

a. *Composition of subnets.* If (z_γ) is a subnet of (y_β), and (y_β) is a subnet of (x_α), then (z_γ) is a subnet of (x_α).

If the two given subnets are Kelley subnets, Willard subnets, or frequent subnets, then then (z_γ) is the same type of subnet of (x_α).

b. Suppose that $(x_\alpha : \alpha \in \mathbb{A})$ is a net in a set X and (x_α) is eventually in some set of the form $E = E_1 \cup E_2 \cup \cdots \cup E_n \subseteq X$. Then there is at least one j such that (x_α) is frequently in E_j. Thus (x_α) has a frequent subnet that takes all its values in E_j.

c. *Definition.* Two nets have the same eventuality filter if and only if each net is a subnet of the other. We shall then say the nets are **AA-equivalent**, or simply **equivalent**.

7.18. Lemma on Common Subnets. Let $(u_\alpha : \alpha \in \mathbb{A})$, $(v_\beta : \beta \in \mathbb{B})$, and $(w_\gamma : \gamma \in \mathbb{C})$ be three nets taking values in a set X. Say the nets have eventuality filters $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{H}$, respectively. Then the following conditions are equivalent:

(A) $F \cap G \cap H$ is nonempty, for every $F \in \mathcal{F}$, $G \in \mathcal{G}$, $H \in \mathcal{H}$.

(B) $\mathcal{M} = \{S \subseteq X : S \supseteq F \cap G \cap H \text{ for some } F \in \mathcal{F},\ G \in \mathcal{G},\ H \in \mathcal{H}\}$ is a proper filter.

(C) The three filters have a common proper superfilter — i.e., there exists a proper filter which contains all three given filters.

(D) The three nets have a common AA subnet — i.e., there exists a net (p_λ) which is an AA subnet of each of the given nets.

(E) The three given nets have a common Willard subnet — i.e., there exists a net $(p_\lambda : \lambda \in \mathbb{L})$ which is a Willard subnet of each of the three given nets. (It is understood that three different functions are used for the monotone mappings φ from $\mathbb{L}$ into $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$.) Furthermore, that net can be chosen so that it is a **maximal common AA subnet** of the three given nets — i.e., so that if (q_μ) is any common AA subnet of the three given nets, then (q_μ) is also an AA subnet of (p_λ).

Note. We have stated the lemma in terms of three nets and three filters to display a typical case. The number 3 may be replaced by any positive integer.

Proof of lemma. The equivalence of (C) and (D) is immediate from our correspondence between AA subnets and superfilters. The implications (C) $\Rightarrow$ (A) $\Rightarrow$ (B) $\Rightarrow$ (C) are easy; the implication (E) $\Rightarrow$ (D) is trivial. It suffices to show that (A)–(D) together imply (E). Note that the filter $\mathcal{M}$ in condition (B) is a a minimum common superfilter — i.e., it is the smallest filter containing all of the given filters. Any net corresponding to it is a maximal common AA subnet of the three given nets. It suffices to exhibit a net $(p_\lambda : \lambda \in \mathbb{L})$ whose eventuality filter is $\mathcal{M}$, such that $(p_\lambda : \lambda \in \mathbb{L})$ is a Willard subnet of each of the three given nets.

For each $(a, b, c) \in \mathbb{A} \times \mathbb{B} \times \mathbb{C}$, define

$$\begin{aligned} T_{a,b,c} &= \{u_\alpha : \alpha \succcurlyeq a\} \cap \{v_\beta : \beta \succcurlyeq b\} \cap \{w_\gamma : \gamma \succcurlyeq c\} \\ &= \{x \in X \ : \ x = u_\alpha = v_\beta = w_\gamma \text{ for some } \alpha \succcurlyeq a,\ \beta \succcurlyeq b,\ \gamma \succcurlyeq c\}. \end{aligned}$$

Then $T_{a,b,c}$ is nonempty, by condition (A). Hence

$$\mathbb{L} \quad = \quad \{(\alpha, \beta, \gamma) \in \mathbb{A} \times \mathbb{B} \times \mathbb{C} \quad : \quad u_\alpha = v_\beta = w_\gamma\}$$

is a frequent subset of $\mathbb{A} \times \mathbb{B} \times \mathbb{C}$, when $\mathbb{A} \times \mathbb{B} \times \mathbb{C}$ is given the product ordering. For each $\lambda = (\alpha, \beta, \gamma)$ in $\mathbb{L}$, define $p_\lambda = u_\alpha = v_\beta = w_\gamma$; the remaining verifications are easy. For the monotone mappings φ from $\mathbb{A} \times \mathbb{B} \times \mathbb{C}$ into $\mathbb{A}, \mathbb{B}, \mathbb{C}$, use the coordinate projections.

7.19. Corollary on equivalent subnets. If (y_β) is an AA subnet of (x_α), then (y_β) is equivalent (in the sense of 7.17.c) to a Willard subnet of (x_α).

Hints: The two given nets have a common AA subnet — namely, (y_β). As in 7.18(E), let (p_λ) be a common Willard subnet and also a maximal common AA subnet of the two given nets. Since (y_β) has the property for which (p_λ) is maximal, (y_β) is an AA subnet of (p_λ). Thus (y_β) and (p_λ) are subnets of each other.

Remarks. A similar result is given by Gähler [1977].

We have seen that every Willard subnet is a Kelley subnet, every Kelley subnet is an AA subnet, and every AA subnet is equivalent to a Willard subnet. Consequently, the three types of subnets can be used interchangeably in many contexts. See especially 15.38.

7.20. Though AA subnets are simpler than Kelley subnets in most respects, Kelley subnets do have at least one advantage, which we now present in two formulations:

(1) Suppose that $f : X \to V$ is some function, $(x_\alpha : \alpha \in A)$ is some net in X, and $(y_\beta : \beta \in B)$ is some Kelley subnet of the net $(f(x_\alpha) : \alpha \in A)$ in V. Then $(x_\alpha : \alpha \in A)$ has a Kelley subnet $(s_\beta : \beta \in B)$ in X such that $f(s_\beta) = y_\beta$ for each β. (Indeed, if $y_\beta = f(x_{\varphi(\beta)})$, take $s_\beta = x_{\varphi(\beta)}$.)

(2) Suppose that $((u_\alpha, v_\alpha) : \alpha \in A)$ is a net in some product of sets $U \times V$; then $(v_\alpha : \alpha \in A)$ is some net in V. Suppose that $(y_\beta : \beta \in B)$ is some Kelley subnet of the net $(v_\alpha : \alpha \in A)$ in V. Then $((u_\alpha, v_\alpha) : \alpha \in A)$ has a Kelley subnet $((p_\beta, q_\beta) : \beta \in B)$ such that $q_\beta = y_\beta$ for each β. (Indeed, if $q_\beta = y_\beta = v_{\varphi(\beta)}$, take $p_\beta = u_{\varphi(\beta)}$.)

These are actually two formulations of the same principle. To see this, observe that if $f, (x_\alpha), (y_\beta)$ are given as in (1), then we can reformulate the problem as in (2) by taking $X = U$ and $(u_\alpha, v_\alpha) = (x_\alpha, f(x_\alpha))$. Conversely, if we are given (u_α, v_α) as in (2), then we can reformulate the problem as in (1) by taking $X = U \times V$ and $x_\alpha = (u_\alpha, v_\alpha)$, and letting $f : X \to V$ be the projection onto the second coordinate.

7.21. Some properties of nets are **subnet hereditary**, in the sense that if a net has the property, then so does every subnet. For instance, we shall see in later chapters that in a topological space, every subnet of a convergent net is convergent.

Likewise, some properties are **supernet hereditary**, in the sense that if a net has the property, then so does every supernet. For instance, in a topological space, the property of *not* being convergent is supernet hereditary.

Many proofs with nets involve such hereditary properties. Consequently, in many proofs it is possible to replace a given net with any convenient subnet, or with any convenient supernet.

Some proofs use the phrase "**we may assume**," particularly in connection with hereditary properties. In many cases, what this means is that by *relabeling*, we may *replace* the given net with some subnet or supernet that has an additional property of interest. See the related discussion in 1.10.

Universal Nets

7.22. *Definition.* A **universal net** (also occasionally known as an *ultranet*) in a set X is a net (x_δ) with the property that for each set $S \subseteq X$, either (i) eventually $x_\delta \in S$ or (ii) eventually $x_\delta \in X \backslash S$.

7.23. *Example.* Let (x_δ) be a net in X. Assume (x_δ) is **eventually constant**; i.e., assume there exists some $z \in X$ such that eventually $x_\delta = z$. Then (x_δ) is a universal net.

Although other universal nets exist, other explicit *examples* of universal nets do *not* exist! That is explained below.

7.24. *Observations.* A net (x_δ) is universal if and only if its eventuality filter is an ultrafilter. If a net is universal, then any AA-equivalent net is also universal; by 7.19, therefore, in the discussions below it does not matter whether we use Willard subnets, Kelley subnets, or AA subnets. A net (x_δ) is eventually equal to some constant x if and only if its eventuality filter is the fixed ultrafilter at x.

Thus, the theory of universal nets is simply a reformulation of the theory of ultrafilters. The Ultrafilter Principle, introduced in 6.32, can be reformulated as

> **(UF3) Universal Subnet Theorem (Tukey, Kelley).** Every net has a subnet that is universal.

Likewise, the Weak Ultrafilter Theorem, presented in 6.33, can be reformulated as

> **(WUF′) Weak Universal Subnet Theorem.** There exists a universal net in $\mathbb{N}$ that is not eventually constant.

As we remarked in 6.33, free ultrafilters are intangibles. The same is therefore also true of universal nets that are not eventually constant. Though we have no explicit examples of these peculiar nets, nevertheless they are useful conceptual tools for some kinds of reasoning.

7.25. *Further properties of universal nets.*

a. If (x_α) is a universal net, then any subnet of (x_α) is AA-equivalent to (x_α) and is also universal.

b. If (x_α) is a universal net in X and x_α is frequently in some set $S \subseteq X$, then x_α is eventually in S.

c. If a net $(x_\alpha : \alpha \in \mathbb{A})$ is not universal, then $\mathbb{A}$ has two disjoint frequent sets $\mathbb{A}_1, \mathbb{A}_2$ such that the resulting frequent subnets $(x_\alpha : \alpha \in \mathbb{A}_j)$ have disjoint ranges.

d. If (x_δ) is a universal net in a set $X = S_1 \cup S_2 \cup \cdots \cup S_n$, then there is at least one j such that eventually $x_\delta \in S_j$. *Hint*: 5.8(E).

e. If (x_δ) is a universal net in a *finite* set X, then (x_δ) is eventually constant.

f. If (x_n) is a universal net that is a *sequence*, then it is eventually constant. *Hint*: If (x_n) has infinite range, then that range can be partitioned into two disjoint infinite sets. Then what?

g. If (x_δ) is a universal net in a set X and $f : X \to Y$ is any function, then $(f(x_\delta))$ is a universal net in Y.

h. Let $(x_\delta : \delta \in \mathbb{D})$ be a net in some set X, and consider its range $R = \{x_\delta : \delta \in \mathbb{D}\}$. Show that (x_δ) is a universal net in X if and only if (x_δ) is a universal net in R. Thus, the universality of a net (x_δ) in a set X does not depend on the choice of X, as long as X is large enough to contain all the points of that net.

More about Subsequences

7.26. Lemma. Let (v_m) and (y_m) be sequences in a set X. Then v is an AA subnet of y if and only if these two conditions are satisfied:

(i) $\text{Range}(v) \setminus \text{Range}(y)$ is a finite subset of X, and

(ii) for each $r \in X$, if $y^{-1}(r)$ is a finite subset of $\mathbb{N}$, then $v^{-1}(r)$ is also a finite subset of $\mathbb{N}$.

Proof. This argument is from Aarnes and Andenæs [1972]. First, suppose that v is an AA subnet of y. Then y is eventually in the set $\text{Range}(y)$, hence v is eventually in that set, hence $v_j \notin \text{Range}(y)$ for only finitely many values of j; this implies (i). For condition (ii), suppose that $y^{-1}(r)$ is a finite set; then y is eventually in $X \setminus \{r\}$; then v is also eventually in that set; hence $v^{-1}(r)$ is also a finite set.

Conversely, suppose conditions (i) and (ii) are satisfied. Let S be a subset of X such that eventually $y \in S$; we are to show that eventually $v \in S$. For each $r \in X \setminus S$, the set $y^{-1}(r)$ is finite; hence the set $v^{-1}(r)$ is finite. The sets $\text{Range}(y) \setminus S$ and $\text{Range}(v) \setminus \text{Range}(y)$ are finite; hence the set $\text{Range}(v) \setminus S$ is finite, and it is a subset of $X \setminus S$. Therefore the set $F = \bigcup_{r \in \text{Range}(v) \setminus S} v^{-1}(r)$ is finite. For k sufficiently large, we have $k \notin F$. For all such k, we have $v_k \in S$.

7.27. Theorem on equivalent subsequences. Let (x_i) and (y_j) be sequences in a set X, and assume that (y_j) is an AA subnet of (x_i). Then (y_j) is AA-equivalent to some subsequence of (x_i).

Proof. This argument is from Aarnes and Andenæs [1972]. Since $\text{Range}(x)$ is an eventual set for $y(\cdot)$, by discarding the first few terms of (y_j) we may assume without loss of generality that $y(\mathbb{N}) \subseteq x(\mathbb{N})$. For each $r \in y(\mathbb{N})$, the set $y^{-1}(r)$ is nonempty; hence the set $x^{-1}(r)$ is nonempty. For such r, define a set $\mathbb{A}_r \subseteq \mathbb{N}$ as follows:

- If $y^{-1}(r)$ is an infinite set, then $x^{-1}(r)$ is also an infinite set; in this case let $\mathbb{A}_r = x^{-1}(r)$.
- If $y^{-1}(r)$ is a finite set, let $\mathbb{A}_r$ be some nonempty finite subset of $x^{-1}(r)$. (Any such set will do for the purposes of this proof. If the reader desires a *canonical* choice of $\mathbb{A}_r$, let $\mathbb{A}_r$ be the singleton whose sole member is the first member of $x^{-1}(r)$.)

In either case we obtain $x(\mathbb{A}_r) = \{r\}$. Now let $\mathbb{A} = \bigcup_{r \in y(\mathbb{N})} \mathbb{A}_r$; then $\mathbb{A}$ is an infinite subset of $\mathbb{N}$. Say its members are, in increasing order,

$$a_1 < a_2 < a_3 < \cdots .$$

Define $v_k = x_{a_k}$; then (v_k) is a subsequence of (x_i). It is clear from the definitions above that $v(\mathbb{N}) = x(\mathbb{A}) = y(\mathbb{N})$. Also, for each $r \in y(\mathbb{N})$, the sets $v^{-1}(r)$ and $y^{-1}(r)$ are both finite or both infinite. By 7.26, v and y are AA subnets of each other.

7.28. (*Optional.*) There are a few minor differences between Aarnes-Andenæs subnets and Kelley subnets; here is one of them. Let X be a given nonempty set. Does every net in X have at least one subnet that is a sequence?

a. No, if we use Kelley subnets. Indeed, take $\mathbb{A} = \mathbb{N}^{\mathbb{N}}$ with the product ordering, and take x to be *any* function from $\mathbb{A}$ into X; then no Kelley subnet of $(x_\alpha : \alpha \in \mathbb{A})$ is a sequence.

Hint: If $(\alpha_1, \alpha_2, \alpha_3, \ldots)$ is a sequence in $\mathbb{A}$, then each α_j is itself a sequence of positive integers. Say $\alpha_j = (m_{1j}, m_{2j}, m_{3j}, \ldots)$. Then there is no j for which $\alpha_j \succcurlyeq (m_{11}+1, m_{22}+1, m_{33}+1, \ldots)$.

b. Yes, if we use Aarnes-Andenæs subnets and X is a finite set. Indeed, by 7.10.e there is at least one $x_0 \in X$ such that frequently $x_\alpha = x_0$. Then the constant sequence $(x_0, x_0, x_0, \ldots)$ is an AA subnet of (x_α).

c. No, if X is an infinite set, regardless of which type of subnets we use. Indeed, let $\mathcal{U}$ be any free ultrafilter on X. (The existence of such an ultrafilter was established in 6.33.) Let (x_α) be a corresponding net; thus (x_α) is a universal net that is not eventually constant. If some sequence (y_m) is a subnet of (x_α), then (y_m) has the same eventuality filter $\mathcal{U}$, hence (y_m) is universal and not eventually constant — contradicting 7.25.f.

7.29. Theorem. Let X be a chain ordered set (e.g., the real line). Then any sequence in X has a monotone subsequence.

Proof (Thurston [1994]). By a maximal element of a sequence we shall mean a maximal element of the range of that sequence. It is easy to see that if s is a sequence that has no maximal element, then s has an increasing subsequence.

Now let $s = (x_1, x_2, x_3, \ldots)$ be a given sequence; we may assume that every subsequence of s has a maximal element. Let $x_{n(1)}$ be a maximal element of s. Let $x_{n(2)}$ be a maximal element of $(x_{n(1)+1}, x_{n(1)+2}, x_{n(1)+3}, \ldots)$. Let $x_{n(3)}$ be a maximal element of $(x_{n(2)+1}, x_{n(2)+2}, x_{n(2)+3}, \ldots)$. Continuing in this fashion, we obtain positive integers $n(1) < n(2) < n(3) < \cdots$ satisfying $x_{n(1)} \geq x_{n(2)} \geq x_{n(3)} \geq \cdots$.

Convergence Spaces

7.30. By a **convergence space** (or *limit space*) we shall mean a set X equipped with a function

$$\lim \quad : \quad \{\text{proper filters on } X\} \quad \longrightarrow \quad \{\text{subsets of } X\}.$$

Any function can be used for lim in this definition, but in most cases of interest the function is determined by some structure already given on X — a topology, an ordering, a measure, etc.

We emphasize that the value of lim is a subset of X. In some convergence spaces (e.g., the one used in college calculus), the set $\lim \mathcal{F}$ contains at most one point of X; such convergence spaces are discussed further in 7.36.

7.31. Whenever $(X, \lim)$ is a convergence space, then we shall extend the function lim in the following ways:

(a) If $\mathcal{B}$ is a filterbase on X, then $\lim \mathcal{B} = \lim \mathcal{F}$ where $\mathcal{F}$ is the filter generated by $\mathcal{B}$.

(b) If (x_α) is a net in X, then $\lim(x_\alpha) = \lim \mathcal{F}$ where $\mathcal{F}$ is the eventuality filter of (x_α).

Note that the resulting "function" $\lim : \{\text{nets in } X\} \to \{\text{subsets of } X\}$ is not a function strictly in the sense of 1.31, since $\{\text{nets in } X\}$ is a proper class, not a set. Note, also, that this function satisfies the following condition:

(*) if (x_α) and (y_β) are nets with the same eventuality filter (i.e., if (x_α) and (y_β) are AA-equivalent), then the set of limits of (x_α) is equal to the set of limits of (y_β).

Conversely, we have this alternate definition:

> A **convergence space** is a set X that is equipped with some function $\lim : \{\text{nets in } X\} \to \{\text{subsets of } X\}$ that satisfies (*).

Indeed, if (*) is satisfied, then (b) defines a corresponding limit function on the collection of all proper filters on X. (In many applications, we verify (*) by verifying a stronger property described below in 7.34.b.)

Thus, in convergence spaces we may use nets and their eventuality filters interchangeably (and use AA subnets and superfilters interchangeably, as well). Each type of object has its advantages, as we noted in 7.13.

Remarks. For more general theories of convergences than those considered in this book, see: Bentley, Herrlich, and Lowen-Colebunders [1970] for categories of convergence spaces; Dolecki and Greco [1986] for algebraic properties of collections of convergence structures; and Gähler [1984] for "convergence spaces" that are more general than "filter convergence spaces." In Kelley's [1955/1975] book, net convergences are considered in great generality, without condition (*) being imposed a priori; see the remarks in 15.10.

7.32. *More notations.* If $\mathcal{F}$ is a proper filter or a net, the expression $z \in \lim \mathcal{F}$ will be read as "z is a **limit** of $\mathcal{F}$." It may also be written as $\mathcal{F} \to z$ and read as "$\mathcal{F}$ **converges** to z." The statement "$\mathcal{F}$ does not converge to z" may be written as $z \notin \lim \mathcal{F}$, or as $\mathcal{F} \not\to z$.

Many variants on these notations can be used for clarification. For instance, for a net $(x_\alpha : \alpha \in \mathbb{A})$, the expression $x_\alpha \to z$ may also be written as "$x_\alpha \to z$ in X as α increases in $\mathbb{A}$." When two or more convergences are being considered, we may use a prefix or subscript or superscript to distinguish them; for instance, we may write

$$z \in \mathcal{T}\text{-}\lim x_\alpha \qquad \text{or} \qquad z \in \lim_{\mathcal{T}} x_\alpha \qquad \text{or} \qquad x_\alpha \xrightarrow{\mathcal{T}} z$$

to indicate that z is a limit of the net (x_α) when we use the convergence function determined by some structure $\mathcal{T}$, rather than some other structure $\mathcal{S}$. Other variants on the notation should be clear from the context; we shall not attempt to list them all here.

7.33. Let $p : X \to Y$ be a mapping from one convergence space into another. We shall say p is **convergence preserving** if it has this property:

> whenever (x_α) is a net converging to a limit x in X, then the net $(p(x_\alpha))$ converges to $p(x)$ in Y

or, equivalently,

> whenever $\mathcal{F}$ is a filter converging to a limit x in X, then the filter $\{S \subseteq Y : p^{-1}(S) \in \mathcal{F}\}$ converges to $p(x)$ in Y.

(*Exercise.* Prove the equivalence.) Observe that the composition of two convergence preserving maps is convergence preserving; this is discussed further in 9.7.

7.34. *Definitions.* Most convergence spaces of interest satisfy both of the properties below; in fact, these properties are satisfied by all the convergence spaces that we shall consider in this book. (Some mathematicians make one or both of these properties a part of their definition of convergence space.)

a. A convergence space is **centered** if it has the property that

> if $\mathcal{U}_z$ is the ultrafilter fixed at z, then $\mathcal{U}_z \to z$,

or, equivalently,

> if (x_α) is a net such that eventually $x_\alpha = z$, then $x_\alpha \to z$.

b. A convergence space is **isotone** if it has this property:

> if $\mathcal{G}$ is a superfilter of $\mathcal{F}$, and $\mathcal{F} \to z$, then $\mathcal{G} \to z$

or, equivalently,

> if (y_β) is a subnet of (x_α), and $x_\alpha \to z$, then $y_\beta \to z$.

In the last sentence, it does not matter which type of subnet we use — Willard, Kelley, or AA — since we have built condition (*) of 7.31 into our definition of convergence space. On the other hand, for AA subnets the isotonicity condition above implies condition (*) of 7.31.

7.35. *Exercise.* Let X be an isotone convergence space. If (x_α) is a universal net and some subnet of (x_α) converges to z, then $x_\alpha \to z$ also.

7.36. A convergence space $(X, \lim)$ is **Hausdorff** if each net or proper filter F has at most one limit — i.e., if each set of the form $\lim F$ contains at most one member.

When $(X, \lim)$ is a Hausdorff convergence space, then $z \in \lim F$ may be rewritten as $z = \lim F$; we say that z is *the* limit of F. (Now the notation should begin to look more like that of college calculus.) In effect, our original limit function — which took values in {subsets of X} — is replaced by a new function, again denoted by "lim," which takes values in X. Thus, we are not asserting that $z = \{z\}$. The distinction between the two different lim functions should be clear in most contexts and should not cause any confusion.

Most convergence spaces or topological spaces in applications are Hausdorff, and so some mathematicians incorporate the Hausdorff condition into other definitions — e.g., they make it a part of their definition of convergence space, compact space, gauge space, completely regular space, topological linear space, or locally convex space. We shall not follow that practice, for many of the concepts in this book are revealed more clearly if Hausdorffness is treated as a separate property. It is often helpful to analyze Hausdorff spaces in terms of other, simpler spaces that are not Hausdorff (see 15.25.d). Throughout this text, Hausdorffness will be assumed only when stated explicitly.

More notation. If X and Y are convergence spaces and Y is Hausdorff, then the equation

$$\lim_{x \to x_0} f(x) = y_0$$

is a condition on x_0, y_0, and f, with the following meaning: Whenever (x_α) is a net in $X \setminus \{x_0\}$ that converges in X to x_0, then $f(x_\alpha) \to y_0$ in Y. Most limits in college calculus are of this form — in some cases with x_0 or y_0 equal to ∞. Making ∞ a member of our convergence space is not particularly difficult; see the discussions in 5.15.f, 5.15.g, 18.24.

Convergence in Posets

7.37. *Remarks.* The two most important kinds of convergences are the topological convergences, studied in Chapter 15, and the order convergences, studied in the remainder of this chapter. The most important type of order convergence needed by analysts is the order convergence in $\mathbb{R}$; that special case should be kept in mind by the reader at all times throughout the remainder of this chapter. However, many of the basic properties of order convergence in $\mathbb{R}$ generalize readily to other settings that are occasionally useful. Thus, we begin our study of order convergence in a setting that has as few hypotheses as possible: the setting of partially ordered sets.

7.38. The literature contains several different, inequivalent definitions of convergence in partially ordered sets. The following one works best for our purposes, despite its complexity. It can be restated in other ways that are sometimes more convenient; see 7.40.d and, in special contexts, 7.41 and 7.45.

Definition. Let $(X, \preccurlyeq)$ be a poset. Let $z \in X$, and let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in X. We shall say that (x_α) is **order convergent** to z (sometimes written $x_\alpha \xrightarrow{o} z$) if

> there exist nonempty sets $S, T \subseteq X$ such that $(S, \preccurlyeq)$ and $(T, \succcurlyeq)$ are directed sets, $\sup(S)$ and $\inf(T)$ both exist in X and are equal to z, and for each fixed $s \in S$ and $t \in T$ we have eventually $s \preccurlyeq x_\alpha \preccurlyeq t$.

(We emphasize that T is to be a directed set when we *reverse* the restriction of the given ordering. Thus, each finite subset of S must have an upper bound in S, and each finite subset of T must have a lower bound in T.)

7.39. *Definitions.* Let $(x_\alpha : \alpha \in \mathbb{A})$ be a net taking values in a partially ordered set $(X, \preccurlyeq)$. We say that (x_α) is **increasing** if

$$\alpha \preccurlyeq \beta \qquad \Rightarrow \qquad x_\alpha \preccurlyeq x_\beta.$$

This may be abbreviated $x_\alpha \uparrow$. We say that (x_α) **increases to a limit z**, denoted $x_\alpha \uparrow z$, if in addition $z = \sup\{x_\alpha : \alpha \in \mathbb{A}\}$.

Analogously, a net $(x_\alpha : \alpha \in \mathbb{A})$ is **decreasing** (written $x_\alpha \downarrow$) if $\alpha \preccurlyeq \beta \Rightarrow x_\alpha \succcurlyeq x_\beta$; the net **decreases to a limit z** (written $x_\alpha \downarrow z$) if in addition $z = \inf\{x_\alpha : \alpha \in \mathbb{A}\}$.

A net is **monotone** if it is increasing or decreasing.

7.40. *Exercises.* Let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in a poset $(X, \preccurlyeq)$, and let $z \in X$. Then:

a. $x_\alpha \uparrow z$ if and only if (x_α) is increasing and $x_\alpha \xrightarrow{o} z$ (in the sense of 7.38).

b. $x_\alpha \downarrow z$ if and only if (x_α) is decreasing and $x_\alpha \xrightarrow{o} z$ (in the sense of 7.38).

c. In a complete lattice, any monotone net converges.

d. *Order convergence in terms of monotone convergence*: $x_\alpha \xrightarrow{o} z$ (defined as in 7.38) if and only if

> there exist nets $(u_\beta : \beta \in \mathbb{B})$ and $(v_\gamma : \gamma \in \mathbb{C})$ such that $u_\beta \uparrow z$ and $v_\gamma \downarrow z$, and for each fixed β and γ we have α-eventually $x_\alpha \in \{x : u_\beta \preccurlyeq x \preccurlyeq v_\gamma\}$.

Hints: For the "if" part, let S and T be the ranges of those nets (u_β) and (v_γ). For the "only if" part, let (u_β) and (v_γ) be given by the identity maps on the sets $\mathbb{B} = S$ and $\mathbb{C} = T$.

e. Order convergence is centered and isotone.

f. *Convergence preserves inequalities.* Suppose $(x_\alpha : \alpha \in \mathbb{A})$ and $(y_\alpha : \alpha \in \mathbb{A})$ are nets based on the same directed set, satisfying $x_\alpha \preccurlyeq y_\alpha$ for all α. If $x_\alpha \xrightarrow{o} x_\infty$ and $y_\alpha \xrightarrow{o} y_\infty$, then $x_\infty \preccurlyeq y_\infty$.

Hint: Let S^x and T^x be two sets that satisfy the conditions in 7.38 that define the convergence $x_\alpha \xrightarrow{o} x_\infty$. Also, let S^y and T^y be two sets that satisfy the conditions in 7.38 that define the convergence $y_\alpha \xrightarrow{o} y_\infty$. Fix any $s^x \in S^x$ and $t^y \in T^y$; then we have eventually $s^x \preccurlyeq x_\alpha \preccurlyeq y_\alpha \preccurlyeq t^y$, and thus $s^x \preccurlyeq t^y$. Use that fact to prove that $\sup(S^x) \preccurlyeq \inf(T^y)$.

g. Order convergence is Hausdorff. Thus, the statement $x_\alpha \xrightarrow{o} z$ may be rewritten as $z = \text{o-}\lim x_\alpha$. *Hint*: Apply the preceding result with $x_\alpha = y_\alpha$.

h. Let $(X, \preccurlyeq)$ and $(Y, \preccurlyeq)$ be posets. (Here we use the same symbol $\preccurlyeq$ for two different partial orderings.) Let $f : X \to Y$ be some function that is sup-preserving and inf-preserving (see 3.22). Then f is also convergence-preserving (see 7.33), if X and Y are equipped with their order convergences.

Hint: First show that f preserves the convergence of monotone sequences — i.e., the convergences described in 7.39; then use 7.40.d. *Remark*: The assumptions cannot be weakened substantially; in 15.45 we give a partial converse.

i. The "squeeze theorem." Suppose $(x_\alpha : \alpha \in \mathbb{A})$, $(y_\alpha : \alpha \in \mathbb{A})$, $(z_\alpha : \alpha \in \mathbb{A})$ are nets based on the same directed set, satisfying $x_\alpha \preccurlyeq y_\alpha \preccurlyeq z_\alpha$ for all α. If $x_\alpha \xrightarrow{o} w$ and $z_\alpha \xrightarrow{o} w$, then also $y_\infty \xrightarrow{o} w$. (*Remark.* Compare with 26.52(E).)

7.41. Theorem on convergence in chains. Let $(X, \leq)$ be a chain. Let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in X, and let $z \in X$. Then $x_\alpha \xrightarrow{o} z$ (that is, order convergence, as defined in 7.38 or characterized in 7.40.d) if and only if these two conditions are satisfied for all σ and τ in X:

(i) if $z > \sigma$, then eventually $x_\alpha > \sigma$, and

(ii) if $z < \tau$, then eventually $x_\alpha < \tau$.

Remarks. Note that condition (i) is satisfied vacuously (i.e., for free) if z happens to be the largest element of X, for then there is *no* element τ that satisfies $z < \tau$. Likewise; condition (ii) is satisfied vacuously if z happens to be the smallest element of X. Considering the examples of $[-\infty, +\infty]$, $[0, +\infty)$, $\mathbb{R}$, we see that some chains have both a largest and smallest element, some chains have one or the other, and some chains have neither.

Proof of equivalence. It is an easy exercise that order convergence implies conditions (i) and (ii); we omit the details. Conversely, assume that (x_α) and z satisfy condition (i) above; we shall find a set T satisfying the conditions of 7.38. (Forming S from (ii) is similar.) If the net (x_α) satisfies eventually $x_\alpha \leq z$, then the singleton $T = \{z\}$ satisfies the requirements for 7.38, and we are done. Assume, therefore, that the net (x_α) does not satisfy eventually $x_\alpha \leq z$. Let $T = \{t \in X \ : \ t > z\}$; we shall show that this set satisfies the requirements. We have frequently $x_\alpha \in T$, and so T is nonempty. From condition (i) we see that for each $t \in T$, eventually $x_\alpha < t$. It suffices to show that $z = \inf(T)$. Clearly z is a lower bound for T; we must show that it is larger than any other lower bound. Suppose, on the contrary, that z' is a lower bound for T and $z' > z$. Then z' is actually a member of T, and thus z' is the smallest element of T. That is, z and z' are adjacent in the ordering — i.e., there is no other element of X between z and z'. Since $z' \in T$, we have eventually $x_\alpha < z'$ and thus eventually $x_\alpha \leq z$, a contradiction. This completes the proof.

7.42. *Proposition (optional).* Suppose X is an infinitely distributive lattice (as defined in 4.23). Then the lattice operations $\vee, \wedge$ are "jointly continuous," in the following sense: If $((x_\alpha, x'_\alpha) : \alpha \in A)$ is a net in $X \times X$ with $x_\alpha \xrightarrow{o} x$ and $x'_\alpha \xrightarrow{o} x'$, then $x_\alpha \vee x'_\alpha \xrightarrow{o} x \vee x'$ and $x_\alpha \wedge x'_\alpha \xrightarrow{o} x \wedge x'$.

Proof. This argument follows Vulikh [1967]. We shall show $x_\alpha \vee x'_\alpha \xrightarrow{o} x \vee x'$; the result for meets is proved analogously. By assumption, there exist nets $(u_\lambda : \lambda \in L)$, $(v_\mu : \mu \in M)$, $(u'_\sigma : \sigma \in S)$, $(v'_\tau : \tau \in T)$ such that

$$u_\lambda \uparrow x, \qquad v_\mu \downarrow x, \qquad u'_\sigma \uparrow x', \qquad v'_\tau \downarrow x',$$

and for each fixed $\lambda, \mu, \sigma, \tau$ we have α-eventually

$$u_\lambda \preccurlyeq x_\alpha \preccurlyeq v_\mu \qquad \text{and} \qquad u'_\sigma \preccurlyeq x'_\alpha \preccurlyeq v'_\tau.$$

Let $L \times S$ and $M \times T$ have the product orderings. Define

$$\widehat{u}_{\lambda,\sigma} = u_\lambda \vee u'_\sigma, \qquad \widehat{v}_{\mu,\tau} = v_\mu \vee v'_\tau.$$

Then for each fixed $\lambda, \mu, \sigma, \tau$ we have α-eventually $\widehat{u}_{\lambda,\sigma} \preccurlyeq x_\alpha \vee x'_\alpha \preccurlyeq \widehat{v}_{\mu,\tau}$. Furthermore, the net $(\widehat{u}_{\lambda,\sigma} : (\lambda, \sigma) \in L \times S)$ is increasing and $(\widehat{v}_{\mu,\tau} : (\mu, \tau) \in M \times T)$ is decreasing. Then

$$\sup_{(\lambda,\sigma)\in L\times S} \widehat{u}_{\lambda,\sigma} = \sup_{\lambda\in L,\sigma\in S} (u_\lambda \vee u'_\sigma) = \left(\sup_{\lambda\in L} u_\lambda\right) \vee \left(\sup_{\sigma\in S} u'_\sigma\right) = x \vee x'$$

by 3.21.m. Use the infinite distributivity of the lattice to prove the middle equality in this string of equations:

$$\inf_{(\mu,\tau)\in M\times T} \widehat{v}_{\mu,\tau} = \inf_{\mu\in M,\tau\in T} (v_\mu \vee v'_\tau) = \left(\inf_{\mu\in M} v_\mu\right) \vee \left(\inf_{\tau\in T} v'_\tau\right) = x \vee x'.$$

Thus $\widehat{u}_{\lambda,\sigma} \uparrow (x \vee x')$ and $\widehat{v}_{\mu,\tau} \downarrow (x \vee x')$, so $x_\alpha \vee x'_\alpha \xrightarrow{o} x \vee x'$.

Convergence in Complete Lattices

7.43. *Remarks on applicability of the theory.* When $(X, \preccurlyeq)$ is a complete lattice, then the preceding characterizations of order convergence can be restated in other forms that are sometimes more convenient. Examples of complete lattices to keep in mind are the extended real line $[-\infty, +\infty]$ and the space $[0,1]^S = \{\text{functions from } S \text{ into } [0,1]\}$ with the product ordering (for any set S).

In applications, one may wish to apply the following results to other posets $(X, \preccurlyeq)$ that are not order complete — e.g., the real line $\mathbb{R}$ or a space such as $C[0,1] = \{\text{continuous functions from } [0,1] \text{ into } \mathbb{R}\}$. Here are two commonly used methods for extending the theory to such spaces: (i) We may work in some larger set $Y \supseteq X$ that is order complete. For instance, $\mathbb{R}$ can be embedded in $[-\infty, +\infty]$, and $C[0,1]$ can be embedded in $[-\infty, +\infty]^{[0,1]}$.

(ii) Alternatively, we may find some subset of X that is order complete and arrange our applications so that everything of interest stays in that subset. For instance, although $\mathbb{R}$ is not order complete, the interval $[a, b]$ is, for any real numbers a, b with $a < b$. More generally, if $(X, \preccurlyeq)$ is *Dedekind complete*, then any set of the form $[a, b] = \{x \in X \ : \ a \preccurlyeq x \preccurlyeq b\}$ is order complete. Thus, the theory of convergences in order complete sets is applicable to a Dedekind complete poset, provided that we restrict our attention to nets that are *eventually bounded.*

7.44. *Definitions.* Let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in a complete lattice $(X, \preccurlyeq)$. Then we may define the related objects

$$s_\alpha = \inf_{\beta \succcurlyeq \alpha} x_\beta \qquad \text{and} \qquad t_\alpha = \sup_{\beta \succcurlyeq \alpha} x_\beta.$$

Observe that $s_\alpha \preccurlyeq x_\alpha \preccurlyeq t_\alpha$.

The net $(s_\alpha : \alpha \in \mathbb{A})$ is increasing; hence it increases to a limit. That limit is called the **liminf** of the given net (x_α), since it is the limit of the infs; it is also called the *lower limit* of the net (x_α). The liminf of the x_α's is denoted $\liminf x_\alpha$, or sometimes $\underline{\lim}\ x_\alpha$. Note that if the given net (x_α) is a sequence, then (s_α) is also a sequence.

The net $(t_\alpha : \alpha \in \mathbb{A})$ is decreasing; hence it decreases to a limit. That limit is called the **limsup** of the given net (x_α), since it is the limit of the sups; it is also called the *upper limit* of the net (x_α). The limsup of the x_α's is denoted $\limsup x_\alpha$, or sometimes $\overline{\lim}\, x_\alpha$. Note that if the given net (x_α) is a sequence, then (t_α) is also a sequence.

Also note that $\liminf x_\alpha \preccurlyeq \limsup x_\alpha$.

7.45. Theorem on convergence in lattices. Let $(X, \preccurlyeq)$ be a complete lattice. Let $z \in X$, and let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in X. Then the following conditions are equivalent.[2]

(A) $x_\alpha \xrightarrow{o} z$ (as defined in 7.38 or as equivalently characterized in 7.40.d).

(B) The net's eventuality filter $\mathcal{F}$ contains a family of intervals $\{[s_\lambda, t_\lambda] : \lambda \in \Lambda\}$ such that $\bigcap_{\lambda\in\Lambda}[s_\lambda, t_\lambda] = \{z\}$.

(C) $\liminf x_\alpha = z = \limsup x_\alpha$.

(D) There exist nets $(s_\alpha : \alpha \in \mathbb{A})$ and $(t_\alpha : \alpha \in \mathbb{A})$ (based on the given directed set $\mathbb{A}$) such that $s_\alpha \uparrow z$ and $t_\alpha \downarrow z$, and $s_\alpha \preccurlyeq x_\alpha \preccurlyeq t_\alpha$ for all α.

Proof. For (C) $\Rightarrow$ (D), define s_α and t_α as in 7.44. It is obvious that (D) implies the condition given in 7.40.d; thus (D) $\Rightarrow$ (A). To prove (A) $\Rightarrow$ (B), suppose (x_α) and z satisfy the conditions in 7.38; then take $\Lambda = S \times T$ — that is, consider the collection of order intervals $\{[s,t] \,:\, s \in S,\ t \in T\}$.

To prove (B) $\Rightarrow$ (C), let $u_\alpha = \inf_{\beta\succcurlyeq\alpha} x_\beta$ and $v_\alpha = \sup_{\beta\succcurlyeq\alpha} x_\beta$ for each $\alpha \in \mathbb{A}$. Then $u_\alpha \preccurlyeq x_\alpha \preccurlyeq v_\alpha$. Let $U = \liminf x_\alpha$ and $V = \limsup x_\beta$; then $u_\alpha \uparrow U$ and $v_\alpha \downarrow V$. Temporarily fix any $\lambda \in \Lambda$. Since $[s_\lambda, t_\lambda] \in \mathcal{F}$, we have $x_\beta \in [s_\lambda, t_\lambda]$ for all β sufficiently large. Therefore $u_\alpha, v_\alpha \in [s_\lambda, t_\lambda]$ for all α sufficiently large. It follows that $U, V \in [s_\lambda, t_\lambda]$. This is valid for every λ. Hence U and V both lie in $\bigcap_{\lambda\in\Lambda}[s_\lambda, t_\lambda] = \{z\}$.

7.46. *Remarks.* In a complete lattice, when a net has a limit, that limit is equal to the liminf and limsup. However, the liminf and limsup exist in any case, whether the limit exists or not. In cases where the limit does not exist or is not known to exist, the liminf and limsup serve as "almost limits," or "pseudo-limits." They possess many of the properties one associates with a limit, and they can be used in place of a limit in many arguments.

For a very different sort of generalized limit, see 12.33.

7.47. *Further properties.* Let $(X, \preccurlyeq)$ be a complete lattice.

a. Suppose $(y_\beta : \beta \in \mathbb{B})$ is a subnet of $(x_\alpha : \alpha \in \mathbb{A})$ in X. Then

$$\liminf x_\alpha \quad\preccurlyeq\quad \liminf y_\beta \quad\preccurlyeq\quad \limsup y_\beta \quad\preccurlyeq\quad \limsup x_\alpha.$$

Hints: Fix any $a \in \mathbb{A}$, and let $t_a = \sup_{\alpha\succcurlyeq a} x_\alpha$. Show that eventually $x_\alpha \preccurlyeq t_a$; hence eventually $y_\beta \preccurlyeq t_a$; hence $\limsup y_\beta \preccurlyeq t_a$. Then what?

[2]Some of these conditions make sense in a more general setting — e.g., if we merely assume that $(X, \preccurlyeq)$ is a poset — and the literature sometimes uses one of these conditions as a definition of order convergence in such a setting. However, in such a setting the several conditions listed here are not all equivalent.

b. Suppose $(x_\alpha : \alpha \in \mathbb{A})$ and $(y_\alpha : \alpha \in \mathbb{A})$ are nets based on the same directed set $\mathbb{A}$, and $x_\alpha \preccurlyeq y_\alpha$ for all α (or for all α sufficiently large). Then

$$\liminf x_\alpha \preccurlyeq \liminf y_\alpha \qquad \text{and} \qquad \limsup x_\alpha \preccurlyeq \limsup y_\alpha,$$

and if both nets possess limits then $\lim x_\alpha \preccurlyeq \lim y_\alpha$.

7.48. *Convergence of sets.* Let (S_α) be a net whose elements are subsets of a set Ω, and let $S \subseteq \Omega$ also. What does $S_\alpha \to S$ mean? There are many different definitions in the literature, not all equivalent. The simplest of these, and perhaps the most frequently useful, is in terms of the ordering in which $S \preccurlyeq T$ means that $S \subseteq T$. That ordering makes $\mathcal{P}(\Omega)$ a complete lattice. Then for any net (S_α), we have

$$\limsup S_\alpha = \bigcap_\beta \bigcup_{\alpha \succcurlyeq \beta} S_\alpha = \{\omega \in \Omega : \text{frequently } \omega \in S_\alpha\},$$

$$\liminf S_\alpha = \bigcup_\beta \bigcap_{\alpha \succcurlyeq \beta} S_\alpha = \{\omega \in \Omega : \text{eventually } \omega \in S_\alpha\}.$$

Then it is always true that $\liminf S_\alpha \subseteq \limsup S_\alpha$, and we define $S_\alpha \to S$ to mean that

$$\liminf S_\alpha = S = \limsup S_\alpha,$$

or equivalently that

$$S \subseteq \liminf S_\alpha \qquad \text{and} \qquad \limsup S_\alpha \subseteq S.$$

That convergence can also be restated in terms of the characteristic functions of the sets; it says

$$\text{for each } \omega \in \Omega, \text{ eventually } 1_{S_\alpha}(\omega) = 1_S(\omega).$$

See also the related result in 15.26.e.

Note that if $\mathcal{S}$ is a σ-algebra of subsets of Ω, and (S_n) is a *sequence* in $\mathcal{S}$, then $\limsup S_n$ and $\liminf S_n$ both lie in $\mathcal{S}$.

7.49. *Remarks.* Although convergence of nets of sets is most often defined as in 7.48 (or equivalently, as in 15.26.e), other definitions are occasionally useful, particularly when the sets have some additional structure:

- A positive charge determines a pseudometric on an algebra of sets, as in 21.9. That pseudometric determines a convergence.
- The Hausdorff metric, defined in 5.18.d, determines a convergence for the nonempty, closed, metrically bounded subsets of a metric space.
- Several different topologies on the collection of closed subsets of a metric space are surveyed by Beer and Lucchetti [1993]. Each topology determines a convergence.

Part B
ALGEBRA

This Page Intentionally Left Blank

Chapter 8

Elementary Algebraic Systems

MONOIDS

8.1. *Definitions.* A **monoid** is a triple $(X, \square, i)$ consisting of a set X, a binary operation $\square$, and a special element $i \in X$ (called the **identity element** of X), satisfying these rules:

$$\begin{aligned} (x \square y) \square z &= x \square (y \square z) && \text{(associative law)} \\ x \square i = x &= i \square x && \text{(identity law)} \end{aligned}$$

for all $x, y, z \in X$. The **fundamental operations** of the monoid are $\square$ (binary) and i (nullary). We may refer to X itself as a monoid if i and $\square$ do not need to be mentioned explicitly. When we disregard i and $\square$, and just consider X as a set, it is called the **underlying set** of the monoid.

Different monoids X and Y generally have different identity elements and different binary operations, but we may use the same symbols i and $\square$ in different monoids if no confusion will result; we may use subscripts (i_X, i_Y, $\square_X$, $\square_Y$) for clarification if necessary.

If $(X, \square, i_X)$ and $(Y, \diamond, i_Y)$ are monoids, a **homomorphism** from X to Y is a mapping $f : X \to Y$ that preserves the fundamental operations — i.e., a mapping such that

$$f(i_X) = i_Y, \qquad\qquad f(x \square x') = f(x) \diamond f(x')$$

for all $x, x' \in X$. If $f : X \to Y$ is a bijective homomorphism, we call f an **isomorphism** of monoids; it is easy to see that $f^{-1} : Y \to X$ is then a homomorphism as well.

In a monoid $(X, \square, i_X)$, a **submonoid** is a subset $S \subseteq X$ that is closed under the fundamental operations of X — i.e., that satisfies $i_X \in S$ and also satisfies $s, t \in S \Rightarrow s \square t \in S$. Thus, it is a subset S that becomes a monoid in its own right when the monoid operations of X are restricted to S.

8.2. *Exercise.* The identity element i in a monoid is uniquely determined. In fact, we don't even need the associative law for that result; if $\square$ is a binary operation on a set X and $i_1, i_2 \in X$ both satisfy the identity law in 8.1, then $i_1 = i_2$.

8.3. *More definitions.* A monoid $(X, \square, i)$ is **commutative** (or *Abelian*) if it also satisfies

$$x \square y = y \square x \qquad \text{(commutative law)}$$

for all $x, y \in X$.

For many commutative monoids, in place of $\Box$ we use the symbol $+$, known as **addition**. Then we say that the operation is *written additively*, and X is an **additive monoid**. In that case the identity element is denoted by "0" and known as **zero** or the **additive identity**. For nonempty subsets S, T of an additive monoid, we write $x + S = \{x + s : s \in S\}$ and $S + T = \{s + t : s \in S, t \in T\}$, as in 2.7. The definition of homomorphism can be restated for additive monoids thus:

$$f(x + x') = f(x) + f(x'), \qquad\qquad f(0) = 0.$$

When these conditions are met we shall say f is an **additive map**. The last equation can be written $f(0_X) = 0_Y$ if some clarification is needed, but usually it is not. *Caution*: Algebraists occasionally use $+$ for a noncommutative operation, but analysts generally do not. *In this book addition will always represent a commutative operation.*

For some monoids — not necessarily commutative — the symbol used in place of $\Box$ is a raised dot $(\cdot)$, known as **multiplication**. Then we say that the operation is *written multiplicatively*, and the monoid is a **multiplicative monoid**. In that case the identity element is denoted by "1" and known as **one** or the **multiplicative identity**. The symbol for multiplication may also be omitted altogether; i.e., we may write $x \cdot y$ instead as xy. Although this looks just like multiplication of real numbers, the reader is cautioned not to assume that it is commutative — see 8.4.e.

8.4. *Examples of monoids.*

a. $(\mathcal{P}(X), \varnothing, \cup)$ and $(\mathcal{P}(X), X, \cap)$ are commutative monoids for any set X.

b. $\mathbb{R}$ is an additive monoid, as are certain subsets of $\mathbb{R}$ — for instance,

$$[0, +\infty), \qquad \mathbb{Z}, \qquad \mathbb{N} \cup \{0\}.$$

c. Measure theory will be introduced briefly in 11.37 and studied in much greater depth in Chapter 21. A measure, or more generally a charge, is a particular type of mapping taking values in a monoid X. In most cases of interest, that monoid X is either $[0, +\infty]$ or a vector space. The choice of X is discussed further in 11.38.

d. Arithmetic in the extended real number system $[-\infty, +\infty]$ was defined in 1.17. The set $[-\infty, +\infty]$ is not an additive monoid, for there is no suitable way to define $(-\infty) + (+\infty)$. However, any other sum of two elements in $[-\infty, +\infty]$ is defined. Consequently, certain subsets of the extended real line are additive monoids — for instance,

$$(-\infty, +\infty], \qquad [-\infty, +\infty), \qquad [0, +\infty], \qquad \mathbb{N} \cup \{0\} \cup \{\infty\}.$$

In $[0, +\infty]$ we can define not only finite sums $x_1 + x_2 + \cdots + x_n$, but also infinite sums $x_1 + x_2 + x_3 + \cdots$; see 10.39.

e. Let X be a set. Then $X^X = \{\text{functions from } X \text{ into } X\}$ is a monoid, with the binary operation being the composition of functions, defined as in 2.3. The identity element of X^X is the identity map $i_X : X \to X$ defined in 2.5.a. Composition of functions is often written "multiplicatively" — i.e., $f \circ g$ is often written simply as fg — but composition of functions generally is not commutative.

Actually, if $(X, \Box, i)$ is any monoid, then X is isomorphic to a submonoid of X^X via the mapping $u \mapsto f_u$, where $f_u : X \to X$ is the mapping defined by $f_u(x) = u \Box x$.

f. Let X be a set. Then $\mathcal{P}(X \times X) = \{\text{subsets of } X \times X\}$ is a monoid, with the binary operation being the composition of relations, defined as in 3.3.e. The identity of this monoid is the diagonal set $I = \{(x,x) : x \in X\}$. Identifying functions with their graphs, we find that X^X (discussed in 8.4.e) is a submonoid of $\mathcal{P}(X \times X)$.

g. Let $\mathcal{A}$ be an *alphabet* — i.e., a collection of symbols that can be distinguished from one another. Let X be the set of all finite strings of symbols made from members of that alphabet — for instance, if $a, b, c \in \mathcal{A}$, then abc and $abac$ and cab are three different members of X. For a binary operation we use **concatenation** — for instance, $abc \circ abac = abcabac$. The empty string — i.e., the string containing no symbols — will also be considered as a member of X; then it is the identity element and X is a monoid. More complicated algebraic systems that are similar to this one are the basis of *formal logic*, studied in Chapter 14.

Groups

8.5. Let $(X, i, \Box)$ be a monoid, with identity element i. If $x \Box y = i$, we say that x is a **left inverse** for y and that y is a **right inverse** for x; these are **one-sided inverses**. If $x \Box y = y \Box x = i$, we say that x and y are **inverses** of each other (or, for emphasis, **two-sided inverses**).

Exercises and examples.

a. A monoid element may have many left inverses (or just one, or none). Similarly for right inverses.

For instance, let $\Omega = \{\text{sequences of real numbers}\}$, and let $X = \Omega^\Omega = \{\text{functions from } \Omega \text{ into } \Omega\}$, with composition for the binary operation. Define $x(r_1, r_2, r_3, \ldots) = (r_2, r_3, r_4, \ldots)$. Also, for each real number p, define $y_p(r_1, r_2, r_3, \ldots) = (p, r_1, r_2, r_3, \ldots)$. Then $x \circ y_p = i$, so x has many right inverses. The element x has no left inverses; this can be proved directly or using 8.5.b. To reverse this example, use the same set X, but use binary operation $\Box$ defined by $u \Box v = v \circ u$.

b. Suppose that $(X, i, \Box)$ is a monoid, $x \in X$, u_l is a left inverse of x, and u_r is a right inverse of x. Then $u_l = u_r$, and x has no other left or right inverses.

Proof. $u_l = u_l \circ i = u_l \circ (x \circ u_r) = (u_l \circ x) \circ u_r = i \circ u_r = u_r$. The same reasoning can be applied if u_r is replaced by any other right inverse of x; thus all the right inverses of x are equal to u_l. Similarly, all the left inverses of x are equal to u_r.

c. Any element of a monoid has at most one inverse. If x has an inverse, that inverse may be denoted x^{-1}. In an additive monoid, that inverse may be denoted $-x$.

8.6. *Definitions.* A **group** is a monoid in which each element has an inverse. We shall restate this definition more directly:

A *group* is a quadruple $(X, \Box, {}^{-1}, i)$ consisting of a set X and three **fundamental operations** that obey certain axioms. The three fundamental operations are a binary operation $(x, y) \mapsto x\Box y$, a unary operation $x \mapsto x^{-1}$, and a nullary operation i — that is, a specially selected element $i \in X$. The axioms are

$$\begin{array}{rclcll} (x\Box y)\Box z & = & x\Box(y\Box z) & & & \text{(associative law)} \\ x\Box i & = & x & = & i\Box x & \text{(identity law)} \\ x\Box x^{-1} & = & i & = & x^{-1}\Box x & \text{(law of inverses)} \end{array}$$

for all $x, y, z \in X$. The group is **commutative** (or **Abelian**) if it also satisfies

$$x\Box y \quad = \quad y\Box x \qquad \text{(commutative law)}$$

for all $x, y \in X$.

A **subgroup** of a group $(X, \Box, {}^{-1}, i)$ is a set $S \subseteq X$ that is closed under the group's fundamental operations — i.e., that includes the identity element and also satisfies

$$x, y \in S \quad \Rightarrow \quad x\Box y,\ x^{-1} \in S.$$

Thus, it is a subset S that becomes a group in its own right when the fundamental operations of X are restricted to S.

The **subgroup generated** by a set $B \subseteq X$ is the smallest subgroup that includes B — i.e., the intersection of all the subgroups that include B; it is the closure of B (in the sense of 4.6) under the fundamental operations of the group.

8.7. *Exercise (optional).* There is some redundancy in our list of axioms for a group — a shorter list would suffice:

Suppose X is a set equipped with a binary operation $\Box$, a unary operation $x \mapsto x^{-1}$, and a special element (i.e., a nullary operation) i, satisfying these axioms:

$$\Box \text{ is associative}, \qquad i\Box x = x, \qquad \text{and} \qquad x^{-1}\Box x = i$$

for all $x \in X$. Show that the set and operations must also satisfy

$$x\Box x^{-1} = i, \qquad x\Box i = x, \qquad (x^{-1})^{-1} = x$$

for all x.

Hint: $\left[(x^{-1})^{-1}\Box x^{-1}\right]\Box(x\Box x^{-1}) = (x^{-1})^{-1}\Box\left[(x^{-1}\Box x)\Box x^{-1}\right]$.

8.8. *More notation.* An **additive group** or **multiplicative group** is a group in which the binary operation is written as $+$ or $\cdot$, respectively.

In a multiplicative, commutative group, the product $x \cdot (y^{-1})$ is also written x/y or $\frac{x}{y}$.

In this book $+$ will only be used for a commutative operation. In an additive group, the inverse of an element x is written as $-x$, and the sum $x + (-y)$ is abbreviated $x - y$. For a nonempty subsets S, T of an additive group we write $-S = \{-s : s \in S\}$ and $S - T = \{s - t : s \in S,\ t \in T\}$, as in 2.7.

8.9. *More definitions.* A **homomorphism** between groups $(X, \Box, {}^{-1}, i_X)$ and $(Y, \diamond, {}^{-1}, i_Y)$ is a mapping $f : X \to Y$ satisfying

$$f(x\Box x') = f(x) \diamond f(x'), \qquad f(x^{-1}) = f(x)^{-1}, \qquad f(i_X) = i_Y,$$

for all $x, x' \in X$. Actually, the second and third equations can be omitted from this definition, for they follow as consequences of the first equation; the proof of this is an easy exercise. Thus, if X and Y are groups, then a mapping $f : X \to Y$ is a homomorphism of groups if and only if it is a homomorphism of monoids. An **isomorphism** of groups is a bijective homomorphism.

When X and Y are additive groups, the last two equations can be rewritten as $f(-x) = -f(x)$ and $f(0) = 0$. A mapping $f : X \to Y$ between additive groups is a homomorphism if and only if it satisfies $f(x_1 + x_2) = f(x_1) + f(x_2)$ for all $x_1, x_2 \in X$; we may then call it an **additive mapping**.

8.10. *Elementary properties and examples of groups.*

a. *Degenerate examples.* The smallest group is a singleton, with obvious operations. All one-element groups are isomorphic to each other. In any group, the subgroup generated by the empty set is the singleton consisting of just the identity element.

The next smallest group contains just two elements, again with obvious operations. All two-element groups are isomorphic to each other. One convenient representation is this: $\{1, -1\}$ is a multiplicative group.

b. Let $(X, \Box, i)$ be a monoid, and let $G = \{x \in X : x \text{ has an inverse}\}$. Then G is a submonoid of X, and in fact G is a group. A particular example of this is given in 8.10.i.

c. In any group $(X, \Box, ^{-1}, i)$, we have

$$i^{-1} = i, \qquad (x^{-1})^{-1} = x, \qquad (x \Box y)^{-1} = y^{-1} \Box x^{-1}.$$

d. The singleton $\{0\}$, the integers $\mathbb{Z}$, the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$ are additive groups, when equipped with their usual addition operation. In fact, $\{0\} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$; each is a subgroup of the next. The set $\mathbb{Z}$ is the subgroup of $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$ generated by the set $\{1\}$.

e. Let r be a positive number. (The values of r most commonly used here are 1 and 2π.) The interval $[0, r)$ can be viewed as an additive group, referred to as **the reals modulo r**. The addition operation for this group is **addition modulo r**, defined as follows: Give $x + y$ its usual meaning when $x + y \in [0, r)$, and let $x + y$ be replaced by $x + y - r$ when $x + y \in [r, 2r)$. The identity element is 0. This group is isomorphic to the *circle group*, discussed in 10.32; consequently $[0, r)$ itself is sometimes referred to as the circle group.

f. The positive reals, $(0, +\infty)$, may be viewed as a commutative group whose binary operation is ordinary multiplication $(\cdot)$ and whose identity element is the number 1. Some subgroups are the positive rational numbers and the set $\{2^k : k \in \mathbb{Z}\}$. The multiplicative group of positive real numbers is isomorphic to the additive group of real numbers, by the mapping $x \mapsto \ln x$.

g. Let X be a set. Then $(\mathcal{P}(X), \triangle, i_{\mathcal{P}(X)}, \varnothing)$ is a commutative group, where $\mathcal{P}(X)$ denotes the power set of X and $\triangle$ denotes symmetric difference. Note that in this group, the inverse operation is the identity map — that is, each member of $\mathcal{P}(X)$ is its own inverse. Hence $(A \triangle C) \triangle (B \triangle C) = A \triangle B$.

Any algebra of subsets of X (defined in 5.25) is a subgroup of $\mathcal{P}(X)$.

h. Let X be an additive group. For $x \in X$ we define

$$0x = 0, \qquad 1x = x, \qquad 2x = x + x, \qquad 3x = x + x + x, \qquad \ldots,$$

and for $n \in \mathbb{N}$ we also define $(-n)x = -(nx)$. In this fashion we define a "multiplication" operation $(n, x) \mapsto nx$, from $\mathbb{Z} \times X$ into X. By induction or any other convincing argument, show that

$$(mn)x = m(nx), \qquad m(x+y) = (mx) + (my), \qquad (m+n)x = (mx) + (nx)$$

for all $m, n \in \mathbb{Z}$ and $x, y \in X$. Also show that $\mathbb{Z}x = \{nx : n \in \mathbb{Z}\}$ is a subgroup of X; it is the subgroup generated by the singleton $\{x\}$.

i. A bijection from a set X onto itself is a **permutation** of X. If X is any nonempty set, then $\mathrm{Perm}(X) = \{\text{permutations of } X\}$ is a group, with the binary group operation given by the composition of functions and with the identity element of the group being the identity function of X. In fact, this is the group of invertible elements obtained from the monoid X^X (see 8.4.e and 8.10.b). If X contains more than two elements, then the group $\mathrm{Perm}(X)$ is not commutative. If n is a positive integer, then the permutation group on a set X containing n elements is also called the **symmetric group of order n**; it is written S_n.

If X is the underlying set of a group $(X, i, \Box)$, then an isomorphism from X onto a subgroup of $\mathrm{Perm}(X)$ is given by $u \mapsto f_u$, where the permutation $f_u : X \to X$ is given by $f_u(x) = u \Box x$.

Sums and Quotients of Groups

8.11. Let $S_1, S_2, \ldots, S_n$ be finitely many subgroups of an additive group X. Then the **sum** of the S_j's is the set

$$S_1 + S_2 + \cdots + S_n \;=\; \{s_1 + s_2 + \cdots + s_n \;:\; s_1 \in S_1,\ s_2 \in S_2,\ \ldots,\ s_n \in S_n\}.$$

More generally, let $\{S_\lambda : \lambda \in \Lambda\}$ be a collection of subgroups of an additive group X. Their sum, $\sum_{\lambda\in\Lambda} S_\lambda$, is defined to be the set of all sums of finitely many elements of $\bigcup_{\lambda\in\Lambda} S_\lambda$. In other words, it is the set of all sums of the form

$$s \;=\; s_1 + s_2 + s_3 + \cdots + s_n,$$

where n is a nonnegative integer and each s_j is a member of some S_λ. Show that

(i) $\sum_{\lambda\in\Lambda} S_\lambda$ is the union of sums of finitely many of the S_λ's.

(ii) $\sum_{\lambda\in\Lambda} S_\lambda$ is the subgroup of X generated by the set $\bigcup_{\lambda\in\Lambda} S_\lambda$.

8.12. Let $S = S_1 + S_2 + \cdots + S_n$ be a sum of finitely many subgroups. The set S is called the **internal direct sum** of the S_j's if it has this further property: Each $s \in S$ can be expressed in one *and only one* way as $s = s_1 + s_2 + \cdots + s_n$, where $s_j \in S_j$. We then write $S = S_1 \oplus S_2 \oplus \cdots \oplus S_n$ or $S = \bigoplus_{j=1}^{n} S_j$. Such a decomposition may be helpful, because it may express a complicated object S in terms of simpler S_j's.

More generally, let $S = \sum_{\lambda\in\Lambda} S_\lambda$ be a sum of arbitrarily many subgroups. We say S is the **internal direct sum** of the S_λ's, and write $S = \bigoplus_{\lambda\in\Lambda} S_\lambda$, if each $s \in S$ can be written in one and only one way as a sum $s = \sum_{\lambda\in\Lambda} s_\lambda$, where each s_λ is a member of S_λ and only finitely many of the s_λ's are nonzero. (The internal direct sum is often called the "direct sum," but it should not be confused with the *external direct sum* described in 9.30.)

If $S = \bigoplus_{\lambda\in\Lambda} S_\lambda$, then we can define mappings $\varphi_\lambda : S \to S_\lambda$ by the rule that $s = \sum_{\lambda\in\Lambda} \varphi_\lambda(s)$; we may call φ_λ the **projection onto S_λ**. (The term "projection" also has other meanings; see 1.34 and 22.45.)

Some basic properties of direct sum decompositions are

a. S is an internal direct sum of the subgroups $\{S_\lambda : \lambda \in \Lambda\}$ if and only if $S = \sum_{\lambda\in\Lambda} S_\lambda$ and $S_\mu \cap \sum_{\lambda\neq\mu} S_\lambda = \{0\}$ for each $\mu \in \Lambda$.

b. Each mapping φ_λ, considered as a map from S into itself, is idempotent (defined in 2.4); it has range S_λ.

c. Each φ_λ, considered as a map from S into either S or S_λ, is additive.

8.13. An important special case is that in which an additive group X itself is the internal direct sum of two subgroups — say S and T. Then we write $X = S \oplus T$. This means that

> each $x \in X$ can be written in one and only one way in the form $s + t$, where $s \in S$ and $t \in T$,

or, equivalently, that

$$S + T = X \text{ and } S \cap T = \{0\}.$$

We shall then say that the subgroups S and T are **additively complementary**, or that they are **additive complements** of each other. (Some mathematicians would simply call these sets "complements" of each other, but in this book we have too many other uses for that term.)

Exercises. Suppose $X = S \oplus T$. Let $\varphi_S : X \to S$ and $\varphi_T : X \to T$ be the projections, defined as in 8.12 — that is, $x = \varphi_S(x) + \varphi_T(x)$ for each x. Show that

a. $\varphi_S + \varphi_T = i_X$ (where i_X is the identity map of X)

b. $\text{Range}(\varphi_S) = \text{Ker}(\varphi_T) = S$ and $\text{Range}(\varphi_T) = \text{Ker}(\varphi_S) = T$.

c. $\varphi_S\varphi_T = \varphi_T\varphi_S = 0$.

d. Conversely, suppose X is an additive group and $p : X \to X$ is an idempotent homomorphism. Let $q = i_X - p$. Show that q is also idempotent, and $X = \text{Range}(p) \oplus \text{Range}(q)$.

8.14. Let G be an additive group, and let H be a subgroup. Define sums of sets as in 8.3. The *cosets* of H are the sets $x+H = \{x+h \;:\; h \in H\}$. Note that any two cosets are either identical or disjoint; thus they form a partition of G. Show that

$$(x+H)+(y+H) \;=\; (x+y)+H, \qquad -(x+H) \;=\; (-x)+H.$$

Let G/H be the set of all cosets of H; show that G/H is an additive group with identity element $0+H$ and with other operations defined as above.

Since the cosets of H form a partition of G, they define an equivalence relation on G by:

$$g_1 \approx g_2 \quad\Longleftrightarrow\quad g_1, g_2 \text{ belong to the same coset} \quad\Longleftrightarrow\quad g_1 - g_2 \in H.$$

The cosets of H are the equivalence classes for this equivalence relation, and G/H is the quotient set (as in 3.11). Consequently, the group G/H is called the **quotient group**. The quotient map $\pi : G \to G/H$ (defined as in 3.11) is given by $\pi(g) = g + H$. It is a group homomorphism from G onto G/H. Note that it satisfies

$$\begin{array}{lll} \pi(\pi^{-1}(B)) = B & \text{for any } B \subseteq G/H, & \text{whereas} \\ \pi^{-1}(\pi(A)) = A + H & \text{for any } A \subseteq G. & \end{array}$$

Algebra books contain a more general theory of quotients, applicable to groups that are not necessarily commutative. However, that theory is more complicated and will not be needed for our purposes.

8.15. Not every quotient group G/H is isomorphic to a subgroup of G.

Example. The circle group $[0,1)$, introduced in 8.10.e, can also be described as the quotient of the additive group $\mathbb{R}$ by the subgroup $\mathbb{Z}$. The circle group is not isomorphic to a subgroup of $\mathbb{R}$. One easy way to show this is to note that 0 and $\frac{1}{2}$ are distinct solutions of $x+x=0$ in $[0,1)$. In the group $\mathbb{R}$, the equation $x+x=0$ has only one solution.

8.16. Not every subgroup of every group has an additive complement. (Contrast 11.30.f.)

Example. $\mathbb{Z}$ is a subgroup of $\mathbb{R}$, but there is no subgroup $G \subseteq \mathbb{R}$ satisfying $\mathbb{R} = \mathbb{Z} \oplus G$. Indeed, show that if G were such a group, it would be isomorphic to $\mathbb{R}/\mathbb{Z}$, and hence isomorphic to $[0,1)$, contradicting the result in 8.15.

8.17. Let $f : X \to Y$ be an additive mapping — i.e., a homomorphism of additive groups. Then the **kernel** of f is the set

$$\mathrm{Ker}(f) \;=\; f^{-1}(0) \;=\; \{x \in X : f(x) = 0\}.$$

A few of its basic properties are:

a. $\mathrm{Ker}(f)$ is a subgroup of X; hence $0 \in \mathrm{Ker}(f)$.

b. $\mathrm{Ker}(f) = \{0\}$ if and only if f is injective.

c. **(Isomorphism Theorem.)** Let $\pi : X \to X/\mathrm{Ker}(f)$ be the quotient map. Then $F(\pi(x)) = f(x)$ defines a group isomorphism $F : X/\mathrm{Ker}(f) \to \mathrm{Ran}(f)$.

d. *Degenerate examples.* Let X be any additive group. Then the identity map $i : X \to X$ has kernel $\{0\}$, and the constant map $x \mapsto 0$ (from X into any additive group) has kernel X.

RINGS AND FIELDS

8.18. *Definitions.* A **ring** is an additive group $(R, 0, +)$ equipped with another associative binary operation $(\cdot)$, called **multiplication**, which distributes over addition on both the left and right:

$$\begin{aligned} w \cdot (x+y) &= (w \cdot x) + (w \cdot y) \qquad \text{and} \\ (x+y) \cdot w &= (x \cdot w) + (y \cdot w) \end{aligned}$$

for all $w, x, y \in R$.

A **ring with unit** also has a special element 1 (**one**), such that $(R, 1, \cdot)$ is a monoid. *Caution*: Some mathematicians work only with rings with unit, and then they may refer to those objects simply as "rings." For a trivial example of a ring without unit, consider the even integers, with the usual operations of addition and multiplication. For a less trivial example of considerable interest to analysts, see 11.4.e.

(Most of the rings used by analysts have additional structure: They are *linear algebras*, as explained in 11.3. However, $\mathbb{Z}$ is an important commutative ring that is not a linear algebra.)

By our definitions, the addition operation in any ring is commutative. A **commutative ring** is a ring in which the multiplication operation is also commutative.

A **field** is a commutative ring with unit, in which $0 \neq 1$ and in which every nonzero element has a multiplicative inverse. Consequently, in fields we are able to perform "ordinary arithmetic" computations. For instance, the student should prove (and explain) that in a field,

$$\frac{w}{x} + \frac{y}{z} = \frac{wz + xy}{xz}.$$

Examples. Some fields with which most readers are informally acquainted are $\mathbb{Q}$ and $\mathbb{R}$; these are introduced formally in 8.22, 10.10, 10.8, and 10.15.

The **fundamental operations** of a ring with unit or a field are those of its additive group (the binary operation $+$, the unary operation $-$, and the nullary operation 0) and those of its multiplicative monoid (the binary operation $\cdot$ and the nullary operation 1). When we talk about fundamental operations and related concepts, then a field will simply be viewed as a particular type of ring with unit. (See the related remarks in 8.54.)

A **homomorphism** of rings with unit is a mapping $f : R \to S$ from one ring into another, which preserves the fundamental operations — i.e., which satisfies

$$\begin{aligned} &f(x_1 + x_2) = f(x_1) + f(x_2), \qquad f(-x) = -f(x), \\ &\qquad\qquad\qquad\qquad f(0) = 0, \qquad f(x_1 x_2) = f(x_1) f(x_2), \qquad f(1) = 1 \end{aligned}$$

for all $x, x_1, x_2 \in R$. All of these conditions are conceptually relevant, but some of them are redundant, i.e., implied by some of the other conditions. A homomorphism of fields will simply mean a homomorphism $f : R \to S$ of rings with unit, where R and S happen to be fields; no additional requirement is imposed on f for this case. However, (*exercise*) it follows from our definition that if $f : R \to S$ is a homomorphism of fields, then f is injective and $f(x^{-1}) = f(x)^{-1}$ for all $x \neq 0$.

8.19. *Some elementary properties.* For all w, x, y, z in a ring R with unit, we have

a. $0 \cdot x = 0 = x \cdot 0$.

b. $(-x) \cdot y = x \cdot (-y) = -(x \cdot y)$.

c. $(-1) \cdot x = -x$. That is, the additive inverse of 1 times any ring element x is the additive inverse of x.

d. There is a unique homomorphism from $\mathbb{Z}$ into the ring R.

e. If $0 = 1$, then $R = \{0\}$. This is the smallest ring.

8.20. *Example: finite rings and fields.* Let m be an integer greater than 1. For integers $x, y \in \mathbb{Z}$, write $x \equiv y \pmod{m}$ if $x - y$ is a multiple of m — that is, if $x - y = km$ for some integer k. We then say that x and y are **congruent modulo m**. It is easy to verify that $\equiv$ is an equivalence relation on $\mathbb{Z}$. The arithmetic operations make sense on the equivalence classes, since

$$x_1 \equiv y_1, \quad x_2 \equiv y_2 \qquad \Rightarrow \qquad x_1 + x_2 \equiv y_1 + y_2, \quad x_1 x_2 \equiv y_1 y_2.$$

The equivalence classes are most often represented by their smallest nonnegative members — i.e., the numbers $0, 1, 2, \ldots, m-1$. Thus we obtain arithmetic operations on the set

$$\mathbb{Z}_m = \{0, 1, 2, 3, \ldots, m-1\},$$

which can be described more directly as follows: to add or multiply two numbers x, y in $\mathbb{Z}_m$, take their ordinary sum or product in $\mathbb{Z}$, and then subtract a suitable multiple of m to obtain an element of $\{0, 1, 2, \ldots, m-1\}$. With these operations, $\mathbb{Z}_m$ is a commutative ring with unit, called **the integers modulo m**. As an illustrative example, below are the addition and multiplication tables for $\mathbb{Z}_6$. Note that, considered as an additive group, $\mathbb{Z}_m$ is the subgroup generated by $\{1\}$ in the group $[0, m)$ of reals modulo m, introduced in 8.10.e.

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

·	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Recall that a **prime number** is one of the numbers 2, 3, 5, 7, 11, etc. — that is, an integer greater than 1 that can only be written as a product of two positive integers if one of those factors is 1. It is an easy exercise to show that the finite ring $\mathbb{Z}_m$ is a *field* if and only if m is a prime number. In particular, $\mathbb{Z}_2 = \{0, 1\}$ is the smallest field; it will be of some importance in the study of Boolean algebras.

Related exercises.

a. In the ring $\mathbb{Z}_6$, we have $2 \cdot 3 = 0$; thus the product of two nonzero elements is zero. In the ring $\mathbb{Z}_4$, we have $2^2 = 0$.

b. The ring $\mathbb{Z}_4$ is not a field, but there does exist a field $\mathbb{F}_4$ containing exactly 4 elements; it is unique up to isomorphism. Find its addition and multiplication tables.

c. In the field $\mathbb{Z}_5$, the squares of the numbers 0, 1, 2, 3, 4 are the numbers 0, 1, 4, 4, 1. More generally, show that if m is an odd prime number, then exactly half of the nonzero members of $\mathbb{Z}_m$ are squares of members of $\mathbb{Z}_m$.

Remarks. Finite fields are not often useful in analysis; we have mentioned them only because they offer very easily understood illustrations of the concept of "field." We shall now state without proof a few more results about finite fields; the proofs of these additional results are beyond the scope of this book, but can be found in more specialized books — see for instance, Lidl and Niederreiter [1983]. Let q be an integer greater than 1. Then there exists a field $\mathbb{F}_q$ containing exactly q elements if and only if q is of the form $q = p^n$ for some prime number p and some positive integer n — in which case the field $\mathbb{F}_q$ is unique (up to isomorphism). Considered as a linear space over $\mathbb{Z}_p$ (see Chapter 11), the field $\mathbb{F}_q$ is isomorphic to $(\mathbb{Z}_p)^n$. The multiplicative group $\mathbb{F}_q \backslash \{0\}$ is isomorphic (as a group) to the additive group $\mathbb{Z}_{q-1}$. The explicit formation of such finite fields — i.e., the computation of their addition and multiplication tables — is a somewhat complicated matter. However, when p is an odd prime, then it is fairly easy to form a field with p^2 elements; a simple method is given in 10.23.b.

8.21. *Example: products.* Suppose that $(R_\lambda : \lambda \in \Lambda)$ is a collection of rings. Then we can make the Cartesian product $P = \prod_{\lambda \in \Lambda} R_\lambda$ into a ring, by defining operations coordinatewise:

$$(f+g)(\lambda) \;=\; f(\lambda) + g(\lambda), \qquad\qquad (fg)(\lambda) \;=\; f(\lambda)\,g(\lambda),$$

etc. The additive identity 0_P is the function that takes the value 0_λ at the λth coordinate. If the R_λ's are rings with unit, then so is P, with multiplicative identity 1_P equal to the function that takes the value 1_λ on the λth coordinate.

The product of two or more fields is not a field, when operations are defined in this fashion, since any element of P with a 0 in at least one component has no multiplicative inverse. However, a different method can sometimes be used to make a product of fields into a field; see 10.22.

8.22. The reader is undoubtedly quite familiar with the field of rational numbers, $\mathbb{Q} = \{m/n \;:\; m, n \in \mathbb{Z},\; n \neq 0\}$. Nevertheless, we shall give a formal construction of it; the same method of construction will subsequently be used to form another, less familiar field.

An **integral domain** is a commutative ring D with the property that

$$\text{whenever } x, y \in D \text{ with } xy = 0, \text{ then at least one of } x, y \text{ is } 0.$$

Of course, any field is an integral domain. The ring $\mathbb{Z} = \{\text{integers}\}$ is an example of an integral domain that is not a field; another example will be given in 8.24. The finite rings $\mathbb{Z}_4$ and $\mathbb{Z}_6$ are not integral domains.

Let D be an integral domain. For pairs (x, m) and (y, n) in $D \times (D \backslash \{0\})$, define $(x, m) \approx (y, n)$ to mean that $xn = ym$. Verify that this is an equivalence relation on $D \times (D \backslash \{0\})$. Define $\mathbb{F}$ to be the set of equivalence classes. Addition and multiplication in $\mathbb{F}$ are defined by

$$(x, m) + (y, n) = (xn + ym,\ mn), \qquad\qquad (x, m) \cdot (y, n) = (xy,\ mn).$$

The reader should verify that these operations are well-defined — i.e., that the definitions above do not depend on the particular choice of representations for the equivalence classes. That is, if $(x_1, m_1) \approx (x_2, m_2)$ and $(y_1, n_1) \approx (y_2, n_2)$, verify that

$$(x_1 n_1 + y_1 m_1,\ m_1 n_1) \approx (x_2 n_2 + y_2 m_2,\ m_2 n_2), \qquad (x_1 y_1,\ m_1 n_1) \approx (x_2 y_2,\ m_2 n_2).$$

With operations so defined, verify that $\mathbb{F}$ is a field. It is called the **field of fractions** of D, or **field of quotients**. The mapping $x \mapsto (x, 1)$ is an embedding of D in $\mathbb{F}$ — that is, an injective ring homomorphism — and so we may view the ring D as a subset of the field $\mathbb{F}$.

Having completed our construction of $\mathbb{F}$, we now switch to conventional notation: The equivalence class containing (x, m) is represented by x/m or $\frac{x}{m}$. Of course, the representation is not unique, since any pair in the equivalence class can be used to form this expression. We urge the reader not to switch to this notation until after completing the construction of $\mathbb{F}$ and the verifications that it requires. The artificiality of the unfamiliar notation (x, m) will make it less likely that we will inadvertently assume some familiar property of $\mathbb{F}$ that has not yet been proved.

In the particular case where the integral domain D is the ring $\mathbb{Z}$, the resulting field of quotients is the **field of rational numbers**; it is denoted by $\mathbb{Q}$.

8.23. *Exercises about* $\mathbb{Q}$.

a. Show that $\text{card}(\mathbb{Q}) = \text{card}(\mathbb{N})$. *Hint*: 2.20.e.

b. There is no $x \in \mathbb{Q}$ satisfying $x^2 = 2$. *Hint*: If $x = p/q$, consider how many factors of 2 there are in p or in q. (We assume familiarity with basic properties of the integers, e.g., the uniqueness of prime factorization.)

c. If $\mathbb{F}$ is a field, there is a unique ring homomorphism from $\mathbb{Q}$ into $\mathbb{F}$.

d. *Example.* There is a unique ring homomorphism $h : \mathbb{Q} \to \mathbb{Z}_5$. With that ring homomorphism, evaluate $h(\frac{2}{3})$. Explain. (Thus we obtain a member of $\{0, 1, 2, 3, 4\}$ which is in a sense "congruent modulo 5" to the fraction 2/3.)

8.24. *Example*: *the ring of polynomials and the field of rational functions.* Let $\mathbb{K}$ be any integral domain (for instance, the integers or the rationals or the reals). Let $S = \{s, t, u, \ldots\}$ be a nonempty (finite or infinite) set of distinct symbols not already used in our description of $\mathbb{K}$ or elsewhere in our language. We write S as $\{s, t, u, \ldots\}$ to display a few typical elements, but we do not require that S be countable or ordered in any fashion. (For the simplest examples, take S to be just a singleton: $S = \{s\}$. However, later we shall have some uses for much larger collections S as well.)

A **monomial** with variables in S and coefficients in $\mathbb{K}$ is any expression such as as^3t^2uv, where $a \in \mathbb{K}$ and $s, t, u, v \in S$ — that is, an element of $\mathbb{K}$ multiplied by finitely many members of S. If the coefficient a is not zero, then the **degree** of the monomial is the sum of the exponents of the variables — for instance, the monomial $as^3t^2uv = as^3t^2u^1v^1$ has degree $3 + 2 + 1 + 1 = 7$.

A **polynomial** with variables in S and coefficients in $\mathbb{K}$ is a sum of finitely many monomials — i.e., any expression such as $as^3 + bst^2 + ct^2 + duv + e$, where $a, b, c, d, e \in \mathbb{K}$ and $s, t, u, v \in S$. The **degree** of the polynomial is the highest degree of any of its monomials; for instance, $as^3 + bst^2 + ct^2 + duv + e$ has degree 3. A **homogeneous polynomial of degree k** is a sum of several monomials of degree k; for instance, $as^3 + bs^2t + cstu + dtu^2 + ev^3$ is a homogeneous polynomial of degree 3.

Addition, multiplication, and equality of polynomials are defined by the usual algebraic rules; we omit the details. The set of all polynomials with variables in S and coefficients in $\mathbb{K}$ is easily seen to form a commutative ring with unit, which we shall denote by $\mathbb{K}[S]$.

Note that each $a \in \mathbb{K}$ may be viewed as a constant polynomial — i.e., a polynomial of degree 0; thus each member of $\mathbb{K}$ may be viewed as a member of $\mathbb{K}[S]$. This mapping from $\mathbb{K}$ into $\mathbb{K}[S]$ is an injective ring homomorphism; thus we may view $\mathbb{K}$ as a subset of $\mathbb{K}[S]$.

When S consists of just one variable — say s — then the ring $\mathbb{K}[S] = \mathbb{K}[\{s\}]$ may be written more briefly as $\mathbb{K}[s]$. Then any polynomial may be written in the form

$$p(s) \quad = \quad a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0$$

where the coefficients a_j are members of $\mathbb{K}$. If $p(s)$ is not the constant function 0, then by dropping any leading zero terms we can choose the representation so that $a_n \neq 0$. Then n is the degree of the polynomial, and a_n is called the **leading coefficient**.

If the ring $\mathbb{K}$ is an integral domain, then so is the ring $\mathbb{K}[S]$. Hence we can form its field of quotients, as in 8.22. That field is called the **field of rational functions** with variables in S and coefficients in $\mathbb{F}$; we shall denote it by $\mathbb{K}(S)$. A member of that field is a **rational function** with variables in S and coefficients in $\mathbb{F}$ — i.e., a quotient of two polynomials. A typical rational function is

$$\frac{as^3 + bst^2 + ct^2 + duv + e}{bt^3 + dst + fuv^3 + g} .$$

Equality between such rational functions and arithmetic operations with such functions are defined in the usual fashion; we omit the details. If S consists of just a single variable s, then the field $\mathbb{K}(S) = \mathbb{K}(\{s\})$ may be written more briefly as $\mathbb{K}(s)$.

8.25. *Blass's Subfield* (*optional*). Define $\mathbb{K}(S)$ as above. Let $\mathcal{B} = \{p/q \in \mathbb{K}(S) : p$ and q are homogeneous polynomials of the same degree$\}$. (For instance, if $S = \{s, t, u\}$ and $\mathbb{K} = \mathbb{R}$, then

$$\frac{3s^3 + \sqrt{2}s^2t + \frac{1}{2}stu + \pi su^2}{17stu - \sqrt[3]{5}st^2 + 6.179t^3}$$

is a typical member of $\mathcal{B}$.) Show that $\mathcal{B}$ is a subfield of $\mathbb{K}(S)$. This field will be mentioned again in 11.29.

Matrices

8.26. *Matrix notation.* Let $\mathbb{K}$ be a ring, and let m and n be positive integers. An ***m*-by-*n* matrix** over $\mathbb{K}$ is a rectangular array

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

with m rows and n columns, where each a_{ij} is an element of $\mathbb{K}$. We say a_{ij} is the **element** (or **component**) in row i and column j. The matrix A given above may be represented more briefly as $A = (a_{ij} : 1 \leq i \leq m;\ 1 \leq j \leq n)$, or still more briefly as (a_{ij}) if no confusion will result.

The **transpose** of an m-by-n matrix A is the n-by-m matrix $A^\top$ obtained by flipping A over diagonally, so that the kth row becomes the kth column and vice versa. Obviously, $(A^\top)^\top = A$. An m-by-n matrix is called a **column matrix** if $n = 1$ (i.e., if it consists of just one column), a **row matrix** if $m = 1$ (i.e., if it consists of just one row), and a **square matrix** if $m = n$. Note that the transpose of a row matrix is a column matrix, and vice versa. For any positive integer p, it is customary[1] to consider elements of $\mathbb{K}^p$ as column matrices when matrices are to be used at all, but to save space on the printed page they are often represented as the transposes of row matrices. Thus the ordered p-tuple $(b_1, b_2, \ldots, b_p)$ can also be written as $[b_1\ b_2\ \cdots\ b_p]^\top$; we emphasize that the representation with parentheses requires commas while the representation with brackets requires that the commas be omitted.

8.27. Matrix multiplication has slightly complicated dimensional requirements. If A is an m-by-n matrix and B is an n-by-p matrix, then we can form their product $AB = R$, an m-by-p matrix:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{m\text{-by-}n} \underbrace{\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}}_{n\text{-by-}p} = \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mp} \end{bmatrix}}_{m\text{-by-}p}$$

defined by this formula: $r_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}$.

In general, multiplication of matrices is not commutative. In fact, when the product AB is defined, the product BA is not necessarily defined. For instance, for the matrices above, we can only define BA if $m = p$, and in that case AB is an m-by-m matrix while BA is an n-by-n matrix. Thus, $AB = BA$ can *only* hold if A and B are square matrices of

[1]Some older algebra books represent members of $\mathbb{K}^p$ as row matrices, but column matrices seem to be the prevailing convention since sometime around 1960.

the same dimension. Even then, $AB = BA$ only holds in an occasional coincidence; it does not hold in general, even if the underlying ring $\mathbb{K}$ is commutative. For instance, if

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

then $AB \neq BA$, provided $\mathbb{K}$ is a ring with unit in which $0 \neq 1$. However, we do have $(AB)^\top = B^\top A^\top$ if the ring $\mathbb{K}$ is commutative.

It is an easy exercise to show that *multiplication of matrices is associative*: $(AB)C = A(BC)$ whenever the dimensions of the matrices match up — i.e., A is an m-by-n matrix, B is an n-by-p matrix, and C is a p-by-q matrix. Hence we may omit the parentheses and write the product simply as ABC. The element in row h, column k of that product is $\sum_{i=1}^{n} \sum_{j=1}^{p} a_{hi} b_{ij} c_{jk}$.

8.28. *Matrices as functions on columns.* An important special case of matrix multiplication is the following: Let A be an m-by-n matrix. Represent elements of $\mathbb{K}^m$ and $\mathbb{K}^n$ by column matrices — i.e., by m-by-1 matrices and by n-by-1 matrices, respectively. Then the mapping $v \mapsto Av$ is an additive map from $\mathbb{K}^n$ into $\mathbb{K}^m$. This type of map plays an important role in the theory of finite-dimensional vector spaces, discussed further in Chapter 11. It is so important that we shall write it out more explicitly here:

$$\text{If } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \text{and } v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\text{then } Av = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}.$$

In particular, the n-by-n matrices act as mappings from $\mathbb{K}^n$ into itself. If $\mathbb{K}$ is a ring with unit, then the n-by-n matrices form a monoid, under the operation of matrix multiplication. The invertible elements of that monoid form a group. An interesting subgroup consists of the **permutation matrices of order n**; these are the n-by-n matrices A that have the following property: Each row contains $n-1$ zeros and 1 one; each column also contains $n-1$ zeros and 1 one. For example, there are six permutation matrices of order 3:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Such a matrix is called a **permutation matrix** for the following reason: If n distinct members of the ring $\mathbb{K}$ are arranged in a column matrix v, then the mapping $v \mapsto Av$ *permutes* those n members — i.e., the column matrix Av consists of the same n members, arranged in some other order (or in the same order, if $A = I$). The group of permutation matrices of order n is isomorphic (as a group) to the symmetric group of order n, introduced in 8.10.i.

8.29. *The ring of matrices.* Addition of m-by-n matrices is defined componentwise:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}.$$

Thus, we can only add two matrices if they have the same dimensions.

With multiplication and addition defined as above, the set $\mathrm{Mat}(n;\mathbb{K}) = \{n\text{-by-}n \text{ matrices over } \mathbb{K}\}$ is a ring; it has additive and multiplicative identities given by

$$0 = 0_n = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \qquad \text{and} \qquad I = I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Here I_n is an n-by-n matrix that has 1s along its main diagonal and 0s elsewhere; it may be written more briefly as (δ_{ij}), where δ is the Kronecker delta. In general the ring $\mathrm{Mat}(n;\mathbb{K})$ is not commutative.

Ordered Groups

8.30. In this book an **ordered monoid** will mean an additive monoid X that is equipped with a partial ordering $\preccurlyeq$ that is **translation-invariant** — i.e., that satisfies

$$x \preccurlyeq y \qquad \Rightarrow \qquad x + u \preccurlyeq y + u$$

for all $x, y, u \in X$. If X is also a group, we shall call it an **ordered group**.

Most of the ordered monoids used by analysts have a great deal more structure — in fact, most of them are Riesz spaces. However, $[0, +\infty]$ is an important ordered monoid that is not even a group.

8.31. *Arithmetic in ordered monoids.* Let S and T be nonempty subsets of an ordered monoid $(X, \preccurlyeq)$. For each of the following equations, show that if the left side exists, then the right side also exists and the two sides are equal:

$$\max(S) + \max(T) \;=\; \max(S+T), \qquad \min(S) + \min(T) \;=\; \min(S+T).$$

Also show that each of the following inequalities holds if both sides exist:

$$\sup(S) + \sup(T) \ \succcurlyeq\ \sup(S+T), \qquad \inf(S) + \inf(T) \ \preccurlyeq\ \inf(S+T).$$

(Compare with 8.33.a.)

8.32. *Proposition.* The sup of a directed family of additive maps is additive. More precisely:
Let M be an additive monoid, and let $(Y, \preccurlyeq)$ be an ordered monoid. Let Φ be a collection of additive maps from M into Y. Assume Φ is directed by the product ordering on Y^M — that is, for each $f_1, f_2 \in \Phi$ there exists $f \in \Phi$ such that

$$f(x) \succcurlyeq f_1(x) \qquad \text{and} \qquad f(x) \succcurlyeq f_2(x) \qquad \text{for all } x \in M.$$

Assume that $h(x) = \sup_{f \in \Phi} f(x)$ exists in Y for each $x \in M$. Then the function $h : M \to Y$ is also additive. (This result will be used in 11.57.)

Hints: The proof of $h(x+x') \preccurlyeq h(x) + h(x')$ is easy — it does not require Φ to be directed; we leave the details as an exercise. For the reverse inequality, show that

$$h(x) + h(x') \ = \ \sup_{f_1, f_2 \in \Phi} [f_1(x) + f_2(x')] \ \leq \ \sup_{f \in \Phi} [f(x) + f(x')] \ = \ h(x+x').$$

8.33. *Arithmetic in ordered groups.* Let $(X, \preccurlyeq)$ be an ordered group. Let S and T be nonempty subsets of X, and let $x, y \in X$. Show that

a. For each of the following equations, if the left side exists, then the right side exists and the two sides are equal:

$$\sup(S) + \sup(T) \ = \ \sup(S+T), \qquad \inf(S) + \inf(T) \ = \ \inf(S+T).$$

b. $x \preccurlyeq y \iff -x \succcurlyeq -y$.

c. *Duality in ordered groups.* For each of the following equations, the left side exists if and only if the right side exists, in which case they are equal:

$$\max(x+S) = x + \max(S), \qquad \sup(x+S) = x + \sup(S),$$

$$\min(x+S) = x + \min(S), \qquad \inf(x+S) = x + \inf(S),$$

$$-\max(S) = \min(-S), \qquad -\sup(S) = \inf(-S).$$

When S contains just two elements, the last equation becomes

$$-(u \vee v) \ = \ (-u) \wedge (-v) \quad \text{or, equivalently,} \quad -(p \wedge q) \ = \ (-p) \vee (-q).$$

From all of these equations it follows that any statements about maxima or suprema can be translated into statements about minima or infima, and vice versa. Such statements occur in pairs; the members of such a pair are said to be **dual** to each other. For brevity, in many cases we mention only one of the two statements. See also 1.7.

d. Let D be a subgroup of X. Then D is sup-dense in X if and only if D is inf-dense in X. *Hint*: $-\sup(S) = \inf(-S)$, etc.

e. Let $f : X \to Y$ be a group homomorphism, where Y is another ordered group. Then f is sup-preserving if and only if f is inf-preserving.
Hint for the "if" part: If $\sigma = \sup(S)$ for some $S \subseteq X$, then

$$\begin{aligned} -f(\sigma) &= -f(\sup(S)) = f(-\sup(S)) \\ &= f(\inf(-S)) = \inf(f(-S)) = \inf(-f(S)) = -\sup(f(S)). \end{aligned}$$

8.34. *More definitions.* If X is an ordered group, then the **positive cone** of X is the set

$$X_+ \quad = \quad \{x \in X : x \succcurlyeq 0\}.$$

Note that the ordering can be recovered from the positive cone: We have $x \succcurlyeq y \iff x - y \in X_+$.

Caution: Elements of the positive cone are *not* necessarily called "positive." In particular, when $X = \mathbb{R}$, then X_+ is the set of all nonnegative real numbers. Thus, 0 is a member of the "positive cone" but it is not a positive number.

8.35. *Exercise (optional).* Let $(X, \preccurlyeq)$ be an ordered group. Then the following conditions are equivalent:

(A) $(X, \preccurlyeq)$ is a directed set.

(B) X_+ generates the group — that is, $X_+ - X_+ = X$.

(C) For each $x \in X$, there is some $p \in X_+$ with $p \succcurlyeq x$.

8.36. In an ordered group, we use the notation $[a, b] = \{x \in X : a \preccurlyeq x \preccurlyeq b\}$. Note that

$$[x_1, y_1] \ + \ [x_2, y_2] \quad \subseteq \quad [x_1 + x_2, y_1 + y_2]$$

in any ordered group. For some examples of the property described below, see 8.37 and 8.38.

Theorem. Let $(X, \preccurlyeq)$ be an ordered group. The following conditions are then equivalent. If one, hence all, are satisfied, we say X has the **Riesz Decomposition Property**.

(A) $[0, u] + [0, v] = [0, u + v]$ whenever $u, v \in X_+$.

(B) If $p_1, p_2, \ldots, p_m \in X$ and $q_1, q_2, \ldots, q_n \in X$ with $p_i \preccurlyeq q_j$ for all i, j, then there exists some $r \in X$ with $p_i \preccurlyeq r \preccurlyeq q_j$ for all i, j.

(C) $[x_1, y_1] + [x_2, y_2] + \cdots + [x_n, y_n] = [x_1 + x_2 + \cdots + x_n, y_1 + y_2 + \cdots + y_n]$ whenever n is a positive integer and the x_i's and y_i's are members of X with $x_i \preccurlyeq y_i$ for all i.

(D) If $x_1, x_2, \ldots, x_m \in X_+$ and $y_1, y_2, \ldots, y_n \in X_+$ with $x_1 + x_2 + \cdots + x_m = y_1 + y_2 + \cdots + y_n$, then there exist some $z_{ij} \in X_+$ for $1 \le i \le m$ and $1 \le j \le n$, such that

$$x_i = \sum_{j=1}^{n} z_{ij} \text{ for all } i \qquad \text{and} \qquad y_j = \sum_{i=1}^{m} z_{ij} \text{ for all } j.$$

Proof of (A) $\Rightarrow$ (B). It suffices to prove (B) for $m = n = 2$; then higher values of m, n follow by induction. Assume, then, that $p_i \preccurlyeq q_j$ for $i, j = 1, 2$. We know that $q_2 - p_1$ lies in $[0, (q_1 - p_1) + (q_2 - p_2)] = [0, q_1 - p_1] + [0, q_2 - p_2]$. Hence $q_2 - p_1 = a_1 + a_2$ for some $a_j \in [0, q_j - p_j]$. Let $r = a_1 + p_1 = -a_2 + q_2$. Then $r = a_1 + p_1 \in [p_1, q_1]$ and $r = q_2 - a_2 \in [p_2, q_2]$.

Proof of (B) $\Rightarrow$ (C). It suffices to prove this for $n = 2$; higher values of n then follow by induction. Let $z \in [x_1 + x_2, y_1 + y_2]$. We know $z - y_1 \preccurlyeq y_2$ and $x_2 \preccurlyeq y_2$ and $z - y_1 \preccurlyeq z - x_1$ and $x_2 \preccurlyeq z - x_1$. Hence there is some r with $z - y_1 \preccurlyeq r$ and $x_2 \preccurlyeq r$ and $r \preccurlyeq y_2$ and $r \preccurlyeq z - x_1$. Hence $z - r \in [x_1, y_1]$ and $r \in [x_2, y_2]$.

Proof of (C) $\Rightarrow$ (D). It suffices to prove (D) for $m = 2$; then higher values follow by induction. If $x_1 + x_2 = y_1 + \cdots + y_n = s$, then $x_1 \in [0, y_1] + \cdots + [0, y_n]$, so $x_1 = z_{11} + \cdots + z_{1n}$ with $z_{1j} \in [0, y_j]$. Now let $z_{2j} = y_j - z_{1j}$.

Proof of (D) $\Rightarrow$ (A). Let $p \in [0, u + v]$. Then $p + q = u + v$ for some $q \succcurlyeq 0$. Decompose both sides of that equation as in (D).

8.37. *A degenerate example.* Any group X can be ordered by this relation: $x \preccurlyeq y$ if and only if $x = y$. We shall refer to this as the **trivial ordering**. Despite its simplicity, this ordering will play an important role in our theory; see 12.32 and 26.53.

Here are a few of its basic properties:

- **a.** It is not a lattice ordering, if X contains more than one element.
- **b.** The positive cone X_+ is just $\{0\}$.
- **c.** The set $[a, b]$ is a singleton if $a = b$, or empty if $a \neq b$.
- **d.** The trivial ordering has the Riesz Decomposition Property.

Lattice Groups

8.38. A **lattice group** is an ordered group whose ordering is a lattice ordering — i.e., an ordered group X that satisfies both of the following conditions:

(i) $x \vee y$ exists for all $x, y \in X$.

(ii) $x \wedge y$ exists for all $x, y \in X$.

Actually, in an ordered group, these statements are dual to each other, and so either implies the other; hence either one of these implies X is a lattice group. Since the ordering of an ordered group is translation-invariant, an even weaker hypothesis is sufficient:

> If X is an ordered group and $\sup\{x, 0\}$ exists for each $x \in X$, then X is a lattice group.

Some examples of lattice groups are given in 11.45 and 11.46.

It is easy to see that any lattice group has the Riesz Decomposition Property — indeed if p_i's and q_j's satisfy the hypotheses of 8.36(B), then $p_1 \vee p_2 \vee \cdots \vee p_m \preccurlyeq q_1 \wedge q_2 \wedge \cdots \wedge q_n$. However, we note that the Riesz Decomposition Property is also enjoyed by some other ordered groups that are not lattice groups; for instance, see 8.37.

8.39. If $(X, \preccurlyeq)$ is a lattice group, then for any $x \in X$ we can define

$$x^+ = x \vee 0, \qquad x^- = (-x) \vee 0, \qquad /x/ = x^+ + x^-.$$

These three objects are elements of the nonnegative cone X_+. They are called the **positive part** of x, the **negative part** of x, and the **absolute value** of x, respectively.

Caution: We have two notions of "absolute value" — the group element $/x/$ defined above, and the nonnegative real number $|x|$ defined in 10.31. Neither of these is a special case of the other. However, the two notions coincide when $X = \mathbb{R}$.

Here our notation is slightly unconventional. In the wider literature, it is customary to represent both kinds of absolute values by the expression $|x|$. However, that convention causes some difficulties for some beginners who are already familiar with the real-valued absolute value of real or complex numbers; they may accidentally attribute some of its properties to this new, unfamiliar "absolute value" of members of a lattice group — e.g., they may inadvertently assume that any two absolute values $/x/, /y/$ are comparable in order. (Those absolute values are *not* necessarily comparable; a simple counterexample is given in 8.41.) Our use of the notation $/x/$ will serve as a constant visual reminder that the "absolute value" being considered is, like x^+ and x^-, a member of some lattice group, not necessarily a member of $\mathbb{R}$. This book will reserve the notation $|x|$ for real-valued absolute values and norms, which are discussed in 10.31 and in Chapter 22. This book's unusual practice may prevent confusion in contexts where *both* types of absolute values are needed — e.g., in 26.55.

8.40. *Examples.* When X is the real line with its usual ordering, then $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$, and $/x/$ is just the usual absolute value $|x|$.

More generally, let Λ be any set. Then the product

$$\mathbb{R}^\Lambda = \{\text{functions from } \Lambda \text{ into } \mathbb{R}\}$$

is a Dedekind complete lattice group (actually a vector lattice), when given the product ordering — that is, when ordered by

$$x \preccurlyeq y \qquad \text{if} \qquad x(\lambda) \leq y(\lambda) \text{ for every } \lambda \in \Lambda.$$

The nonnegative cone is then

$$(\mathbb{R}^\Lambda)^+ = \{x \in \mathbb{R}^\Lambda \,:\, x(\lambda) \geq 0 \text{ for all } \lambda \in \Lambda\}.$$

The lattice operations $\vee, \wedge$ are defined pointwise on Λ, as in 11.45. We also have the functions

$$x^+(\lambda) = \max\{x(\lambda), 0\}, \qquad x^-(\lambda) = \max\{-x(\lambda), 0\}, \qquad /x/(\lambda) = |x(\lambda)|,$$

where the last expression is the usual (real-valued) absolute value of the real number $x(\lambda)$. We emphasize that $/x/ = |x(\cdot)|$ is an element of $\mathbb{R}^\Lambda$ — that is, a *function* from Λ into $\mathbb{R}$ — whereas $|x(\lambda)|$ is a real number. The pointwise formulas given above for $x \vee y$, $x \wedge y$, $\sup(S)$, $\inf(S)$, x^+, x^-, $/x/$ remain valid in many important subsets of $\mathbb{R}^\Lambda$; some of these are given in 11.46.

However, the pointwise formulas for the lattice operations are *not* valid in some other subsets of $\mathbb{R}^\Lambda$. Some vector lattices have appreciably more complicated formulas for the sup, inf, absolute value, etc. See for instance 4.21 and 11.47.

8.41. If x and y are members of a lattice group X, we do *not* necessarily have $/x/ \preccurlyeq /y/$ or $/y/ \preccurlyeq /x/$.

Example. Let $X = \mathbb{R}^\mathbb{R}$, with the product ordering. Then for any function x, the function $/x/$ is defined by $/x/(t) = |x(t)|$. Observe that if $x(t) = t$ and $y(t) = t - 1$, then $/x/(0) < /y/(0)$ and $/x/(1) > /y/(1)$. Thus, neither $/x/ \preccurlyeq /y/$ nor $/y/ \preccurlyeq /x/$ is valid. (Here we use $\preccurlyeq$ to compare members of X and $\leq$ to compare members of $\mathbb{R}$.)

8.42. *Arithmetic in lattice groups.* Let X be a lattice group, and let $x, y, z \in X$. Show that

a. $\vee$ and $\wedge$ are translation-invariant. That is,

$$\begin{aligned} (x+z) \vee (y+z) &= (x \vee y) + z, \\ (x+z) \wedge (y+z) &= (x \wedge y) + z. \end{aligned}$$

This can also be described as: Addition distributes over $\vee$ and $\wedge$.

b. *Sum decomposition.* $x + y = (x \vee y) + (x \wedge y)$. *Hint*: Use translation-invariance and 8.33.c.

c. $x^- = (-x)^+ = -(x \wedge 0)$ and $x^+ = (-x)^-$.

d. $x \succcurlyeq 0 \iff x^+ = x \iff x^- = 0 \iff /x/ = x$.

e. $/x/ = x^+ + x^- = /-x/ = x^+ \vee x^- = x \vee (-x) \vee 0$.

Remark. In 11.50 we'll see that if X is a *vector lattice*, then $/x/ = x \vee (-x)$.

f. Jordan Decomposition. $x = x^+ - x^-$. *Hint*: Sum decomposition and 8.33.c.

g. $x^+ \wedge x^- = 0$. *Hint*: By translation invariance,

$$x^+ \wedge x^- = [(x^+ - x^-) + x^-] \wedge x^- = [(x^+ - x^-) \wedge 0] + x^- = (x \wedge 0) - (x \wedge 0).$$

h. *Uniqueness of the Jordan Decomposition.* If $x = u - v$ where $u \wedge v = 0$, then $u = x^+$ and $v = x^-$.

Hints: Show $u = x + v \succcurlyeq x$ and $u \succcurlyeq 0$, hence $u \succcurlyeq x^+$. Then $p = u - x^+ = v - x^- \succcurlyeq 0$. On the other hand, $u = p + x^+ \succcurlyeq p$ and $v = p + x^- \succcurlyeq p$, hence $0 = u \wedge v \succcurlyeq p$.

i. $/x/ = 0 \iff x = 0$.

j. $-/x/ \preccurlyeq -(x^-) \preccurlyeq x \preccurlyeq x^+ \preccurlyeq /x/$.

k. $/x/ \preccurlyeq y$ if and only if both $-y \preccurlyeq x \preccurlyeq y$ and $y \succcurlyeq 0$.

Remark. In 11.50 we'll see that if X is a vector lattice, then $/x/ \preccurlyeq y \iff -y \preccurlyeq x \preccurlyeq y$.

l. $2(x \vee y) = x + y + /x - y/$ and $2(x \wedge y) = x + y - /x - y/$.
Hint: Immediate from 8.42.b. *Remark.* In a vector lattice, it follows immediately that $x \vee y = \frac{1}{2}(x+y) + \frac{1}{2}/x - y/$ and $x \wedge y = \frac{1}{2}(x+y) - \frac{1}{2}/x-y/$.

m. $(x+y)^+ \preccurlyeq x^+ + y^+$ and $(x+y)^- \preccurlyeq x^- + y^-$.

n. Triangle Inequality. $/x+y/ \preccurlyeq /x/ + /y/$.

o. $/u^+ - v^+/ \preccurlyeq /u-v/$ and $/u^- - v^-/ \preccurlyeq /u-v/$, and likewise

$$\Big/ /u/ - /v/ \Big/ \quad \preccurlyeq \quad /u-v/.$$

Hint: Let $f(t)$ be any of the functions t^+, t^-, or $/t/$. By the preceding exercises, $f(x+y) - f(x) \preccurlyeq f(y)$. Apply this once with $x = v$, $y = u - v$ and once with $x = u$, $y = v - u$, to prove $/f(u) - f(v)/ \preccurlyeq /u-v/$.

p. If $x \in X$ and n is a positive integer, then

$$0 \vee x \vee (2x) \vee (3x) \vee \cdots \vee (nx) \quad = \quad \underbrace{x^+ + x^+ + \cdots + x^+}_{n \text{ summands}}.$$

(Here it is understood that $0x = x$ and $nx = x + x + x + \cdots + x$ is the sum of n x's, as in 8.10.h.) *Hint*: Use induction on n, with

$$\bigvee_{j=0}^{n+1}(jx) = \bigvee_{j=0}^{n}[(j+1)x \vee jx] = \bigvee_{j=0}^{n}[jx + (x \vee 0)] = \left(\bigvee_{j=0}^{n}(jx)\right) + (x \vee 0).$$

Remark. In a Riesz space this formula simplifies to $n(x^+) = (nx)^+$; see 11.50.

q. A set $S \subseteq X$ is said to be **solid** if it satisfies:

$$/x/ \preccurlyeq /y/ \quad \text{and} \quad y \in S \qquad \Rightarrow \qquad x \in S.$$

(In particular, the empty set is solid.) Note that the union of any collection of solid sets is solid. Thus solid sets are "Moore open sets," i.e., their complements form a collection of Moore closed sets in the sense of 4.3.

For any set $T \subseteq X$, let $\operatorname{sk}(T)$ be the union of all the solid subsets of T. Show that $\operatorname{sk}(T)$ is solid; it is called the **solid kernel** of T. It is the largest solid set contained in T; thus it is a sort of "Moore interior" (dual to a Moore closure). Show also that

$$\operatorname{sk}(T) \quad = \quad \Big\{x \in X \,:\, [-/x/, /x/] \subseteq T\Big\} \quad = \quad \bigcup_{[-u,u]\subseteq T} [-u, u].$$

8.43. Theorem on distributivity. Let X be a lattice group. Then X is *distributive*; that is,

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \qquad \text{and} \qquad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

for every $x, y, z \in X$. In fact, we can make a stronger assertion: X is *infinitely distributive.* That is, for any $x \in X$ and any nonempty set $S \subseteq X$,

(i) $x \wedge \sup(S) = \sup\{x \wedge s : s \in S\}$ and

(ii) $x \vee \inf(S) = \inf\{x \vee s : s \in S\}$

where each equation is interpreted in this sense: Whenever the left side of the equation exists, then the right side also exists and the two sides are equal.

Proof. It suffices to prove (i); then (ii) will follow by duality. To prove (i), assume $\sigma = \sup(S)$ exists, and let $T = \{x \wedge s : s \in S\}$; we are to show that $x \wedge \sigma$ is the supremum of T. It is certainly an upper bound for T, since $x \wedge s \preccurlyeq x \wedge \sigma$ for each $s \in S$. To show that it is the *least* upper bound, let τ be any upper bound for T; we are to show $\tau \succcurlyeq x \wedge \sigma$. For each $s \in S$ we know $\tau \succcurlyeq x \wedge s = (x+s) - (x \vee s)$, hence

$$\tau + (x \vee \sigma) - x \quad \succcurlyeq \quad \tau + (x \vee s) - x \quad \succcurlyeq \quad s.$$

Take the supremum on the right; thus $\tau + (x \vee \sigma) - x \succcurlyeq \sigma$. Add $x - (x \vee \sigma)$ to both sides to prove $\tau \succcurlyeq x \wedge \sigma$.

8.44. *Convergence in lattice groups.* Let $(X, \preccurlyeq)$ be a lattice group. Since X is infinitely distributive, the conclusions of 7.42 are applicable. Show also that

a. $x_\alpha \xrightarrow{o} x$ (as defined in 7.38 and 7.40.d) if and only if there exists a set $S \subseteq X$ with these three properties:

(i) S is directed downward — i.e., for each $s_1, s_2 \in S$ there exists $s \in S$ with $s \preccurlyeq s_1 \wedge s_2$.

(ii) $0 = \inf(S)$.

(iii) For each $s \in S$, we have eventually $/x_\alpha - x/ \preccurlyeq s$.

b. The lattice group operations are continuous, in the following sense: Suppose (x_α, y_α) is a net in $X \times X$, with $x_\alpha \xrightarrow{o} x$ and $y_\alpha \xrightarrow{o} y$. Then

$$-x_\alpha \xrightarrow{o} -x, \qquad /x_\alpha/ \xrightarrow{o} /x/, \qquad x_\alpha^+ \xrightarrow{o} x^+, \qquad x_\alpha + y_\alpha \xrightarrow{o} x + y.$$

If X is a vector lattice, then we can also conclude $cx_\alpha \xrightarrow{o} cx$ for every real number c.

8.45. *Proposition.* Let X and Y be lattice groups, and let $f : X \to Y$ be a group homomorphism — i.e., an additive map. Then the following conditions are equivalent:

(A) f is a lattice homomorphism — that is, $f(u \vee v) = f(u) \vee f(v)$ and $f(u \wedge v) = f(u) \wedge f(v)$ for all $u, v \in X$.

(B) $f(x^+) = (f(x))^+$ for all $x \in X$.

Proof. For (A) $\Rightarrow$ (B), note that $x^+ = x \vee 0$ and $f(0) = 0$. For (B) $\Rightarrow$ (A), compute

$$\begin{aligned} f(u \vee v) &= f((u-v) \vee 0 + v) = f\left((u-v)^+\right) + f(v) \\ &= (f(u-v))^+ + f(v) = (f(u) - f(v)) \vee 0 + f(v) = f(u) \vee f(v). \end{aligned}$$

Thus f preserves sups. By duality (8.33.c), since f is a group homomorphism, it also preserves infs.

8.46. *Observation.* Any lattice homomorphism is order-preserving. (*Hint*: 3.21.f.)

However, an order-preserving group homomorphism between lattice groups is not necessarily a lattice homomorphism. *Example.* Let $X = C[0,2\pi] = \{$continuous functions from $[0,2\pi]$ into $\mathbb{R}\}$, and let $Y = \mathbb{R}$. Define $f : X \to Y$ by $f(x) = \int_0^{2\pi} x(t)\,dt$. Now consider the function $u(t) = \sin(t)$. We have $f(u) = 0$, hence $(f(u))^+ = 0$. On the other hand,

$$u^+(t) \quad = \quad \max\{0, u(t)\} \quad = \quad \begin{cases} \sin(t) & \text{when } 0 \le t \le \pi \\ 0 & \text{when } \pi \le t \le 2\pi, \end{cases}$$

hence $f(u^+) = \int_0^{\pi} \sin(t)\,dt = 2$.

Universal Algebras

8.47. We shall now study certain ideas that can be applied simultaneously to lattices, monoids, groups, Abelian groups, rings, lattice groups, etc. This material is taken from McKenzie, McNulty, and Taylor [1987] and other books on varieties or universal algebras.

An **arity function**, or **type**, for algebraic systems is a function τ, defined on any nonempty set J, taking values in $\{0,1,2,3,\ldots\}$. That function's domain, J, may be finite or infinite; both cases will be important in our applications.

Concerning the range of τ: In most examples in the literature and in all examples in this book, the arity function τ actually maps the set J into the set $\{0,1,2\}$, but that restriction is not required for the theory; in principle other values are possible. (Some algebraists permit τ to also take infinite values; an algebraic system is then called **infinitary**. However, we shall only consider **finitary** algebraic systems — i.e., those in which each $\tau(j)$ is finite. Some of our results will use this assumption.)

Let τ be an arity function. An **algebraic system of arity τ** is a set X equipped with a collection of functions $\varphi_j : X^{\tau(j)} \to X$ (for $j \in J$). Thus the jth function, φ_j, is a $\tau(j)$-ary operation on X (see 1.40). The functions φ_j are called the **fundamental operations** of X. Another term for "algebraic system" is **universal algebra**.

To be precise, we should denote the system by an expression such as $\big(X, J, \tau, \{\varphi_j\}\big)$ — i.e., the fundamental operations are part of the definition of the algebraic system; different algebraic systems may be built from the same underlying set X by attaching different fundamental operations. However, in practice we often refer to X itself as the algebraic system, with the choices of $J, \tau, \{\varphi_j\}$ understood. The notation involving τ, φ_j's, etc., is helpful for our present purposes — i.e., developing an abstract theory of algebraic systems — but it is seldom used in the context of individual algebraic systems; see the remarks at the end of 8.53.

If X and X' are algebraic systems with the same arity function τ, then their jth fundamental operations φ_j and φ'_j may be quite different, but at least they have the same arity — i.e., they are both $\tau(j)$-ary operations; thus they behave alike in certain important respects. When no confusion will result, we may drop the primes, and use one symbol for both φ_j and φ'_j — for instance, we commonly use the same symbol + in different commutative groups.

On the other hand, in some introductory discussions such as this one it will be helpful to use different symbols (such as φ_j, φ'_j or $+, \oplus$) to distinguish between the operations of different algebraic systems.

8.48. (For examples see 8.52.) Let X and X' be algebraic systems with the same arity function τ and corresponding fundamental operations φ_j and φ'_j. A **homomorphism** from X into X' is a mapping $f : X \to X'$ that preserves the fundamental operations — i.e., that satisfies

$$f\Big(\varphi_j(x_1, x_2, \ldots, x_{\tau(j)})\Big) = \varphi'_j\Big(f(x_1),\ f(x_2),\ \ldots,\ f(x_{\tau(j)})\Big)$$

for all $j \in J$ and all $x_1, x_2, \ldots \in X$. We may call this a **homomorphism of arity τ** to emphasize the particular arity being used. This generalizes the definitions of lattice homomorphism, monoid homomorphism, group homomorphism, and ring homomorphism, given in 4.26, 8.1, 8.9, and 8.18.

Note that this definition does not involve any additional properties that may be enjoyed by the algebraic systems X and X'. For instance, φ_1 is *commutative* if it is a binary operation satisfying $\varphi_1(x_1, x_2) = \varphi_1(x_2, x_1)$, but this additional information is not relevant in determining whether f is a homomorphism. A function $f : X \to Y$, from one monoid into another, is a monoid homomorphism if and only if it satisfies $f(x_1 \,\square_X\, x_2) \;=\; f(x_1) \,\square_Y\, f(x_2)$, regardless of whether one or both monoids are commutative.

8.49. *Exercise.* If $f : X \to Y$ is an isomorphism (i.e., a bijective homomorphism) from one algebraic system of arity τ onto another, then $f^{-1} : Y \to X$ is also a homomorphism.

8.50. Let τ be an arity function. Our main interest lies not in all algebraic systems of type τ, but just in those algebraic systems that satisfy a given collection of *identities*, as explained below.

Let X be an algebraic system of arity τ. A **term** in X is an n-ary operation on X that is formed by composing finitely many of the fundamental operations, finitely many times. For instance, if φ_1 is a 1-ary operation and φ_2 is a 2-ary operation, then the function

$$p(w, x, y, z) \quad = \quad \varphi_2\left(\varphi_1(x)\,,\ \varphi_2\Big(y\,,\ \varphi_2\big(x\,,\ \varphi_1(z)\big)\Big)\right)$$

is a term in the algebraic system. Note that the right side does not depend on w; this illustrates that a term is not required to depend on all of its arguments. The identity map $x \mapsto x$ will be considered a term; it is the composition of *no* fundamental operations.

Note that our method of specifying a term depends only on the arities of the φ_j's (i.e., the values of $\tau(j)$) and on the order of composition of the φ_j's, not on other information about X or the φ_j's. For instance, if $\tau(1) = 1$ and $\tau(2) = 2$, then we can define a term by $\varphi_2(x, \varphi_1(z))$ but not by $\varphi_2(x, \varphi_1(z), w)$ — regardless of other properties that may or may not be enjoyed by the functions φ_1 and φ_2. Hence corresponding compositions of fundamental operations can be used to define corresponding terms in different algebraic systems, provided they are of the same arity τ. By a "term of arity τ" we shall mean a *method of specifying a term.* The method does not refer to any particular algebraic system X; it specifies a corresponding term for each algebraic system of arity τ.

An **equational axiom**, or **identity**, for algebras X of arity τ is a condition on X of the form

$$p(x_1, x_2, \ldots, x_n) \;=\; q(x_1, x_2, \ldots, x_n) \qquad \text{for all } x_1, x_2, \ldots, x_n \in X,$$

where p and q are terms of arity τ. Such a condition is satisfied by some algebraic systems of arity τ and not by others.

For instance, let τ be an arity function such that $\tau(1) = 2$ — i.e., such that φ_1 is a binary operation. Then the **commutative law** for φ_1 is the equational axiom $\varphi_1(x_1, x_2) = \varphi_1(x_2, x_1)$. This equational axiom is satisfied by the binary operation of a commutative group such as $(\mathbb{R}, +)$, but not by the binary operation of a noncommutative group such as $\mathrm{Perm}(X)$ — see 8.10.i.

Let τ be an arity, and let $\mathfrak{I}$ be a collection of identities compatible with τ. By an **algebraic system of variety** $(\boldsymbol{\tau}, \mathfrak{I})$ we shall mean a algebraic system X of arity τ that satisfies all the identities in $\mathfrak{I}$. Examples will be given starting in 8.52.

8.51. *Proposition* (*optional*). The equational variety $(\tau, \mathfrak{I})$ is a *complete theory*, in the following sense: Let n be a nonnegative integer, and let p and q be any terms of arity τ, each taking n arguments. Consider the equation

$$p(x_1, x_2, \ldots, x_n) \;=\; q(x_1, x_2, \ldots, x_n) \qquad \text{for all } x_1, x_2, \ldots, x_n \in X \qquad (*)$$

(not necessarily belonging to $\mathfrak{I}$). Then

> equation $(*)$ is a **semantic theorem** in $(\tau, \mathfrak{I})$, in the sense that it is satisfied by every algebraic system of type $(\tau, \mathfrak{I})$,

if and only if

> equation $(*)$ is a **syntactic theorem** in $(\tau, \mathfrak{I})$, in the sense that it can be deduced from the identities that belong to $\mathfrak{I}$ by using finitely many substitutions.

Remarks. Thus, for any equation $(*)$, we can find either a proof (as in 8.7) or a counterexample (as in 8.27, which shows by example that not every ring is commutative).

Sketch of proof. We shall omit most of the proof, since it is not needed later in this book; it can be found in more detail in Johnstone [1987] and in other textbooks. Obviously, any syntactic theorem is also a semantic theorem. To prove any semantic theorem is a syntactic theorem, the main idea is this: Call two terms α and β "equivalent" if the equation $\alpha = \beta$ is a syntactic theorem in $(\tau, \mathfrak{I})$; this is an equivalence relation on terms. The quotient set — i.e., the set of all equivalence classes — can be made into an algebraic system Ξ of type $(\tau, \mathfrak{I})$ in a natural way. Since $p = q$ is a semantic theorem, it is satisfied by Ξ; hence $p = q$ is a semantic theorem.

Optional exercise. Carry out this argument in detail for some particularly simple variety — e.g., the variety of monoids, described in 8.53. *Related discussion*: see 14.58.

Examples of Equational Varieties

8.52. Let τ be defined on the set $J = \{1, 2\}$ by the values $\tau(1) = \tau(2) = 2$. Then an algebraic system of arity τ means a set X equipped with two binary fundamental operations.

Let X and Y be two such algebraic systems, with fundamental operations denoted by $\vee$ and $\wedge$. Then a mapping $f : X \to Y$ is a homomorphism (of arity τ) if and only if it satisfies

$$f(x_1 \vee x_2) = f(x_1) \vee f(x_2), \qquad f(x_1 \wedge x_2) = f(x_1) \wedge f(x_2)$$

for all $x_1, x_2 \in X$.

A *lattice* is an algebraic system with this arity function τ, which also satisfies the equational axioms L1–L3 of 4.20. Thus, lattices make up the equational variety $(\tau, \{\text{L1}, \text{L2}, \text{L3}\})$. Some algebraic systems of arity τ are lattices, and some are not. A *lattice homomorphism* is a homomorphism of arity τ between two lattices.

In many contexts we describe a lattice in terms of its ordering $\preccurlyeq$, but for purposes of this chapter we must instead describe a lattice in terms of its fundamental operations $\vee$, $\wedge$. Most other kinds of ordered sets — posets, chains, directed sets, etc. — cannot be described in an analogous fashion, and so they do not form equational varieties.

8.53. Let τ be the function defined on $J = \{0, 1\}$ by the values $\tau(0) = 0$ and $\tau(1) = 2$. Then an algebraic system of arity τ is a set X equipped with one nullary operation φ_0 (i.e., a specially selected constant member of X) and one binary operation φ_1. A homomorphism from one algebraic system of this arity to another is a mapping $f : X \to X'$ that satisfies

$$f(\varphi_0) = \varphi_0' \qquad \text{and} \qquad f\Big(\varphi_1(x_1, x_2)\Big) = \varphi_1'\Big(f(x_1), f(x_2)\Big)$$

for all $x_1, x_2 \in X$.

A *monoid* is an algebraic system with this arity τ whose two fundamental operations φ_0, φ_1 satisfy these three axioms:

$$\varphi_1\Big(x, \varphi_1(y, z)\Big) = \varphi_1(\varphi_1(x, y), z), \qquad \text{(associative law)}$$

$$\varphi_1(\varphi_0, x) = x, \qquad \varphi_1(x, \varphi_0) = x. \qquad \text{(identity laws)}$$

A *group* is an algebraic system of arity σ defined on $\{0, 1, 2\}$ by $\tau(0) = 0$, $\tau(1) = 2$, $\tau(2) = 1$ — that is, with the same fundamental operations as monoids, plus a unary operation φ_2 — and that satisfies the equational axioms above and also these two equations:

$$\varphi_1\Big(x, \varphi_2(x)\Big) = \varphi_0 = \varphi_1\Big(\varphi_2(x), x\Big). \qquad \text{(inverse laws)}$$

A homomorphism of arity σ means a homomorphism $f : X \to X'$ of arity τ that also satisfies $f(\varphi_2(x)) = \varphi_2'(f(x))$, regardless of whether X, X' satisfy the equational axioms for a group. A monoid or group is commutative if it satisfies the equational axioms listed above plus this axiom:

$$\varphi_1(x_1, x_2) = \varphi_1(x_2, x_1). \qquad \text{(commutative law)}$$

Of course, the various properties of monoids and groups take a simpler appearance if we use the notations introduced earlier in this chapter:

$$\varphi_0 = i, \qquad \varphi_1(x_1, x_2) = x_1 \Box x_2, \qquad \varphi_2(x) = x^{-1}.$$

Thus, we prefer those notations when we are working solely with monoids or groups. The notation $\varphi_0, \varphi_1, \varphi_2$ is advantageous only when we are trying to see how monoids and groups fit into a more general theory of algebraic systems.

8.54. Rings with unit were introduced in 8.18. A ring with unit is an algebraic system with arity function given by the table below and satisfying certain identities that we shall not list here. Rings with unit form an equational variety. Attaching one more equational axiom, we obtain commutative rings with unit, another equational variety.

j	0	1	2	3	4
$\tau(j)$	0	2	1	0	2
φ_j	0	$+$	$-$	1	$\cdot$

Boolean rings will be studied in 13.13 and thereafter. A Boolean ring is a ring with unit, in which each element satisfies $x^2 = x$. Thus, Boolean rings form an equational variety; we simply add one more identity to the list of identities for rings with unit. The fundamental operations of a Boolean ring are the fundamental operations of a ring with unit: $(0, 1, -, \cdot, +)$.

Boolean lattices are another equational variety, described in 13.1, with rather different fundamental operations $0, 1, \complement, \vee, \wedge$, satisfying rather different equational axioms. However, in a certain sense, Boolean rings and Boolean lattices are different views of the same objects: Boolean rings and Boolean lattices can be transformed into each other, as described in 13.14. The terms "Boolean ring," "Boolean lattice," and "Boolean algebra" are used interchangeably in some of the literature, but in this book we distinguish between the ring and lattice viewpoints.

A field X has a multiplicative inverse operation $x \mapsto x^{-1}$, but that operation is only defined on $X \setminus \{0\}$, not on all of X. Consequently we cannot view fields as an equational variety (unless we replace our definitions with much more complicated definitions, as some mathematicians do). Instead we shall simply view a field as a particularly interesting member of the variety of commutative rings with unit.

8.55. Let $\mathbb{F}$ be a field. In Chapter 11 we shall introduce $\mathbb{F}$-linear spaces. These form an equational variety, but their arity is a little more complicated to describe. The operations of an $\mathbb{F}$-linear space X are the operations of an additive group, together with the operation of scalar multiplication. In most contexts, scalar multiplication is thought of as a mapping $m : (c, x) \mapsto cx$, from $\mathbb{F} \times X$ into X. However, to fit scalar multiplication into our theory of universal algebras, we prefer to think of scalar multiplication as a collection of many unary operations $m_c : x \mapsto cx$. We have one mapping from X into X for each $c \in \mathbb{F}$. If $\mathbb{F}$

is an infinite field (such as the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$), then we have *infinitely many* of these unary operations.

We obtain the equational variety of $\mathbb{F}$-linear algebras (defined as in 11.3) by adding one more fundamental operation (vector multiplication) governed by a few more identities.

8.56. Lattice groups were introduced in 8.38. They are an equational variety, with the fundamental operations of additive groups plus the fundamental operations of a lattice, and with appropriate equational axioms.

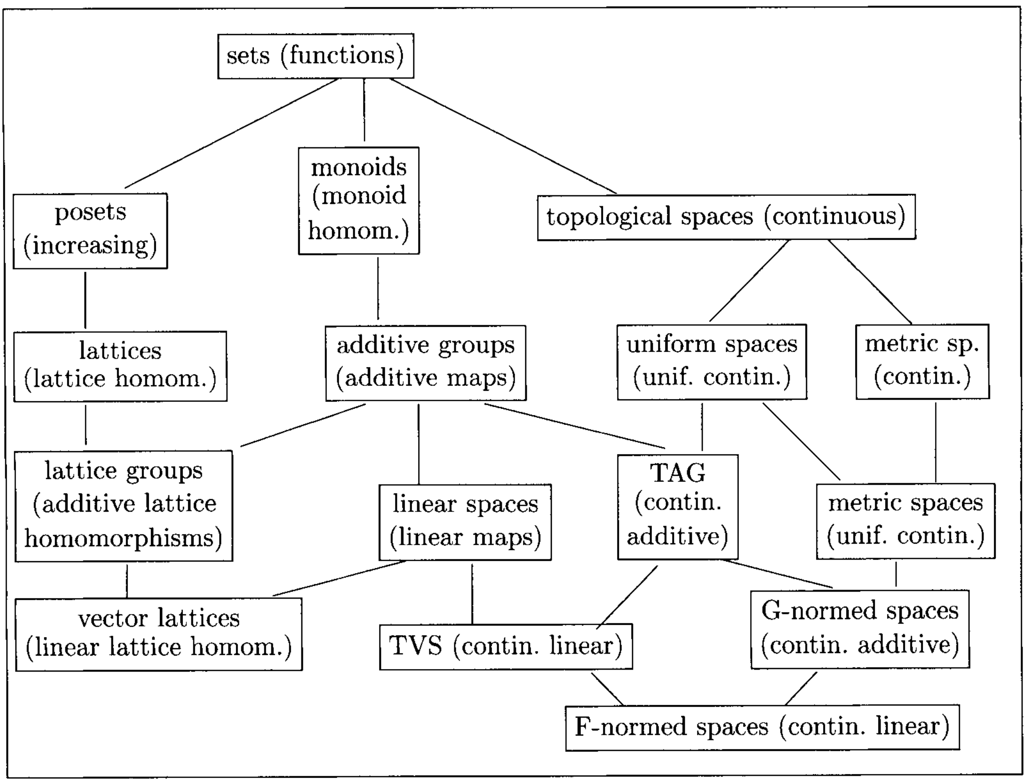
Part of our definition of a lattice group was the translation-invariance of the ordering, $x \preccurlyeq y \Rightarrow x + z \preccurlyeq y + z$, introduced in 8.30. However, an implication is not an equational axiom, and a partial ordering is not a binary operator; the condition $x \preccurlyeq y \Rightarrow x+z \preccurlyeq y+z$ is not permitted as an ingredient in our theory of universal algebras. How can the condition be reformulated? We can dispense with $\preccurlyeq$, replacing statements of the form $x \preccurlyeq y$ with corresponding statements of the form $x \vee y = y$. Thus, the translation-invariance of the ordering can be restated as the equational axiom $(x + z) \vee (y + z) = (x \vee y) + z$ (introduced in 8.42.a).

Vector lattices and lattice algebras will be introduced in 11.44. They are equational varieties, with the fundamental operations of vector spaces or algebras together with $\vee$, $\wedge$.

Ordered monoids, ordered groups, and ordered vector spaces, introduced in 8.30 and 11.44, are *not* equational varieties, since their orderings cannot be described in terms of fundamental operations.

Chapter 9

Concrete Categories



9.1. *Preview.* The chart above shows some of the most basic categories that we shall consider in this book. (An additional chart at the beginning of Chapter 22 shows some more advanced categories.) The components of a *category* are (i) its *objects* — sets with additional structure — and (ii) its *morphisms* — mappings between those sets, which (in most cases of interest) preserve that additional structure in at least one direction. Morphisms are indicated in parentheses in the chart; for instance, "topological spaces (continuous)" is included in the chart to indicate the category whose objects are topological spaces and whose morphisms are continuous maps between those spaces.

Precise definitions will be given in 9.3, and examples will be given in some detail starting in 9.6. Some of the categories mentioned in this chapter are not introduced formally until later; this chapter may be considered as a preview of those categories. The line segments in the chart indicate natural relations between categories via forgetful functors (discussed in 9.34). Some, but not all, of these forgetful functors are given by the inclusion of a subcategory in a category (discussed in 9.5).

The category theory being introduced here is based loosely on the theory of Eilenberg and Mac Lane. It should not be confused with Baire category theory, an unrelated topic introduced elsewhere in this book. The Eilenberg-Mac Lane theory was originally developed mainly for applications in algebraic topology (discussed briefly in 9.33); recently it has also been useful in the abstract theory of computer programs. However, most theorems of Eilenberg-Mac Lane category theory are irrelevant to the purposes of this book and will be omitted. The *language* of the Eilenberg-Mac Lane theory is useful to us, but we shall take the liberty of modifying that language slightly to make it more useful for the purposes of analysts; thus some of our definitions differ slightly from the definitions to be found in books on category theory.

Some other introductions to category theory can be found in Herrlich and Strecker [1979], Mac Lane [1971], and Mac Lane and Birkhoff [1967].

9.2. *Introductory discussion.* We say that two objects X and Y are **isomorphic** if there is a correspondence between them that preserves (in both directions) all the structure currently of interest. Such a mapping is then called an **isomorphism**. Different branches of mathematics, being concerned with different kinds of structures — order, algebraic, topological, uniform, etc. — have different meanings for the terms "isomorphic" and "isomorphism." (This multiplicity of meanings may confuse some beginners.) However, most meanings of isomorphic and isomorphism can be subsumed by one abstract meaning developed in this chapter; see particularly 9.14.

If two objects A and B are isomorphic, then they differ only in their labeling and are essentially two different representations of the *same* object. They can be used interchangeably, provided that we are willing to relabel everything else that they interact with. The "essence" of the objects is the part of them that does not depend on the particular choice of representation. This interchangeability is the heart of mathematics (and, indeed, of all abstract thinking); for instance, the "essence" of the number 4 does not depend on whether we are dealing with four apples or four airplanes.

When two objects X and Y are isomorphic, we may sometimes *identify* X and Y, and treat them as equal, because for most practical purposes they are the "same" set. We may even write $X = Y$, if this will not cause confusion. More generally, suppose X is isomorphic to a subset of a set Y; then we may identify X with that subset and write $X \subseteq Y$. A structure-preserving map from one set into another is sometimes called an **embedding**, although this term has more specific meanings in some contexts.

Different categories — groups, topological spaces, etc. — have different properties, so ultimately they must be studied separately. However, there are analogies between the most elementary properties of these different categories — e.g., between subgroups and topological subspaces, or between products of groups and products of topological spaces. These analogies may help the beginner through the unavoidable plethora of definitions and elementary propositions.

Definitions and Axioms

9.3. Following are precise definitions. Some readers will find it helpful to glance ahead to the examples, which begin in 9.6.

A **concrete category** consists of a collection of *objects* and a collection of *morphisms*.

An **object** is a pair $(X, \mathcal{S})$ consisting of a set X and some additional structure $\mathcal{S}$ on X (such as a preordering or a σ-algebra); then X is called the **underlying set** of the object $(X, \mathcal{S})$. The nature of the "additional structure" will vary from one category to another, but the meaning of this term will be clear in particular categories. We will often refer to X itself as the object if the choice of $\mathcal{S}$ is clear or does not need to be mentioned explicitly, but it is understood that $\mathcal{S}$ is still part of the object. One set X may give rise to several different objects, by being equipped with several different structures — for instance, two different preordered sets may contain the same points. Two objects $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ are considered *equal* (as objects) if $X = Y$ and $\mathcal{S} = \mathcal{T}$.

To define morphisms, consider triples

$$(X, \mathcal{S}) \quad \xrightarrow{f} \quad (Y, \mathcal{T})$$

consisting of two objects of the given category and a function $f : X \to Y$ whose domain and codomain are the underlying sets X and Y of those two objects. The collection of all such triples forms a class that is usually larger than what we want. Some subclass will be specified as the collection of **morphisms** for the category; the specified subclass must satisfy two axioms noted below. When $f : (X, \mathcal{S}) \to (Y, \mathcal{T})$ is a morphism, we call $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ its **domain** and **codomain**, respectively. A morphism is also sometimes known as an *arrow*.

The collection of morphisms must satisfy these two axioms:

(i) **(Compositions)** Any composition of two morphisms is a morphism. In other words, if A, B, C are objects and $f : A \to B$ and $g : B \to C$ are morphisms, then $g \circ f : A \to C$ is a morphism.

(ii) **(Identity)** For each object $A = (X, \mathcal{S})$, the identity map i_X on the underlying set X is a morphism from A to A.

In most categories of interest, the class of morphisms is chosen so that the morphisms preserve the structure of the objects in at least one direction, but that preservation is *not* explicitly built into the two axioms listed above. In principle, it is possible to create a category in which the morphisms are entirely unrelated to the additional structure — but that would not be a particularly interesting category.

The literature of category theory sometimes refers to categories by their objects. For instance, when topological spaces are used for objects, continuous maps are almost invariably used for morphisms, so we may refer to the "category of topological spaces." However, in general, a particular choice of objects does not force upon us a particular description of morphisms, nor vice versa. Thus, we may refer to the "category of metric spaces with continuous maps," or the "category of metric spaces with uniformly continuous maps;" these are two different categories.

Whenever we discuss two or more objects and/or morphisms together, it should be understood that all the objects and/or morphisms being considered are in the *same* category, unless specified otherwise.

9.4. A set X may be made into an object in more than one way, by equipping it with different structures $\mathcal{S}_1, \mathcal{S}_2$ — for instance, different topologies or different orderings. Whether or not a function $f : X \to Y$ is a morphism depends on what structures $\mathcal{S}, \mathcal{T}$ we attach to X and Y; the function f may be a morphism for one choice of structures but not for another choice.

If the identity map i_X is a morphism from $(X, \mathcal{S})$ into $(X, \mathcal{T})$, we say that $\mathcal{S}$ is **stronger** (or *finer*) than $\mathcal{T}$, or that $\mathcal{T}$ is **weaker** (or *coarser*) than $\mathcal{S}$. Thus any structure is stronger than itself. Here mathematical language differs from everyday English, which would prohibit anything from being stronger than itself; see 1.4.

Clearly, the relation "stronger than" is a preordering (i.e., transitive and reflexive) on the collection of all structures on X. In many categories of interest (but not all), this preordering is also antisymmetric and thus a partial ordering — i.e., in many categories of interest, if the identity map $i_X : X \to X$ is a morphism in both directions between $(X, \mathcal{S})$ and $(X, \mathcal{T})$, then $\mathcal{S} = \mathcal{T}$. For instance, if two topologies are both stronger than each other, then they are equal; see the last paragraph of 9.8. In some categories, if *either* structure is stronger than the other, then they are equal; see the last paragraph of 9.11.

This terminology — "stronger," "weaker," etc. — will also be applied to any devices (metrics, gauges, etc.) that are used to define the structure of a category. For instance, let d and e be metrics that determine topologies $\mathcal{T}_d$ and $\mathcal{T}_e$ and uniformities $\mathcal{U}_d$ and $\mathcal{U}_e$ on a set X. We shall say that d is **topologically stronger** than e if $\mathcal{T}_d$ is stronger than $\mathcal{T}_e$ (that is, if $\mathcal{T}_d \supseteq \mathcal{T}_e$); we shall say that d is **uniformly stronger** than e if $\mathcal{U}_d$ is stronger than $\mathcal{U}_e$. Two metrics are **topologically equivalent** or **uniformly equivalent** if they determine the same topology or the same uniformity.

This syntactic convention also applies to other devices than metrics — e.g., it also applies to gauges. It even applies to uniformities: one uniformity $\mathcal{U}$ is topologically stronger than another uniformity $\mathcal{U}'$ if it determines a stronger topology — i.e., if $\mathcal{T}_{\mathcal{U}} \supseteq \mathcal{T}_{\mathcal{U}'}$, where the topologies are defined as in 5.33.

If the context is understood, then we may omit mentioning the category — e.g., we may simply say d is **stronger** than e or d is **equivalent** to e. This omission is made most often for topological structure — i.e., if no other meaning is evident, then "stronger" usually means "topologically stronger."

9.5. Let $\mathfrak{S}$ and $\mathfrak{K}$ be categories. We say that $\mathfrak{S}$ is a **subcategory** of $\mathfrak{K}$ if these two conditions are satisfied:

(i) $\{\mathfrak{S}\text{-objects}\} \subseteq \{\mathfrak{K}\text{-objects}\}$. More precisely, every $\mathfrak{S}$-object can also be viewed as a $\mathfrak{K}$-object — perhaps via some change of description, as in 9.10 — and the mapping from $\mathfrak{S}$-objects to $\mathfrak{K}$-objects is injective.

(ii) Whenever A and B are objects of $\mathfrak{S}$ (and hence also objects of $\mathfrak{K}$), then every $\mathfrak{S}$-morphism from A into B is also a $\mathfrak{K}$-morphism from A into B.

We say $\mathfrak{S}$ is a **full subcategory** of $\mathfrak{K}$ if condition (i) is satisfied, as well as the following strengthened version of condition (ii):

(ii′) Whenever A and B are objects of $\mathfrak{S}$ (and hence also objects of $\mathfrak{K}$), then the $\mathfrak{S}$-morphisms from A into B are the same as the $\mathfrak{K}$-morphisms from A into B.

Examples will be given below; see particularly 9.10.

Examples of Categories

9.6. The simplest category is the **category of sets**, in which the objects are sets (without any additional structure specified) and the morphisms are functions.

For an isomorphism in this category, we might use a bijection between two sets. Then two sets are isomorphic if they have the same cardinality.

9.7. A category can be formed by taking **convergence spaces** for objects and convergence preserving maps for morphisms; see 7.33.

9.8. Inverse image categories. The categories of measurable spaces, topological spaces, and uniform spaces differ in their deeper properties, but they are quite similar in their most elementary properties. In each of these categories, an object is a pair $(X, \mathcal{S})$ consisting of a set X and a collection $\mathcal{S}$ of specially designated sets — a σ-algebra or topology $\mathcal{S} \subseteq \mathcal{P}(X)$, or a uniformity $\mathcal{S} \subseteq \mathcal{P}(X \times X)$. In each of these categories, a morphism is a mapping with respect to which the inverse image of a specially designated set is also a specially designated set. This is explained in greater detail below.

Topological spaces form the objects of a category. In this category, a morphism $f : (X, \mathcal{S}) \to (Y, \mathcal{T})$ is a map $f : X \to Y$ with the property that $T \in \mathcal{T} \;\Rightarrow\; f^{-1}(T) \in \mathcal{S}$ — i.e., for which the inverse image of an open set is an open set. Such functions are called **continuous** maps. Some elementary examples of continuous maps are given in 15.17.

Measurable spaces form the objects of a category. In this category, a morphism $f : (X, \mathcal{S}) \to (Y, \mathcal{T})$ is a mapping $f : X \to Y$ with the property that $T \in \mathcal{T} \;\Rightarrow\; f^{-1}(T) \in \mathcal{S}$ — i.e., for which the inverse image of a measurable set is a measurable set. Such functions are called **measurable** mappings. In some contexts the choices of $\mathcal{S}$ and $\mathcal{T}$ are understood and do not need to be mentioned — one may refer to a measurable mapping from X to Y — but we emphasize that the meaning of "measurable mapping" does nevertheless depend very much on the choices of $\mathcal{S}$ and $\mathcal{T}$. In most of the theory developed in later chapters, the codomain Y is a topological space and $\mathcal{T}$ is the σ-algebra of Borel subsets of Y, but no such restriction will be imposed in the more general theory developed in this chapter. (We remark that even greater restrictions are imposed in applied mathematics. In that context, a "measurable mapping" usually means a measurable mapping from an open subset of $\mathbb{R}^m$ equipped with its Lebesgue measurable subsets to an open subset of $\mathbb{R}^n$ equipped with its Borel subsets. This class of mappings is not closed under composition; thus, the "measurable mappings" of applied mathematics do *not* form the morphisms of a category.)

Uniform spaces form the objects of a category. In this category, a morphism $f : (X, \mathcal{S}) \to (Y, \mathcal{T})$ is a mapping $f : X \to Y$ with the property that

$$T \in \mathcal{T} \quad \Rightarrow \quad \{(x_1, x_2) \in X \ : \ (f(x_1), f(x_2)) \in T\} \ \in \ \mathcal{S}$$

— i.e., for which the inverse image of a vicinity is a vicinity. Such functions are called **uniformly continuous** mappings. Such functions are studied further, and examples are given, in Chapter 18.

It will sometimes be convenient to adopt a notation that makes these three categories look more alike. For any set X, let us define

$$\widehat{X} \quad = \quad \begin{cases} X & \text{for measurable or topological spaces} \\ X \times X & \text{if we are working with uniform spaces.} \end{cases}$$

For any mapping $f : X \to Y$, define a mapping $\widehat{f} : \widehat{X} \to \widehat{Y}$ by taking

$$\widehat{f} \quad = \quad \begin{cases} f & \text{for measurable or topological spaces} \\ f \times f & \text{if we are working with uniform spaces,} \end{cases}$$

where $(f \times f) : (X \times X) \to (Y \times Y)$ is defined by $(f \times f)(x_1, x_2) = (f(x_1), f(x_2))$. With these conventions, an object (in any of the three categories) consists of a pair $(X, \mathcal{S})$ where $\mathcal{S}$ is a collection of subsets of $\widehat{X}$ satisfying certain axioms, and a **morphism** $f : (X, \mathcal{S}) \to (Y, \mathcal{T})$ is a mapping $f : X \to Y$ with the property that

$$T \in \mathcal{T} \quad \Rightarrow \quad \widehat{f}^{-1}(T) \in \mathcal{S}.$$

It is easy to verify that if a mapping $f : X \to Y$ satisfies

$$T \in \mathcal{G} \quad \Rightarrow \quad \widehat{f}^{-1}(T) \in \mathcal{S}$$

for some generating collection of sets $\mathcal{G} \subseteq \mathcal{T}$, then f is a morphism; here "generating collection" is defined as in 5.23.b, 5.26.e, and 5.37.

For each of our inverse image categories, a structure $\mathcal{S}$ on a set X is **stronger** than another structure $\mathcal{T}$ on the same set X precisely when $\mathcal{S} \supseteq \mathcal{T}$. Thus, the *stronger* structure is represented by the *larger* set.

We note a few important examples of subcategories that will be studied in later chapters: Use metric spaces for objects, and use uniformly continuous maps for morphisms; the resulting category is a subcategory of either the topological spaces or the uniform spaces. Further subcategories are obtained by further restricting the choice of morphisms: Use

Hölder continuous maps, Lipschitzian maps, or nonexpansive maps; we shall see in later chapters that these classes of maps are closed under composition.

Some functional analysis books gloss over the distinction between topological spaces and uniform spaces, because in the setting of topological vector spaces — or more generally, in the setting of topological Abelian groups — the two kinds of structures are nearly interchangeable: There is a one-to-one correspondence between topologies and "nice" uniformities, as shown in 26.37. However, we shall maintain a distinction between topological and uniform spaces, because this facilitates understanding and because occasionally one wants to apply these concepts in some context other than that of topological Abelian groups.

9.9. A **pointed topological space** is a topological space X with a particular point $x_0 \in X$ selected; that point is called the **base point** of the space. A category can be formed with objects consisting of pointed topological spaces, and with morphisms consisting of continuous maps that preserve the base points. This category will be important in 9.33.

9.10. Categories of ordered sets can be formed in many ways. Perhaps the simplest way is to use preordered sets for objects and to use increasing mappings for morphisms. Many important subcategories of this category can be obtained by using a smaller collection of objects — e.g., chains or complete lattices — and/or by using some smaller collection of morphisms — e.g., sup-preserving maps.

In the category of preordered sets with increasing mappings, the statement "$\preccurlyeq$ is stronger than $\sqsubseteq$" (as defined in 9.4) means that $x \preccurlyeq y \quad \Rightarrow \quad x \sqsubseteq y$; equivalently, it means $\mathrm{Graph}(\preccurlyeq) \subseteq \mathrm{Graph}(\sqsubseteq)$. Thus the *stronger* preordering is represented by the *smaller* set, in contrast with the situation described in the last paragraph of 9.8.

Exercise (*optional*). Let X and Y be preordered sets, equipped with their lower set topologies (defined as in 5.15.d). Show that a function $f : X \to Y$ is increasing if and only if it is continuous (defined in 9.8). Conclude that preordered sets (with increasing mappings for morphisms) are a full subcategory of topological spaces (with continuous mappings for morphisms).

Caution: Although we may view each preordered set $(X, \preccurlyeq)$ as a topological space $(X, \mathcal{S})$, note that the ordering is not *equal* to the topology. Indeed, the ordering $\preccurlyeq$ (or its graph) is a subset of $X \times X$, whereas the topology $\mathcal{S}$ is a subset of $\mathcal{P}(X)$. Thus, we *change our description* of the object when we go from $(X, \preccurlyeq)$ to $(X, \mathcal{S})$.

9.11. Algebraic categories. Let τ be an arity function (defined in 8.47). The universal algebras of type τ can be used for the objects of a category, with the homomorphisms of type τ (defined in 8.48) for the morphisms.

However, generally we are interested in a full subcategory of that category, obtained as follows: Let $\mathcal{I}$ be a collection of identities compatible with τ; then the algebraic systems of variety $(\tau, \mathcal{I})$ (defined in 8.50) can be used as the objects for a category. Examples are the category of lattices, the category of monoids, the category of groups, the category of Abelian groups, the category of lattice groups, and the category of rings.

In an algebraic category,

($*$) if $f : X \to Y$ is a bijection and a morphism, then $f^{-1} : Y \to X$ is also a morphism.

This invertibility property is not shared by most other kinds of categories studied in this book.

In an algebraic category, it is not particularly meaningful to discuss whether one structure on a set is "stronger" than another. Indeed, if $\mathcal{S}$ and $\mathcal{T}$ are two structures on a set X, and the identity map i_X is a morphism in *either* direction between $(X, \mathcal{S})$ and $(X, \mathcal{T})$, then in fact $\mathcal{S} = \mathcal{T}$. (That follows from $(*)$.) Thus, if one structure is stronger than another, then they are equal.

9.12. *Remarks: overview of categories.* There is some overlap between our general classes of categories; for instance, lattices may be viewed as algebraic systems $(X, \vee, \wedge)$ or as preordered sets $(X, \preccurlyeq)$. Each viewpoint has its advantages, depending on what properties and structures we wish to study.

Early chapters of this book are devoted to the simplest categories. Later chapters will introduce more complicated and specialized categories. Many of these are "hybrid categories," combining structures of two simpler categories and also imposing some condition of compatibility between the two structures. For instance, a topological linear space has both a topology and a linear structure, which must be compatible in that the vector space operations are jointly continuous.

Viewing an object in different categories may yield different kinds of information about that object. Following are a few examples.

Every metric determines a topology (see 5.15.g), but there are also some topologies that are not determined by any metric. The topologies that can be determined by a metric are called *metrizable topologies.* Thus, metrizable topological spaces form a full subcategory of topological spaces (both equipped with continuous maps for morphisms).

On the other hand, one topology $\mathcal{T}$ may be determined by two different metrics d, d'. Thus the category of metrizable topological spaces $(X, \mathcal{T})$ is slightly different from the category of metric spaces (X, d) (both equipped with continuous maps), since (X, d) and (X, d') are different objects in the latter category.

Different questions arise naturally in these slightly different categories. For instance:

- A theorem of Banach states that any strict contraction self-mapping of a complete metric space has a fixed point. This is a statement about metric spaces. If we replace the metric with another metric that yields the same topology, the self-mapping may no longer be a strict contraction. (Meyers' converse in 19.47 considers the effects of such a replacement.)

- A theorem of Baire states that in a topologically complete space (i.e., a topological space $(X, \mathcal{T})$ whose topology can be given by various metrics, at least one of which is complete), the intersection of countably many open dense sets is dense. This is a statement about metrizable spaces. If we replace a metric with an equivalent metric, the open sets and the dense sets are unaffected.

This kind of distinction is also displayed in a chart in 18.1.

9.13. *Nonconcrete categories (optional).* Concrete categories will suffice for the purposes of this book, but the reader should be aware that the Eilenberg-Mac Lane theory deals with other kinds of categories as well. A **category** consists of certain collections of mathematical devices called **objects** (not necessarily sets) and **morphisms** (not necessarily functions), satisfying certain rules listed below. Each morphism is represented in the form $f : A \to B$, where f is the name of the morphism and A and B are objects called the **domain** and **codomain** of the morphism, respectively. Morphisms must satisfy these rules:

(i) If $f : A \to B$ and $g : B \to C$ are morphisms, then there exists a morphism $g \circ f : A \to C$, called the **composition** of f and g.

(ii) Composition of morphisms is **associative**: $f \circ (g \circ h) = (f \circ g) \circ h$.

(iii) For each object A, there exists a morphism i_A, called the **identity morphism** of A, satisfying

$$g \circ i_A = g \text{ for any morphism } g \text{ with domain } A, \text{ and}$$

$$i_A \circ f = f \text{ for any morphism } f \text{ with codomain } A.$$

It is an easy exercise to show that the identity morphism is unique — i.e., if i_A and i'_A both satisfy condition (iii), then $i_A = i'_A$.

We mention two examples of nonconcrete categories:

a. Let $(S, \preccurlyeq)$ be any preordered set; for objects take elements of S; for morphisms take ordered pairs (x, y) satisfying $x \preccurlyeq y$. Note that in this category, for any two objects x and y there is at most one morphism from x to y.

b. Let X and Y be topological spaces. Let $[0,1]$ have its usual topology, and let $[0,1] \times X$ have the product topology (discussed elsewhere in this chapter and in Chapter 15).

Let $f, g : X \to Y$ be continuous mappings. A **homotopy** from f to g is a continuous mapping $h : [0,1] \times X \to Y$ such that

$$h(0, x) = f(x), \qquad h(1, x) = g(x) \qquad\qquad \text{for all } x \in X.$$

If such a mapping exists, we say f and g are **homotopy-equivalent**. The reader should verify that this is an equivalence relation on the collection of all continuous mappings from X into Y.

A category can be formed by taking topological spaces for objects and homotopy equivalence classes for morphisms. This category is typical of the ones used in algebraic topology. See also 9.33.

9.14. *More definitions.* In any category — concrete or not — an *isomorphism* between two objects A and B is a morphism $f : A \to B$ for which there exists another morphism $g : B \to A$ such that $g \circ f = i_A$ and $f \circ g = i_B$. This makes precise the definition given in 9.2.

In a concrete category, this definition can also be restated as follows: An isomorphism between two objects $A = (X, \mathcal{S})$ and $B = (Y, \mathcal{T})$ is a bijection $f : X \to Y$ such that both $f : A \to B$ and $f^{-1} : B \to A$ are morphisms.

In any category, an **automorphism** of an object A is an isomorphism from A onto A. Clearly the identity map of A is an automorphism, but there may be others as well. For instance, the translation mapping $\varphi : x \mapsto x + 3$, from $\mathbb{R}$ into $\mathbb{R}$, is an automorphism of metric spaces (but it is not an automorphism, or even a morphism, of additive groups, since it does not preserve the identity).

The inverse of any automorphism is another automorphism. It is easy to see that the automorphisms of A form a (not necessarily Abelian) *group*, with composition of morphisms for the group's binary operation. The automorphism group of A is often denoted by $\mathrm{Aut}(A)$. This notation does not indicate what category is being used. We could indicate it with subscripts. For instance, let φ be the translation map mentioned in the previous paragraph; then $\varphi \in \mathrm{Aut}_{\text{metric spaces}}(\mathbb{R})$ but $\varphi \notin \mathrm{Aut}_{\text{additive groups}}(\mathbb{R})$. However, usually the choice of category is clear from the context, and so the subscripts are not needed.

Initial Structures and Other Categorical Constructions

9.15. *Definition.* Let X be a set, let $\{(Y_\lambda, \mathcal{T}_\lambda) : \lambda \in \Lambda\}$ be a collection of objects in some category, and for each λ let $\varphi_\lambda : X \to Y_\lambda$ be some given mapping. Then an **initial structure** determined by the φ_λ's and $\mathcal{T}_\lambda$'s is a structure $\mathcal{S}$ that makes $(X, \mathcal{S})$ into an object with this property:

> Let $(W, \mathcal{R})$ be any object in the category, and let $f : W \to X$ be any function. Then f is a morphism from $(W, \mathcal{R})$ into $(X, \mathcal{S})$ if and only if each of the compositions $\varphi_\lambda \circ f$ is a morphism from $(W, \mathcal{R})$ into $(Y_\lambda, \mathcal{T}_\lambda)$

— or, equivalently, an object with these two properties:

> (i) Each of the mappings $\varphi_\lambda : X \to Y_\lambda$ is a morphism. (ii) Let $(W, \mathcal{R})$ be any object in the category, and let $f : W \to X$ be any function. Suppose that each of the compositions $\varphi_\lambda \circ f$ is a morphism from $(W, \mathcal{R})$ into $(Y_\lambda, \mathcal{T}_\lambda)$. Then f is a morphism from $(W, \mathcal{R})$ into $(X, \mathcal{S})$.

Exercises.

a. Prove the equivalence stated above.

b. If $\mathcal{S}$ is an initial structure on X determined by the φ_λ's and $\mathcal{T}_\lambda$'s, then $\mathcal{S}$ is weaker than any other structure on X that makes the φ_λ's into morphisms. (For this reason it is sometimes called the **weak structure**.)

c. If $\mathcal{S}_1$ and $\mathcal{S}_2$ are two initial structures on X determined by the φ_λ's and $\mathcal{T}_\lambda$'s, then each of $\mathcal{S}_1, \mathcal{S}_2$ is weaker than the other.

Hence there is at most one initial structure determined by the φ_λ's and $\mathcal{T}_\lambda$'s, in many categories of interest to us — particularly, in our inverse image categories and algebraic categories. It is the *weakest* structure that makes the φ_λ's into morphisms.

Further remarks. Let any X and $(Y_\lambda, \mathcal{T}_\lambda)$'s and φ_λ's be given. Then the initial structure does not always exist; in our algebraic categories, this will be evident from our discussion in 9.20. However, the initial structure does always exist in our inverse image categories; we shall prove that in 9.16.

The definition of initial structure given above is admittedly rather complicated. Simpler, equivalent definitions are available in some categories — e.g., for topological spaces (see 15.24) and for uniform spaces (see 18.9.f).

9.16. *Proposition.* Initial structures exist in the categories of measurable spaces, topological spaces, and uniform spaces.

More precisely: Let $\{(Y_\lambda, \mathcal{T}_\lambda) : \lambda \in \Lambda\}$ be a collection of objects in one of those categories. Let X be a set, and let some mappings $\varphi_\lambda : X \to Y_\lambda$ be given. Then there exists an initial structure $\mathcal{S}$ on X, determined by those spaces and mappings. In fact, $\mathcal{S}$ is the structure generated by the collection of sets

$$\mathcal{G} \quad = \quad \bigcup_{\lambda \in \Lambda} \widehat{\varphi_\lambda}^{-1}(\mathcal{T}_\lambda) \quad = \quad \left\{\widehat{\varphi_\lambda}^{-1}(T) \ : \ \lambda \in \Lambda \text{ and } T \in \mathcal{T}_\lambda\right\},$$

where mappings $\widehat{\varphi_\lambda} : \widehat{X} \to \widehat{Y}_\lambda$ are defined as in 9.8.

Proof. In the category of measurable spaces or topological spaces, $\mathcal{G}$ can be used to generate a structure, since *any* collection of sets can be used to generate a structure. In the category of uniform spaces, each $\widehat{\varphi_\lambda}^{-1}(\mathcal{T}_\lambda)$ is a preuniformity, by 5.40.c; hence $\mathcal{G}$ is a preuniformity, by 5.38.a; hence $\mathcal{G}$ can be used to generate a structure.

Let $\mathcal{S}$ be the structure on X generated by $\mathcal{G}$. We shall show that $\mathcal{S}$ is an initial structure. It is clear that each φ_λ is a morphism from $(X, \mathcal{S})$ into $(Y_\lambda, \mathcal{T}_\lambda)$, since $\mathcal{S} \supseteq \mathcal{G} \supseteq \widehat{\varphi_\lambda}^{-1}(\mathcal{T}_\lambda)$. Now suppose that $(W, \mathcal{R})$ is some object and $g : W \to X$ is some mapping such that each composition $\varphi_\lambda \circ g$ is a morphism; we must show that g itself is a morphism. Let $\Sigma = \{S \subseteq \widehat{X} : \widehat{g}^{-1}(S) \in \mathcal{R}\}$; it suffices to show that $\Sigma \supseteq \mathcal{S}$.

We first show that $\Sigma \supseteq \mathcal{G}$. Fix any $\lambda \in \Lambda$ and fix any $T \in \mathcal{T}_\lambda$. Then $\varphi_\lambda \circ g : (W, \mathcal{R}) \to (Y, \mathcal{T}_\lambda)$ is a morphism; hence $\widehat{g}^{-1}(\widehat{\varphi_\lambda}^{-1}(T)) = (\widehat{\varphi_\lambda} \circ \widehat{g})^{-1}(T)$ is a member of $\mathcal{R}$; hence $\widehat{\varphi_\lambda}^{-1}(T)$ is a member of Σ. Thus $\Sigma \supseteq \mathcal{G}$.

In the category of measurable spaces or topological spaces, Σ is a structure on X, by 5.40.b. Since it is a structure containing $\mathcal{G}$, it also contains the structure generated by $\mathcal{G}$; thus $\Sigma \supseteq \mathcal{S}$.

For the category of uniform spaces, we know $\mathcal{R}$ is a uniformity on W, hence a filter on $W \times W$, hence Σ is a filter on $X \times X$ by 5.40.b. Also Σ contains $\mathcal{G}$, which is a preuniformity. Hence Σ contains the smallest filter that contains $\mathcal{G}$ — that is, Σ contains $\mathcal{S}$.

9.17. *An important special case.* Let $\{\mathcal{S}_\lambda : \lambda \in \Lambda\}$ be a collection of structures on a set X, in the category of measurable spaces, topological spaces, or uniform spaces. Then the **supremum** of the $\mathcal{S}_\lambda$'s is the smallest structure that contains $\bigcup_{\lambda \in \Lambda} \mathcal{S}_\lambda$; it is equal to the initial structure determined by the identity mappings $i_\lambda : X \to (X, \mathcal{S}_\lambda)$.

9.18. Products. Let $\{(Y_\lambda, \mathcal{T}_\lambda) : \lambda \in \Lambda\}$ be a collection of objects in some category, and let $X = \prod_{\lambda \in \Lambda} Y_\lambda$ be the product of the underlying sets. By a **product structure** on X we shall mean an initial structure (defined as in 9.15) determined by the coordinate projections $\pi_\lambda : X \to Y_\lambda$.

Products always exist in our inverse image categories; this is just a special case of 9.16. As we noted in 9.10, preordered sets (with increasing maps for morphisms) may be viewed as a full subcategory of topological spaces; hence products also exist in the category of preordered sets. *Exercise* (*optional*): Show that the product of preordered sets, defined in this fashion, is the same as the product defined in 3.9.j.

Although initial structures do not always exist in our algebraic categories (as we shall see in 9.20), nevertheless product structures always exist. Fundamental operations are defined coordinatewise. For instance, if U, V, W are lattices, then lattice operations are defined on $U \times V \times W$ by

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} \vee \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} u_1 \vee u_2 \\ v_1 \vee v_2 \\ w_1 \vee w_2 \end{bmatrix}, \qquad \begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} \wedge \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} u_1 \wedge u_2 \\ v_1 \wedge v_2 \\ w_1 \wedge w_2 \end{bmatrix}.$$

This construction generalizes readily to products of any number of factors in any equational variety. Say $X = Y_\alpha \times Y_\beta \times Y_\gamma \times \cdots$ is a product of algebraic systems of type $(\tau, \mathfrak{I})$; suppose $\tau(j) = n$; we shall now describe the action of the n-ary operation Φ_j of X in terms of the n-ary operations $\varphi_{\alpha j}, \varphi_{\beta j}, \varphi_{\gamma j}, \ldots$ of the factor spaces. The function Φ_j acts on n-tuples $(x_1, x_2, \ldots, x_n) \in Y^n$. Say $x_i = (y_{i\alpha}, y_{i\beta}, y_{i\gamma}, \ldots)$ where $y_{i\alpha} \in Y_\alpha$, $y_{i\beta} \in Y_\beta$, etc. Then

$$\Phi_j\left(\begin{bmatrix} y_{1\alpha} \\ y_{1\beta} \\ y_{1\gamma} \\ \vdots \end{bmatrix}, \begin{bmatrix} y_{2\alpha} \\ y_{2\beta} \\ y_{2\gamma} \\ \vdots \end{bmatrix}, \ldots, \begin{bmatrix} y_{n\alpha} \\ y_{n\beta} \\ y_{n\gamma} \\ \vdots \end{bmatrix}\right) = \begin{bmatrix} \varphi_{\alpha j}(y_{1\alpha},\, y_{2\alpha},\, \ldots,\, y_{n\alpha}) \\ \varphi_{\beta j}(y_{1\beta},\, y_{2\beta},\, \ldots,\, y_{n\beta}) \\ \varphi_{\gamma j}(y_{1\gamma},\, y_{2\gamma},\, \ldots,\, y_{n\gamma}) \\ \vdots \end{bmatrix}.$$

We leave it to the ambitious reader to unwind all the notation and verify that this formula does indeed satisfy the definition given in 9.15, and verify that X is an algebraic system of type $(\tau, \mathfrak{I})$. It is sometimes called the **direct product**.

Some of our "hybrid" categories also have product objects. For instance, if $((Y_\lambda, \mathcal{T}_\lambda) : \lambda \in \Lambda)$ is a collection of topological vector spaces, then the product topological structure and the product vector space structure on the product set $X = \prod_{\lambda \in \Lambda} Y_\lambda$ are compatible with each other and thus yield a product topological vector space; see 26.20.a.

This does not work in some other categories. For instance, a product of finitely many normed spaces is a normed space, but a product of infinitely many normed spaces is a topological vector space that cannot be equipped with a norm; see 27.7.c.

9.19. *Exercise.* A product of morphisms is a morphism.

More precisely: For each λ, suppose that $f_\lambda : X_\lambda \to Y_\lambda$ is a morphism. Define a mapping f from $P = \prod_{\lambda \in \Lambda} X_\lambda$ into $Q = \prod_{\lambda \in \Lambda} Y_\lambda$ by taking

$$f\Big((x_\alpha,\, x_\beta,\, x_\gamma,\, \ldots)\Big) \quad = \quad \Big(f_\alpha(x_\alpha),\, f_\beta(x_\beta),\, f_\gamma(x_\gamma),\, \ldots\Big)$$

if $\Lambda = \{\alpha, \beta, \gamma, \ldots\}$. Assume that P and Q are equipped with product structures. Then f is a morphism.

9.20. *Discussion of codomains.* Let X be a subset of a set Y; let $i : X \xrightarrow{\subseteq} Y$ be the inclusion map. Then any function $f : W \to X$ is the "same" as the composition

$i \circ f : W \to Y$, insofar as it has the same domain and the same values $f(w) = i(f(w))$; it differs only in our designation of the codomain. See 2.5.c.

Of course, if we attach structures $\mathcal{R}, \mathcal{S}, \mathcal{T}$ to the sets W, X, Y, making them into different objects in a category, then the difference between f and $i \circ f$ may be more substantial: perhaps one is a morphism while the other is not. That depends on our choices of the additional structures $\mathcal{R}, \mathcal{S}, \mathcal{T}$.

Definition. Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be objects in some category. We say that $(X, \mathcal{S})$ is a **subobject** of $(Y, \mathcal{T})$ if these two objects are related by this condition:

> Let $(W, \mathcal{R})$ be any object in the category, and let $f : W \to X$ be any function. Then $(W, \mathcal{R}) \xrightarrow{f} (X, \mathcal{S})$ is a morphism if and only if the composition $(W, \mathcal{R}) \xrightarrow{i \circ f} (Y, \mathcal{T})$ is a morphism.

In other words, $\mathcal{S}$ is the initial structure on X determined by the inclusion map $i : X \xrightarrow{\subseteq} Y$ and the structure $\mathcal{T}$.

Further remarks. In our inverse image categories, if $(Y, \mathcal{T})$ is any object, then every subset $X \subseteq Y$ has a unique subobject structure $\mathcal{S}$; that is just a special case of 9.16. In fact, it follows easily from 9.16 that the subobject structure is $\mathcal{S} = \{\widehat{X} \cap T : T \in \mathcal{T}\}$. It is sometimes called the **trace** of $\mathcal{T}$ on X. In the category of topological spaces, $\mathcal{S}$ is the same as the *relative topology*, which we introduced in 5.15.e.

Subobjects in algebraic categories are studied in more detail below.

9.21. *Basic properties of subalgebras.* We consider the category consisting of the algebraic systems of some type $(\tau, \mathcal{I})$, with homomorphisms of type τ. (For examples, think of the category of lattices, the category of rings, or the category of lattice groups.) In this category, a subobject is called a **subalgebra**. Let X be an object. Show that

a. Y is a subalgebra of X if and only if Y is a subset of X that is closed under the fundamental operations of X. It then follows that Y itself is also an algebraic system of type $(\tau, \mathcal{I})$, whose fundamental operations are the restrictions to Y of the fundamental operations of X.

Except in degenerate cases, not every subset of X is closed under the fundamental operations. Thus, in an algebraic category, initial structures do not always exist.

b. X is a subalgebra of itself.

c. The intersection of any collection of subalgebras of X is a subalgebra of X. Thus, the subobjects of an algebraic system Y form a Moore collection. (We mentioned this for sublattices in 4.21.)

d. The intersection of all the subalgebras containing some given set $T \subseteq X$ is the smallest subalgebra containing T; it is called the subalgebra **generated** by T. It is the closure of T under the fundamental operations; all of those operations are finitary. Hence the closure operator is an *algebraic closure operator*, as defined in 4.8(A). Thus, the operator $S \mapsto \mathrm{cl}(S)$ is an algebraic closure, if $\mathrm{cl}(S)$ is the submonoid, subgroup, subring, etc., generated by a set $S \subseteq X$, where X is a given monoid, group, ring, etc.

e. The empty set is a subalgebra of X if and only if none of the fundamental operations of X is nullary. Thus, the empty set is a subalgebra when we consider lattices, but that category is atypical. Most of the algebraic systems considered in this book have a nullary operation — i.e., a special element, usually denoted 0 or 1 — and so *the empty set is not a subalgebra in most algebraic systems considered in this book.*

f. Suppose $f : X \to Y$ is a homomorphism of arity τ. Then:

(i) The set $f(X) = \mathrm{Range}(f)$ is a subalgebra of Y.

(ii) More generally, if $S \subseteq X$ is a subalgebra of X, then $f(S)$ is a subalgebra of Y.

(iii) f is uniquely determined by its values on any subset $S \subseteq X$ that generates X.

(iv) If T is a subalgebra of Y, then $f^{-1}(T)$ is a subalgebra of X.

(v) If X satisfies identities $\mathfrak{I}$, then so does the subalgebra $f(X) \subseteq Y$ (regardless of whether Y satisfies those identities).

g. The product of subalgebras is a subalgebra. That is: Let $\prod_{\lambda \in \Lambda} Y_\lambda$ be the direct product of some algebraic systems Y_λ of some variety, and let W_λ be a subalgebra of Y_λ for each λ; then the set $\prod_{\lambda \in \Lambda} W_\lambda$ is a subalgebra of X.

Hint: It is closed under the fundamental operations, since those act separately on each coordinate.

h. Let $f : X \to Y$ be a function from one algebraic system to another of the same arity. Then f is a homomorphism if and only if $\mathrm{Graph}(f)$ is a subalgebra of $X \times Y$.

Varieties with Ideals

9.22. *Remarks.* Among the algebraic categories studied in this book, the category of lattices is atypical. Most equational varieties of interest to us have an *addition* operation $(+)$, which plays a special role among the various fundamental operations. It is the basis for a theory of ideals and quotients developed below.

Our presentation is based on Kurosh [1965]; what we call an object in a "ideal-supporting variety" is what Kurosh calls an "Ω-group." However, we assume that "addition" is commutative. Kurosh and other algebraists do not make that assumption; consequently they have a slightly more general and more complicated theory of ideals and quotients.

9.23. *Definitions.* Let $(\tau, \mathfrak{I})$ be a variety. We shall say that $(\tau, \mathfrak{I})$ is an **ideal-supporting variety** if the following two further conditions are satisfied:

(i) Included among the fundamental operations of the category are operations of addition $(+)$, minus $(-)$, and zero (0) (respectively binary, unary, and nullary), satisfying identities that make $(X, +, -, 0)$ into an additive group.

(There may or may not also be other fundamental operations and other identities.)

(ii) Whenever φ is one of the fundamental operations of the algebraic system and φ is not nullary, then $\varphi(0, 0, 0, \ldots, 0) = 0$. (This is an identity. It may be included as a member of $\mathfrak{I}$. However, in many cases of interest it does not have to be assumed explicitly, because it follows as a consequence from other identities in $\mathfrak{I}$.)

Some examples: the varieties of additive groups, rings, commutative rings, lattice groups, vector lattices, and $\mathbb{F}$-linear spaces (for any field $\mathbb{F}$) are all ideal-supporting. The varieties of monoids and lattices are not ideal-supporting.

In an ideal-supporting variety $(\tau, \mathfrak{I})$, when clarification is needed, we may distinguish between *group homomorphisms* (which preserve $0, +, -$ but not necessarily any other fundamental operations) and τ-homomorphisms (which preserve all the fundamental operations).

9.24. *Elementary observations.* Suppose $(\tau, \mathfrak{I})$ is an ideal-supporting variety, and X is an object in this variety. Then:

a. $\{0\}$ is an object in this category — i.e., an algebraic system of type $(\tau, \mathfrak{I})$. (However, although $\{0\}$ is a subgroup of X, we do not assert that $\{0\}$ is a subalgebra.)

b. The mapping $x \mapsto 0$, from X into $\{0\}$, is a τ-homomorphism.

c. If 0 is the only nullary operation, then the constant mapping $0 : X \to Y$ defined by $x \mapsto 0$ is a τ-homomorphism from any object X into any object Y.

d. If $f : X \to Y$ is a τ-homomorphism, then f is also a group homomorphism. Hence we may define the subgroup $\mathrm{Ker}(f) = f^{-1}(0)$ and the quotient group $X/\mathrm{Ker}(f)$ as in 8.14 and 8.17. (However, we do not assert that these additive groups are objects of the category $(tau, \mathfrak{I})$.)

9.25. *Definition and proposition.* Let $(\tau, \mathfrak{I})$ be an ideal-supporting variety. Let X be an algebraic system of the variety $(\tau, \mathfrak{I})$. Let $S \subseteq X$ be an additive subgroup of X. Then the following conditions are equivalent. If any (hence all) of them is satisfied, we say S is an **ideal** in X.

(A) S is the kernel of some τ-homomorphism $f : X \to Y$, for some object Y in the category.

(B) For each integer $n \geq 0$, for each fundamental operation φ that is n-ary, and for each $x_1, x_2, \ldots, x_n \in X$, the set S is closed under the n-ary operation $\psi_{\varphi, x_1, x_2, \ldots, x_n} : X^n \to X$ defined by

$$\begin{aligned} &\psi_{\varphi, x_1, x_2, \ldots, x_n}(s_1, s_2, \ldots, s_n) \\ &\qquad = \varphi(x_1 + s_1,\, x_2 + s_2,\, \ldots,\, x_n + s_n) \;-\; \varphi(x_1, x_2, \ldots, x_n). \end{aligned}$$

(This condition is trivially satisfied for $n = 0$, since then φ is a constant and $\varphi - \varphi = 0$ is a member of the subgroup S.)

(C) The quotient group X/S can be made into an object of the variety $(\tau, \mathfrak{I})$ (called the **quotient object**), and the quotient map $\pi : X \to X/S$ can be made into a τ-homomorphism, with the fundamental operations $\widehat{\varphi}$ on X/S defined in terms of the given fundamental operations φ on X, as follows:

$$\widehat{\varphi}\Big(\pi(x_1), \pi(x_2), \ldots, \pi(x_n)\Big) \;=\; \pi\Big(\varphi(x_1, x_2, \ldots, x_n)\Big),$$

if φ is n-ary. (The quotient object has different names in different categories — quotient group, quotient ring, quotient vector space, quotient algebra, etc.)

Remarks. In these conditions we do *not* assert that S is necessarily an object in the category.

We might say S is an *ideal in the algebra* X, to distinguish this from the "ideal of sets" introduced in 5.2. The two notions of "ideal" coincide in the context of the Boolean algebra $\mathcal{P}(\Omega)$; see 13.17.d.

The last equation in (C) is admittedly complicated. It may be easier to understand if we mention a typical example: In the category of rings, addition and multiplication in X/S are operations $\boxed{+}$ and $\boxed{\cdot}$ defined by

$$\pi(x_1)\,\boxed{+}\,\pi(x_2) \;=\; \pi(x_1 + x_2), \qquad \pi(x_1)\,\boxed{\cdot}\,\pi(x_2) \;=\; \pi(x_1 \cdot x_2).$$

For both of these equations, some verification is needed: One must show that $\pi(x_1 + x_2)$ and $\pi(x_1 \cdot x_2)$ do not depend on the particular choice of representatives x_1, x_2 from equivalence classes — i.e., one must show that

$$\pi(x_1) = \pi(x_1'), \quad \pi(x_2) = \pi(x_2') \qquad \Rightarrow$$
$$\pi(x_1 + x_2) = \pi(x_1' + x_2'), \quad \pi(x_1 \cdot x_2) = \pi(x_1' \cdot x_2').$$

Verifications of this sort follow from the proof of (B) $\Rightarrow$ (C), below.

Proof of (B) $\Rightarrow$ (C). We must verify that the functions $\widehat{\varphi}$ are well-defined by the formula in (C) — i.e., we must show that

$$\pi(x_i) = \pi(x_i') \text{ for all } i \quad \Rightarrow \quad \pi\Big(\varphi(x_1, \ldots, x_n)\Big) = \pi\Big(\varphi(x_1', \ldots, x_n')\Big).$$

But that is just a restatement of (B). Obviously $\widehat{\varphi}$ is n-ary, and thus X/S is an algebraic system of arity τ. By our definition of the $\widehat{\varphi}$'s, it follows that π is a homomorphism of algebraic systems. By 9.21.f(v), it now follows that X/S satisfies all the identities $\mathfrak{I}$.

Proof of (C) $\Rightarrow$ (A). If the quotient map $\pi : X \to X/S$ is a homomorphism, then S is its kernel.

Proof of (A) $\Rightarrow$ (B). Assume $S = \mathrm{Ker}(f)$. Under the hypotheses of (B) we have $(x_i + s_i) - x_i \in S$, and therefore $f(x_i + s_i) = f(x_i)$. Now

$$\begin{aligned} 0 &= \varphi\Big(f(x_1 + s_1), \ldots, f(x_n + s_n)\Big) - \varphi\Big(f(x_1), \ldots, f(x_n)\Big) \\ &= f\Big(\varphi(x_1 + s_1, \ldots, x_n + s_n)\Big) - f\Big(\varphi(x_1, \ldots, x_n)\Big) \\ &= f\Big(\varphi(x_1 + s_1, \ldots, x_n + s_n) - \varphi(x_1, \ldots, x_n)\Big) \end{aligned}$$

and therefore $\varphi(x_1 + s_1, \ldots, x_n + s_n) - \varphi(x_1, \ldots, x_n)$ is in $\mathrm{Ker}(f) = S$.

9.26. *Further examples.*

a. If X is an object in an ideal-supporting category, then $\{0\}$ and X are ideals in X.

Any ideal in X other than X itself is called a **proper ideal**. A **maximal ideal** is a proper ideal that is not contained in any other proper ideal.

b. Ideals, subalgebras, and subgroups are all the same thing in the category of additive groups.

c. Let X be a ring with unit. Then the same set X, with the same addition and 0 and multiplication, may also be viewed as a ring "without unit" — i.e., an object in the category of rings — by forgetting that its member "1" has some special property. Observe that X has the same ideals in either category. This may not be entirely obvious from definitions 9.25(A) or 9.25(C), but it is easy to see from 9.25(B).

d. In the category of rings or in the category of rings with unit, an ideal in a ring X is an additive subgroup $S \subseteq X$ that satisfies

$$s \in S, \quad x \in X \qquad \Rightarrow \qquad sx, xs \in S.$$

In the category of rings with unit, note that a ring X is the only ideal in X that contains 1; hence it is the only ideal that is also a subring (i.e., subalgebra in the category of rings with unit).

e. The ring $\mathbb{Z}_m$ (introduced in 8.20) can also be described as the quotient of the ring $\mathbb{Z}$ by the ideal $m\mathbb{Z} = \{mz : z \in \mathbb{Z}\}$.

f. Let $\mathbb{R}[x]$ be the ring of all polynomials in one variable x, with coefficients in $\mathbb{R}$, with multiplication given pointwise — that is, $(fg)(x) = f(x) \cdot g(x)$. Let $\mathbb{R}^{\mathbb{R}} = \{$functions from $\mathbb{R}$ into $\mathbb{R}\}$; this is also a ring with unit. Then:

(i) $\mathbb{R}[x]$ is an ideal in itself, but not in $\mathbb{R}^{\mathbb{R}}$, for the category of rings. The inclusion map $i : \mathbb{R}[x] \xrightarrow{\subseteq} \mathbb{R}^{\mathbb{R}}$ is a homomorphism. Thus, *the homomorphic image of an ideal is not necessarily an ideal.* Contrast this with the result for subalgebras noted in 9.21.f(i).

(ii) The set of all polynomials of degree ≤ 1 is not an ideal in $\mathbb{R}[x]$, for the category of rings. However, it is an additive subgroup, and thus it is an ideal when we consider the category of additive groups. Thus, *whether a subset S is an ideal in an algebraic system X may depend on what category we use in considering S and X.*

g. Let $\mathfrak{A} = (\tau, \mathfrak{I})$ and $\mathfrak{B} = (\tau, \mathfrak{J})$ be two ideal-supporting varieties, with the same arity function τ — i.e., with the same fundamental operations, but possibly with different sets of equational axioms. Suppose X is an algebraic system that satisfies both sets of equational axioms — i.e., X is an object in both categories. Then the $\mathfrak{A}$-ideals in X are the same as the $\mathfrak{B}$-ideals in X. (This is may not be obvious from 9.25(A) or 9.25(C), but it is immediately evident from 9.25(B), since that condition only involves the fundamental operations of X, not the other objects of the category.)

Example. A Boolean ring is a ring X with unit, which satisfies $x^2 = x$ for all $x \in X$. Show that we obtain the same ideals in a Boolean ring, whether we view it in the category of rings or the category of rings with unit or the category of Boolean rings. (Boolean rings will be studied further in Chapter 13 and thereafter.)

9.27. *Proposition on ideals in lattice groups.* Let X be a lattice group, and let $S \subseteq X$ be an additive subgroup. Then the following conditions are equivalent.

(A) S is an ideal in X, as defined in 9.25. (Use 9.25(B) here.)

(B) S is solid — that is, $s \in S,\ /x/ \preccurlyeq /s/ \ \Rightarrow\ x \in S$.

(C) Whenever $s, s' \in S$ and $x, x' \in X$, then $\Big[(x+s) \vee (x'+s')\Big] - (x \vee x') \in S$.

(D) Whenever $t \in S$ and $u \in X$, then $(u \vee t) - (u \vee 0) \in S$.

(E) Whenever $s \in S$ and $u \in X$, then $(u+s)^+ - u^+ \in S$.

(F) $s \in S \iff /s/ \in S$, and moreover, whenever $s \in S$ and $x \in X$ satisfy $0 \preccurlyeq x \preccurlyeq s$, then $x \in S$.

Taking $x = x' = 0$ in condition (C), we note this corollary: *In the category of lattice groups, any ideal is also a sublattice.*

Proof of equivalence. We begin by considering what 9.25(B) looks like in the category of lattice groups. Any additive subgroup is closed under the operation ψ determined by the mappings $\varphi(x) = 0$ or $\varphi(x) = -x$ or $\varphi(x, x') = x + x'$. The two remaining fundamental operations in a lattice group are $\vee$ and $\wedge$; taking these binary operations for φ yields the functions

$$\begin{aligned}
\psi_1(s, s') &= \Big[(x+s) \vee (x'+s')\Big] - (x \vee x'), \\
\psi_2(s, s') &= \Big[(x+s) \wedge (x'+s')\Big] - (x \wedge x').
\end{aligned}$$

Thus, an additive subgroup S is an ideal if and only if it is closed under these two binary operations for every choice of $x, x' \in X$. But after some changes of sign, one of these functions is dual to the other, by 8.33.c. This proves (A) $\iff$ (C).

Proofs of (C) $\iff$ (D) $\iff$ (E) follow from translation-invariance of the lattice operations, plus the fact that S is an additive group. Proofs of (F) $\iff$ (B) follow from elementary considerations about the absolute value function. For (D) $\Rightarrow$ (F), use $u = s$ to show $s \in S \iff /s/ \in S$. For (F) $\Rightarrow$ (E), use 8.42.o.

9.28. *Further properties of ideals.* Let X be an object in an ideal-supporting variety. Then:

a. (Isomorphism Theorem.) If $f : X \to Y$ is a homomorphism in that category, then

$$X/\mathrm{Ker}(f) \qquad \text{is isomorphic to} \qquad \mathrm{Ran}(f)$$

by the mapping $F(\pi(x)) = f(x)$ — thus generalizing 8.17.c.

b. The ideals are the sets closed under the finitary operations $\psi_{\varphi,x_1,\ldots,x_n}$ defined in 9.25(B). Hence our earlier results about Moore closures in 4.6 and our earlier results about algebraic closures in 4.8 are applicable. Using those results or by a direct argument, show that:

(i) Any intersection of ideals in X is an ideal.

(ii) For any set $B \subseteq X$, there is a smallest ideal in X that contains B. It is the intersection of the ideals that contain B. It is called the **ideal generated by B**.

c. If S_λ $(\lambda \in \Lambda)$ are ideals in X, then the sum $\sum_{\lambda\in\Lambda} S_\lambda$ (defined in 8.11) is also an ideal; in fact, it is the ideal generated by $\bigcup_{\lambda\in\Lambda} S_\lambda$.

d. The intersection of a subalgebra and an ideal is an ideal.

Proof. Let A be an algebraic system in some ideal-supporting variety; let S be a subalgebra of A, and let I be an ideal in A. We shall show $S \cap I$ is an ideal in S. Let $j : S \xrightarrow{\subseteq} A$ be the inclusion homomorphism; let $h : A \to B$ be a homomorphism with kernel equal to I. Then the composition $S \xrightarrow{j} A \xrightarrow{h} B$ has kernel equal to $S \cap I$.

e. If $f : X \to Y$ is a homomorphism and $T \subseteq Y$ is an ideal, then $f^{-1}(T)$ is an ideal in X.

f. A product of ideals is an ideal. In other words, if E_λ is an ideal in X_λ for each λ, then $\prod_{\lambda\in\Lambda} E_\lambda$ is an ideal in $\prod_{\lambda\in\Lambda} X_\lambda$.

Hint: E_λ is the kernel of some homomorphism $f_\lambda : X_\lambda \to Y_\lambda$. Show that $\prod_{\lambda\in\Lambda} E_\lambda$ is the kernel of the homomorphism $f = \prod_{\lambda\in\Lambda} f_\lambda : \prod_{\lambda\in\Lambda} X_\lambda \to \prod_{\lambda\in\Lambda} Y_\lambda$ defined as in 9.19.

9.29. Let $X = \prod_{\lambda\in\Lambda} Y_\lambda$ be a product of algebraic systems in some ideal-supporting variety $(\tau, \mathcal{I})$; let X be equipped with the product structure. Let $\mathcal{G}$ be a filter of sets on the set Λ. Then $\Gamma = \{g \in X : g^{-1}(0) \in \mathcal{G}\}$ is an ideal in the algebra X.

Proof. It is easy to verify that Γ is an additive subgroup of X. We shall show that Γ also satisfies 9.25(B). Let $\varphi_\lambda : Y_\lambda^n \to Y_\lambda$ $(\lambda \in \Lambda)$ and $\Phi : X^n \to X$ be corresponding n-ary fundamental operations. Let any functions $g_1, g_2, \ldots, g_n \in \Gamma$ and $f_1, f_2, \ldots, f_n \in X$ be given; we are to show that the function

$$h = \Phi(f_1 + g_1, \ \ldots, \ f_n + g_n) - \Phi(f_1, \ldots, f_n)$$

belongs to Γ. Unwind the notation, as in 9.18; then the function $h : \Lambda \to Y$ is defined by

$$h(\lambda) = \varphi_\lambda\Big(f_1(\lambda) + g_1(\lambda), \ \ldots, \ f_n(\lambda) + g_n(\lambda)\Big) - \varphi_\lambda\Big(f_1(\lambda), \ldots, f_n(\lambda)\Big).$$

From this it follows easily that $\bigcap_{j=1}^n g_j^{-1}(0) \subseteq h^{-1}(0)$. Now, each set $g_j^{-1}(0)$ belongs to the filter $\mathcal{G}$, hence $h^{-1}(0)$ belongs to $\mathcal{G}$. Thus $h \in \Gamma$.

9.30. *Corollary.* Let $X = \prod_{\lambda\in\Lambda} Y_\lambda$ be a product of algebraic systems in some ideal-supporting variety $(\tau, \mathfrak{I})$, equipped with the product structure. Then the set

$$\coprod_{\lambda\in\Lambda} Y_\lambda \quad = \quad \left\{ f \in \prod_{\lambda\in\Lambda} Y_\lambda \quad : \quad f(\lambda) \neq 0 \text{ for only finitely many } \lambda\text{'s} \right\}$$

is an ideal in the algebra X; this is immediate from 9.29 using the cofinite filter. We shall call this ideal the **external direct sum** of the Y_λ's. (In some categories it coincides with the *coproduct.*) Of course, when Λ is finite, then the external direct sum is the same as the product.

Further properties. For each $\lambda \in \Lambda$, an injective homomorphism $j_\lambda : Y_\lambda \to X$ can be defined by $j_\lambda(v) = (0, 0, \ldots, 0, 0, v, 0, 0, \ldots, 0, 0)$ — that is, put v in the λth component and zeros elsewhere. Then $j_\lambda(Y_\lambda)$ is a subgroup of X that is isomorphic to Y_λ. Also note that $\pi_\lambda \circ j_\lambda$ is the identity map of Y_λ and $\pi_\mu \circ j_\lambda : Y_\lambda \to Y_\mu$ is the zero map if $\mu \neq \lambda$. Show that

$$\coprod_{\lambda\in\Lambda} Y_\lambda \quad = \quad \bigoplus_{\lambda\in\Lambda} j_\lambda(Y_\lambda)$$

(where $\oplus$ represents an internal direct sum, as defined in 8.12). Thus, the external direct sum of the Y_λ's is the internal direct sum of a collection of groups that are isomorphic to the Y_λ's. If we gloss over the distinction between isomorphism and equality, then the external direct sum of the Y_λ's is "the same as" the internal direct sum of the Y_λ's.

Caution: In the wider literature, internal direct sums and external direct sums are often used interchangeably; both are referred to simply as direct sums.

FUNCTORS

9.31. Loosely speaking, a *functor* is a morphism in the category of categories. A little more precisely, a functor is a mapping from one category into another, sending objects to objects and morphisms to morphisms and preserving the "relevant structure." In this context the relevant structure involves such things as the compositions of morphisms.

To be entirely precise, a **covariant functor** preserves compositions and arrow directions; a **contravariant functor** reverses compositions and arrow directions. Thus, suppose that

$$p : X \to Y \qquad \text{and} \qquad u = v \circ w$$

are a typical morphism and a typical composition of morphisms in category $\mathfrak{A}$. Then a covariant functor $\mathbf{F} : \mathfrak{A} \to \mathfrak{B}$ yields

$$\mathbf{F}(p) : \mathbf{F}(X) \to \mathbf{F}(Y) \qquad \text{and} \qquad \mathbf{F}(u) = \mathbf{F}(v) \circ \mathbf{F}(w)$$

in category $\mathfrak{B}$, whereas a contravariant functor $\mathbf{G} : \mathfrak{A} \to \mathfrak{B}$ yields

$$\mathbf{G}(p) : \mathbf{G}(Y) \to \mathbf{G}(X) \qquad \text{and} \qquad \mathbf{G}(u) = \mathbf{G}(w) \circ \mathbf{G}(v)$$

in category $\mathfrak{B}$.

The **reduced power functor** $S \mapsto {}^*S$ will be discussed starting in 9.37; note particularly 9.50.a. This functor is covariant; it is usually represented with an asterisk on the left. Do not confuse it with the contravariant **exponential functor** $S \mapsto S^*$, described in 9.55 below; this functor is usually represented with an asterisk on the right.

9.32. *Some elementary examples of functors.* The **covariant power set functor** is a functor from the category of sets to itself. This functor sends each set X to the set $\mathcal{P}(X)$ and sends each mapping $f : X \to Y$ to the forward image map $f : \mathcal{P}(X) \to \mathcal{P}(Y)$ defined in 2.7.

The **contravariant power set functor** is another functor from the category of sets to itself. This functor also sends each set X to the set $\mathcal{P}(X)$, but it sends each mapping $f : X \to Y$ to the inverse image map $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ defined in 2.8.

9.33. (*Optional.*) We now specialize slightly the notion developed in 9.13.b. Let (X, x_0) be a pointed topological space (defined as in 9.9). Consider all paths in X that begin and end at x_0 — that is, all continuous functions

$$f : [0,1] \to X \qquad \text{that satisfy} \qquad f(0) = f(1) = x_0.$$

Call two such paths f, g equivalent if there exists a homotopy from f to g that preserves the endpoints — i.e., if there exists a continuous function $h : [0,1] \times [0,1] \to X$ that satisfies

$$h(0,t) = f(t), \qquad h(1,t) = g(t), \qquad h(s,0) = h(s,1) = x_0$$

for all $s, t \in [0,1]$. It is easy to verify that the equivalence classes form a group, under the operation of "composition" —

- to compose two paths, follow one and then the other;
- the inverse any path is the same path run backward.

This group, denoted $\pi_1(X, x_0)$, is called the **Poincaré fundamental group** of the pointed space (X, x_0).

If $\varphi : (X, x_0) \to (Y, y_0)$ is a morphism of pointed topological spaces (defined as in 9.9), then we can define a mapping between the fundamental groups,

$$\pi_1(\varphi) \quad : \quad \pi_1(X, x_0) \quad \to \quad \pi_1(Y, y_0),$$

as follows: If $f : [0,1] \to X$ is a member of some equivalence class that is, in turn, a member of $\pi_1(X, x_0)$, then $\varphi \circ f : [0,1] \to Y$ is a member of some corresponding equivalence class that is a member of $\pi_1(Y, y_0)$. It is not hard to verify that this mapping is well defined — i.e., that it does indeed preserve equivalence — and furthermore, this mapping is a group homomorphism.

Thus π_1 is a *covariant functor* from the category of pointed topological spaces to the category of groups. Using this functor, we can transform some questions about topological spaces into corresponding questions about groups. That is one of the basic ideas of **algebraic topology**. It will not be pursued further in this book, however.

9.34. Many covariant functors can be described as **forgetful functors**. A forgetful functor is in use when we go from one category to another by *forgetting* part of the relevant structure. For instance, a lattice is a special type of preordered set, and a lattice homomorphism is a special type of increasing map. Any theorem about increasing maps between preordered sets can also be applied to the special case of lattice homomorphisms between lattices.

In forgetting some structure, we permit some change in the description of the objects. For instance, we noted in 9.10 that any preordered set $(X, \preccurlyeq)$ may be viewed as a topological space $(X, \mathcal{S})$, but $\preccurlyeq$ is not equal to $\mathcal{S}$.

If $\mathcal{A}$ is a subcategory of $\mathcal{B}$, then the inclusion $\mathcal{A} \xrightarrow{\subseteq} \mathcal{B}$ is a forgetful functor. Not every forgetful functor is of this form, however; see the two examples below.

9.35. *Preview.* We now describe two especially important forgetful functors that will be important in later chapters.

Every uniform structure determines a topology (5.33), and any uniformly continuous map is also continuous (18.9.c). Thus there is a forgetful functor from uniform spaces (with uniformly continuous maps) to topological spaces (with continuous maps). This forgetful functor is not given by the inclusion of a subcategory, since different uniformities on a set may determine the same topology — for instance, see 18.9.d and 19.11.e. However, this forgetful functor has the interesting property that it *preserves the formation of initial objects.* That is, if $\mathcal{U}$ is the initial uniformity determined on a set X by a collection of mappings $\varphi_\lambda : X \to (Y_\lambda, \mathcal{U}_\lambda)$, then the resulting uniform topology $\mathcal{T}(\mathcal{U})$ is equal to the initial topology determined by the maps $\varphi_\lambda : X \to (Y_\lambda, \mathcal{T}(\mathcal{U}_\lambda))$. (The proof of that equality may be easier to prove in 18.9.g, after we have developed a few more tools.)

Every topology determines a Borel σ-algebra (see 5.26.e), and it is easy to verify that any continuous map is measurable when its domain and codomain are equipped with the Borel σ-algebras (see 21.2.a). Thus we obtain a forgetful functor from topological spaces (with continuous maps) to measurable spaces (with measurable maps). This forgetful functor is not given by a subcategory inclusion, since different topologies may yield the same σ-algebra — for instance, the discrete topology on $\mathbb{N}$ and the two lower set topologies given in 5.15.d all yield the discrete σ-algebra. The forgetful functor from topological spaces to measurable spaces sometimes does *not* preserve the formation of initial objects; an example of that fact is given in 21.8(iii).

9.36. *Other functors.* The functors that take any poset to its sup completion, any Tychonov space to its Stone-Čech compactification, or any separated uniform space to its separated uniform completion, are examples of inclusions of **reflective subcategories**. That topic will not be discussed here; it can be found in Herrlich and Strecker [1979].

The Reduced Power Functor

9.37. *Preview.* In the next few pages we shall develop a "junior version" of nonstandard analysis. This simplified approach is less powerful than the customary treatment, but it avoids the conceptual difficulties of sets of sets of sets and avoids the formal study of mathematical languages — a study that is second nature to logicians but may seem quite foreign to many analysts. Our junior version, which may seem more natural to analysts, is adequate for a few minor applications including a "construction" of the hyperreal number system $^*\mathbb{R}$ in 10.19 and an explanation of limits in terms of infinitesimals in 10.37; this will give the reader a quick taste of what nonstandard analysis is like. In Chapter 14 we shall sketch some of the remaining ingredients of the customary approaches to nonstandard analysis, but that sketch will rely on some results and intuition developed in the next few pages.

9.38. *Preview of the Transfer Principle.* In the next few pages we shall show how, given any set S, function f, or relation R, we can construct a corresponding set, function, or relation $^*S, {}^*f, {}^*R$ in a "larger universe." The **Transfer Principle** states that any *suitably worded* statement without stars is true if and only if the corresponding statement with stars is true. For instance,

$$S = T \qquad \text{if and only if} \qquad {}^*S = {}^*T,$$

and therefore the mapping $S \mapsto {}^*S$ is injective. Likewise, we shall show in 9.45.h that

$$T = S_1 \cap S_2 \cap \cdots \cap S_n \qquad \text{if and only if} \qquad {}^*T = {}^*S_1 \cap {}^*S_2 \cap \cdots \cap {}^*S_n$$

for any positive integer n. However, the Transfer Principle only applies to "suitably worded" statements, not to all statements. For instance, our observation about finite intersections does not extend to infinite intersections — see 9.46.a. To make precise this notion of "suitably worded statements," we will need to analyze our language; that logical analysis will be carried out in part in Chapter 14.

In the next few pages we develop some basic properties of the star mapping by purely ad hoc methods, without use of the Transfer Principle. These ad hoc methods are sometimes a bit tedious; the Transfer Principle would be a helpful shortcut. Some readers may prefer to glance through a text on nonstandard analysis, master the Transfer Principle, and *then* proceed through the next few pages.

9.39. Let Λ be a nonempty set. For the discussions below we may refer to Λ as the **index set**, or *domain*. We shall consider many functions from Λ to various sets. For simplicity, we shall disregard the codomains of these functions. Any two functions that are defined on Λ and agree at every point of Λ will be viewed as the "same" function, as in 2.5.c. The particular choice of the codomain does not matter, provided that it is *sufficiently large for our applications*; any larger set will do just as well. Thus, we will be concerned with sets such as $S^\Lambda, T^\Lambda, U^\Lambda$ (the functions from Λ into S, T, or U) for various choices of sets S, T, U, but we will not be concerned with a larger set containing all of S, T, U.

9.40. Let Λ be the index set, as indicated above. Let $\mathcal{F}$ be a proper filter on Λ, and let $\mathcal{I}$ be the proper ideal that is dual to $\mathcal{F}$ — that is, $\mathcal{I} = \{\Lambda \setminus F : F \in \mathcal{F}\}$. Our choices of $\Lambda, \mathcal{F}, \mathcal{I}$ will be held fixed throughout the discussion. For our junior version of nonstandard analysis, $\mathcal{F}$ will usually be a free ultrafilter, but other choices of $\mathcal{F}$ are also of some interest; see for instance 21.17. (The existence of free ultrafilters was discussed in 6.33.)

Two functions g, h defined on Λ will be said to be **$\mathcal{F}$-equivalent**, or to **agree $\mathcal{F}$-almost everywhere**, if the statement $g = h$ is satisfied $\mathcal{F}$-almost everywhere in the sense of 5.3 — i.e., if the set

$$\{\lambda \in \Lambda \ : \ g(\lambda) = h(\lambda)\}$$

is "*large*" in the sense that it is an element of $\mathcal{F}$ — or, equivalently, if the set

$$\{\lambda \in \Lambda \ : \ g(\lambda) \neq h(\lambda)\}$$

is "*small*" in the sense that it is a member of $\mathcal{I}$. It is easy to verify that this is an equivalence relation on the set $\Omega^\Lambda = \{\text{functions from } \Lambda \text{ into } \Omega\}$, for any codomain Ω. If we do not specify a codomain Ω, then $\mathcal{F}$-equivalence is an equivalence relation on the *proper class of all functions* that are defined on Λ.

For the present discussion, let $\pi(g)$ denote the equivalence class containing a function g.

9.41. Let $\Lambda, \mathcal{F}, \mathcal{I}$ be as above, and let S be any set. Then the set of equivalence classes

$$^*S \quad = \quad \{\pi(g) \ : \ g \in S^\Lambda\} \quad = \quad \{\pi(f) \ : \ f(\lambda) \in S \text{ for almost all } \lambda\}$$

is called the **reduced power** of S. (Here "for almost all λ" means for all λ in some member of $\mathcal{F}$, as in 9.40.) In other words, $\alpha \in {}^*S$ if and only if

> α is an equivalence class, at least one member of which is a function whose range is a subset of S.

When the choices of Λ and $\mathcal{F}$ need to be mentioned explicitly, then the reduced power $^*S$ can be written instead as $S^\Lambda/\mathcal{F}$ or as $S^\Lambda/\mathcal{I}$. Usually that notation is not needed, however, for most interesting results are obtained when we hold Λ and $\mathcal{F}$ fixed and consider what happens as S is varied.

When the filter $\mathcal{F}$ is a free ultrafilter, then the reduced power $^*S$ is called the **ultrapower** of S. *Remark.* This notion of "ultrapower" should not be confused with the *Banach space ultrapower*, a related but slightly different object that is often used when techniques of nonstandard analysis are applied in the study of Banach spaces. A brief introduction to Banach space ultrapowers can be found in Coleman [1987].

9.42. For any point $s \in S$, let c_s be the constant function taking the value s — i.e. the function defined by $c_s(\lambda) = s$ for all $\lambda \in \Lambda$. Then it is clear that $\pi(c_s) \in {}^*S$. Moreover, it is easy to see that the mapping $s \mapsto \pi(c_s)$ is injective — i.e., if $s \neq t$, then the equivalence classes $\pi(c_s)$ and $\pi(c_t)$ are distinct. We may identify each point s with the resulting equivalence class $\pi(c_s)$; thus we may consider S as a subset of $^*S$.

We shall see, in exercises below, that $^*S$ inherits many of the properties of S, and thus it is a sort of "enlarged copy" of S. The reduced power construction is a simplified version of the **nonstandard enlargement** construction used in nonstandard analysis.

9.43. *Remarks.* The reduced power construction is used in substantially different ways, with different intuition and syntactic conventions, in at least two parts of analysis:

(i) Nonstandard analysis will be introduced briefly in 14.63. In that context, $\mathcal{F}$ is usually a free ultrafilter, and elements of $^*S$ are discussed much as though they were elements of S — i.e., *points* in some set slightly larger than S. For instance, elements of $^*\mathbb{R}$ are treated as some sort of generalized "numbers."

(ii) In the theory of measure and integration, reduced powers also arise naturally. Let $\mathcal{J}$ be the collection of null sets for some complete positive measure μ on a set Λ; this is discussed in 21.17. Then $\mathcal{J}$ is a σ-ideal, but it is generally not a maximal ideal, since not every subset of Λ is necessarily a null set or the complement of a null set. Thus, the dual filter $\mathcal{F}$ is generally not an ultrafilter. In this context, members of $^*\mathbb{R}$ are sometimes called **real random variables**; more generally, members of $^*X$ may be called ***X*-valued random variables**. In this context, elements of $^*X$ are discussed much as though they were elements of X^Λ — i.e., *functions* defined on Λ. For instance, elements of the Lebesgue spaces $L^p(\Lambda, \mathcal{S}, \mu)$ are equivalence classes of functions, but they are often discussed as if they were functions. This is, admittedly, an abuse of notation — the elements of $L^p(\Lambda, \mathcal{S}, \mu)$ are not *really* functions. Occasionally the distinction between functions and their equivalence classes becomes important; then the distinction is pointed out. But the blurring of that distinction, quite common in the literature, is convenient and usually harmless because the quotient map $\pi : X^\Lambda \to {}^*X$ preserves most (not quite all) of the structures and operations that are of interest; see particularly 9.53.

9.44. *Exercise: When are S and $^*S$ different?* We have seen that $S \subseteq {}^*S$; when do we have $S \neq {}^*S$ also?

a. Suppose $\mathcal{F}$ is the fixed ultrafilter at some point $\lambda_0 \in \Lambda$. Then $^*S = S$ for all sets S. Then the operation $S \mapsto {}^*S$ brings us nothing new; this case is of little interest to us.

b. Suppose $\mathcal{F}$ is a proper filter, but *not* an ultrafilter. Show that $^*S = S$ if S is the empty set or a singleton, but $^*S \neq S$ if S contains two or more points.

Hint: If $A, \complement A$ are nonempty proper subsets of Λ that do not belong to $\mathcal{F}$, and x, y are distinct members of S, show that

$$f(u) \quad = \quad \begin{cases} x & \text{if } u \in A \\ y & \text{if } u \notin A \end{cases}$$

defines a function $f : \Lambda \to S$ that is not equivalent to a constant function.

c. If $\mathcal{F}$ is a free ultrafilter and S is a finite set, then $^*S = S$. *Hint*: 5.8(E).

d. If $\mathcal{F}$ is a free ultrafilter and $\mathrm{card}(S) \geq \mathrm{card}(\Lambda)$, then $^*S \neq S$. *Hint*: There exists an injective mapping $i : \Lambda \to S$; then i is not equivalent to a constant mapping.

e. *Corollary.* If $\Lambda = \mathbb{N}$ and $\mathcal{F}$ is a free ultrafilter on $\mathbb{N}$, then $^*S \neq S$ for every infinite set S. *Hint*: Here we use the fact (established in 6.27) that any infinite set S satisfies $\mathrm{card}(S) \geq \mathrm{card}(\mathbb{N})$.

9.45. *Further properties of reduced powers of sets.* The list of properties below, and much of the other material in this subchapter, is based on Robinson and Zakon [1969].

Assume $\mathcal{F}$ is a proper filter on Λ (not necessarily an ultrafilter). Let S, T, and S_1, S_2, S_3, ... and S_α $(\alpha \in A)$ be sets. Then:

a. $^*\varnothing = \varnothing$.

b. $^*S \subseteq {^*T} \iff S \subseteq T$.

c. The * mapping is injective: If $S \neq T$, then $^*S \neq {^*T}$.

d. If $S \subseteq T$, then $S = T \cap {^*S}$.

e. $^*(S \setminus T) \subseteq (^*S) \setminus (^*T)$.

f. If $\mathcal{F}$ is an ultrafilter, then $^*(S \setminus T) = (^*S) \setminus (^*T)$.

g. Intersections and unions satisfy these inclusions:

$$^*\left(\bigcap_{\alpha \in A} S_\alpha\right) \subseteq \bigcap_{\alpha \in A} (^*S_\alpha) \qquad \text{and} \qquad ^*\left(\bigcup_{\alpha \in A} S_\alpha\right) \supseteq \bigcup_{\alpha \in A} (^*S_\alpha).$$

h. For any positive finite integer n,

$$^*(S_1 \cap S_2 \cap \cdots \cap S_n) \quad = \quad {^*S_1} \cap {^*S_2} \cap \cdots \cap {^*S_n}.$$

i. If $\mathcal{F}$ is an ultrafilter, then also

$$^*(S_1 \cup S_2 \cup \cdots \cup S_n) \quad = \quad {^*S_1} \cup {^*S_2} \cup \cdots \cup {^*S_n}.$$

9.46. *Examples.* For both of the examples below, let $\Lambda = \mathbb{N}$; define $f : \mathbb{N} \to \mathbb{N}$ by taking $f(n) = n$. A filter $\mathcal{F}$ will be specified below; let $\pi(f)$ be the equivalence class containing the function f.

a. Without additional assumptions, the inclusions in 9.45.g cannot be strengthened to equalities; we now show this by examples. Let $\mathcal{F}$ be any filter on $\mathbb{N}$ that includes the cofinite filter. Then

$$\pi(f) \in {^*S_1} \cap {^*S_2} \cap {^*S_3} \cap \cdots \quad \text{but} \quad S_1 \cap S_2 \cap S_3 \cdots = \varnothing, \quad \text{if } S_j = \mathbb{N} \setminus \{j\} \text{ for all } j$$

(compare with 9.54). Also,

$$\pi(f) \in {^*(S_1 \cup S_2 \cup S_3 \cup \cdots)}, \quad \pi(f) \notin {^*S_1} \cup {^*S_2} \cup {^*S_3} \cup \cdots \quad \text{if } S_j = \{j\} \text{ for all } j.$$

b. The conclusions of 9.45.f and 9.45.i may not be valid if we do not assume $\mathcal{F}$ is an ultrafilter; we now show this with examples. Let $\mathcal{F}$ be the cofinite filter. Then

$$\pi(f) \in {^*(\mathbb{N})} \setminus {^*(S_1)}, \qquad \pi(f) \notin {^*(\mathbb{N} \setminus S_1)} \qquad \text{if } S_1 = \{1, 3, 5, 7, \ldots\};$$

$$\pi(f) \in {^*(S_1 \cup S_2)}, \qquad \pi(f) \notin {^*S_1} \cup {^*S_2} \qquad \text{if } S_2 = \{2, 4, 6, 8, \ldots\}.$$

9.47. *Reduced power of a finite products of sets.* Let $\mathcal{F}$ be a filter on Λ, and let $f_1, f_2, \ldots, f_n$ be finitely many functions defined on Λ. Then an n-tuple of functions $(f_1, f_2, \ldots, f_n)$ may also be viewed as an n-tuple-valued function. We shall use the two viewpoints interchangeably.

Observe that two n-tuple-valued functions $f = (f_1, f_2, \ldots, f_n)$ and $g = (g_1, g_2, \ldots, g_n)$ are *equivalent* in the sense of 9.40 if and only if the set $\{\lambda \in \Lambda : f(\lambda) = g(\lambda)\}$ belongs to

$\mathcal{F}$ — that is, if and only if the set $\bigcap_{j=1}^{n}\{\lambda \in \Lambda : f_j(\lambda) = g_j(\lambda)\}$ belongs to $\mathcal{F}$. Since $\mathcal{F}$ is a filter, this condition holds if and only if each of the n sets $\{\lambda \in \Lambda : f_j(\lambda) = g_j(\lambda)\}$ belongs to $\mathcal{F}$. In other words, $(f_1, f_2, \ldots, f_n)$ is equivalent to $(g_1, g_2, \ldots, g_n)$ if and only if f_1 is equivalent to g_1, f_2 is equivalent to g_2, ..., and f_n is equivalent to g_n. Therefore, an equivalence class of n-tuples can be represented as an n-tuple of equivalence classes. It is easy to verify that

$$^*(S_1 \times S_2 \times \cdots \times S_n) \quad = \quad {}^*S_1 \times {}^*S_2 \times \cdots \times {}^*S_n$$

for any n sets $S_1, S_2, \ldots, S_n$.

9.48. What about an *infinite* product of sets? Not all of the reasoning in the preceding section generalizes readily. Let's see what goes wrong.

Let $\mathcal{F}$ be a filter on Λ, and let $f_1, f_2, f_3, \ldots$ be infinitely many functions defined on Λ. Then a sequence of functions $(f_1, f_2, f_3, \ldots)$ may also be viewed as a sequence-valued function. We may use the two viewpoints interchangeably.

Observe that two sequence-valued functions $f = (f_1, f_2, f_3, \ldots)$ and $g = (g_1, g_2, g_3, \ldots)$ are equivalent in the sense of 9.40 if and only if the set $\{\lambda \in \Lambda : f(\lambda) = g(\lambda)\}$ belongs to $\mathcal{F}$ — that is, if and only if the set $\bigcap_{j=1}^{\infty}\{\lambda \in \Lambda : f_j(\lambda) = g_j(\lambda)\}$ belongs to $\mathcal{F}$. Since $\mathcal{F}$ is a filter, this condition implies, but is not necessarily implied by, the condition that each of the sets $\{\lambda \in \Lambda : f_j(\lambda) = g_j(\lambda)\}$ belongs to $\mathcal{F}$. In other words, if $(f_1, f_2, f_3, \ldots)$ is equivalent to $(g_1, g_2, g_3, \ldots)$, then each f_j is equivalent to g_j, but not necessarily conversely. An equivalence class of sequences is not the same thing as a sequence of equivalence classes.

For instance, let $\Lambda = \mathbb{N}$, and let $\mathcal{F}$ be the cofinite filter on $\mathbb{N}$. Let f_j be the constant function 0, and let $g_j : \mathbb{N} \to \mathbb{N}$ be defined by $g_j(k) = \delta_{jk}$, where δ is the Kronecker delta (defined in 2.2.d). Then for each j, we see that f_j is equivalent to g_j since they agree everywhere on $\mathbb{N}$ except at one point. But $f = (f_1, f_2, f_3, \ldots)$ is not equivalent to $g = (g_1, g_2, g_3, \ldots)$ since they agree nowhere on $\mathbb{N}$ — indeed, $f(n)$ and $g(n)$ differ in their nth coordinate.

9.49. *Reduced powers of functions.* How do we extend functions? For instance, we would like to extend the function $\sin : \mathbb{R} \to \mathbb{R}$ to a function $^*\sin : {}^*\mathbb{R} \to {}^*\mathbb{R}$; how is this accomplished?

Let $p : X \to Y$ be a function from one set to another. There are a couple of natural methods for defining a reduced power $^*p : {}^*X \to {}^*Y$; fortunately they yield the same result.

(A) One method is to identify a function with its graph. Then a function is a set of ordered pairs, no two of which have the same first element. Show that if $\mathrm{Gr}(p) \subseteq X \times Y$ is the graph of a function $p : X \to Y$, then $^*(\mathrm{Gr}(p)) \subseteq {}^*X \times {}^*Y$ is the graph of a function from $^*X$ into $^*Y$, which we shall denote by $^*p$. Thus $^*(\mathrm{Gr}(p)) = \mathrm{Gr}(^*p)$. Note that, since $\mathrm{Gr}(p) \subseteq \mathrm{Gr}(^*p)$, the function $^*p$ is an extension of the function p — that is, we have $X \subseteq {}^*X$, and $^*p(x) = p(x)$ for every $x \in X$.

(B) Another method is by this rule: If $p : X \to Y$ is some function, we wish to define a function $(^*p) : {}^*X \to {}^*Y$ by specifying its value on each $\xi \in {}^*X$. Any $\xi \in {}^*X$ may be written in the form $\pi(f)$ for some function $f : \Lambda \to \Omega$ (where

Ω is some sufficiently large codomain), and $\pi : \Omega^\Lambda \to {}^*\Omega$ is the quotient map taking functions to their equivalence classes, for some sufficiently large codomain Ω. Show that the mapping $f \mapsto p \circ f$ respects the equivalence relation on X^Λ — that is, $\pi(f_1) = \pi(f_2) \Rightarrow \pi(p \circ f_1) = \pi(p \circ f_2)$ (see 3.12). Hence a function $({}^*p) : {}^*X \to {}^*Y$ is well defined by the rule

$$({}^*p)(\pi(f)) \quad = \quad \pi(p \circ f) \qquad \text{for} \quad f \in X^\Lambda.$$

Show that these two definitions yield the same function ${}^*p$.

When p is some familiar function, then it is customary to write ${}^*p$ without the star. For instance, the extension of sin, which would naturally be written ${}^*\sin$, is customarily written sin instead.

9.50. *Further properties of reduced powers of functions.*

a. The taking of reduced powers preserves identity maps — i.e., if i_X is the identity map of X, then ${}^*(i_X)$ is equal to the identity map of the set ${}^*X$. Also, the taking of reduced powers preserves composition of functions; that is, ${}^*(p \circ q) = ({}^*p) \circ ({}^*q)$ for any functions $q : W \to X$ and $p : X \to Y$. From these two facts it follows that the taking of reduced powers is a *covariant functor* from the category of sets to the category of sets; that term was introduced in 9.31.

b. The reduced power ${}^*p : {}^*X \to {}^*Y$ is injective or surjective if and only if the mapping $p : X \to Y$ has that property.

c. Let $f : X \to Y$ and let $S \subseteq X$ and $T \subseteq Y$. If $f(S) \subseteq T$, then $({}^*f)({}^*S) \subseteq {}^*T$.

Hints: The hypothesis can be restated as: $\mathrm{Gr}(f) \cap (S \times Y) \subseteq X \times T$. Using several results of the last few pages, we can show that $\mathrm{Gr}({}^*f) \cap ({}^*S \times {}^*Y) \subseteq {}^*X \times {}^*T$.

9.51. *Reduced powers of relations.* How do we extend relations? For instance, we know $3 < 5$; we would like a corresponding notion for members of ${}^*\mathbb{R}$.

Let R be a binary relation on a set X. There are a couple of natural methods for defining a binary relation ${}^*R$ on the set ${}^*X$; fortunately they yield the same result.

(A) One method is to identify the relation with its graph — i.e., to work with the set $\mathrm{Gr}(R) \subseteq X \times X$. Then we can take its reduced power, ${}^*(\mathrm{Gr}(R)) \subseteq {}^*(X \times X) = ({}^*X) \times ({}^*X)$. It is the graph of a binary relation on ${}^*X$, which we naturally call ${}^*R$. Thus ${}^*(\mathrm{Gr}(R)) = \mathrm{Gr}({}^*R)$.

(B) For functions $f, g : \Lambda \to X$, say that $\pi(f)\ {}^*R\ \pi(g)$ if and only if the statement $f\ R\ g$ is $\mathcal{F}$-true in the sense of 5.3 — i.e., if and only if the set $\{\lambda \in \Lambda : f(\lambda)\ R\ g(\lambda)\}$ is a member of $\mathcal{F}$. Show that this makes ${}^*R$ well defined on ${}^*X$ — i.e., show that if $\pi(f_1) = \pi(f_2)$ and $\pi(g_1) = \pi(g_2)$, then

$$\{\lambda \in \Lambda \ : \ f_1(\lambda)\ R\ g_1(\lambda)\} \in \mathcal{F} \iff \{\lambda \in \Lambda \ : \ f_2(\lambda)\ R\ g_2(\lambda)\} \in \mathcal{F}.$$

Show that these two definitions yield the same binary relation ${}^*R$ on the set ${}^*X$.

9.52. *Further properties of reduced powers of relations.*

a. The restriction of $^*R$ to the set $X \subseteq {}^*X$ is precisely R. That is: $x\ R\ y$ if and only if $x, y \in X$ and $x\ {}^*R\ y$. *Hint*: Use characterization (A), above, together with 9.45.d.

b. If R has any of the following properties, then so does $^*R$: reflexive, irreflexive, transitive, symmetric, antisymmetric, preorder, partial order, equivalence relation, lattice.

c. If $p : (X, \preccurlyeq) \to (Y, \sqsubseteq)$ is an order-preserving map from one preordered set into another, then so is $^*p\ :\ ({}^*X,\ {}^*\preccurlyeq) \to ({}^*Y,\ {}^*\sqsubseteq)$.

d. Suppose $a, b \in X$, and $S = \{x \in X : aRx \text{ and } xRb\}$. Then

$$^*S \quad = \quad \{\xi \in {}^*X \ \ : \ \ a\ {}^*R\ \xi \text{ and } \xi\ {}^*R\ b\}.$$

In particular, the enlargement of an interval (a, b) in $\mathbb{R}$ is the corresponding interval (a, b) in $^*\mathbb{R}$.

e. If $\mathcal{F}$ is an ultrafilter on Λ and $(S, \leq)$ is a chain, then $({}^*S,\ {}^*\leq)$ is a chain.

Hint: If $g, h \in S^\Lambda$, then the three sets

$$\{\lambda\ : g(\lambda) < h(\lambda)\}, \qquad \{\lambda\ : g(\lambda) = h(\lambda)\}, \qquad \{\lambda\ : g(\lambda) > h(\lambda)\}$$

form a partition of Λ, so exactly one of them is a member of $\mathcal{F}$.

f. *Example.* In the preceding result, we cannot omit the assumption that $\mathcal{F}$ be an ultrafilter. In fact, if S is a chain containing at least two elements and $\mathcal{F}$ is a proper filter on Λ but not an ultrafilter, then $^*X$ *cannot be* a chain.

Proof. There exist sets A and B that partition Λ, such that neither A nor B is a member of $\mathcal{F}$. By relabeling, we may assume that two of the elements of S are called 0 and 1, and that $0 < 1$. Then the equivalence classes of the characteristic functions of the sets A and B are elements $\alpha, \beta \in {}^*S$ such that none of the conditions $\alpha \,{}^*\!< \beta$, $\alpha = \beta$, or $\alpha \,{}^*\!> \beta$ holds.

g. *Remark.* The reduced power of a complete or Dedekind complete ordering may inherit that completeness property (as in 21.42), or it may not (as in 10.19).

9.53. *Reduced powers of algebraic systems.* Let X be an algebraic system of an ideal-supporting type $(\tau, \mathfrak{I})$, let Λ be a set, and let $\mathcal{F}$ be a filter of subsets of Λ. Then by 9.29,

$$N \quad = \quad \{g \in X^\Lambda \ \ : \ \ g^{-1}(0) \in \mathcal{F}\}$$

is an ideal in the product algebra $X^\Lambda = \{\text{functions from } \Lambda \text{ into } X\}$. Hence we can form the quotient X^Λ/N as in 9.25(C); it is another algebraic system of variety $(\tau, \mathfrak{I})$. It is easy to verify (*exercise*) that this quotient X^Λ/N is the same thing as the reduced power $^*X = X^\Lambda/\mathcal{F}$ defined in 9.41, and the fundamental operations of the algebraic system X^Λ/N (defined as in 9.25(C)) are the same as the reduced powers $^*\varphi$ of the fundamental operations φ of X (defined as in 9.49).

We may embed X in $^*X$, by the method described in 9.41. (That is, any $x \in X$ is mapped to the equivalence class of the constant function from Λ into X whose constant value is x.) The embedding is an injective homomorphism, and so X is a subalgebra of $^*X$. For instance, if X is a ring, then $^*X$ is a ring, X is a subring of $^*X$, and the inclusion map $X \xrightarrow{\subseteq} {}^*X$ is a ring homomorphism.

Elements of $^*X$ are equivalence classes of functions from Λ into X. However, as we remarked in 9.43, elements of $^*X$ are sometimes discussed as if they were elements of X^Λ or elements of X. These styles of discussion are feasible largely because each of these maps is a homomorphism:

- the quotient map $\pi : X^\Lambda \to X^\Lambda/N$,
- the coordinate projections $\pi_\lambda : X^\Lambda \to X$, and
- the inclusion $X \xrightarrow{\subseteq} {}^*X$

Thus, each of these maps preserves the fundamental operations, and therefore preserves a great deal of the relevant structure.

9.54. The Ultrafilter Principle, introduced in 6.32, is equivalent to the following principle, which is similar to a principle of nonstandard analysis:

> **(UF4) Enlargement (Concurrence, Idealization) Principle.** Let Ω be a set. Then it is possible to choose an index set Λ and a free ultrafilter $\mathcal{F}$ on Λ, such that the resulting ultrapowers $^*S = S^\Lambda/\mathcal{F}$ have this property: Whenever $\mathcal{E}$ is a proper filter on a subset of Ω, then $\bigcap_{E\in\mathcal{E}}\ {}^*E$ is nonempty.

We emphasize that it is possible to make a single choice of Λ and $\mathcal{U}$ that works for *all* choices of $\mathcal{E}$. It may be helpful to compare this principle with the following characterization of compact topological spaces: they are spaces in which, whenever $\mathcal{E}$ is a proper filter, then $\bigcap_{E\in\mathcal{E}} \mathrm{cl}(E)$ is nonempty. The equivalence of (UF1) and (UF4) is similar to a result proved by Lutz and Goze [1981].

Proof of (UF1) $\Rightarrow$ (UF4). We may assume Ω is infinte, by replacing it with a larger set if necessary.

Let Φ be the family of all proper filters on subsets of Ω. We shall use $\Lambda = \Omega^\Phi =$ {functions from Φ into Ω}. For each $\mathcal{E} \in \Phi$ and each $E \in \mathcal{E}$, consider the set

$$\Lambda_{\mathcal{E},E} \quad = \quad \{\lambda \in \Lambda \quad : \quad \lambda(\mathcal{E}) \in E\}.$$

It is easy to verify that the collection $\mathcal{S} = \{\Lambda_{\mathcal{E},E} \ : \ \mathcal{E} \in \Phi, E \in \mathcal{E}\}$ is a filter subbase — i.e., it has the finite intersection property. Hence, by Cartan's Ultrafilter Principle, there exists an ultrafilter $\mathcal{F}$ on Λ such that $\mathcal{F} \supseteq \mathcal{S}$. We shall show that this $\mathcal{F}$ has the required property.

Indeed, let any proper filter $\mathcal{E} \in \Phi$ be given. Define a function $\varepsilon : \Lambda \to \Omega$ by taking $\varepsilon(\lambda) = \lambda(\mathcal{E})$ for each $\lambda \in \Lambda$, and let $\xi \in {}^*\Omega$ be the equivalence class of the function ε. We shall show that $\xi \in \bigcap_{E\in\mathcal{E}}\ {}^*E$. (The remainder of the proof is just a matter of unwinding the notation; the reader may find it easier to proceed on his or her own instead of reading further.) For every $E \in \mathcal{E}$, we have

$$\{\lambda \in \Lambda : \varepsilon(\lambda) \in E\} \quad = \quad \{\lambda \in \Lambda \ : \ \lambda(\mathcal{E}) \in E\} \quad = \quad \Lambda_{\mathcal{E},E} \quad \in \quad \mathcal{F}.$$

Thus the condition $\varepsilon(\cdot) \in E$ is satisfied "almost everywhere," so $\xi \in {}^*E$.

If $\mathcal{F}$ is not a free ultrafilter, then it is fixed — whence $^*S = S$ for every set S, by 9.44.a. But since Ω is infinite, it has some free ultrafilter $\mathcal{E}$, by 6.33. Then $\varnothing = \bigcap_{E \in \mathcal{E}} E = \bigcap_{E \in \mathcal{E}} {}^*E \neq \varnothing$, a contradiction. Thus the ultrafilter $\mathcal{F}$ must be free.

Proof of (UF4) $\Rightarrow$ (UF1). Let $\mathcal{E}$ be a proper filter on a set Ω; we wish to extend it to an ultrafilter. Let Λ and $\mathcal{F}$ be as in (UF4); let $\xi \in \bigcap_{E \in \mathcal{E}} {}^*E$; let $\varepsilon : \Lambda \to \Omega$ be any function whose equivalence class is ξ. Let $\mathcal{D}$ be the filter on Ω generated by the filterbase $\{\varepsilon(F) : F \in \mathcal{F}\}$. It is easy to verify that $\mathcal{D}$ is an ultrafilter on Ω and that $\mathcal{D} \supseteq \mathcal{E}$.

Exponential (Dual) Functors

9.55. A few categories that we shall consider in later chapters have functors that we shall now describe, called **exponential functors** or **dual functors**. These categories satisfy five hypotheses, listed below as (H1), (H2), (H3), (H4), and (H5).

Let $\mathfrak{C}$ be a given category, and let Δ be some particular object in that category. Some commonly used choices of $\mathfrak{C}$ and Δ are listed in the table below. In the table, 2 stands for the set $\{0, 1\}$, $\mathbb{F}$ stands for a scalar field (generally $\mathbb{R}$ or $\mathbb{C}$), and $\mathbb{T}$ stands for the circle group (see 10.32).

Objects of $\mathfrak{C}$	Morphisms	Δ
sets	functions	2
Boolean algebras	Boolean homomorphisms	2
Boolean spaces	continuous maps	2
Tychonov spaces	continuous maps	$[0, 1]$
vector spaces	linear maps	$\mathbb{F}$
Riesz spaces	order bounded linear maps	$\mathbb{R}$
topological vector spaces	continuous linear maps	$\mathbb{F}$
Banach spaces	continuous linear maps	$\mathbb{F}$
Pontryagin groups	continuous homomorphisms	$\mathbb{T}$

For each object X in $\mathfrak{C}$, define the set

$$X^* \quad = \quad \{\mathfrak{C}\text{-morphisms from } X \text{ into } \Delta\}.$$

For each morphism $f : X \to Y$ in the category $\mathfrak{C}$, define a mapping $f^* : Y^* \to X^*$ by the rule $f^*(\lambda) = \lambda \circ f$ for all $\lambda \in Y^*$, as in the diagram below.

$$\begin{array}{ccccc} & X & \xrightarrow{f} & Y & \\ \lambda \circ f = f^*(\lambda) \in X^* & & \searrow \quad \swarrow & & \lambda \in Y^* \\ & & \Delta & & \end{array}$$

We sometimes refer to X^* and f^* as the **duals** or **adjoints** of X and f. For a simple, concrete example, see 11.22.d.

In the categories where duals are useful, the following hypotheses are satisfied:

(H1) For each object X in $\mathfrak{C}$, the elements of X^* separate the points of X. That is, for any two distinct points $x_1, x_2 \in X$, there exists at least one morphism $f : X \to \Delta$ such that $f(x_1) \neq f(x_2)$.

(H2) There is some natural way to attach structures to the dual sets X^*, making them objects in another category $\mathfrak{C}^*$, and making the dual functions $f^* : Y^* \to X^*$ into morphisms in that category.

It is easy to verify that the rules $X \mapsto X^*$ and $f \mapsto f^*$ reverse arrows and compositions; thus they define a contravariant functor from $\mathfrak{C}$ into $\mathfrak{C}^*$. We may refer to it as the **exponential functor**, **dual functor**, or **adjoint functor**, though each of those terms has other meanings as well.

In the most interesting instances of this theory, the bidual category $\mathfrak{C}^{**}$ is identical to the original category $\mathfrak{C}$, and that fact is very important in the theory sketched below.

Actually, in most cases of interest, $\mathfrak{C}$ and $\mathfrak{C}^*$ are the same category, but that could be viewed as mere coincidence; it is actually irrelevant to the theory developed below. Moreover, we have $\mathfrak{C} \neq \mathfrak{C}^*$ in at least one important application: The categories of Boolean spaces and Boolean algebras are dual to each other.

Elementary example. If $\mathfrak{C}$ and $\mathfrak{C}^*$ are both the category of sets and $\Delta = \{0, 1\}$, then $f \mapsto f^*$ is the inverse image functor defined in 9.32.

9.56. Much of the interest in exponential functors stems from the fact that some properties of X and f correspond to *dual* properties of X^* and f^*. We can switch back and forth between either setting and its dual, working with whichever properties are more convenient. For instance, show that

> if one of the functions $f : X \to Y$ or $f^* : Y^* \to X^*$ is surjective, then the other is injective.

(*Hint*: Use (H1).) In some categories, though not all, a converse can be proved:

> if one of the functions $f : X \to Y$ or $f^* : Y^* \to X^*$ is injective, then the other is surjective.

Exercise. Prove that this converse is valid in the category of sets — i.e., where any set is an object, and any function is a morphism; assume Δ is a set containing two or more points.

9.57. Assume $\mathfrak{C}$ is a category that has a dual functor, mapping into some category $\mathfrak{C}^*$. Suppose that the category $\mathfrak{C}^*$ also has a dual functor, which maps into some category $\mathfrak{C}^{**}$. Also assume that

(H3) The categories $\mathfrak{C}$ and $\mathfrak{C}^*$ have special objects Δ with the *same* underlying set (perhaps with different structures attached).

We shall denote both of these special objects by the same symbol Δ. By composing the two contravariant, dual functors, we obtain a covariant functor from $\mathfrak{C}$ into $\mathfrak{C}^{**}$, called the **bidual functor**:

$$X \mapsto X^* \mapsto X^{**}, \qquad f \mapsto f^* \mapsto f^{**}.$$

This functor has some further properties of interest, which we now describe.

Each $\lambda \in X^*$ was defined as a function with argument $x \in X$ and value $\langle \lambda, x\rangle = \lambda(x) \in \Delta$. Let now us change our viewpoint, and instead view x as the function, with λ for the argument. Then x acts as a mapping $T_x = \langle \cdot, x\rangle : X^* \to \Delta$, called the **evaluation map** at x. In the categories of interest, this further hypothesis is satisfied:

(H4) If X is an object in $\mathfrak{C}$ and $x \in X$, then the mapping $T_x = \langle \cdot, x\rangle : X^* \to \Delta$ is a morphism in the category $\mathfrak{C}^*$, and thus T_x is a member of the set X^{**}.

By (H1), if $x \neq x'$ then the mappings T_x and $T_{x'}$ are distinct. Thus $x \mapsto T_x$ is an injective mapping from X into X^{**}, which we may view as an inclusion — i.e., we may view the underlying set of X as a *subset* of the underlying set of X^{**} (without regard to the additional structures attached to those sets). The inclusion map $T : X \xrightarrow{\subseteq} X^{**}$ is sometimes known as the **canonical embedding of X in its bidual**. In many categories, the bidual functor has this further property:

(H5) The categories $\mathfrak{C}$ and $\mathfrak{C}^{**}$ are the same, and X is a subobject of X^{**} in that category; thus the canonical embedding $X \xrightarrow{\subseteq} X^{**}$ is a morphism.

9.58. *Exercise.* If $f : X \to Y$ is a morphism, then the function $f^{**} : X^{**} \to Y^{**}$ is an extension of f — that is, $\mathrm{Graph}(f^{**}) \supseteq \mathrm{Graph}(f)$.

Hint: Let $S : X \xrightarrow{\subseteq} X^{**}$ and $T : Y \xrightarrow{\subseteq} Y^{**}$ be the canonical embeddings. What must be verified is $f^{**}(S_x) = T_{f(x)}$ — or more concretely, $[f^{**}(S_x)](\lambda) = T_{f(x)}(\lambda)$ for each $\lambda \in Y^*$.

9.59. For some objects X, the canonical embedding $x \mapsto T_x$ turns out to be *surjective* — i.e., the inclusion morphism $X \xrightarrow{\subseteq} X^{**}$ is actually an isomorphism $X \xrightarrow{\simeq} X^{**}$. Such an object X will be called **reflexive**.

In some categories $\mathfrak{C}$, *every* object is reflexive (and so the term "reflexive" is not commonly used in those categories). In such a category we have $X = X^{**}$ and $f = f^{**}$ for all objects X and morphisms f. The canonical embedding $x \mapsto T_x$, from X to X^{**}, is then called the **canonical isomorphism**. This isomorphism is established:

- for the category of Hausdorff locally convex topological linear spaces with weak topologies, by 28.12.e.
- for the Banach spaces of type $L^p(\mu)$, with $1 < p < \infty$, by 28.50.
- for the category of Pontryagin groups (i.e., locally compact Hausdorff Abelian groups), by the Pontryagin Duality Theorem 26.44.
- for the categories of Boolean algebras and Boolean spaces, by the topological version of the Stone Representation Theorem (see 17.44 and the sections following it).

On the other hand, in some categories we can easily establish that *no* object is reflexive. For instance, in the category of sets (without additional structure) and in the category of infinite-dimensional vector spaces (without topology), we can prove that X^* is strictly larger than X (see 2.20.l and 11.36), and therefore X^{**} cannot equal X.

In still other categories, some objects are reflexive while others are not, and reflexivity may be linked to other, important properties. For instance, a Banach space is reflexive if and only if its closed unit ball is weakly compact; see 28.41.

Chapter 10

The Real Numbers

10.1. *Preview.* A significant part of the history of mathematics is the successive extension of number systems — especially, the inclusions $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.

Our language still reflects the resistance with which some of these extensions originally were met. The ancient Greeks were reluctant to admit that the universe could not be explained in terms of ratios of whole numbers; even today, the word "irrational" means an element of $\mathbb{R}\backslash\mathbb{Q}$ but also means "crazy." Other terms occasionally used for an irrational number are "radical" or "surd" (an abbreviation of "absurd"). Centuries later, when mathematicians began to use complex numbers to analyze polynomial equations, they were reluctant to admit that -1 could have a square root. A "real" number could not be such a square root; such a square root must be "imaginary," and this name stuck, too. Mathematical nomenclature was not so disparaging a few decades ago when nonstandard analysis gave a rigorous foundation for the use of infinitesimals; the new numbers in ${}^*\mathbb{R}\backslash\mathbb{R}$ were simply called "nonstandard" — a rather neutral term, by comparison.

Usually, a "number" means an element of a field. The two fields most commonly used in analysis are the real number system $\mathbb{R}$ and the complex number system $\mathbb{C}$. We shall introduce both of these fields formally in this chapter, though we have assumed some informal familiarity with $\mathbb{R}$ in earlier chapters.

Dedekind Completions of Ordered Groups

10.2. *Remarks and definition.* If X is an ordered group other than $\{0\}$, then X cannot have a greatest element (*easy exercise*). Hence X cannot be "complete," in the sense of 3.23. The closest X can come to such a condition is Dedekind completeness. For this reason, in the context of ordered groups, a "complete ordered group" generally means a Dedekind complete ordered group.

10.3. Let $(X, \preccurlyeq)$ be an ordered group. For any $x \in X$ and $K \subseteq \mathbb{Z}$, let $Kx = \{kx : k \in K\}$, where the multiplication is defined as in 8.10.h. Then the following two conditions are equivalent to each other. If one, hence both, of them are satisfied, we say X is **integrally closed**.

(A) Whenever the set $\mathbb{N}x$ is bounded above, then $x \preccurlyeq 0$.

(B) Whenever the set $\mathbb{Z}_+x = \{0, x, 2x, \ldots\}$ is bounded above, then $x \preccurlyeq 0$.

Also, the following two conditions are equivalent to each other. If one, hence both, are satisfied, we say X is **Archimedean**.

(C) Whenever the set $\mathbb{Z}x$ is bounded above, then $x = 0$.

(D) $\{0\}$ is the only subgroup of X that has an upper bound.

Furthermore, conditions (A)(B) imply conditions (C)(D). If the ordering $\preccurlyeq$ on X is a lattice ordering, then all four conditions are equivalent.

Proof. To show (A) $\iff$ (B), note that the set $\{x, 2x, 3x, \ldots\}$ is bounded above by β if and only if the set $\{0, x, 2x, \ldots\}$ is bounded above by $\beta - x$. The proof of (C) $\iff$ (D) is easy; we omit the details. To show (B) $\Rightarrow$ (C), note that if $\mathbb{Z}x$ is bounded above, then $\mathbb{Z}_+x$ and $\mathbb{Z}_+(-x)$ are both bounded above, hence $x \preccurlyeq 0$ and $-x \preccurlyeq 0$, hence $x = 0$.

Finally, to prove (C) $\Rightarrow$ (B) when X is lattice ordered, suppose $\mathbb{Z}_+x$ is bounded above by β. By 8.42.p show that $\mathbb{Z}_+(x \vee 0)$ is bounded above by $\beta \vee 0$. On the other hand, $-(x \vee 0) \preccurlyeq 0$; by adding show that $-n(x \vee 0) \preccurlyeq 0$ for $n \in \mathbb{N}$. Thus the subgroup $\mathbb{Z}(x \vee 0)$ is bounded above. By (C), then, $x \vee 0 = 0$; hence $x \preccurlyeq 0$.

10.4. *Some basic properties and examples.*

a. Any subgroup of an integrally closed group is integrally closed.

b. Any Dedekind complete, ordered group is integrally closed.

Hints: Suppose $x \in X$ with $\mathbb{Z}_+x$ bounded above. Let $\beta = \sup(\mathbb{Z}_+x)$. Show that $\beta - x$ is also an upper bound for $\mathbb{Z}_+x$; hence $\beta \preccurlyeq \beta - x$.

c. *Examples.* The groups $\mathbb{Z}$ and $\mathbb{R}$ are Dedekind complete; $\mathbb{Q}$ is not. All three of these groups are integrally closed.

The group $\mathbb{Z}^2$ with the lexicographical ordering (see 3.44.a) is an ordered group that is not integrally closed.

Another example of an ordered group that is not integrally closed will be given in 10.19.

10.5. Theorem: Completion of a Group. Let $(D, \preccurlyeq)$ be an ordered group, and let X be a Dedekind completion of D (as in 4.33 and 4.34). Then

X can be made into an ordered group in which D is a subgroup

if and only if D is integrally closed. Furthermore, if those conditions are satisfied, then the group operations on X must satisfy

$$\begin{aligned} \xi + \eta &= \sup\{x + y \;:\; x, y \in D,\ x \preccurlyeq \xi,\ y \preccurlyeq \eta\}, && \text{(a1)} \\ &= \inf\{x + y \;:\; x, y \in D,\ x \succcurlyeq \xi,\ y \succcurlyeq \eta\}, && \text{(a2)} \end{aligned}$$

$$\begin{aligned} -\xi &= \sup\{-x \;:\; x \in D,\ x \succcurlyeq \xi\}, && \text{(b1)} \\ &= \inf\{-x \;:\; x \in D,\ x \preccurlyeq \xi\}. && \text{(b2)} \end{aligned}$$

Remarks. Our proof is based on that in Fuchs [1963]. However, we assume D is commutative, whereas Fuchs does not impose that restriction. Fuchs mentions Krull, Lorenzen, Clifford, Everett, and Ulam as contributors to this theorem.

Proof of theorem. If D has such a group completion X, then X is Dedekind complete, hence integrally closed (by 10.4.b); hence D is integrally closed (by 10.4.a). Equations (a1) through (b2) follow from 8.33 and the fact that D is sup- and inf-dense in X.

Conversely, suppose D is integrally closed. Let X be a Dedekind completion of D. Define operations on X by (a1) and (b1); we shall show that these make X into an ordered group that has D as a subgroup. At the end of this proof we shall show the validity of (a2) and (b2) as well.

With definition (a1) it is easy to see that $+$ extends the addition operation of D, and that $+$ is commutative. The beginner is cautioned not to assume too much just on the basis of notation: Although we use the symbol "$+$," our proof must not rely on any as-yet-unestablished properties of addition in X. In particular, we must not yet use associativity or subtraction (the existence of additive inverses) in X. However, we can freely use associativity and subtraction in the given group D.

Our first step will be to show that addition in X is associative. Let any $\xi, \eta, \zeta \in X$ be given. Let $p \in D$. Then each of the following statements is equivalent to the next:

$p \succcurlyeq (\xi + \eta) + \zeta$

$p \succcurlyeq u + z$ whenever $u, z \in D$ and $u \preccurlyeq \xi + \eta$ and $z \preccurlyeq \zeta$

$p - z \succcurlyeq u$ whenever $u, z \in D$ and $u \preccurlyeq \sup\{x + y \ : \ x, y \in D,\ x \preccurlyeq \xi,\ y \preccurlyeq \eta\}$ and $z \preccurlyeq \zeta$

$p - z \succcurlyeq \sup\{x + y \ : \ x, y \in D,\ x \preccurlyeq \xi,\ y \preccurlyeq \eta\}$ whenever $z \in D$ and $z \preccurlyeq \zeta$

$p - z \succcurlyeq x + y$ whenever $x, y, z \in D$ and $x \preccurlyeq \xi,\ y \preccurlyeq \eta,\ z \preccurlyeq \zeta$

$p \succcurlyeq x + y + z$ whenever $x, y, z \in D$ and $x \preccurlyeq \xi,\ y \preccurlyeq \eta,\ z \preccurlyeq \zeta$.

The last statement is symmetric in ξ, η, ζ; hence the first statement is not affected by a permuting of those three terms. Thus addition in X (defined as in (a1)) is associative. Since D is sup-dense in X, the addition in X also satisfies $\xi + 0 = 0 + \xi = \xi$. Thus we have established that $(X, 0, +)$ is an additive monoid.

Define $-\xi$ as in (b1); the mapping $x \mapsto -x$ from D into D is thus extended to a mapping from X into X. To show that $(X, +, -, 0)$ is an additive group, fix any $\xi \in X$, and let $\gamma = \xi + (-\xi)$; we must show that $\gamma = 0$. Observe that

$$v \preccurlyeq -\xi \quad \Leftrightarrow \quad -v \succcurlyeq \xi, \qquad \text{for } v \in D,$$

and hence

$$\gamma = \xi + (-\xi) = \sup\{x - v : \ x, v \in D,\ x \preccurlyeq \xi \preccurlyeq v\}.$$

From this it follows immediately that $\gamma \preccurlyeq 0$. To show that $\gamma \succcurlyeq 0$, we shall apply 3.21.g. Fix any $u \in D$ with $u \succcurlyeq \gamma$; it suffices to show that $u \succcurlyeq 0$. From $u \succcurlyeq \gamma$ we conclude, successively, that

$x, v \in D$, $x \preccurlyeq \xi \preccurlyeq v$ implies $u \succcurlyeq x - v$;

for each $v \in D$ with $v \succcurlyeq \xi$, we have $x \in D,\ x \preccurlyeq \xi \ \Rightarrow\ x \preccurlyeq u + v$;

for each $v \in D$ with $v \succcurlyeq \xi$, we have $\xi \preccurlyeq u + v$;

$\mathbb{N}(u)$ is bounded below;

$\mathbb{N}(-u)$ is bounded above.

By assumption D is integrally closed; hence $u \succcurlyeq 0$.

Thus X is an additive group, when equipped with the operations $+$ and $-$ defined as in (a1) and (b1). The ordering is translation-invariant (as defined in 8.30); this follows trivially from our definition of addition in X. Hence X is in fact an ordered group, with D as a subgroup. Therefore $\inf(-S) = -\sup(S)$ for any set $S \subseteq X$; now (a2) and (b2) follow from (a1) and (b1).

10.6. Suppose the conditions of the preceding theorem are satisfied. Suppose, also, that Q is another ordered group, and $f : D \to Q$ is a sup-preserving group homomorphism, and $F : X \to Q$ is a sup-preserving extension of f. Then F is also a group homomorphism.

Proof. It suffices to show that F preserves addition. Define sets L_ζ as in 4.31; then $\zeta = \sup(L_\zeta)$. Then for any $\xi, \eta \in X$ we have

$$\begin{aligned}&\{x + y \ :\ x, y \in D,\ x \preccurlyeq \xi,\ y \preccurlyeq \eta\}\\ &\qquad = \quad \{x \in D \ :\ x \preccurlyeq \xi\} + \{y \in D \ :\ y \preccurlyeq \eta\} \quad = \quad L_\xi + L_\eta.\end{aligned}$$

Hence

$$\begin{array}{rcll}F(\xi + \eta) & = & F(\sup(L_\xi + L_\eta)) & \text{by (a1) in 10.5}\\ & = & \sup(f(L_\xi + L_\eta)) & \text{since } F \text{ is sup-preserving}\\ & = & \sup(f(L_\xi) + f(L_\eta)) & \text{since } f \text{ is additive on } D\\ & = & \sup(f(L_\xi)) + \sup(f(L_\eta)) & \text{by 8.33}\\ & = & F(\sup(L_\xi)) + F(\sup(L_\eta)) & \text{since } F \text{ is sup-preserving}\\ & = & F(\xi) + F(\eta) & \text{since } D \text{ is sup-dense in } X.\end{array}$$

Ordered Fields and the Reals

10.7. *Definitions.* A **chain ordered ring** is a ring R equipped with an ordering $\leq$ such that

(i) $(R, \leq)$ is a chain;

(ii) the ordering is translation-invariant — that is, $x \leq y \ \Rightarrow\ x + u \leq y + u$ for all $x, y, u \in X$; and

(iii) $x, y \geq 0 \Rightarrow xy \geq 0$.

If R is also a field, we shall call it a **chain ordered field**. (Some mathematicians call these an *ordered ring* and an *ordered field*, respectively, but that is not specific enough for the purposes of this book.)

Some basic examples. The rational number system $\mathbb{Q}$ (with its usual ordering) is clearly a chain ordered field; we shall assume informal familiarity with that fact, but it also follows from a construction presented in 10.11. The real number system $\mathbb{R}$, introduced in the next paragraph, is a chain ordered field. Certain other subsets of $\mathbb{R}$ are also chain ordered fields — for instance, $\{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$, which is introduced in 10.23.a. In 10.12 we present a chain ordered field that is not contained in $\mathbb{R}$.

10.8. We now define the real number system $\mathbb{R}$ to be a *Dedekind complete, chain ordered field* (or, in the terminology of some mathematicians, a *complete ordered field*). To make sense of this definition, we shall show that (i) there exists a Dedekind complete, chain ordered field, and (ii) any two such fields are isomorphic; thus, there is only *one* real number system. We shall prove those facts in 10.15.

Discussion. Intuitively, we usually think of the real number system as a model for the set of all points on a Euclidean straight line. However, that description has certain drawbacks. It does not determine $\mathbb{R}$ uniquely, for it also fits $^*\mathbb{R}$ quite well. Also, the geometric description does not translate readily into usable algebraic axioms.

We also think of reals as "infinite decimal expansions" such as $3.14159265358979323\cdots$. In grade school we learn, informally, how to perform arithmetic operations with such expansions. A formal theory of such expansions is sketched in 10.44 and 10.45. Perhaps this view of the real number system is the most concrete and the most useful for purposes of real-world applications — in physics, engineering, etc.

However, in advanced analysis we usually consider the decimal expansions to be just *representations* for numbers, not the numbers themselves. Those numbers have other representations (in binary, in ternary, in hexadecimal, etc.). In the development of abstract theory, what we really need are not concrete representations such as $3.14159265358979323\ldots$, but the essential *properties* of the real numbers, which are used to prove theorems. That $\mathbb{R}$ is a *field* means that we can do ordinary arithmetic; that it is *chain ordered* means that inequalities work the way they should; that it is *Dedekind complete* means that we can take sups, infs, and limits. Analysts often take these ideas for granted and forget how complicated a structure the real number system is.

Actually, we shall prove the existence of $\mathbb{R}$ in several different ways. The proof in 10.15.d is fairly detailed; other proofs are sketched briefly in 10.45 and 19.33.c. All of the constructions are somewhat complicated and nonintuitive — they represent a real number as a set of rational numbers, or a pair of sets of rational numbers, or a set of pairs of rational numbers, etc. The theorem on the uniqueness of the reals, in 10.15.e, tells us that these constructions of $\mathbb{R}$ from $\mathbb{Q}$ all yield the same result. Any one of these constructions is sufficient, and it does not matter which one we use. After we have proved the existence of a Dedekind complete, chain ordered field by representing it in terms of rational numbers, we may discard that representation; we may return to thinking of real numbers as indivisible, primitive objects like the points on a line.

10.9. *A few basic properties.*

a. If $\mathbb{F}$ is a chain ordered field, then $0 < x < y \quad \Rightarrow \quad 0 < \dfrac{1}{y} < \dfrac{1}{x}$.

b. If $\mathbb{F}$ is a chain ordered field, then $\mathbb{F}$ has no greatest or least element.

c. If $\mathbb{F}$ is a chain ordered field, then $x^2 = -1$ has no solution x in $\mathbb{F}$.

d. If R is a chain ordered ring with unit, other than $\{0\}$, then the unique homomorphism from $\mathbb{Z}$ to R (noted in 8.19.d) is injective and order-preserving. Thus R contains an isomorphic copy of $\mathbb{Z}$.

Hint: Let 1_R denote the multiplicative identity of R. Show that $1_R > 0$, and hence $1_R + 1_R + \cdots + 1_R$ (the sum of finitely many such terms) is also positive.

e. If $\mathbb{F}$ is a chain ordered field, then the unique ring homomorphism from $\mathbb{Q}$ into $\mathbb{F}$ (noted in 8.23.c) is injective and order-preserving. Thus $\mathbb{F}$ contains a uniquely determined isomorphic copy of $\mathbb{Q}$. Identifying various sets with their isomorphic copies when no confusion will result, we may write $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{F}$; we shall follow this convention in results below.

f. No finite field can be a chain ordered field.

10.10. *An abstract construction of chain ordered fields.* Let D be an integral domain, and let $\mathbb{F}$ be the resulting field of fractions, as in 8.22. Suppose that some ordering $\sqsubseteq$ is given on D, making it a chain ordered ring. Define an ordering $\preccurlyeq$ on $\mathbb{F}$ as follows: For $m, n, x, y \in D$ with $m, n \sqsupset 0$, define

$$\frac{x}{m} \succcurlyeq \frac{y}{n} \qquad \text{to mean} \qquad xn \sqsupseteq ym.$$

(The reader should verify that this ordering does not depend on the choice of the representatives of the equivalence classes.) Show that $\mathbb{F}$ is then a chain ordered field. Moreover, considering D as a subset of $\mathbb{F}$, show that the ordering on $\mathbb{F}$ is an extension of the ordering on D.

Two important particular cases of this construction are given in the next two sections.

10.11. *Example.* When the integral domain D is $\mathbb{Z} = \{\text{integers}\}$, then the field of fractions is $\mathbb{Q}$, the field of **rational numbers**; the ordering given in 10.10 is just the usual ordering.

10.12. *A non-Archimedean example.* Let $(\mathbb{A}, \leq)$ be a chain ordered ring that is also an integral domain. (An example to keep in mind for now is $\mathbb{A} = \mathbb{Z}$; later we may reconsider this construction with $\mathbb{A} = \mathbb{R}$.) Let x be a variable. Let D be $\mathbb{A}[x]$, the ring of polynomials in the one variable x with coefficients in $\mathbb{A}$; then D is an integral domain (see 8.24). The resulting field $\mathbb{F}$ of fractions is $\mathbb{A}(x)$, the **field of rational functions** in the one variable x with coefficients in $\mathbb{A}$.

On $D = \mathbb{A}[x]$, we now define this ordering: $p \sqsupset q$ will mean that the leading coefficient of the polynomial $p - q$ (defined in 8.24) is strictly greater than 0. Verify that this makes $(\mathbb{A}[x], \sqsubseteq)$ into a chain ordered ring. Hence our construction in 10.10 makes $(\mathbb{A}(x), \preccurlyeq)$ into a chain ordered field. This example can be found in various algebra books; another source is Lightstone and Robinson [1975].

A few observations about this field will be useful later in this chapter:

a. If p and q are polynomials, the degree of p is greater than the degree of q, and the leading coefficient of p is positive, then $p \sqsupset q$.

b. If p and q are polynomials other than 0, the **degree** of the rational function p/q will mean the difference $\deg(p) - \deg(q)$, and the leading coefficient of p/q will mean the quotient of the leading coefficients of p and q. Show that if r, s are rational functions with positive leading coefficients, and $\deg(r) > \deg(s)$, then $r \sqsupset s$.

c. The function $p(x) = x$ is strictly greater than every constant function k. Thus, the sequence $1, 2, 3, \ldots$ is bounded above.

d. The sequence $1, x, x^2, x^3, \ldots$ is not bounded above — i.e., there does not exist a rational function $r(x)$ that satisfies $r(x) \sqsupseteq x^n$ for all nonnegative integers n. (Contrast this result with 10.20.c.)

10.13. *Definition and exercise.* Let $\mathbb{F}$ be a chain ordered field (as defined in 10.7 — hence $\mathbb{F}$ is also lattice ordered). Then $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{F}$ as noted in 10.9.e.

Then the following conditions are equivalent; a chain ordered field $\mathbb{F}$ possessing one (hence all) of these conditions is said to be an **Archimedean** field.

(A) $\mathbb{F}$ is Archimedean in the sense of 10.3; that is, $\{0\}$ is the only additive subgroup of $\mathbb{F}$ that is bounded above by an element of $\mathbb{F}$.

(B) $\mathbb{N}$ does not have an upper bound in $\mathbb{F}$.

(C) The set $\{\frac{1}{n} : n \in \mathbb{N}\}$ has infimum (in $\mathbb{F}$) equal to 0.

(D) For each $c \in \mathbb{F}$, the set $\{m \in \mathbb{Z} : m > c\}$ has a lowest element.

(E) Between any two elements of $\mathbb{F}$ there is an element of $\mathbb{Q}$. (This is sometimes called the **Density Property**.)

(F) $\mathbb{Q}$ is sup-dense and inf-dense in $\mathbb{F}$ (see 4.31).

Hints for the equivalence proof: It is fairly easy to see that conditions (A) through (D) are equivalent. To show that those conditions imply (E), let $\alpha, \beta \in \mathbb{F}$ be given with $\alpha < \beta$. By (C) and (D), there exist $n \in \mathbb{N}$ with $\beta - \alpha > 1/n$, and $m \in \mathbb{Z}$ with $m > n\alpha \geq m - 1$. Show that m/n lies between α and β.

It is easy to see that (E) implies (F).

To see that (F) implies (C), let $S = \{\frac{1}{n} : n \in \mathbb{N}\}$ and $T = \{q \in \mathbb{Q} : q > 0\}$. Since every element of either of these sets is less than some member of the other set, we have $\inf(S) = \inf(T)$; but $\inf(T) = 0$ since $\mathbb{Q}$ is inf-dense in $\mathbb{F}$.

This presentation is based partly on Davis [1977].

10.14. (*Optional.*) Let $\mathbb{F}$ be a chain ordered field. Say that a sequence (x_n) is *Cauchy* in $\mathbb{F}$ if for each ε in $\mathbb{F}$ with $\varepsilon > 0$ there exists a positive integer M such that

$$j, k \geq M \qquad \Rightarrow \qquad -\varepsilon < x_j - x_k < \varepsilon.$$

(Cauchy sequences will be studied in another setting in Chapter 19.)

Proposition. Let $\mathbb{F}$ be a chain ordered field. Then (A) $\mathbb{F}$ is Archimedean if and only if (B) each bounded, monotone sequence in $\mathbb{F}$ is Cauchy. (Also see related results in 10.17.)

Hints: For (B) $\Rightarrow$ (A), it suffices to show that the set $S = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$ has infimum equal to 0. Suppose that some $b > 0$ is a lower bound for S. By the Cauchy criterion, there is some positive integer M such that

$$j, k \geq M \qquad \Rightarrow \qquad \frac{1}{j} - \frac{1}{k} < b.$$

Since $\frac{1}{2M} \in S$, we have $b \leq \frac{1}{2M}$. Use $j = M$ and $k = 2M$ to obtain a contradiction.

For (A) $\Rightarrow$ (B), let (x_n) be a bounded, monotone sequence. We may assume that (x_n) is increasing (why?) and that $x_1 \geq 0$ (why?); thus $0 \leq x_1 \leq x_2 \leq x_3 \leq \cdots \leq b$ for some $b \in \mathbb{F}$. Suppose (x_n) is not Cauchy. Then there exist some $\varepsilon > 0$ in $\mathbb{F}$ and some positive integers $n_1 < p_1 < n_2 < p_2 < \cdots$ such that $x_{p_j} - x_{n_j} > \varepsilon$ for all j. For any positive integer k, show that

$$b \geq x_{p_k} - 0 \geq (x_{p_k} - x_{n_k}) + \cdots + (x_{p_2} - x_{n_2}) + (x_{p_1} - x_{n_1}) > k\varepsilon.$$

Thus the set $\{1, 2, 3, \ldots\}$ is bounded above by b/ε.

10.15. *Examples and theorems about Archimedean fields.*

a. The field constructed in 10.12 is not Archimedean, by 10.12.c.

b. $\mathbb{Q}$ is an Archimedean field.

c. The multiplicative group $\{x \in \mathbb{Q} : x > 0\}$ is an Archimedean chain ordered group (as in 10.3).

d. Existence of the reals. There exists a Dedekind complete, chain ordered field. In fact, if $\mathbb{F}$ is any Archimedean field, then the Dedekind completion of $\mathbb{F}$ is a Dedekind complete, chain ordered field.

Hints: Let $\mathbb{R}$ denote the Dedekind completion of $\mathbb{F}$. The completion is unique up to order isomorphism, by 4.38. Show that $\{\xi \in \mathbb{R} : \xi > 0\}$ is the (also unique) Dedekind completion of the multiplicative group $\{x \in \mathbb{F} : x > 0\}$. Use 10.5 to define addition and additive inverses in $\mathbb{R}$ and to define multiplication and multiplicative inverses in $\{\xi \in \mathbb{R} : \xi > 0\}$; thus $\mathbb{F}$ and $\{x \in \mathbb{F} : x > 0\}$ are subgroups of the groups $(\mathbb{R}, +)$ and $(\{\xi \in \mathbb{R} : \xi > 0\}, \cdot)$ respectively. Extend the definition of multiplication to other products of real numbers by $\xi \cdot \eta = [\mathrm{sgn}(\xi)][\mathrm{sgn}(\eta)]\ |\xi|\ |\eta|$. Show that this makes $\mathbb{R}$ into a Dedekind complete, chain ordered field.

The construction of the reals by cuts was published by Dedekind in 1872.

e. Uniqueness of the reals. Let $\mathbb{R}_1$ and $\mathbb{R}_2$ be two Dedekind complete, chain ordered fields. Then $\mathbb{R}_1$ and $\mathbb{R}_2$ contain ring-isomorphic copies of $\mathbb{Q}$, and there is an isomorphism (of rings with unit) from $\mathbb{R}_1$ onto $\mathbb{R}_2$ that leaves elements of $\mathbb{Q}$ fixed and preserves order.

Hints: By 10.13(E), $\mathbb{Q}$ is both sup-dense and inf-dense in $\mathbb{R}_1$. Hence $\mathbb{R}_1$ is a Dedekind completion of $\mathbb{Q}$. Similarly for $\mathbb{R}_2$. By the uniqueness of completions (4.38), there is a unique order isomorphism from $\mathbb{R}_1$ onto $\mathbb{R}_2$ that leaves $\mathbb{Q}$ fixed. By 10.6, that isomorphism preserves sums. Applying 10.6 to the multiplicative groups of positive elements, we see that that isomorphism also preserves products.

f. (*Optional.*) Let $\mathbb{F}$ be an ordered field. Show that $\mathbb{F}$ is Archimedean if and only if (after relabeling by isomorphism) we have $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$, in which case $\mathbb{Q}$ is both sup-dense and inf-dense in $\mathbb{F}$, and $\mathbb{R}$ is the Dedekind completion of $\mathbb{F}$.

10.16. *Remarks.* Our proof of the uniqueness of $\mathbb{R}$ depends on our use of conventional language and logic. If we change our rules of inference — e.g., if we restrict ourselves to first order language and logic, as is common in nonstandard analysis — then there may be many different models of the real line (though they may be indistinguishable except through the use of higher-order language and logic). See 14.68.

10.17. (*Optional.*) Let $\mathbb{F}$ be a chain ordered field. Let $\mathbb{F}$ be equipped with the order interval topology (see 5.15.f) and the resulting convergence (see 7.41 and 15.41). Then it can be shown that the following conditions are equivalent.

(A) $\mathbb{F}$ is Dedekind complete and thus is the real line.

(B) If (x_n) is any monotone, bounded sequence in $\mathbb{F}$, then (x_n) has a limit in $\mathbb{F}$.

(C) $\mathbb{F}$ is Archimedean, and every Cauchy sequence in $\mathbb{F}$ (defined as in 10.14) has a limit in $\mathbb{F}$.

(D) $\mathbb{F}$ is connected (defined as in 5.12).

(E) For any $a, b \in \mathbb{F}$ with $a \leq b$, the set $[a, b] = \{x \in \mathbb{F} : a \leq x \leq b\}$ is compact (defined as in 17.2).

(F) For any $a, b \in \mathbb{F}$ with $a \leq b$, the set $[a, b] = \{x \in \mathbb{F} : a \leq x \leq b\}$ is pseudocompact (defined as in 17.26.a).

We shall not prove the equivalence. These conditions and others are proved equivalent by Artmann [1988]; that exposition is based in part on Steiner [1966]. Artmann's book also gives an example of a non-Archimedean field in which every Cauchy sequence converges. Thus Dedekind completeness is *not* the same thing as Cauchy completeness.

THE HYPERREAL NUMBERS

10.18. By the **hyperreal line** (or the **hyperreal number system**) we shall mean any non-Archimedean, chain ordered field $\mathbb{H}$ that contains $\mathbb{R}$ as a subfield; the members of $\mathbb{H}$ are called **hyperreal numbers**.

Strictly speaking, there are many hyperreal lines. We gave one construction in 10.12; two more constructions are given in 10.19 and 10.20. However, usually we work with just one such field at a time, and so it is convenient to call that field "the" hyperreal number system while we are working with it.

Let $\mathbb{H}$ be a hyperreal line; thus $\mathbb{Z} \subsetneqq \mathbb{Q} \subsetneqq \mathbb{R} \subsetneqq \mathbb{H}$. Then:

a. Elements of $\mathbb{R}$ are called **real numbers**, or sometimes (for emphasis) **standard real numbers**.

b. A hyperreal number ξ is called **bounded** if $-r < \xi < r$ for some real number r; otherwise ξ is **unbounded**. (Other terms commonly used in place of "bounded" are *limited* and *hyperfinite*.)

Clearly, any real number is bounded. Show that some hyperreal number α is unbounded. Then $\pm\alpha, \pm 2\alpha, \pm 3\alpha, \ldots$ are different unbounded hyperreal numbers, and $\alpha \cdot \alpha$ is yet another one. The set of positive unbounded hyperreal numbers has no largest or smallest member. The set of bounded hyperreals is a commutative ring with unit.

c. A hyperreal number ξ is called **infinitesimal** if $-r < \xi < r$ for every positive real number r. Show that 0 is the only real infinitesimal. Show that a nonzero hyperreal number ξ is infinitesimal if and only if $1/\xi$ is unbounded. The set of positive infinitesimal numbers has no largest or smallest member.

Every positive real number is an upper bound for the set of infinitesimals. Show that the set of infinitesimals does not have a least upper bound in $\mathbb{H}$. This illustrates the fact (which we already knew) that $\mathbb{H}$ is not Dedekind complete.

The set of all infinitesimals is an ordered ring (without unit). (Some mathematicians exclude 0 when they define infinitesimal, but that definition has the disadvantage that the resulting set of infinitesimals does not have such a nice algebraic structure.)

d. Two hyperreal numbers are said to be **infinitely close** (or *infinitesimally close*) if their difference is an infinitesimal.

Let ξ be a bounded hyperreal number. Show that there is one and only one real number r that is infinitely close to ξ. That number r is called the **standard part** of ξ; we may abbreviate it by $\mathrm{std}(\xi)$.

Hint: To show that there is at least one such number, use the Dedekind completeness of $\mathbb{R}$ to prove that there is a real number $r = \inf\{s \in \mathbb{R} : s > \xi\}$; then show it has the required properties.

e. Show that

$$\{\text{bounded hyperreals}\} \quad = \quad \{\text{real numbers}\} \quad \oplus \quad \{\text{infinitesimals}\}$$

is an internal direct sum decomposition of one additive group into two subgroups. Show that

$$\mathrm{std} \quad : \quad \{\text{bounded hyperreals}\} \quad \to \quad \{\text{real numbers}\}$$

is an isotone map (for the ordering) and a ring homomorphism.

Thus, nestled around each real number r there are infinitely many bounded hyperreal numbers, all infinitely close to that real number r. In some books, some of these hyperreal numbers are denoted by $r + \varepsilon$, $r - \varepsilon$, $r + \delta$, $r - \delta$, etc.; a picture of a microscope is sometimes used to suggest their closeness to r.

f. *Remark.* It can be shown that the *smallest* hyperreal line is the field $\mathbb{R}(x)$, of rational functions in one variable with real coefficients, constructed as in 10.12. Indeed, if we take the real line and adjoin some element x that is infinitely large, then x acts as a transcendental over $\mathbb{R}$ and hence acts algebraically as an indeterminate — i.e., as a variable. The resulting field generated by $\mathbb{R} \cup \{x\}$ must then be $\mathbb{R}(x)$. This is discussed by Fleischer [1967]. Fleischer also points out that, although we may not be able to extend quite as many functions in the setting of this field as we did in 9.49, at least we can extend some functions *constructively*. See also the related remarks in 10.20.c.

10.19. *Ultrapowers of the reals.* Let $\mathcal{F}$ be a proper filter on a set Λ. Define the reduced power ${}^*\mathbb{R} = \mathbb{R}^\Lambda/\mathcal{F}$ and its arithmetical operations and ordering as in 9.41, 9.49, and 9.51. Then ${}^*\mathbb{R}$ is a ring with unit by 9.53 since $\mathbb{R}$ is a ring with unit. In fact, ${}^*\mathbb{R}$ is a commutative lattice algebra since $\mathbb{R}$ is. (Similarly, the *hypernatural numbers* ${}^*\mathbb{N}$ inherit some of the properties of $\mathbb{N}$.)

Recall from 9.52.e and 9.52.f that ${}^*\mathbb{R}$ is chain ordered if and only if $\mathcal{F}$ is an ultrafilter. Show that

a. ${}^*\mathbb{R}$ is a field if and only if $\mathcal{F}$ is an ultrafilter.

Hints: If $\mathcal{F}$ is not an ultrafilter, then Λ can be partitioned into sets Λ_1 and Λ_2, neither of which is an element of $\mathcal{F}$. Let α_1 and α_2 be their characteristic functions. Show that neither α_1 nor α_2 is equivalent to 0, but their product $\alpha_1\alpha_2$ is 0.

b. Suppose $\Omega, \Lambda, \mathcal{F}$ satisfy the conditions of the Enlargement Principle (9.54), and $\mathbb{R} \subseteq \Omega$. Then ${}^*\mathbb{R} = \mathbb{R}^\Lambda/\mathcal{F}$ is a non-Archimedean, chain ordered field.

Hint: Let $\mathcal{E} = \{S \subseteq \mathbb{R} : S \supseteq (n, +\infty)$ for some positive integer $n\}$. Then $\mathcal{E}$ is a proper filter on $\mathbb{R}$. Show that any member of $\bigcap_{E\in\mathcal{E}} {}^*E$ is an upper bound for $\mathbb{N}$.

10.20. Assume that $\mathcal{F}$ is a free ultrafilter on the set $\Lambda = \mathbb{N}$; hence ${}^*\mathbb{R} = \mathbb{R}^\mathbb{N}/\mathcal{F}$ is a chain ordered field. Show that

a. ${}^*\mathbb{R}$ is non-Archimedean. In fact, the equivalence class of the sequence $(1, 2, 3, \ldots)$ is an upper bound for $\mathbb{N}$ in ${}^*\mathbb{R}$.

b. Different constructions may yield slightly different hyperreal number systems. For instance, one of the sequences $\alpha = (1, 0, 1, 0, 1, \ldots)$ or $\beta = (0, 1, 0, 1, 0, \ldots)$ is equivalent to the real number 0 and the other is equivalent to the real number 1. We can choose which is which, if we want to, by specifying $\mathcal{F}$ in more detail. Let $\mathcal{C} = \{$cofinite subsets of $\mathbb{N}\}$, let $E = \{$even numbers$\} = \{2, 4, 6, \ldots\}$, and let $F = \{$odd numbers$\} = \{1, 3, 5, \ldots\}$. Using 5.5.i and 6.33, we can obtain free ultrafilters $\mathcal{E}, \mathcal{F}$ on $\mathbb{N}$ that satisfy $\mathcal{E} \supseteq \{E\} \cup \mathcal{C}$ and $\mathcal{F} \supseteq \{F\} \cup \mathcal{C}$. In the hyperreal line $\mathbb{R}^\mathbb{N}/\mathcal{E}$ we have $\pi(\alpha) = 0$ and $\pi(\beta) = 1$; in the hyperreal line $\mathbb{R}^\mathbb{N}/\mathcal{F}$ we have $\pi(\alpha) = 1$ and $\pi(\beta) = 0$.

c. Every countable set $S \subseteq {}^*\mathbb{R}$ is order bounded.

Hints: Let $S = \{\pi(f_1), \pi(f_2), \pi(f_3), \ldots\}$, where each f_n is a function from $\mathbb{N}$ into $\mathbb{R}$. Define functions $u, v : \mathbb{N} \to \mathbb{R}$ by taking

$$u(k) = \min\{f_1(k), f_2(k), \ldots, f_k(k)\}, \qquad v(k) = \max\{f_1(k), f_2(k), \ldots, f_k(k)\}.$$

For each $j \in \mathbb{N}$, show that $C_j = \{k \in \mathbb{N} : u(k) \leq f_j(k) \leq v(k)\}$ is cofinite, hence a member of $\mathcal{F}$, and thus $\pi(u) \leq \pi(f_j) \leq \pi(v)$. Therefore $\pi(u)$ and $\pi(v)$ are lower and upper bounds for S. This result can be found in Takeuchi [1984] and elsewhere.

Further remark. Contrasting the result above with 10.12.d, we see that $\mathbb{R}^\Lambda/\mathcal{F}$ is *not* isomorphic to the minimal hyperreal line $\mathbb{R}(x)$ discussed in 10.18.f. This observation is taken from Fleischer [1967].

10.21. (*Optional.*) Let $\mathcal{F}$ be a free ultrafilter on a set Λ, and define the ultrapower $^*\mathbb{R} = \mathbb{R}^\Lambda/\mathcal{F}$. Does it follow that $^*\mathbb{R}$ is non-Archimedean? It does under certain additional hypotheses, as in 10.19.b and 10.20.a, but in general the answer is not clear. However, in general (i.e., for any free ultrafilter $\mathcal{F}$) the following conditions are equivalent:

(A) $^*\mathbb{R}$ is Archimedean.

(B) $^*\mathbb{R} = \mathbb{R}$.

(C) $^*S = S$ for every set $S \subseteq \mathbb{R}$.

(D) $^*\mathbb{N} = \mathbb{N}$.

(E) Whenever $\mathcal{S} = \{S_1, S_2, S_3, \ldots\}$ is a countably infinite partition of Λ, then a unique member of $\mathcal{S}$ belongs to $\mathcal{F}$.

(F) Whenever $F_1, F_2, F_3, \ldots$ is a sequence in $\mathcal{F}$, then $\bigcap_{n=1}^{\infty} F_n \in \mathcal{F}$. In other words, the ideal of sets $\{\Lambda \setminus F : F \in \mathcal{F}\}$ is a σ-ideal.

Proof of equivalence. For (A) $\Rightarrow$ (B), observe that $\mathbb{R} \subseteq {}^*\mathbb{R}$ in any case, and an ordered field $\mathbb{F}$ is Archimedean if and only if $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$ (see 10.15.f). For (B) $\Rightarrow$ (C), use 9.45.d. Implication (C) $\Rightarrow$ (D) is obvious. For (D) $\Rightarrow$ (E), define a function $f : \Lambda \to \mathbb{N}$ by taking $f(\lambda) = n$ when $\lambda \in S_n$. Since f is equivalent to some constant k, the set S_k is a member of $\mathcal{F}$. For (E) $\Rightarrow$ (F), suppose that no member of $\mathcal{S}$ belongs to $\mathcal{F}$. Then all the sets $F_n = \Lambda \setminus S_n$ belong to $\mathcal{F}$. Hence their intersection belongs to $\mathcal{F}$ — but that intersection is empty, a contradiction. For (F) $\Rightarrow$ (A), suppose $^*\mathbb{R}$ is not Archimedean. Then $\mathbb{N}$ is bounded above by some $\zeta \in {}^*\mathbb{R}$. Then ζ is the equivalence class of some function $f : \Lambda \to \mathbb{R}$, such that the set $F_n = \{\lambda \in \Lambda : n \leq f(\lambda)\}$ is a member of $\mathcal{F}$ for each $n \in \mathbb{N}$. Then $\bigcap_{n=1}^{\infty} F_n$ is nonempty, a contradiction.

Further remarks. Do there exist any filters $\mathcal{F}$ satisfying those conditions (A)–(F) above? That is a famous problem in set theory. To discuss it we need a few definitions:

An ultrafilter $\mathcal{F}$ on a set Λ is said to be **α-complete** if, whenever $\mathcal{C} \subseteq \mathcal{F}$ with $\mathrm{card}(\mathcal{C}) \leq \alpha$, then $\bigcap_{C \in \mathcal{C}} C$ is a member of $\mathcal{F}$. A **measurable cardinal** is a cardinal α with this property: α is uncountable and there exists a free ultrafilter $\mathcal{F}$ on some set with cardinality α such that $\mathcal{F}$ is β-complete for every cardinality $\beta < \alpha$. A **ω-measurable cardinal** is a cardinal α with this property: α is uncountable and there exists a free ultrafilter $\mathcal{F}$ on some set Λ with cardinality α such that $\mathcal{F}$ is $\mathrm{card}(\mathbb{N})$-complete.

Clearly, a filter satisfying condition (F) above exists if and only if an ω-measurable cardinal exists; and it is shown by Bell and Slomson [1969] that a ω-measurable cardinal exists if and only if a measurable cardinal exists.

However, the existence or nonexistence of a measurable cardinal cannot be proved in conventional set theory. More precisely, the following results are known: (i) The consistency of ZF implies the consistency of ZF + AC + "there does not exist a measurable cardinal." (ii) The axiom system ZF + AC + "there exists a measurable cardinal" is empirically consistent, but its consistency is not implied by Con(ZF). These results can be found in Kunen [1980] and other books on formal logic.

Quadratic Extensions and the Complex Numbers

10.22. Let $\mathbb{F}$ be a field, and suppose q is an element of $\mathbb{F}$ that is not a square — i.e., assume $q \in \mathbb{F}$ and suppose there is no solution $x \in \mathbb{F}$ for the equation $x^2 = q$. (Examples are given in 8.20.c, 8.23.b, and 10.9.c.) Let $\mathbb{F}(\sqrt{q})$ represent the set $\mathbb{F} \times \mathbb{F}$, equipped with binary operations defined as follows:

$$\text{addition:} \qquad (a_1, b_1) + (a_2, b_2) \;=\; (a_1 + a_2,\ b_1 + b_2)$$

$$\text{multiplication:} \qquad (a_1, b_1)(a_2, b_2) \;=\; (a_1 a_2 + q b_1 b_2,\ a_1 b_2 + a_2 b_1)$$

where the expressions $a_1 a_2 + q b_1 b_2$, etc., are computed using the arithmetic rules of $\mathbb{F}$. Verify that $\mathbb{F}(\sqrt{q})$ is a field when equipped with these binary operations; the additive and multiplicative identities are $(0,0)$ and $(1,0)$ and

$$\text{the multiplicative inverse of } (a,b) \text{ is } \quad \left(\frac{a}{a^2 - qb^2}, \frac{-b}{a^2 - qb^2}\right)$$

when $(a,b) \neq (0,0)$. (This makes sense since $(a,b) \neq (0,0) \Rightarrow a^2 - qb^2 \neq 0$.)

Furthermore, the mapping $a \mapsto (a,0)$ is an injective homomorphism from $\mathbb{F}$ into $\mathbb{F}(\sqrt{q})$. Thus we may view $\mathbb{F}$ as a subset of $\mathbb{F}(\sqrt{q})$. If we write $(a,0)$ as a and $(0,b)$ as $b\sqrt{q}$, then (a,b) may be written as $a + b\sqrt{q}$, with all the usual rules of arithmetic being preserved. Then $\mathbb{F}(\sqrt{q})$ has additive identity 0 and multiplicative identity 1. We have extended our original field to a larger field in which the equation $x^2 = q$ does have a solution; the solution is the "number" $(0,1) = \sqrt{q}$.

Exercise. Show that the only other solution of $x^2 = q$ in $\mathbb{F}(\sqrt{q})$ is the "number" $(0,-1) = -\sqrt{q}$. (The beginner is urged to use the ordered pair notation (a,b) rather than the more familiar notation $a + b\sqrt{q}$, to reduce the likelihood of assuming something that has not already been proved.)

10.23. *Examples.*

a. $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is a subfield of $\mathbb{R}$.

b. Let m be an odd prime. As we noted in 8.20.c, there are elements $q \in \mathbb{Z}_m$ for which $x^2 = q$ has no solution in $\mathbb{Z}_m$. Fix any such q. Then $\mathbb{Z}_m(\sqrt{q})$ is a field containing exactly m^2 elements. *Exercise.* Construct addition and multiplication tables for the field with 9 elements.

10.24. The **complex numbers** are the quadratic extension field $\mathbb{R}(\sqrt{-1})$, formed from $\mathbb{R}$ by the construction of 10.22; this field is usually denoted by $\mathbb{C}$. The complex number $\sqrt{-1}$ is usually written as i. Thus complex numbers can be written as $x + iy$ or $x + yi$, where $x, y \in \mathbb{R}$. The complex number $\alpha = x + iy$ has **real part**, **imaginary part**, and **complex conjugate** defined by

$$\operatorname{Re}\alpha = x, \qquad \operatorname{Im}\alpha = y, \qquad \overline{\alpha} = x - iy.$$

(*Caution*: Some mathematicians use overlines for other purposes than complex conjugation — e.g., set complementation or topological closure.)

10.25. The spaces $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic, when considered as real vector spaces: They yield the same results for addition and for multiplication by a real number.

Any complex-valued function of a complex variable can be rewritten as a $\mathbb{R}^2$-valued function of a variable in $\mathbb{R}^2$. If $w = f(z)$, we may write $z = x + iy$ and $w = u + iv = f(x + iy)$, where u, v, x, y are real. The letters f, u, v, w, x, y, z are customarily used in the literature in precisely this arrangement.

Example. If $w = f(z) = z^3$, then

$$u + iv \quad = \quad (x + iy)^3 \quad = \quad x^3 + 3ix^2y - 3xy^2 - iy^3,$$

so $u(x, y) = x^3 - 3xy^2$ and $v(x, y) = 3x^2y - y^3$.

10.26. *Exercise.* The 2-by-2 matrices of the form $\begin{bmatrix} x & y \\ -y & x \end{bmatrix}$ with $x, y \in \mathbb{R}$, equipped with matrix addition and multiplication, form a field that is isomorphic to $\mathbb{C}$; the matrix above corresponds to the complex number $x + iy$.

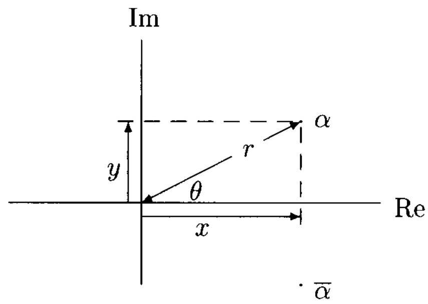
More generally, let $\mathbb{F}$ be a field, and let q be an element of $\mathbb{F}$ that has no square root in $\mathbb{F}$. Then the 2-by-2 matrices of the form $\begin{bmatrix} x & y \\ qy & x \end{bmatrix}$ with $x, y \in \mathbb{F}$, equipped with matrix addition and multiplication, form a field that is isomorphic to the quadratic extension field $\mathbb{F}(\sqrt{q})$, with the matrix above corresponding to $x + y\sqrt{q}$.

10.27. For many purposes it is convenient to represent complex numbers as points in the plane, with the real part being the distance to the right of the origin, and the imaginary part being the distance up from the origin; then $\overline{\alpha}$ is the reflection of α in the horizontal coordinate axis. This representation is sometimes known as an **Argand diagram**. See the following illustration. This illustration also shows the polar coordinate representation: If r is the distance from 0 to α and θ is the angle from the positive real axis to the line between 0 and α, then $\alpha = r\cos\theta + ir\sin\theta$. The real number θ is sometimes called the *argument* of the complex number α.

Historical remarks. Calculations with complex numbers were performed long before such numbers were properly understood or fully accepted. For instance, Cardan showed that the quadratic equation $x(10 - x) = 40$ has the two solutions $5 + \sqrt{-15}$ and $5 - \sqrt{-15}$, without any clear understanding of what such numbers could mean. Euler, around 1750, wrote that such computations are a method for showing that the equation $x(10 - x) = 40$ has *no*

Argand diagram

$$\alpha = x + iy$$
$$= r(\cos\theta + i\sin\theta)$$



solutions. Some mathematicians took this attitude: Of course there is no square root of a negative number; but if there *were* such a thing, what algebraic properties should it have? Around 1800, writings of Argand and Gauss gave our present geometrical interpretation of complex numbers as points in the plane. Finally, in 1830, William R. Hamilton published a paper explaining complex numbers in terms of ordered pairs; probably that is the simplest starting point for mathematicians learning about complex numbers today. Some accounts of the history of this subject are given by Kline [1990] and Tietze [1965].

10.28. The rules for addition and multiplication of complex numbers are the same as the rules given in 10.22, with q taken to be -1. The rule for addition is fairly simple; it is the same as the addition of vectors in $\mathbb{R}^2$ (see Chapter 11). The rule for multiplication,

$$(a_1 + ib_1)(a_2 + ib_2) \quad = \quad (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1), \qquad (*)$$

is more complicated, and may seem rather arbitrary to the beginner. However, it becomes much more natural when interpreted geometrically with polar coordinates. Let $\text{cis}(\theta)$ denote $\cos(\theta) + i\sin(\theta)$. Using $(*)$ and some basic trigonometric identities, verify that the product of the complex numbers $r_1\text{cis}(\theta_1)$ and $r_2\text{cis}(\theta_2)$ is the complex number $(r_1r_2)\text{cis}(\theta_1 + \theta_2)$. Thus, to multiply two complex numbers, we multiply the radii and add the angles. Arithmetic with complex numbers may be viewed as transformations of the plane: Addition is a translation, while multiplication is a rotation and stretching.

It follows that

$$[r\text{cis}(\theta)]^n \quad = \quad r^n\text{cis}(n\theta)$$

for integers n. This is known as **De Moivre's formula.**

10.29. Reversing the process described above, we find that the nth roots of any complex number $r\text{cis}(\theta)$ are

$$r^{1/n}\text{cis}\left[\frac{\theta + 2\pi j}{n}\right] \qquad (j = 0, 1, 2, \ldots, n-1).$$

These are points equally spaced along a circle centered at 0 with radius $r^{1/n}$.

The reader is cautioned that familiar properties of $\sqrt{x}$ or $\sqrt[n]{x}$ for positive real numbers x do not always extend to complex numbers x. Indeed, the notations $\sqrt{x}$ and $\sqrt[n]{x}$ are too ambiguous and imprecise for some computations with complex numbers. For instance, $\sqrt{p}\sqrt{q} = \sqrt{pq}$ is valid for $p, q > 0$, but the computation $-1 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1$ is clearly incorrect. However, the reader who makes an error of this sort is in good company: Euler made some similar mistakes, in the years before complex numbers were well understood.

10.30. *Polynomial equations.* The discussion in 10.29 shows that for any complex number α other than 0, we have n distinct solutions z to the polynomial equation $z^n - \alpha = 0$. What about other polynomial equations? In 17.36 we shall prove that every nonconstant polynomial (in a single complex variable, with complex coefficients) has at least one complex root; in fact, counting multiplicities of multiple roots, every polynomial of degree n has exactly n complex roots. Power series and analytic functions, which are like "polynomials of infinite degree," have a more complicated theory, which will be considered briefly in 22.23 and 25.27.

We can actually give *formulas* for the roots of the simplest polynomials:

a. The quadratic equation $az^2 + bz + c = 0$ (with $a \neq 0$) has solutions given by the **quadratic formula**, $z = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. Every student learns this in high school — at least, when a, b, c are real numbers and $b^2 - 4ac \geq 0$. Actually, we can apply this formula with any complex numbers a, b, c (with $a \neq 0$), since 10.29 tells us how to find square roots of $b^2 - 4ac$. The quadratic formula yields two distinct complex solutions z, or one solution repeated if $b^2 - 4ac = 0$.

b. An analogous formula, involving square roots and cube roots, can be given for the **cubic** equation $az^3 + bz^2 + cz + d = 0$ (with $a \neq 0$), but it is a bit more complicated. It was published by Cardan in 1545. Divide through by a; thus we may assume $a = 1$. Substitute $z = w - \dfrac{b}{3}$. Some cancellation occurs. The resulting equation can be rewritten in the more convenient form $w^3 + 3ew - 2f = 0$, with known constants

$$e = \frac{1}{3}c - \frac{1}{9}b^2, \qquad f = -\frac{1}{27}b^3 + \frac{1}{6}bc - \frac{1}{2}d.$$

Now make another substitution, taking $w = \zeta - e\zeta^{-1}$. Again some cancellation occurs. The resulting equation $\zeta^3 - e^3\zeta^{-3} - 2f = 0$ can be rewritten as $\zeta^6 - 2f\zeta^3 - e^3 = 0$. This is a quadratic in ζ^3. Solve it as in 10.30.a; thus $\zeta^3 = f \pm \sqrt{f^2 + e^3}$. Then find ξ, as in 10.29. This leads, finally, to

$$z = \sqrt[3]{f + \sqrt{f^2 + e^3}} + \sqrt[3]{f - \sqrt{f^2 + e^3}} - \frac{b}{3}.$$

This looks like six values, since each complex number other than 0 has two square roots and three cube roots, but there is some repetition and we only end up with three distinct solutions. This complicated procedure is seldom used in applications; numerical approximation methods do not depend on the use of these formulas.

c. A still more complicated formula or method yields the solution of the **quartic** equation $z^4+az^3+bz^2+cz+d=0$. This problem was solved by Cardan's student and published in Cardan's book in 1545. Here is one description of the method: By completing the square, we may rewrite the given equation as

$$\left(z^2+\frac{1}{2}az\right)^2 \quad = \quad \left(\frac{1}{4}a^2-b\right)z^2-cz-d.$$

For any constant r (to be specified), we have

$$\left(z^2+\frac{1}{2}az+r\right)^2 \quad = \quad \left(\frac{1}{4}a^2-b+2r\right)z^2 \;+\; (ar-c)z \;+\; (r^2-d),$$

which we shall rewrite as $\left(z^2+\frac{1}{2}az+r\right)^2=Az^2+Bz+C$.

Now, an expression of the form Az^2+Bz+C, with constants A,B,C is a perfect square — i.e., an expression of the form $A(z-D)^2$ — if the constants A,B,C satisfy $B^2=4AC$, and in that case we have $D=-\dfrac{C}{2B}$. If we can choose a constant r to satisfy this condition, then we will have

$$\left(z^2+\frac{1}{2}az+r\right)^2=A(z-D)^2,$$

which can be rewritten as $z^2+\frac{1}{2}az+r \;=\; \pm\sqrt{A}\,(z-D)$ — that is, two quadratic equations, which we can solve for z as in 10.30.a.

It remains only to find a value of r that satisfies $B^2=4AC$. This equation, written in more detail, is $(ar-c)^2=4\left(\frac{1}{4}a^2-b+2r\right)(r^2-d)$ — a third-degree equation for r, which we can solve as in 10.30.b.

d. Some **quintic** (or fifth-degree) equations have solutions that can be expressed in terms of fifth roots. For instance, one of the solutions of $x^5+20x+32=0$ is the number

$$\begin{aligned}
&-\frac{1}{5}\sqrt[5]{2500\sqrt{5}+250\sqrt{50-10\sqrt{5}}-750\sqrt{50+10\sqrt{5}}}\\
&+\frac{1}{5}\sqrt[5]{2500\sqrt{5}-250\sqrt{50+10\sqrt{5}}-750\sqrt{50-10\sqrt{5}}}\\
&+\frac{1}{5}\sqrt[5]{2500\sqrt{5}+250\sqrt{50+10\sqrt{5}}+750\sqrt{50-10\sqrt{5}}}\\
&-\frac{1}{5}\sqrt[5]{2500\sqrt{5}-250\sqrt{50-10\sqrt{5}}+750\sqrt{50+10\sqrt{5}}}
\end{aligned}$$

(taken from [Wolfram 1994]). Examples like this could lead one to expect that the general fifth-degree equation, like the equations of lower degree, could be solved by a formula in terms of radicals. Mathematicians sought such a formula for many years. But in 1826 Abel proved that such a formula is impossible, and a few years later Galois developed a theory that describes exactly when a polynomial is solvable by radicals. For instance, the polynomial $2x^5 - 2x + 1$ has Galois group S_5, which is not solvable, so the roots of $2x^5 - 2x + 1$ cannot be represented in terms of radicals.

Still, we know by the Fundamental Theorem of Algebra (17.36) that every fifth-degree polynomial equation with complex coefficients has five complex roots. If they cannot be represented using radicals, how can the roots be represented? Radicals are not enough; more functions are needed. In 1844 Eisenstein solved quintic equations in terms of radicals and what we shall call the *Eisenstein function* in the paragraph below. In 1858 Hermite, Kronecker, and Brioschi solved quintic equations in terms of radicals and elliptic modular functions; in 1877 Klein solved quintic equations in terms of radicals and the hypergeometric function. For additional information about some of these solutions, see Wolfram [1994] or Shurman [1995].

For most choices of a, the quintic equation $x^5 + x = a$ cannot be solved by radicals. However, some basic properties of the polynomial $p(x) = x^5 + x$ are easy to figure out: That polynomial has derivative $p'(x) = 5x^4 + 1 \geq 1$, and so p is strictly increasing and gives a bijection from $\mathbb{R}$ onto $\mathbb{R}$. Let the inverse of that function p be denoted by $Eis(x)$; we shall call it the **Eisenstein function**. Then it can be shown that a solution of $x^5 + ax^4 + bx^3 + cx^2 + dx + e = 0$ is given by $x = Eis\Big(q(a,b,c,d,e)\Big)$, where $q(\cdot)$ is a function of five variables that *can* be expressed entirely in terms of radicals (i.e., in terms of nth roots for $n \leq 5$, together with sums, differences, products, and quotients). The function q can be expressed in closed form, but the formula is extremely long, and we shall not give it here. That formula can be produced by methods described by Stillwell [1995].

e. Nowadays, when one wants to solve a polynomial equation of degree higher than two, generally one uses a numerical iterative scheme on an electronic computer. For instance, one such scheme is Newton's Method, which can be found in every modern textbook on calculus. These numerical schemes do not yield exact solutions, but they yield solutions to as much accuracy as one wishes; 10 decimal places of accuracy is more accuracy than any engineering problem will ever require. However, roots produced by a numerical scheme may have no apparent rhyme or reason; they may seem to be arranged entirely at random. The formulas of Cardan, Hermite, Eisenstein, et al. are of interest because they reveal the *pattern* of the roots — i.e., the relationships between the roots and the other numbers present in the problem. That is important for theoretical purposes and ultimately has some effect on engineering problems as well.

ABSOLUTE VALUES

10.31. *Definitions.* Let X be a field (not necessarily ordered). By an **absolute value** on

X we mean a mapping $|\cdot| : X \to [0, +\infty)$ satisfying

$$|x| = 0 \iff x = 0 \qquad \text{(positive-definiteness)}$$

$$|xy| = |x|\,|y| \qquad \text{(multiplicativeness)}$$

$$|x + y| \leq |x| + |y| \qquad \text{(subadditivity)}$$

for all $x, y \in X$. These properties imply also $|1| = 1$ (*exercise*). An absolute value is also known as a *modulus* or *magnitude* or *value* or *valuation.*

The absolute value thus defined on fields should not be confused with the absolute value defined in 8.39 for lattice groups. Fortunately, the two notions do coincide in the case where the field or lattice group is $\mathbb{R}$ (or any subfield of $\mathbb{R}$); in that case $|x|$ is just $\max\{x, -x\}$, the **usual absolute value on** $\mathbb{R}$. It will always be used on $\mathbb{R}$, except when some other arrangement is specified.

10.32. *The absolute value of a complex number.* For real numbers x and y, define $|x + iy| = \sqrt{x^2 + y^2}$. Show that

a. $|\alpha| = \sqrt{(\operatorname{Re}\alpha)^2 + (\operatorname{Im}\alpha)^2}$ for any complex number α.

b. If $\alpha = r \operatorname{cis}\theta$ for some real number θ and positive number r (as in 10.28), then $|\alpha| = r$. Hence $|\alpha\beta| = |\alpha||\beta|$, by our observations in 10.28.

c. $|\alpha| = |\overline{\alpha}|$ and $\alpha\overline{\alpha} = |\alpha|^2$.

d. $|\alpha + \beta|^2 = |\alpha|^2 + 2\operatorname{Re}\alpha\overline{\beta} + |\beta|^2 \leq |\alpha|^2 + 2|\alpha\overline{\beta}| + |\beta|^2 = (|\alpha| + |\beta|)^2$.

e. $|\ |$ is an absolute value.

It is the **usual absolute value on** $\mathbb{C}$. It will always be used on $\mathbb{C}$, except if some other arrangement is specified. The usual topology and uniform structure on $\mathbb{C}$ are given by the metric $d(\alpha, \beta) = |\alpha - \beta|$. *Remark.* This topology and uniform structure are the same as those of $\mathbb{R}^2$; see 18.18 and 22.11. However, $\mathbb{C}$ and $\mathbb{R}^2$ have different differentiable structures; see 25.8.

f. The set $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ is a commutative group whose operation is the multiplication of complex numbers. It is often called the **circle group**, since it is geometrically a circle, $x^2 + y^2 = 1$, in the complex plane. It is isomorphic (as a group) to the additive group introduced in 8.10.e. In fact, the mapping $\theta \mapsto \operatorname{cis}(\theta)$ (defined in 10.28) is an isomorphism from the additive group $[0, 2\pi)$ onto the multiplicative group $\mathbb{T}$. Also, the mapping $\theta \mapsto \operatorname{cis}(\theta)$ is a group homomorphism (not an isomorphism) from the additive group $\mathbb{R}$ onto the multiplicative group $\mathbb{T}$.

10.33. Let $\mathbb{F}$ be a field, not necessarily contained in $\mathbb{R}$ or containing $\mathbb{R}$. (For instance, $\mathbb{F}$ could be one of the finite fields discussed in 8.20.) For clarity in the discussion below, let 1 denote the multiplicative identity of $\mathbb{R}$, and let e denote the multiplicative identity of $\mathbb{F}$. For any $n \in \mathbb{N}$, let us denote $ne = e + e + \cdots + e$ (the sum of n e's). Let $|\ \ |$ be an absolute value on the field $\mathbb{F}$. Then the following conditions are equivalent; if any (hence all) are satisfied we say that the absolute value is **non-Archimedean**.

(A) The set $\{|ne| : n \in \mathbb{N}\}$ is bounded in $\mathbb{R}$.

(B) $|ne| \le 1$ for every $n \in \mathbb{N}$.

(C) $|a + b| \le \max\{|a|, |b|\}$ for all $a, b \in \mathbb{F}$.

(D) The metric $d(u, v) = |u - v|$ satisfies the ultrametric inequality $d(u, v) \le \max\{d(u, w), d(v, w)\}$.

Proof of equivalence. The proofs of (C) $\iff$ (D) and (C) $\Rightarrow$ (B) $\Rightarrow$ (A) are easy; it suffices to prove (A) $\Rightarrow$ (C). Let $r = \sup\{|ne| : n \in \mathbb{N}\}$. Let any $a, b \in \mathbb{F}$ be given; let $s = \max\{|a|, |b|\}$. For each $n \in \mathbb{N}$, observe that

$$|a+b|^n \;=\; |(a+b)^n| \;=\; \left|\sum_{j=0}^{n} \binom{n}{j} a^{n-j} b^j\right| \;\le\; \sum_{j=0}^{n} r s^n \;=\; (n+1) r s^n.$$

Hence $|a + b| \le \sqrt[n]{(n+1)r}\, s$. Take limits as $n \to \infty$ to obtain $|a + b| \le s$; this completes the proof. This proof follows Rooij [1978].

Examples of non-Archimedean valuations.

a. The **discrete absolute value** (or *Kronecker absolute value*) on any field $\mathbb{F}$ is given by

$$|x| \;=\; \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x \neq 0. \end{cases}$$

Obviously this satisfies condition (C) given above.

b. Let p be a prime number — i.e., one of the numbers 2, 3, 5, 7, 11, ... Any nonzero rational number can be expressed in the form $m/(np^r)$, where r, m, n are integers, with m and n nonzero and not divisible by p. Define $|m/(np^r)|_p = p^r$; this is the **p-adic absolute value** on $\mathbb{Q}$. Completions of metric spaces will be studied in Chapter 19; the completion of the metric space $(\mathbb{Q}, |\ \ |_p)$ is the system of *p-adic numbers*, which are used in algebraic number theory and in the study of topological groups. An introduction to this subject is given by Bachman [1964].

c. Let $\mathbb{F}$ be a field, let x be a variable, and let $\mathbb{F}(x)$ be the field of rational functions in the variable x with coefficients in $\mathbb{F}$ (see 9.28). Each nonzero element can be written in the form $r(x) = x^m p(x)/q(x)$ where m is an integer, $p(x)$ and $q(x)$ are polynomials, and neither p nor q has a factor of x. Then let $|r| = 2^{-m}$. Verify that this yields a non-Archimedean valuation on $\mathbb{F}(x)$. Instead of 2^{-m} we could use c^{-m} for any constant $c > 1$. For further reading, see also Narici and Beckenstein [1990].

10.34. *Intentional ambiguity.* In later chapters, the main fields we shall use are $\mathbb{R}$ and $\mathbb{C}$. We shall sometimes state a theorem involving a field $\mathbb{F}$ that may be $\mathbb{R}$ or $\mathbb{C}$, without specifying which of these fields is intended. This intentional ambiguity permits us to cover both cases simultaneously. For any $\alpha \in \mathbb{F}$, we shall let $\operatorname{Re}\alpha$, $\operatorname{Im}\alpha$, $\overline{\alpha}$, and $|\alpha|$ be the real and imaginary parts of α, the complex conjugate of α, and the absolute value of α, respectively. This notation is applicable (albeit unnecessarily complicated) even when $\mathbb{F} = \mathbb{R}$; in that case we have $\operatorname{Re}\alpha = \overline{\alpha} = \alpha$ and $\operatorname{Im}\alpha = 0$. Likewise, the expression $|\alpha|$ has the same value regardless of whether the field being used is $\mathbb{R}$ or $\mathbb{C}$, since the absolute value function on $\mathbb{R}$ is just the restriction of the absolute value function on $\mathbb{C}$.

See also the related discussions in 11.1, 26.5, 26.21, and 27.30.

10.35. Clarkson's inequality for scalars. Let $p, q \in (1, +\infty)$ with $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Then for any complex numbers ξ, η,

$$|\xi + \eta|^p + |\xi - \eta|^p \quad \leq \quad 2\left(|\xi|^q + |\eta|^q\right)^{p/q} \qquad \text{if} \quad p \geq 2 \geq q,$$

and the reverse of this inequality holds if $p \leq 2 \leq q$. (This result will be used in Chapter 22 to prove that the Banach spaces $L^p(\mu)$ are uniformly convex; other properties of uniform convexity are studied in that chapter and in Chapter 28.)

Proof. Our presentation is based on that of Weir [1974]. Define

$$\varphi(t) \quad = \quad \left(1 + \sqrt[p]{t}\right)^p + \left(1 - \sqrt[p]{t}\right)^p \qquad (0 < t < 1).$$

Then compute the first two derivatives and simplify to:

$$\begin{aligned} \varphi'(t) \quad &= \quad \left[\left(1 + \sqrt[p]{t}\right)^{p-1} - \left(1 - \sqrt[p]{t}\right)^{p-1}\right] t^{(1/p)\,-1} \\ \varphi''(t) \quad &= \quad \left(\frac{1}{p} - 1\right)\left[\left(1 + \sqrt[p]{t}\right)^{p-2} - \left(1 - \sqrt[p]{t}\right)^{p-2}\right] t^{(1/p)\,-2}. \end{aligned}$$

For most of the remaining steps, we shall give inequalities only for $p \geq 2 \geq q$; the reversed inequalities are then valid when $p \leq 2 \leq q$. Observe that $\varphi''(t) \leq 0$ (assuming $p \geq 2 \geq q$). Hence, by integrating,

$$\varphi(a) \quad \leq \quad \varphi(b) + (a - b)\varphi'(b) \qquad \text{for } a, b \in (0, 1).$$

Substituting $a = r^p$ and $b = r^{pq}$ and simplifying, we obtain

$$(1 + r)^p + (1 - r)^p \quad \leq \quad 2\left(1 + r^q\right)^{p/q}$$

for $r \in (0, 1)$. Taking limits, we find that this is also valid for $r \in [0, 1]$.

(This paragraph can be omitted if we wish to consider only real numbers for scalars.) Next we claim that if ζ is any complex number with $|\zeta| \leq 1$, then

$$|1 + \zeta|^p + |1 - \zeta|^p \quad \leq \quad 2\Big|1 + |\zeta|^q\Big|^{p/q}$$

— again, with inequality reversed if $p \leq 2 \leq q$. To establish this inequality, note that ζ can be represented in the form $\zeta = r\cos\theta + i\,r\sin\theta$ for some real number θ, with $r = |\zeta| \in [0, 1]$. Holding r fixed, define $\psi(\theta) = |1 + \zeta|^p + |1 - \zeta|^p$; it suffices to show that $\psi(\theta) \leq \psi(0)$ for all θ. Note that ψ is periodic with period π, since ζ and $-\zeta$ switch places when we increase θ by π; hence it suffices to consider $\theta \in [0, \pi]$. Observe that

$$\begin{aligned} |1 + \zeta|^2 \quad &= \quad (1 + \zeta)\overline{(1 + \zeta)} \quad = \quad 1 + 2r\cos\theta + r^2 \qquad \text{and} \\ |1 - \zeta|^2 \quad &= \quad (1 - \zeta)\overline{(1 - \zeta)} \quad = \quad 1 - 2r\cos\theta + r^2. \end{aligned}$$

This yields the representation

$$\psi(\theta) \quad = \quad \left(1 + 2r\cos\theta + r^2\right)^{p/2} \ + \ \left(1 - 2r\cos\theta + r^2\right)^{p/2},$$

and it is then easy to compute

$$\psi'(\theta) \quad = \quad pr\Big\{ - |1+\zeta|^{p-2} + |1-\zeta|^{p-2}\Big\}\sin\theta.$$

In the interval $0 \le \theta \le \pi$ we have $\sin\theta \ge 0$. Also, we have

$$\begin{array}{lllll} |1+\zeta| & \geq & |1-\zeta| & \text{when} \quad 0 \le \theta \le \pi/2, & \text{and} \\ |1+\zeta| & \leq & |1-\zeta| & \text{when} \quad \pi/2 \le \theta \le \pi. & \end{array}$$

Hence ψ assumes a maximum at $\theta = 0$ and $\theta = \pi$ (or a minimum, if $p \le 2 \le q$). This completes the proof of the claim.

Finally, to prove the theorem, let any complex numbers ξ, η be given. We may assume that at least one of ξ, η is nonzero; without loss of generality we may assume $|\xi| \le |\eta|$. Then we substitute $\zeta = \xi/\eta$.

Convergence of Sequences and Series

10.36. *Remarks.* Both $\mathbb{R}$ and $[-\infty, +\infty]$ are chains; hence their natural convergences are defined as in 7.41. The extended real line $[-\infty, +\infty]$ has the further advantage that it is order complete; hence its convergence can also be described as in 7.45.

It is sometimes easier to work in $[-\infty, +\infty]$, since any net in that space has a limsup and a liminf. Questions about convergence in $\mathbb{R}$ can be restated as questions about convergence in $[-\infty, +\infty]$, since the convergence in $\mathbb{R}$ is just the restriction of the convergence in $[-\infty, +\infty]$.

On the other hand, it is sometimes easier to work in $\mathbb{R}$, because that space has a simpler metric and simpler arithmetic. Questions about convergence in $[-\infty, +\infty]$ can be restated as questions about convergence in a bounded subset of $\mathbb{R}$, via the following observation: The mapping $\theta \mapsto \tan\theta$ is an order isomorphism from $\left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$ onto $[-\infty, +\infty]$.

Any net in $[-\infty, +\infty]$ has both a lim inf and a lim sup in $[-\infty, +\infty]$. Any *bounded* net in $\mathbb{R}$ has both a lim inf and a lim sup in $\mathbb{R}$. Not every bounded net in $\mathbb{R}$ has a limit; for instance, the sequence $0, 1, 0, 1, 0, 1, \ldots$ has no limit. (However, every bounded net in $\mathbb{R}$ has *generalized* limits, in a sense described in 12.33.)

10.37. (*Optional.*) Notions of calculus, such as limits, can be described in terms of infinitesimals; Newton and Leibniz had something like this in mind when they invented calculus. (The epsilon-delta approach now widely used in calculus books was not developed until many decades after Newton and Leibniz.)

Assume that $\mathcal{F}$ is a free ultrafilter on $\mathbb{N}$; hence ${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}}/\mathcal{F}$ is a non-Archimedean field, as in 10.20.a. Let $f : \mathbb{R} \to \mathbb{R}$ be some function; let p, L be some real numbers. Show that the right-hand limit $f(p+) = \lim_{a\downarrow 0} f(p+a)$ exists and equals L, if and only if for each positive infinitesimal α the hyperreal number ${}^*f(p+\alpha)$ is infinitely close to L.

(The number a is understood to vary through *real* values, not hyperreal values.) Applying that result twice, we obtain:

$\lim_{a\to 0} f(p+a)$ exists and equals L if and only if for each nonzero infinitesimal α the hyperreal number ${}^*f(p+\alpha)$ is infinitely close to L.

Proof. It suffices to prove the first equivalence. First suppose that $f(p+) = L$. Let any positive real number ε be given. Then there is some positive real number δ such that

$$a \in \mathbb{R}, \quad 0 < a < \delta \qquad \Rightarrow \qquad L - \varepsilon < f(p+a) < L + \varepsilon.$$

It follows from 9.50.c and 9.52.d that

$$\alpha \in {}^*\mathbb{R}, \quad 0 < \alpha < \delta \qquad \Rightarrow \qquad L - \varepsilon < {}^*f(p+\alpha) < L + \varepsilon.$$

In particular, if α is a positive infinitesimal, then $0 < \alpha < \delta$ is satisfied for *every* positive real number δ; hence $-\varepsilon < {}^*f(p+\alpha) - L < \varepsilon$ is satisfied for *every* positive real number ε; hence ${}^*f(p+\alpha)$ is infinitely close to L.

Conversely, suppose that ${}^*f(p+\alpha)$ is infinitely close to L for every positive infinitesimal α. We wish to show that $f(p+) = L$. Let (a_n) be any sequence of positive real numbers decreasing to 0; it suffices to show that $f(p+a_n) \to L$. Let any real number $\varepsilon > 0$ be given; it suffices to show that $|f(p+a_n) - L| < \varepsilon$ for all n sufficiently large. Define a function $A : \mathbb{N} \to \mathbb{R}$ by taking $A(n) = a_n$; let α be the equivalence class of that function; then α is a positive infinitesimal. Since ${}^*f(p+\alpha)$ is infinitely close to L, we have in particular $L - \varepsilon < {}^*f(p+\alpha) < L + \varepsilon$. That is, $L - \varepsilon < f(p + A(n)) < L + \varepsilon$ for all but finitely many values of n. This completes the proof.

10.38. (This result can be postponed; it will not be needed until 28.37.)

Let X be a nonempty set. A sequence (f_j) in $\mathbb{R}^X$ is a **Pryce sequence** if

$$\sup_{x\in X} \liminf_{j\to\infty} f_j(x) \quad = \quad \sup_{x\in X} \limsup_{j\to\infty} f_j(x).$$

(The liminf and limsup take their values in $[-\infty, +\infty]$.) It is easy to show that any subsequence of a Pryce sequence is also a Pryce sequence.

Pryce Selection Theorem. Every sequence in $\mathbb{R}^X$ has a subsequence that is a Pryce sequence.

Proof. We follow the presentation of König [1986]. Throughout this proof, both subscripts and superscripts will be used as indices; superscripts will not denote exponentiation or composition. Also, for brevity, for any function $u : X \to [-\infty, +\infty]$,

$$\overline{u} \qquad \text{will denote the number} \qquad \sup_{x\in X} u(x).$$

Let the given sequence be (f_j). For $j = 1, 2, 3, \ldots$, let $g_j^0 = f_j$. For each $n = 1, 2, 3, \ldots$, recursively define a point $x^n \in X$ and a subsequence $(g_j^n : j \in \mathbb{N})$ of the sequence $(g_j^{n-1} : j \in \mathbb{N})$ as follows: Let

$$p^{n-1}(x) \;=\; \liminf_{j\to\infty}\, g_j^{n-1}(x), \qquad\qquad q^{n-1}(x) \;=\; \limsup_{j\to\infty}\, g_j^{n-1}(x)$$

for each $x \in X$, with values in $[-\infty, +\infty]$. Then choose $x^n \in X$ according to the value of $\overline{q^{n-1}}$, as follows:

If $\overline{q^{n-1}} = -\infty$, let x^n be any point of X.

If $\overline{q^{n-1}} \in \mathbb{R}$, choose some $x^n \in X$ satisfying $q^{n-1}(x^n) > \overline{q^{n-1}} - \frac{1}{n}$.

If $\overline{q^{n-1}} = +\infty$, choose some $x^n \in X$ satisfying $q^{n-1}(x^n) > n$.

Let $\tau^n = q^{n-1}(x^n)$. Since $\tau^n = \limsup_{j\to\infty} g_j^{n-1}(x^n)$, some subsequence $(g_j^n : j \in \mathbb{N})$ of the sequence $(g_j^{n-1} : j \in \mathbb{N})$ satisfies $\tau^n = \lim_{j\to\infty} g_j^n(x^n)$. This completes the recursive definition. Since the sequence $(g_j^n(x^n) : j \in \mathbb{N})$ is convergent, we have $\tau^n = p^n(x^n) = q^n(x^n)$.

Now let (h^n) be a diagonal subsequence of the g's — that is, let h^n be a member of the sequence $(g_j^n : j \in \mathbb{N})$, and then let h^{n+1} be some later member of the sequence $(g_j^n : j \in \mathbb{N})$, chosen so that h^{n+1} also belongs to the subsequence $(g_j^{n+1} : j \in \mathbb{N})$. For each n, it follows that $(h^n, h^{n+1}, h^{n+2}, h^{n+3}, \ldots)$ is a subsequence of $(g_j^n : j \in \mathbb{N})$, and therefore $\tau_n = \lim_{j\to\infty} h^j(x^n)$. In particular, (h^n) is a subsequence of the originally given sequence (f_j); we shall show that (h^n) is a Pryce sequence. Define

$$p(x) \;=\; \liminf_{j\to\infty}\, h^j(x), \qquad\qquad q(x) \;=\; \limsup_{j\to\infty}\, h^j(x);$$

then $p \leq q$ and it suffices to show that $\overline{p} \geq \overline{q}$. We may assume that $\overline{q} > -\infty$ and $\overline{p} < +\infty$.

Since (h^j) is a subsequence of $(g_j^n : j \in \mathbb{N})$, which is in turn a subsequence of $(g_j^{n-1} : j \in \mathbb{N})$, we have

$$p_0 \;\leq\; p_1 \;\leq\; p_2 \;\leq\; \cdots \;\leq\; p \;\leq\; q \;\leq\; \cdots \;\leq\; q_2 \;\leq\; q_1 \;\leq\; q_0.$$

Hence the numbers $\overline{q^{n-1}}$ are bounded below by the number $\overline{q}$, which is not $-\infty$. Also,

$$q^{n-1}(x^n) \quad=\quad \tau^n \quad=\quad p^n(x^n) \quad\leq\quad \overline{p^n} \quad\leq\quad \overline{p} \quad<\quad +\infty.$$

Since $\overline{p} < +\infty$, we have $q^{n-1}(x^n) \leq n$ for all n sufficiently large. For those n, our definition of x^n tells us that $\overline{q^{n-1}}$ cannot be $+\infty$. Thus for all n sufficiently large we have $\overline{q^{n-1}}$ finite, and therefore (by our definition of x^n)

$$-\infty \;<\; \overline{q} \;\leq\; \overline{q^{n-1}} \;<\; \frac{1}{n} + q^{n-1}(x^n) \;\leq\; \frac{1}{n} + \overline{p} \;<\; +\infty.$$

Taking limits yields $\overline{q} \leq \overline{p}$.

10.39. *Convergence of infinite series.* Let $a_1, a_2, a_3, \ldots$ be complex numbers (or, in particular, real numbers). Then the expression

$$\sum_{k=1}^{\infty} a_k \qquad \text{or} \qquad a_1 + a_2 + a_3 + \cdots$$

is called a **series** (or an *infinite series*). That expression also represents the limit of the sequence

$$a_1, \qquad a_1 + a_2, \qquad a_1 + a_2 + a_3, \qquad a_1 + a_2 + a_3 + a_4, \qquad \ldots$$

— that is, the value of $\lim_{n\to\infty} \sum_{k=1}^{n} a_k$ — if that limit exists. (We say the series is the "limit of the partial sums.") When the limit exists, we say the series $\sum_{k=1}^{\infty} a_k$ **converges**. When the limit fails to exist, we say the series **diverges**.

The definitions above all generalize readily to the case where $a_1, a_2, a_3, \ldots$ are members of any monoid equipped with a Hausdorff convergence structure (see 7.36). In 22.20 we consider the case where X is any Banach space; in Chapter 26 we consider the case where X is a topological vector space.

For infinite series of real numbers, it is customary to extend the definition a little further: When $a_1, a_2, a_3, \ldots$ are real numbers, then $\sum_{k=1}^{\infty} a_k$ is understood to mean the limit of $a_1 + a_2 + \cdots + a_n$, not just in $\mathbb{R}$, but in the extended real line $[-\infty, +\infty]$. When the limit happens to be $+\infty$, some mathematicians say that the series **diverges to infinity**. (Some mathematicians say that the series **converges to infinity**, but we shall not follow that terminology in our discussion of infinite series.) Similar terminology applies for $-\infty$.

The *sequence* $(a_k) = (a_1, a_2, a_3, \ldots)$ should not be confused with the *series* $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots$. For instance, the sequence $(2^{-k}) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots)$ converges to 0, while the series $\sum_{k=1}^{\infty} 2^{-k} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$ converges to 1 by the result in 10.41.d.

10.40. When all the a_k's are nonnegative real numbers, then $\sum_{k=1}^{\infty} a_k$ always exists in $[0, +\infty]$ — that is, the series always converges to a finite number or diverges to $+\infty$. We may abbreviate these two cases by saying simply that $\sum_{k=1}^{\infty} a_k < \infty$ or that $\sum_{k=1}^{\infty} a_k = \infty$.

More generally, let $(a_\lambda : \lambda \in \Lambda)$ be any parametrized collection of members of $[0, +\infty]$. Then we define the sum $\sum_{\lambda\in\Lambda} a_\lambda$ to mean the supremum of all sums of the form $\sum_{\lambda\in L} a_\lambda$ for finite sets $L \subseteq \Lambda$. The supremum exists, since $[0, +\infty]$ is order complete. Again, the order of the terms does not affect the summation. *Exercise.* When $\Lambda = \mathbb{N}$, then this definition is equivalent to the one given earlier for $\sum_{k=1}^{\infty} a_k$.

Actually, we are mainly interested in the countable case, because (*exercise*) if $\sum_{\lambda\in\Lambda} a_\lambda$ is finite, then at most countably many of the a_λ's are nonzero. *Hint*: Show that for each positive integer m, the set $\Lambda_m = \{\lambda \in \Lambda : a_\lambda > \frac{1}{m}\}$ is finite.

10.41. *Some basic properties of convergent series.*

a. Suppose $\sum_{j=1}^{\infty} a_j$ and $\sum_{j=1}^{\infty} b_j$ are convergent series of real or complex numbers with finite sums, and k is any constant. Then $\sum_{j=1}^{\infty}(a_j + b_j)$ and $\sum_{j=1}^{\infty}(ka_j)$ are also convergent series, with sums equal to $\left(\sum_{j=1}^{\infty} a_j\right) + \left(\sum_{j=1}^{\infty} b_j\right)$ and $k\sum_{j=1}^{\infty} a_j$, respectively.

b. If $a_j \leq b_j$ for all j, then $\sum_{j=1}^{\infty} a_j \leq \sum_{j=1}^{\infty} b_j$. In particular, if $0 \leq a_j \leq b_j$ and $\sum b_j$ is convergent, then $\sum a_j$ is convergent.

c. If $\sum_{j=1}^{\infty} a_j$ is a convergent series of real or complex numbers, then $\lim_{j\to\infty} a_j = 0$.

On the other hand, if $\lim_{j\to\infty} a_j = 0$, it does not follow that $\sum_{j=1}^{\infty} a_j$ is convergent — for example, consider the harmonic series in 10.41.f below.

d. *Geometric series.* Show that $1 + r + r^2 + \cdots + r^n = (1 - r^{n+1})/(1 - r)$ if $r \neq 1$, and hence

$$\sum_{j=0}^{\infty} r^j \quad = \quad 1 + r + r^2 + r^3 + \cdots \quad = \quad \begin{cases} 1/(1-r) & \text{if } |r| < 1 \\ \text{divergent} & \text{if } |r| \geq 1. \end{cases}$$

e. *Integral test.* If $f : [1, +\infty) \to [0, +\infty)$ is a decreasing function, then $\sum_{j=1}^{\infty} f(j)$ and $\int_1^{\infty} f(x)\,dx$ are both finite or both infinite. In fact, $\sum_{j=2}^{n+1} f(j) \leq \int_1^{n+1} f(x)dx \leq \sum_{j=1}^{n} f(j)$.

f. *Corollary.* The series $\sum_{j=1}^{\infty} j^{-p}$ converges for real numbers $p > 1$ and diverges if $0 < p \leq 1$.

In particular, the **harmonic series** $\sum_{j=1}^{\infty} \frac{1}{j} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$ diverges. However, it diverges rather slowly — i.e., to make $\sum_{j=1}^{n} \frac{1}{j}$ moderately large, we must make n incredibly enormous. In fact, the integral test tells us that $\sum_{j=1}^{n} \frac{1}{j}$ is approximately equal to $\ln n$. For instance, when n is a trillion, then $\sum_{j=1}^{n} \frac{1}{j}$ is only approximately equal to $\ln(10^{12}) = 12\ln(10) \approx 27.63$. When n is a **googol**, or 10^{100}, then $\sum_{j=1}^{n} \frac{1}{j}$ is still only about $\ln(10^{100}) = 100\ln(10) \approx 230.26$.

The harmonic series is a sort of "borderline case" — it often takes delicate calculations to decide the convergence or divergence of series that are similar to the harmonic series. The series $\sum_{j=2}^{\infty} \frac{1}{j\ln j}$ and $\sum_{j=2}^{\infty} \frac{1}{j(\ln j)(\ln\ln j)}$ also diverge, even more slowly. On the other hand, the series $\sum_{j=2}^{\infty} \frac{1}{j(\ln j)^2}$ converges.

g. *Alternating series test.* If $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$ and $\lim_{n\to\infty} a_n = 0$, then the series $a_1 - a_2 + a_3 - a_4 + \cdots$ converges. We omit the rather elementary proof (which can be found in most calculus books), since we shall prove a stronger result in 22.21.

h. Let $t_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n$. Show that $t_n - t_{n+1} = \int_n^{n+1} \frac{1}{x}\,dx - \frac{1}{n+1} > 0$. Hence the sequence $t_1, t_2, t_3, \ldots$ is bounded and decreasing. It therefore converges to a limit, which is called **Euler's constant**. That limit is approximately $0.577215664901532\ldots$; it is commonly denoted by γ.

10.42. When we add up only finitely many numbers, or add up infinitely many nonnegative numbers, then it does not matter in what order we add them; the result is the same. However, when we add up infinitely many numbers, some positive and some negative, then changing the order of the terms may affect the answer. For instance, using 10.41.h, show that $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots = \ln 2$, but

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \frac{1}{13} + \frac{1}{15} - \frac{1}{8} \cdots \quad = \quad \frac{3}{2}\ln 2.$$

If we change the order of the terms a bit more, we obtain the series

$$1-\frac{1}{2}+\underbrace{\frac{1}{3}}_{\text{1 odd term}}-\frac{1}{4}+\underbrace{\frac{1}{5}+\frac{1}{7}}_{\text{2 odd terms}}-\frac{1}{6}+\underbrace{\frac{1}{9}+\frac{1}{11}+\frac{1}{13}+\frac{1}{15}}_{\text{4 odd terms}}-\frac{1}{8}$$
$$+\underbrace{\frac{1}{17}+\frac{1}{19}+\frac{1}{21}+\frac{1}{23}+\frac{1}{25}+\frac{1}{27}+\frac{1}{29}+\frac{1}{31}}_{\text{8 odd terms}}-\frac{1}{10}+\cdots,$$

which converges to $+\infty$. *Hints*: Observe that

$$\frac{1}{3}>\frac{1}{4},\qquad \frac{1}{5}+\frac{1}{7}>\frac{1}{8}+\frac{1}{8},\qquad \frac{1}{9}+\frac{1}{11}+\frac{1}{13}+\frac{1}{15}>\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16},$$

etc. Also, show that

$$-\frac{1}{4}>\frac{\ln(1)-\ln(2)}{2},\qquad -\frac{1}{6}>\frac{\ln(2)-\ln(3)}{2},\qquad -\frac{1}{8}>\frac{\ln(3)-\ln(4)}{2},\qquad \ldots$$

Other rearrangements of this series yield other sums. In fact, it can be proved that *any* number L in $[-\infty,+\infty]$ can be obtained as the sum of a suitable rearrangement of the series above. (*Hints*: Obtain $L=-\infty$ in a fashion analogous to the method used above for the sum $L=+\infty$. Now consider any finite number L. Take just enough positive terms to get a partial sum that is greater than L; then take just enough negative terms to get a partial sum that is less than L; then just enough positive terms...; etc.)

Thus, it is erroneous and misleading to say that $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots$ is simply the "sum" of the a_k's. To be more precise, we must say that $\sum_{k=1}^{\infty} a_k$ is the sum of the a_k's *in the specified order*; this is reflected in our definition $\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty}(a_1+a_2+\cdots+a_n)$. Different orderings of the a_k's yield different partial sums $s_n = a_1 + a_2 + \cdots + a_n$ and thus different sequences (s_n), which may have different limits. Intuitively, it may be helpful to view a series this way: The numbers in $a_1+a_2+a_3+\cdots$ are not all added "simultaneously;" rather, the leftmost terms are added earlier than the terms occurring farther to the right. See the related results in 23.26.

10.43. *Example.* We shall now show that the series $\sum_{n=1}^{\infty}\frac{1}{n}|\sin(nx)|$ diverges, for any real number x that is not a multiple of π. (However, in 22.22 we shall show that the series $\sum_{n=1}^{\infty}\frac{1}{n}\sin(nx)$ converges for any real number x.)

Proof. Since $|\sin(n(x+\pi))| = |\sin(nx)|$, we may translate x by any multiple of π; thus we may assume that $-\pi/2 < x < \pi/2$ and $x \neq 0$. Since $|\sin(-nx)| = |\sin(nx)|$, we may assume that $0 < x < \pi/2$. Choose a positive integer M large enough so that $(M-1)x > 2\pi$.

Consider any M consecutive integers $k+1, k+2, \ldots, k+M$. Since $(k+M)x-(k+1)x > 2\pi$, the angles $(k+1)x, \ldots, (k+M)x$ go a bit more than once around a circle. Those angles cannot skip across the interval $(\frac{1}{4}\pi, \frac{3}{4}\pi)$ (modulo 2π) without taking a value in that interval, since that interval has width $\pi/2$, which is larger than x. Thus at least one of those angles lies in the interval $(\frac{1}{4}\pi, \frac{3}{4}\pi)$ (modulo 2π), and so at least one of the numbers $\sin((k+1)x), \ldots, \sin((k+M)x)$ is larger than $\frac{1}{2}\sqrt{2}$. Hence for any nonnegative integer j

we have

$$\max\left\{\frac{\sin(Mj+1)}{Mj+1}, \frac{\sin(Mj+2)}{Mj+2}, \ldots, \frac{\sin(Mj+M)}{Mj+M}\right\} \quad > \quad \frac{\frac{1}{2}\sqrt{2}}{Mj+1}.$$

Therefore

$$\sum_{n=1}^{\infty}\frac{|\sin(n)|}{n} \quad = \quad \sum_{j=0}^{\infty}\sum_{p=1}^{M}\frac{|\sin(Mj+p)|}{Mj+p} \quad \geq \quad \sum_{j=0}^{\infty}\frac{\frac{1}{2}\sqrt{2}}{Mj+1},$$

which diverges to ∞ since the harmonic series does.

10.44. *Decimals from real numbers.* Let $D = \{0, 1, 2, \ldots, 9\}$. For each sequence $\sigma = (s_1, s_2, s_3, s_4, \ldots)$ in $D^{\mathbb{N}}$, let

$$h(\sigma) \quad = \quad \frac{s_1}{10} + \frac{s_2}{10^2} + \frac{s_3}{10^3} + \frac{s_4}{10^4} + \cdots \quad = \quad \sum_{j=1}^{\infty}\frac{s_j}{10^j}\ .$$

Since the s_j's are nonnegative, the partial sums are an increasing sequence, and it is easy to see that they are bounded above; hence the series converges to a finite real number $h(\sigma)$. Then the expression "$0.s_1s_2s_3\cdots$" is called the the **decimal representation** of the number $h(\sigma)$. Show that

a. $0 \leq h(\sigma) \leq 1$.

b. By a *decimal rational* we shall mean a number of the form $m/10^k$, where m and k are integers. Show that any decimal rational $m/10^k$ in $(0, 1)$ is equal to $h(\sigma)$ for exactly two different sequences σ: one that is all 0s after a certain point, and another that is all 9s after a certain point. (For instance, $3.279999\ldots = 3.280000\ldots$.)

c. Any other real number $r \in (0, 1)$ (i.e., not a decimal rational) is equal to $h(\sigma)$ for exactly one sequence σ.

d. Note that there are only countably many decimal rationals in $[0, 1]$. Use this to show that $\mathrm{card}([0, 1]) = \mathrm{card}(D^{\mathbb{N}})$.

e. We evolved the decimal representation system because we each have ten fingers. But mathematically, there is nothing special about the number ten. An analogous system, which might develop on a planet where the people have b fingers for some integer $b > 1$, would use representations of the form

$$\frac{s_1}{b} + \frac{s_2}{b^2} + \frac{s_3}{b^3} + \frac{s_4}{b^4} + \cdots$$

with $s_j \in \{0, 1, 2, \ldots, b-1\}$. Thus $\mathrm{card}([0, 1]) = \mathrm{card}(\{0, 1, 2, \ldots, b-1\}^{\mathbb{N}})$.

In particular, we could take $b = 2$. Thus $\mathrm{card}([0, 1]) = \mathrm{card}(2^{\mathbb{N}})$.

f. Cardinality of the reals. Conclude that $\mathrm{card}(\mathbb{R}) = \mathrm{card}(2^{\mathbb{N}})$.

The number system $\mathbb{C}$, defined in 10.24, has a natural bijection to $\mathbb{R} \times \mathbb{R}$; conclude that also $\mathrm{card}(\mathbb{C}) = \mathrm{card}(2^{\mathbb{N}})$.

10.45. *Real numbers from decimals (optional).* In the preceding section, we considered $\mathbb{R}$ as already known — i.e., as defined in 10.8 and constructed in 10.15.d — and we studied decimal expansions as infinite series in $\mathbb{R}$. Historically, decimal expansions predate the abstract ideas of a Dedekind complete, chain ordered field. We could actually construct a Dedekind complete, chain ordered field by using formal decimal expansions; these ideas were published by Stolz in 1886.

We assume some familiarity with $\mathbb{Q}$, but not with $\mathbb{R}$. Define $\mathbb{R}$ to be the set of all infinite sequences of the form

$$(z, y_1, y_2, y_3, y_4, \ldots) \;\in\; \mathbb{Z} \times \{0, 1, 2, \ldots, 9\}^{\mathbb{N}},$$

where we identify a sequence ending in infinitely many 0s with the corresponding sequence ending in infinitely many 9s (as in 10.44.b). Such a sequence will be represented, as usual, by "$z + .y_1y_2y_3y_4 \ldots$." Its *rational truncations* are the finite sequences of symbols

$$z, \qquad z + .y_1, \qquad z + .y_1y_2, \qquad z + .y_1y_2y_3, \qquad z + .y_1y_2y_3y_4, \qquad \text{etc.}$$

where "$z + .y_1y_2 \cdots y_k$" represents the rational number $z + \frac{y_1}{10} + \frac{y_2}{100} + \cdots + \frac{y_k}{10^k}$.

Define the ordering of $\mathbb{R}$ in the usual lexicographical fashion. Then it is easy to show that $\mathbb{R}$ is chain ordered and Dedekind complete. Define the arithmetic operations $(+)$ and $(\cdot)$ first for rational truncations, in the obvious fashion. Then the sum of two numbers $a, b \in \mathbb{R}$ is defined to be the sup of the sums of the rational truncations of a and b. The product of two *positive* numbers $a, b \in \mathbb{R}$ is the sup of the products of their rational truncations. The product of two not-necessarily-positive numbers is defined in terms of the products of positive numbers; we omit the details. Zealous readers can verify that $\mathbb{R}$, defined in this fashion, is a complete ordered field. This approach is developed in greater detail in various other sources — for instance, Abian [1981], Dienes [1957], and Ritt [1946].

10.46. *Constructible numbers.* The constructivists' notion of "number" is a bit different from the mainstream mathematicians' notion. For a constructivist, a number is acceptable if it can be approximated arbitrarily closely and some estimates can be given for how fast the approximations are converging. Thus, numbers such as $\sqrt{2}$ and π are perfectly acceptable, for we have formulas (albeit complicated) for approximating these numbers to as many decimal places as we may wish. (See also 6.7.) However, the constructivists' notion of "number" has a few surprising consequences.

For instance, recall from the footnote in 6.4 that **Goldbach's Conjecture** asserts that for each integer $k > 1$,

($*$) the number $2k$ can be written as the sum of two prime numbers.

No proof or counterexample for this proposition has yet been found. Although we do not know whether ($*$) is true for every k, we can easily test it for any particular k, and thus we can evaluate the number defined by

$$x_k \;=\; \begin{cases} 0 & \text{if } (*) \text{ is true for this } k \\ 1 & \text{if } (*) \text{ is false for this } k. \end{cases}$$

Let us also define $x_1 = 0$. We can evaluate x_k for as many k's as we wish. (So far, all known x_k's are 0. Perhaps someday someone will find a k for which $x_k = 1$, or will prove that all the x_k's are 0.) Now define

$$\Gamma \quad = \quad -\frac{x_1}{10} + \frac{x_2}{100} - \frac{x_3}{1000} + \cdots + (-1)^k \frac{x_k}{10^k} + \cdots .$$

(To show that this series converges to a real number, either use results about Cauchy sequences in Chapter 19, or prove that the liminf and limsup of the partial sums differ by less than 10^{-n} for any n.) We shall refer to this number Γ as the "Goldbach number." We don't know the "exact value" of Γ yet. Nevertheless, constructivists would say that Γ is indeed a "real number," since we have an algorithm that can "find" Γ as accurately as we wish — i.e., given any $\varepsilon > 0$, we can compute an approximation Γ' satisfying $|\Gamma - \Gamma'| < \varepsilon$. The sign of the Goldbach number is related to the Goldbach Conjecture:

- $\Gamma = 0$, if the conjecture is true;
- $\Gamma > 0$, if the conjecture is false and the first counterexample (i.e., the first contradiction to $(*)$) occurs when k is even;
- $\Gamma < 0$, if the conjecture is false and the first counterexample occurs when k is odd.

We don't yet know which of those three cases holds; it is possible that we will never know. This mysterious quality may make some classical mathematicians reluctant to accept Γ as a "real number." It leads constructivists to conclude that, *even if we know two real numbers* (0 *and* Γ) *to arbitrarily high accuracy, we still may be unable to tell which of the numbers is larger.* This makes plausible our assertion in 6.6 that the Trichotomy Law for real numbers is not constructively provable.

We shall encounter the Goldbach number Γ again in 15.48.

Chapter 11

Linearity

Linear Spaces and Linear Subspaces

11.1. *Definitions.* Let $\mathbb{F}$ be any field. An **$\mathbb{F}$-linear space** is a set V equipped with operations 0, $-$, $+$, which make it an additive group, and also equipped with another mapping called **scalar multiplication**, from $\mathbb{F} \times V$ into V, satisfying certain rules noted below. The elements of V are called **vectors**. The elements of $\mathbb{F}$ are then called the **scalars**; we refer to $\mathbb{F}$ as the **scalar field**.

For any vector v and scalar c, the result of the scalar multiplication of c and v is called their **product**. It is usually written as $c \cdot v$ or as cv; generally the raised dot is included only for clarification or emphasis. The rules satisfied by scalar multiplication are:

(i) $1 \cdot v = v$,

(ii) $\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$,

(iii) $\alpha \cdot (u + v) = (\alpha \cdot u) + (\alpha \cdot v)$,

(iv) $(\alpha + \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v)$,

for all $\alpha, \beta \in \mathbb{F}$ and $u, v \in V$. The second rule is a sort of associativity of multiplication, although it should be noted that two different kinds of multiplication are involved: scalar times scalar and scalar times vector. The last two rules assert the distributivity of multiplication over addition; they can also be described as asserting the additivity of the mapping $v \mapsto \alpha \cdot v$ (for fixed scalar α) and the mapping $\alpha \mapsto \alpha \cdot v$ (for fixed vector v).

The same symbol "0" will be used for the additive identities of the scalar field $\mathbb{F}$ and the various linear spaces; it should be clear from the context just which additive identity is meant by any "0."

An $\mathbb{F}$-linear space may be called a **linear space**, or a **vector space**, if the choice of the scalar field $\mathbb{F}$ is clear or does not need to be mentioned explicitly. Whenever we work with several linear spaces at once, it will be understood that all the linear spaces are over the same scalar field $\mathbb{F}$ (unless some other arrangement is specified) — e.g., the discussion may apply to several vector spaces over $\mathbb{R}$ or to several vector spaces over $\mathbb{C}$, but we do not mix the two types unless that is mentioned explicitly. Whenever possible, we prefer not to specify what scalar field is being used, so that we can apply our results to many different scalar fields. See also the related discussion in 10.34.

11.2. *Some basic properties.*

a. $0 \cdot v = 0$ for any vector v; that is, the field's additive identity times any vector v yields the linear space's additive identity, and

b. $(-1) \cdot v = -v$ for any vector v; that is, the field's -1 times any vector v yields the additive inverse of v.

11.3. *More definitions.* A **linear algebra** over a field $\mathbb{F}$ is a set X equipped with 0, $+$, and two multiplication operations $\otimes$ and $*$, such that

(i) X with $0, +, *$ is a ring. (The operation $*$ may be called the **ring multiplication**; in some contexts it is referred to as the **vector multiplication**.

(ii) X with $0, +, \otimes$ is a linear space over some field $\mathbb{F}$ (and $\otimes$ is the multiplication of scalars times vectors, often called the **scalar multiplication**).

(iii) The two multiplication operations satisfy this compatibility rule: $c \otimes (x * y) = (c \otimes x) * y = x \otimes (c * y)$ for all scalars c and vectors x, y.

Such an object X is simply called an "algebra" in some of the older literature; we might refer to it as an **algebra in the classical sense**. For clarification we might call X an *algebra over* $\mathbb{F}$. (Perhaps a better term would be **linear ring**, or **$\mathbb{F}$-linear ring**.) If $(X, 0, +, *)$ is a ring with unit 1, then the resulting linear algebra is called an *algebra with unit*, or a **unital algebra**.

The linear algebra is said to be **commutative** if its ring multiplication is commutative — i.e., if $x * y = y * x$ for all $x, y \in X$.

Of course, we have used the symbols $\otimes$ and $*$ in this introductory discussion only for emphasis. Usually, the multiplication operations are both written as a raised dot $(\cdot)$ or indicated by juxtaposition — i.e., the product of x and y (with either type of multiplication) is usually denoted $x \cdot y$ or xy.

Most of the rings used by analysts are linear algebras over the field $\mathbb{R}$ or $\mathbb{C}$. Boolean algebras, studied in Chapter 13 and thereafter, can be viewed as algebras over the finite field $\mathbb{Z}_2 = \{0, 1\}$.

11.4. *Examples.*

a. Any field $\mathbb{F}$ is a commutative unital algebra over itself.

b. More generally, if $\mathbb{F}$ is a field and n is a positive integer, then $\mathbb{F}^n = \{n\text{-tuples of elements of } \mathbb{F}\}$ is a commutative unital algebra over $\mathbb{F}$. Elements of $\mathbb{F}^n$ are customarily represented in the form $v = (v_1, v_2, \ldots, v_n)$ using parentheses and commas, or as n-by-1 column matrices, or as the transposes of 1-by-n row matrices: $[v_1\ v_2\ \cdots\ v_n]^\top$; see

8.26. The vector operations on $\mathbb{F}^n$ are defined coordinatewise:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}, \qquad c\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix},$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 \\ x_2y_2 \\ \vdots \\ x_ny_n \end{bmatrix}$$

etc., for any vectors $x, y \in \mathbb{F}^n$ and scalar $c \in \mathbb{F}$.

c. Still more generally, any product $P = \prod_{\lambda\in\Lambda} X_\lambda$ of $\mathbb{F}$-linear spaces can be made into an $\mathbb{F}$-linear space, with operations defined coordinatewise:

$$(f+g)(\lambda) \;=\; (f(\lambda)) + (g(\lambda)), \qquad (c\cdot f)(\lambda) \;=\; c\cdot(f(\lambda))$$

for all $f, g \in P$, $\lambda \in \Lambda$, and scalars c. If the X_λ's are $\mathbb{F}$-(unital) algebras, then P is also an $\mathbb{F}$-(unital) algebra, with vector multiplication

$$(fg)(\lambda) \;=\; (f(\lambda))(g(\lambda)).$$

It is commutative if the X_λ's are all commutative.

In particular, when all the X_λ's are equal to one space X, we see that $X^\Lambda = \{\text{functions from } \Lambda \text{ into } X\}$ is a linear space or a linear algebra.

The product vector space takes a more intuitively appealing form if we write $\Lambda = \{\alpha, \beta, \gamma, \ldots\}$. (Here we follow the convention of 1.32: it is not assumed that Λ is ordered or countable.) Then we have

$$\begin{bmatrix} x_\alpha \\ x_\beta \\ x_\gamma \\ \vdots \end{bmatrix} + \begin{bmatrix} y_\alpha \\ y_\beta \\ y_\gamma \\ \vdots \end{bmatrix} = \begin{bmatrix} x_\alpha+y_\alpha \\ x_\beta+y_\beta \\ x_\gamma+y_\gamma \\ \vdots \end{bmatrix}, \qquad c\begin{bmatrix} x_\alpha \\ x_\beta \\ x_\gamma \\ \vdots \end{bmatrix} = \begin{bmatrix} cx_\alpha \\ cx_\beta \\ cx_\gamma \\ \vdots \end{bmatrix},$$

and for linear algebras

$$\begin{bmatrix} x_\alpha \\ x_\beta \\ x_\gamma \\ \vdots \end{bmatrix}\begin{bmatrix} y_\alpha \\ y_\beta \\ y_\gamma \\ \vdots \end{bmatrix} = \begin{bmatrix} x_\alpha y_\alpha \\ x_\beta y_\beta \\ x_\gamma y_\gamma \\ \vdots \end{bmatrix}.$$

d. Let $\mathbb{F}$ be a field, let n be a positive integer, and let X be the set of all n-by-n matrices over $\mathbb{F}$. Then $\mathbb{F}$ is a unital algebra, with vector multiplication given by the multiplication of matrices (as defined in 8.27). This algebra is not commutative if $n > 1$.

Preview. More generally, if X is a linear space, then the linear operators from X into X form a noncommutative unital algebra with ring multiplication given by composition of operators. If X is a topological vector space, we may also consider the *continuous* linear operators; it is another unital algebra.

e. Let G be a locally compact Abelian group equipped with its Haar measure, and let $L^1(G)$ be defined accordingly — see 26.45. It can be shown that $L^1(G)$ is a commutative algebra, generally *not* unital, with ring multiplication defined by the **convolution** operation $(f * g)(t) \quad = \quad \int_G f(t-s)g(s)\,ds$.

f. Another important algebraic system can be described as follows: Let $X = \mathbb{R}^3$ be equipped with the usual vector space operations, as in 11.4.b. The **cross product** of two vectors is defined by

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \times \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \quad = \quad \begin{bmatrix} y_1 z_2 - z_1 y_2 \\ z_1 x_2 - x_1 z_2 \\ x_1 y_2 - y_1 x_2 \end{bmatrix}.$$

This multiplication is **anticommutative**: it satisfies $x \times y = -y \times x$, and consequently $x \times x = 0$. In particular, we have

$$\begin{array}{ccc} \mathbf{i} \times \mathbf{j} = \mathbf{k}, & \mathbf{j} \times \mathbf{k} = \mathbf{i}, & \mathbf{k} \times \mathbf{i} = \mathbf{j}, \\ \mathbf{j} \times \mathbf{i} = -\mathbf{k}, & \mathbf{k} \times \mathbf{j} = -\mathbf{i}, & \mathbf{i} \times \mathbf{k} = -\mathbf{j}, \end{array}$$

where

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The cross product is *not* associative; for instance, $\mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = -\mathbf{j} \neq 0 = (\mathbf{i} \times \mathbf{i}) \times \mathbf{j}$. Consequently, $\mathbb{R}^3$ is *not* a linear algebra when the cross product is used for vector multiplication.

Several more examples are given in 11.45 and 11.46, and in Chapter 22 and thereafter.

11.5. *Definitions.* Let X be an $\mathbb{F}$-linear space, and let $S \subseteq X$. A **linear combination** of elements of S is an expression of the form

$$t \quad = \quad \alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_n s_n,$$

where n is a nonnegative integer, the s_j's are elements of S, and the α_j's are elements of $\mathbb{F}$. We permit $n = 0$, with the convention that the sum of no elements of X is 0.

A **linear subspace** of X is a subset $S \subseteq X$ with the property that any linear combination of elements of S is also an element of S. Equivalently, it is a nonempty set that is closed under scalar multiplication by all scalars and under the binary operation of addition. (Thus, it is a subalgebra in the variety of $\mathbb{F}$-linear spaces — see 8.55 and 9.21.)

11.6. *Basic properties.* Let X be an $\mathbb{F}$-linear space. Prove the following results, either directly or by using results of 9.21.

a. The whole vector space X is a linear subspace of itself.

b. Any intersection of linear subspaces is a linear subspace.

c. Let $T \subseteq X$. Then there exists a smallest linear subspace containing T; it is the intersection of all the linear subspaces containing T. It is also equal to the set of all linear combinations of elements of T. It is called the **span** (or *linear span*) of T; it may be abbreviated $\mathrm{span}(T)$. A linear subspace S is said to be **spanned** by T if $S = \mathrm{span}(T)$. (This is a special type of Moore closure, but the term "closure" generally is not used in this context.)

d. If S is a linear subspace of X, then S becomes a linear space in its own right, when the vector space operations of X (addition, scalar multiplication, additive inverse, 0) are restricted to S.

e. $\{0\}$ is a linear space over any scalar field. It is also a linear subspace of any linear space. It is contained in the span of any set. In fact, $\{0\}$ is the span of the empty set. The empty set is *not* a linear space.

f. The definition of linear subspace depends on the choice of the scalar field $\mathbb{F}$. For instance, the set $S = \{u \in \mathbb{C} : \mathrm{Re}(u) = 0\}$ is a linear subspace of $\mathbb{C}$ when we take $\mathbb{R}$ for the scalar field, but not when we take $\mathbb{C}$ for the scalar field.

g. If S and T are linear subspaces of X and $c \in \mathbb{F}$, then the sets

$$S + T = \{s + t : s \in S,\ t \in T\}, \qquad cS = \{cs : s \in S\}$$

are linear subspaces also.

h. If S_λ ($\lambda \in \Lambda$) are linear subspaces of X, then the sum $\sum_{\lambda \in \Lambda} S_\lambda$ (defined in 8.11) is also a linear subspace; in fact, it is the span of $\bigcup_{\lambda \in \Lambda} S_\lambda$.

i. Let $(Y_\lambda : \lambda \in \Lambda)$ be an indexed set of $\mathbb{F}$-linear spaces. The **external direct sum** of the Y_λ's is the set

$$\coprod_{\lambda \in \Lambda} Y_\lambda = \left\{ f \in \prod_{\lambda \in \Lambda} Y_\lambda \ : \ f(\lambda) \text{ is nonzero for at most finitely many } \lambda\text{'s} \right\}.$$

This is a linear subspace of the product $\prod_{\lambda \in \Lambda} Y_\lambda$. Of course, if Λ is a finite set, then the external direct sum is *equal* to the product.

The external direct sum described above is a special case of the external direct sum defined in 9.30. *Caution*: Some mathematicians call this the "direct sum;" see the remarks in 9.30.

An important special case is that in which all the Y_λ's are equal to one vector space Y. Then the external direct sum $\coprod_{\lambda \in \Lambda} Y$ is equal to the set of all functions $f : \Lambda \to Y$ that vanish on all but finitely many λ's.

Specializing further: Let $\mathbb{F}$ be the scalar field; then $\coprod_{n \in \mathbb{N}} \mathbb{F}$ is the linear space consisting of all sequences of scalars that have only finitely many nonzero terms.

j. If $\mathbb{F}$ is any field, then $\mathbb{F}^{\mathbb{F}} = \{\text{functions from } \mathbb{F} \text{ into itself}\}$ is a linear space. (In fact, it is a *commutative algebra*; see 11.3.) For each positive integer n, let $P_n = \{$polynomials of degree at most n, in one variable, with coefficients in $\mathbb{F}\}$; this is a linear subspace of $\mathbb{F}^{\mathbb{F}}$. The set $Q_n = \{$polynomials of degree exactly $n\}$ is not a linear space, since it is not closed under addition.

k. *Preview.* Let $\mathbb{F}$ be the scalar field (either $\mathbb{R}$ or $\mathbb{C}$). Then $\mathbb{F}^{(0,1)} = \{$functions from $(0,1)$ into $\mathbb{F}\}$ is a linear space. Following are some linear subspaces of $\mathbb{F}^{(0,1)}$, of types that will be studied later in this book:

$$\begin{array}{rcl} B & = & \{\text{bounded functions}\}, \\ BC & = & \{\text{bounded continuous functions}\}, \\ BUC & = & \{\text{bounded, uniformly continuous functions}\}, \\ \text{Lip} & = & \{\text{Lipschitzian functions}\}, \\ C_0^\infty & = & \{\text{smooth functions vanishing at endpoints}\}. \end{array}$$

All of the relevant terms are defined later in this book. *Exercise for more advanced readers*: Show that $\mathbb{F}^{(0,1)} \supsetneqq B \supsetneqq BC \supsetneqq BUC \supsetneqq \text{Lip} \supsetneqq C_0^\infty$.

LINEAR MAPS

11.7. *Definitions.* An $\mathbb{F}$**-linear map** is a mapping $f : X \to Y$ from one $\mathbb{F}$-linear space into another that satisfies

$$f(x+x') \quad = \quad f(x)+f(x') \qquad \text{and} \qquad f(cx) = cf(x)$$

for all $x, x' \in X$ and $c \in \mathbb{F}$. We may omit the prefix "$\mathbb{F}$" and simply refer to a **linear map**, if no confusion will result. However, we emphasize that the choice of $\mathbb{F}$ is part of the definition. For instance, the map $\alpha \mapsto \overline{\alpha}$, from $\mathbb{C}$ into itself, is $\mathbb{R}$-linear but not $\mathbb{C}$-linear.

The $\mathbb{F}$-linear maps are just the homomorphisms (as defined in 8.48) for the variety of all $\mathbb{F}$-linear spaces. The *category of* $\mathbb{F}$*-linear spaces* has $\mathbb{F}$-linear spaces for objects and $\mathbb{F}$-linear maps for morphisms.

Mathematicians often omit the parentheses in writing linear maps — i.e., if f is linear then $f(x)$ may be written as fx. As usual, operations written multiplicatively are performed before operations written additively, if no parentheses dictate otherwise. Thus, an expression such as $fx + u$ is understood to mean $(f(x)) + u$, not $f(x+u)$.

A **bilinear map** is a mapping $f : X \times Y \to Z$ from the product of two linear spaces into a linear space, such that

$$\begin{array}{l} f(x,\cdot) : Y \to Z \text{ is linear for each fixed } x \in X, \text{ and} \\ f(\cdot,y) : X \to Z \text{ is linear for each fixed } y \in Y. \end{array}$$

11.8. It is easy to see that if X and Y are $\mathbb{F}$-linear spaces, then

$$\text{Lin}(X,Y) \quad = \quad \{\mathbb{F}\text{-linear maps from } X \text{ into } Y\}$$

is a linear subspace of Y^X.

A linear map from a vector space into the scalar field is also called a **linear functional**. The **linear dual** of an $\mathbb{F}$-linear space is the linear space

$$\text{Lin}(X,\mathbb{F}) \quad = \quad \{\text{linear maps from } X \text{ into } \mathbb{F}\}.$$

When the context is clear, $\mathrm{Lin}(X, \mathbb{F})$ may be called the **dual** of X and denoted more briefly by $\mathrm{Lin}(X)$ or by X^*. The reader is cautioned that "dual" and "X^*" have other meanings in other contexts; see 9.55.

11.9. *Examples and further properties.* Prove the following, either directly or as specializations of results about homomorphisms between algebraic systems.

a. Any linear map f is a homomorphism of additive groups. Hence it satisfies

$$f(0) = 0, \qquad \{0\} \subseteq f^{-1}(0), \qquad f \text{ is injective } \iff \{0\} = f^{-1}(0).$$

b. If $f : X \to Y$ is a linear map, then $\mathrm{Graph}(f)$ is a linear subspace of $X \times Y$.

c. (This example requires some familiarity with calculus.) The set

$$C[0,1] \quad = \quad \{\text{continuous functions from } [0,1] \text{ into } \mathbb{R}\}$$

is a linear subspace of $\mathbb{R}^{[0,1]}$. Let any $g \in C[0,1]$ be fixed; then a linear functional $L_g : C[0,1] \to \mathbb{R}$ can be defined by the Riemann integral

$$L_g(f) \quad = \quad \int_0^1 f(t)g(t)\,dt \qquad (f \in C[0,1]).$$

This example will be generalized substantially in later chapters.

d. If $f : X \to Y$ is a linear map and S is a linear subspace of X, then $f(S)$ is a linear subspace of Y.

e. If $f : X \to Y$ is a linear map and T is a linear subspace of Y, then $f^{-1}(T)$ is a linear subspace of X.

f. A **linear isomorphism** is a linear map that is bijective. Show that if $f : X \to Y$ is a linear isomorphism, then $f^{-1} : Y \to X$ is also linear, and hence is also a linear isomorphism.

g. The identity map $i : X \to X$ is linear; its kernel is $\{0\}$.

h. If X and Y are linear spaces, then the constant mapping from X to Y that sends all elements to 0 is a linear map; its kernel is X.

i. Let S be a linear subspace of X. Define a relation on X by $x_1 \approx x_2$ if $x_1 - x_2 \in S$. This is an equivalence relation. Let X/S be the quotient space — i.e., the set of all equivalence classes — and let $\pi : X \to X/S$ be the quotient map. Then X/S is a linear space, with operations defined by

$$\pi(x_1) + \pi(x_2) = \pi(x_1 + x_2), \qquad c\pi(x) = \pi(cx)$$

for $x_1, x_2, x \in X$ and $c \in \mathbb{F}$, and the quotient map is a linear map.

j. **(Isomorphism Theorem.)** $X/\mathrm{Ker}(f)$ is isomorphic to $\mathrm{Range}(f)$, by the mapping $F(\pi(x)) = f(x)$.

k. Let A be an m-by-n matrix over a field $\mathbb{F}$. Represent elements of the vector spaces $\mathbb{F}^m$ and $\mathbb{F}^n$ as column vectors. Then the map $v \mapsto Av$, defined as in 8.28, is a linear map from $\mathbb{F}^n$ into $\mathbb{F}^m$.

11.10. *Proposition.* Let X and Y be linear spaces. Let $S \subseteq X$, and let $f : S \to Y$ be some function. Then f can be extended to a linear map $F : \text{span}(S) \to Y$ if and only if f has this property:

> whenever $s_1, s_2, \ldots, s_m$ are elements of S and $a_1, a_2, \ldots, a_m$ are scalars such that $a_1 s_1 + a_2 s_2 + \cdots + a_m s_m = 0$, then f satisfies $a_1 f(s_1) + a_2 f(s_2) + \cdots + a_m f(s_m) = 0$.

Moreover, if f satisfies that condition, then the extension F is unique, for it must satisfy

$$F(a_1 s_1 + a_2 s_2 + \cdots + a_m s_m) \quad = \quad a_1 f(s_1) + a_2 f(s_2) + \cdots + a_m f(s_m). \qquad (*)$$

Proof. If $(a_1 s_1 + \cdots + a_m s_m) - (b_1 t_1 + \cdots + b_n t_n) = 0$, with the s_i's and t_j's in S, then $[a_1 f(s_1) + \cdots + a_m f(s_m)] - [b_1 f(t_1) + \cdots + b_n f(t_n)] = 0$ by our hypothesis on f. Hence the formula $(*)$ does indeed define a function F. Obviously that function is linear.

Further observation (optional). Let Φ be the collection of all graphs of functions f from subsets of X into Y, such that f can be extended to a linear map on span(Dom(f)). Then Φ has finite character (see 3.46).

11.11. *Real linear versus complex linear.* In applications, the most important fields are $\mathbb{R}$ and $\mathbb{C}$. A linear space over the scalar field $\mathbb{R}$ is a **real linear space**; a linear space over the scalar field $\mathbb{C}$ is a **complex linear space**.

In some parts of functional analysis — e.g., vector lattices or nonlinear functional analysis — there are relatively few benefits from working with complex scalars. Consequently some mathematicians simplify their notation by only considering real linear spaces. For many purposes, this limitation is without loss of generality, since every complex linear space can also be viewed as a real linear space. Indeed, since $\mathbb{R} \subseteq \mathbb{C}$, we can replace the scalar multiplication $(\cdot) : \mathbb{C} \times V \to V$ with its restriction $(\cdot) : \mathbb{R} \times V \to V$.

Complex scalars are important for some areas of functional analysis and its applications, particularly spectral theory and mathematical physics. Consequently some mathematical books and papers only consider complex linear spaces. This limitation is without much loss of generality, for every real linear space X can be viewed as a subset of a complex linear space, as we now show:

Let X be any real linear space. Then on the vector space $X \times X$ we can define the vector operations

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) \quad &= \quad (x_1 + x_2, y_1 + y_2), \qquad \text{and} \\ (a + ib) \cdot (x_1, y_1) \quad &= \quad (ax_1 - by_1, ay_1 + bx_1) \end{aligned}$$

for (x_1, y_1) and (x_2, y_2) in $X \times X$ and $a, b \in \mathbb{R}$. These definitions make $X \times X$ a complex linear space, called the **complexification** of X. We may denote it by $X + iX$ and its element (x, y) by $x + iy$. Note that X is isomorphic to the subset $\{(x, 0) : x \in X\}$.

This construction seems rather cumbersome, but in most applications the complexification arises naturally. Indeed, $\mathbb{C}$ (introduced in 10.24) is just the complexification of $\mathbb{R}$, and for any set Λ the linear space $\mathbb{C}^\Lambda$ is the complexification of $\mathbb{R}^\Lambda$. Here is another example:

Let Λ be a topological space; then a function $f : \Lambda \to \mathbb{C}$ is continuous if and only if it is of the form $f = u + iv$, where u, v are real-valued continuous functions.

In 11.30.e we shall see that *any* complex linear space can be viewed as the complexification of a real linear space — though not necessarily in a constructive fashion.

11.12. *Bohnenblust-Sobczyk Correspondence.* Any complex vector space V may also be viewed as a real vector space (by "forgetting" how to multiply by scalars), but the two viewpoints give us two different collections of linear functionals. How are the two collections related?

Recall that a mapping $f : V \to Z$, from one $\mathbb{F}$-linear space into another, is an $\mathbb{F}$-linear map if it is additive and satisfies $f(cv) = cf(v)$ for all $v \in V$ and $c \in \mathbb{F}$. A **real linear functional** on V is an $\mathbb{R}$-linear map from V into $\mathbb{R}$; a **complex linear functional** on V is a $\mathbb{C}$-linear map from V into $\mathbb{C}$.

Now suppose V is a complex linear space. Show that

a. If f is a complex linear functional on V, then $g_1(v) = \operatorname{Re} f(v)$ and $g_2(v) = \operatorname{Im} f(v)$ are real linear functionals on V with $g_2(v) = -g_1(iv)$, and $f = g_1 + ig_2$.

b. Conversely, if g_1 is a real linear functional on V, then $f(v) = g_1(v) - ig_1(iv)$ is a complex linear functional on V.

c. These transformations give a bijection $f \leftrightarrow g_1$ between the real linear and complex linear functionals on V.

d. (*Optional.*) Generalize the preceding argument. If V is a complex linear space and X is a real linear space with complexification $X + iX$, then there is a bijection between complex linear maps $f : V \to X + iX$ and real linear maps $g_1 : V \to X$. Also, if T and X are real linear spaces, then any real linear map from T into X extends uniquely to a complex linear map from $T + iT$ into $X + iX$.

Linear Dependence

11.13. *Definitions.* A set $S \subseteq X$ is **linearly dependent** if we can write

$$0 \quad = \quad c_1 s_1 + c_2 s_2 + \cdots + c_n s_n,$$

where n is a positive integer, the c_i's are *nonzero* scalars, and the s_i's are distinct elements of S. If 0 cannot be expressed in this fashion, S is **linearly independent**.

11.14. *Observations.* In any linear space:

a. $\varnothing$ is a linearly independent subset.

b. Any subset containing 0 is linearly dependent.

c. If v is a nonzero vector then the singleton $\{v\}$ is a linearly independent set.

d. If b and c are distinct scalars and v is any vector, then any set containing both bv and cv is linearly dependent.

e. A set S is linearly dependent if and only if some point $s \in S$ is in $\operatorname{span}(S \backslash \{s\})$.

f. (*Optional.*) A set S is linearly independent if and only if each finite subset of S is linearly independent. Thus, the collection of all linearly independent sets has finite character (see 3.46).

11.15. *Example.* For each $r \in \mathbb{R}$, let $v(r) = (1, r, r^2, r^3, \ldots)$. Then $\{v(r) : r \in \mathbb{R}\}$ is a linearly independent subset of the real linear space $\mathbb{R}^{\mathbb{N}} = \{\text{sequences of reals}\}$.

Hint: Suppose $a_1 v(r_1) + a_2 v(r_2) + \cdots + a_n v(r_n) = 0$ for some scalars $a_1, a_2, \ldots, a_n$ and some distinct real numbers $r_1, r_2, \ldots, r_n$. Show that $a_1 p(r_1) + a_2 p(r_2) + \cdots + a_n p(r_n) = 0$ for every polynomial p. By considering the Lagrange polynomials (2.2.e), show $a_1 = a_2 = \cdots = a_n = 0$.

11.16. *Common Kernel Lemma.* Let k be a positive integer, let X be a linear space, and let λ_0, λ_1, λ_2, ..., λ_k be elements of X^*, the linear dual of X (defined in 11.8). Then:

(P_k) $\bigcap_{i=1}^{k} \operatorname{Ker}(\lambda_i) \subseteq \operatorname{Ker}(\lambda_0)$ if and only if $\lambda_0 \in \operatorname{span}\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$.

(Q_k) $\lambda_1, \lambda_2, \ldots, \lambda_k$ are linearly independent elements of X^* if and only if there exist vectors $x_1, x_2, \ldots, x_k \in X$ such that $\lambda_j(x_i) = \delta_{ij}$ (where δ is the Kronecker delta).

Hints: The "if" parts are obvious. We shall prove the "only if" parts by induction on k, showing $Q_m \Rightarrow P_m \Rightarrow Q_{m+1}$. The proof of Q_1 is trivial.

To prove $Q_m \Rightarrow P_m$, first note that we can omit any element of the set $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ that is a linear combination of the other elements of that set; hence we may assume that that set is linearly independent. Now show that if $x_1, x_2, \ldots, x_m$ are as in Q_m and $x \in X$, then $\sum_{j=1}^{m} \lambda_j(x) x_j - x \in \bigcap_{i=1}^{m} \operatorname{Ker}(\lambda_i)$.

To prove $P_m \Rightarrow Q_{m+1}$, let $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_m$ be linearly independent elements of $\operatorname{Lin}(X)$. By symmetry (explain), it suffices to establish the existence of x_0 satisfying $\lambda_i(x_0) = \delta_{i0}$. If no such x_0 exists, show $\bigcap_{i=1}^{m} \operatorname{Ker}(\lambda_i) \subseteq \operatorname{Ker}(\lambda_0)$.

11.17. *Definitions.* Let X be a linear space, and let $B \subseteq X$. Then the following conditions are equivalent. If one (hence all) of them is satisfied, we say B is a **basis** for X — or, to be more specific, a **vector basis** or **linear basis**. (Some mathematicians also call it a **Hamel basis**, but other mathematicians reserve that term for a narrower meaning indicated in 11.30.c.)

(A) For each nonzero vector $x \in X$, there is one and only one way (except for changing the order of the summation) to write $x = c_1 s_1 + c_2 s_2 + \cdots + c_n s_n$, with n equal to a positive integer, with the c_i's equal to nonzero scalars, and with the s_i's equal to distinct elements of B.

(B) B is linearly independent and $\operatorname{span}(B) = X$.

(C) B is a maximal linearly independent subset of X, i.e., a linearly independent set that is not included in any other linearly independent set.

(D) B is a minimal spanning set for X; that is, $\operatorname{span}(B) = X$ and B does not contain some other set A satisfying $\operatorname{span}(A) = X$.

Later we shall use the Axiom of Choice to prove that every linear space V has a vector basis and that any two vector bases for V have the same cardinality. That cardinality is called the **dimension** of V; it is written $\dim(V)$. The linear space is said to be **finite-dimensional** or **infinite-dimensional** according as $\dim(V)$ is finite or infinite. When V is finite-dimensional, then $\dim(V)$ is a nonnegative integer.

11.18. *Examples and observations.*

a. $\{(1,0),(0,1)\}$ and $\{(1,0),(-1,1)\}$ are two different vector bases for $\mathbb{F}^2$.

b. Let X be the degenerate linear space $\{0\}$, which contains just the one vector 0. Then the empty set is a vector basis for X.

c. Let X be a linear space, and let $S \subseteq X$. Then S is linearly independent if and only if S is a vector basis for $\mathrm{span}(S)$.

d. Let X and Y be linear spaces, let B be a vector basis for X, and let $f \in Y^B$. Then f extends uniquely to a linear map from X into Y. Thus, there is an isomorphism between the linear spaces Y^B and $\mathrm{Lin}(X;Y) = \{$linear maps from X into $Y\}$. In particular, $\mathbb{F}^B$ is isomorphic to the linear dual of X (defined in 11.8).

e. If X is an $\mathbb{F}$-linear space with vector basis B, then X is isomorphic to $\coprod_{b\in B}\mathbb{F}$ — that is, the external direct sum of B copies of the scalar field $\mathbb{F}$ (defined in 11.6.i). An isomorphism i from the external direct sum onto X is given by $i(f) = \sum_{b\in B} f(b)\cdot b$. (This sum makes sense since $f(b)$ is a scalar, b is a vector, and $f(b)$ vanishes for all but finitely many b.) For each $b \in B$, we have $i^{-1}(b)$ equal to $1_{\{b\}}$, the characteristic function (defined on B) of the singleton $\{b\}$; such functions form a vector basis for the external direct sum.

Further Results in Finite Dimensions

11.19. Let n be a positive integer. For each $j \in \{1,2,\ldots,n\}$, define the vector

$$\eta_j \quad = \quad (0,\ldots,0,1,0,\ldots,0) \quad = \quad [0 \;\cdots\; 0\; 1\; 0 \;\cdots\; 0]^{\top} \;\in\; \mathbb{F}^n,$$

where η_j has a 1 in the jth position and 0s elsewhere. It is easy to see that any vector $v = (v_1,v_2,\ldots,v_n) \in \mathbb{F}^n$ can be written in one and only one way as $v = c_1\eta_1 + c_2\eta_2 + \cdots + c_n\eta_n$ for scalars $c_1,c_2,\ldots,c_n$; indeed, we must take $c_j = v_j$ for each j. Hence $\{\eta_1,\eta_2,\ldots,\eta_n\}$ is a vector basis for $\mathbb{F}^n$; it is called the **standard basis** for $\mathbb{F}^n$. The vector η_j is called the jth **standard basis vector** for $\mathbb{F}^n$.

11.20. *Matrices as linear maps.* Every linear map from $\mathbb{F}^n$ into $\mathbb{F}^m$ is uniquely representable as an m-by-n matrix, in the sense of 11.9.k.

Indeed, let $\{\eta_1,\eta_2,\ldots,\eta_n\}$ be the standard vector basis for $\mathbb{F}^n$, as in 11.19. Let $A : \mathbb{F}^n \to \mathbb{F}^m$ be a linear map. Then the values of A are determined by its values on the η_j's; indeed, if $v = c_1\eta_1 + c_2\eta_2 + \cdots + c_n\eta_n$ for some scalars $c_j \in \mathbb{F}$, then $Av = c_1A\eta_1 + c_2A\eta_2 + \cdots + c_nA\eta_n$.

It follows that A is represented by the m-by-n matrix whose columns are the vectors $A\eta_j$ $(j = 1, 2, \ldots, n)$.

11.21. If $v, w \in \mathbb{F}^n$, then $v^\top w$ is the product of a 1-by-n matrix and a n-by-1 matrix. Thus it is a 1-by-1 matrix, which we may view as a scalar. We shall call this the **scalar product** of the vectors v and w. A couple of its basic properties are:

a. The scalar product is symmetric — that is, $v^\top w = w^\top v$.

b. In a product of matrices $C = AB$, the component c_{ik} is the scalar product of the ith row of A with the kth column of B.

11.22. For each fixed $v \in \mathbb{F}^n$, define a mapping $f_v : \mathbb{F}^n \to \mathbb{F}$ by $w \mapsto v^\top w$. Then:

a. f_v is a linear functional on $\mathbb{F}^n$; that is, f_v is a member of $(\mathbb{F}^n)^*$, the linear dual of $\mathbb{F}^n$.

b. The mapping $f : v \mapsto f_v$ is a linear map from $\mathbb{F}^n$ into $(\mathbb{F}^n)^*$.

c. The mapping $f : v \mapsto f_v$ is a bijection from $\mathbb{F}^n$ onto $(\mathbb{F}^n)^*$. (Thus $\mathbb{F}^n$ is isomorphic to its own linear dual.)

Hints: Any member of $(\mathbb{F}^n)^*$ is a linear map from $\mathbb{F}^n$ into $\mathbb{F}^1$, hence (by 11.20) representable as f_v for some n-by-1 matrix v; thus f is surjective. To show it is injective, note that if $v = (v_1, v_2, \ldots, v_n) \neq 0$, then $v_j \neq 0$ for at least one j; hence $f_v(\eta_j) \neq 0$, where η_j is the jth standard basis vector; hence $f_v \neq 0$.

d. *The dual of a linear map.* Suppose X and Y are linear spaces over the scalar field $\mathbb{F}$, and $A : X \to Y$ is some linear map. Then we may define a dual map $A^* : Y^* \to X^*$ by $A^*(f) = f \circ A$, as in 9.55. Show that if X and Y are finite dimensional, and A is represented by a matrix, then A^* is represented by the transpose of that matrix, introduced in 8.26.

More precisely: Let $v \mapsto f_v$ be the bijection from $\mathbb{F}^n$ onto $(\mathbb{F}^n)^*$ described above, and let $w \mapsto g_w$ be the analogous bijection from $\mathbb{F}^m$ onto $(\mathbb{F}^m)^*$. Suppose $A : \mathbb{F}^n \to \mathbb{F}^m$ is some linear map, represented by an m-by-n matrix which we shall also denote by A. Define a corresponding map $A^* : (\mathbb{F}^m)^* \to (\mathbb{F}^n)^*$ by this rule: $[A^*(g_w)](x) = g_w(A(x))$ for any $x \in \mathbb{F}^n$ — that is, $A^*(g_w)$ is the composition

$$A^*(g_w) \quad = \quad g_w \circ A \quad : \quad \mathbb{F}^n \xrightarrow{A} \mathbb{F}^m \xrightarrow{g_w} (\mathbb{F}^m)^*.$$

Show that $A^*(g_w) = f_{A^\top w}$, where $A^\top$ is the transpose of the matrix A — that is, $A^*(g_w) = f_v$, where $v = A^\top w$.

11.23. Lemma. Let A be an m-by-n matrix over $\mathbb{F}$. Then the n columns of A are linearly independent elements of $\mathbb{F}^m$ if and only if there exists an n-by-m matrix B such that $BA = I_n$. (We then say B is a **left inverse** for A.)

Proof. For the "if" part, suppose $BA = I_n$, but the columns $A_1, A_2, \ldots, A_n$ are linearly dependent vectors in $\mathbb{F}^m$ — i.e., suppose $c_1 A_1 + \cdots + c_n A_n$ is equal to 0_m, the zero vector in $\mathbb{F}^m$, for some scalars $c_1, \ldots, c_n$ that are not all 0. Let $c = [c_1 \; c_2 \; \cdots \; c_n]^\top$; infer that Ac is equal to 0_m. But then $c = I_n c = BAc = B0_m = 0_n$, a contradiction.

For the "only if" part, assume the columns $A_1, A_2, \ldots, A_n$ are linearly independent vectors in $\mathbb{F}^m$. View them as elements of the linear dual of $\mathbb{F}^m$, by their action in the scalar product (see 11.22). By the Common Kernel Lemma, there are vectors $b_1, b_2, \ldots, b_n \in \mathbb{F}^m$ such that $b_i^\top A_j = \delta_{ij}$. Take the 1-by-$m$ matrices $b_i^\top$ to be the rows of a matrix B; then $BA = I_n$.

11.24. *Remarks.* By taking transposes in the preceding result, we obtain this dual result:

> The *rows* of a matrix A are linearly independent if and only if A has a *right* inverse, i.e., a matrix C such that $AC = I$.

It can also be proved, though not so easily, that the rows of a *square* matrix are linearly independent if and only if its columns are linearly independent. This slightly deeper result can be proved using determinants or other more advanced methods, but we shall not need it. It implies that a linear operator $f : V \to V$, from a finite-dimensional linear space into itself, has a left inverse if and only if it has a right inverse. That conclusion is not valid in infinite-dimensional linear spaces, as we see from the example in 8.5.a.

11.25. Proposition. In any linear space V over the field $\mathbb{F}$, if $w_1, w_2, \ldots, w_{n+1}$ are $n+1$ vectors in the span of some n vectors $y_1, y_2, \ldots, y_n$, then the vectors $w_1, w_2, \ldots, w_{n+1}$ are linearly dependent.

Proof. We first prove this in the special case where $V = \mathbb{F}^n$ and the y_i's are the basis vectors η_i; that is, we first show that any $n+1$ vectors in $\mathbb{F}^n$ are linearly dependent. Assume, to the contrary, that $w_1, w_2, \ldots, w_n, w_{n+1}$ are linearly independent. View $w_1, w_2, \ldots, w_n$ as the columns of an n-by-n matrix W; then there exists an n-by-n matrix B such that $BW = I$. Let $Bw_{n+1} = c = [c_1\ c_2\ \cdots\ c_n]^\top$. Show that $w_{n+1} = c_1 w_1 + \cdots + c_n w_n$, proving that $w_1, w_2, \ldots, w_{n+1}$ are linearly dependent.

Now, for an arbitrary linear space V, assume that $w_1, w_2, \ldots, w_{n+1}$ lie in the span of $\{y_1, y_2, \ldots, y_n\}$. Then

$$w_i \quad = \quad \sum_{j=1}^{n} a_{ij} y_j \qquad (i = 1, 2, \ldots, n+1)$$

for some scalars a_{ij}. Let $A_i = [a_{i1}\ a_{i2}\ \cdots\ a_{in}]^\top$. Then the A_i's are $n+1$ vectors in $\mathbb{F}^n$, so they are linearly dependent. Hence $c_1 A_1 + \cdots + c_{n+1} A_{n+1} = 0$ for some scalars c_i that are not all 0. Now show that $c_1 w_1 + \cdots + c_{n+1} w_{n+1} = 0$.

11.26. Corollary. Let X be a linear space. Assume that X can be spanned by some finite subset of X. Let n be the smallest number of vectors that span X. Then X has at least one vector basis, any vector basis for X contains exactly n vectors, and X is isomorphic to $\mathbb{F}^n$. (We then say that X is **finite dimensional**, and we call n the **dimension** of the vector space X.)

11.27. *Example.* Let $\mathbb{F}$ be a field. The set of all polynomials of degree $\le n$, in one variable x, with coefficients in $\mathbb{F}$, is a linear space with dimension $n+1$, when the vector operations are defined in the obvious fashion. One vector basis is $\{1, x, x^2, \ldots, x^n\}$.

Choice and Vector Bases

11.28. *Remarks.* By using the Axiom of Choice, we shall now obtain further results about vector bases and cardinality in infinite-dimensional linear spaces. These results provide representations of linear spaces, which may be conceptually helpful to the beginner. However, the results below are optional; they will not be needed later except for pathological examples. These results have little practical value in applied mathematics or in functional analysis, because (1) the Axiom of Choice and its consequences are nonconstructive, and (2) the vector basis of an infinite-dimensional *topological* linear space generally has little connection with the topology of that space.

11.29. Many formulations of the Axiom of Choice were introduced in Chapter 6. We now state three more equivalents of Choice:

> **(AC16) Vector Basis Theorem (strong form).** Let X be a linear space over a field $\mathbb{F}$. Suppose that I is a linearly independent subset of X, G is a generating set (that is, $\operatorname{span}(G) = X$), and $I \subseteq G$. Then $I \subseteq B \subseteq G$ for some vector basis B.
>
> **(AC17) Vector Basis Theorem (intermediate form).** Let X be a linear space over a field $\mathbb{F}$, and let G be a subset of X that generates X (that is, $\operatorname{span}(G) = X$). Then X has a vector basis B contained in G.
>
> **(AC18) Vector Basis Theorem (weak form).** Every linear space has a vector basis.

Obviously (AC16) $\Rightarrow$ (AC17) $\Rightarrow$ (AC18).

Proof of (AC5) or (AC7) $\Rightarrow$ (AC16). Use 11.17(C).

Proof of (AC17) $\Rightarrow$ (AC2). This proof is due to Halpern [1966]. Let $\{S_\alpha : \alpha \in A\}$ be a nonempty set of nonempty disjoint sets; we wish to prove that there is a set S_0 consisting of exactly one element from each S_α. To this end, we shall construct not only a suitable vector space, but also a suitable scalar field.

Let $S = \bigcup_{\alpha \in A} S_\alpha$. Let $\mathbb{E}$ be a field disjoint from S; this can always be accomplished by relabeling. Let $\mathbb{F} = \mathbb{E}[S]$ be the field of rational functions with coefficients in $\mathbb{E}$ and variables in S (see 8.24). Form the external direct sum

$$\coprod_{\alpha \in A} \mathbb{F} \quad = \quad \{f \in \mathbb{F}^A \ : \ f(\alpha) \neq 0 \text{ for at most finitely many } \alpha \in A\}$$

as in 11.6.i; then Φ is a linear space over $\mathbb{F}$. For each $s \in S$ and $\alpha \in A$, let

$$g_s(\alpha) \quad = \quad \begin{cases} s & \text{if } s \in S_\alpha \\ 0 & \text{if } s \notin S_\alpha. \end{cases}$$

Then $g_s \in \Phi$ and in fact the set $G = \{g_s : s \in S\}$ spans Φ. Let $B \subseteq G$ be a vector basis for Φ over $\mathbb{F}$; then $B = \{g_s : s \in S_0\}$ for some set $S_0 \subseteq S$. For each $\alpha \in A$, the characteristic function of the singleton $\{\alpha\}$ is an element of Φ and thus in the span of B. Hence there is at least one $g_s \in B$ that does not vanish at α, and thus S_0 meets S_α. To show that S_0 meets each S_α in at most one point, suppose t, u are distinct members of $S_0 \cap S_\alpha$. Then $tg_u = ug_t$, contradicting the fact that the set $\{g_t, g_u\}$ is linearly independent.

Proof of (AC18) $\Rightarrow$ (MC). (Recall that (MC) was stated in 6.15.) This proof, given by Blass [1984], is similar but somewhat longer, and so we shall omit it. We remark that it uses Blass's subfield (8.25).

11.30. *Corollaries of the Vector Basis Theorem.*

a. Any $\mathbb{F}$-linear space can be represented as an external direct sum of copies of $\mathbb{F}$ (see 11.18.e).

b. If V and W are linear spaces over $\mathbb{F}$, $I \subseteq V$ is linearly independent, and $f : I \to W$ is any function, then f can be extended to a linear function from V into W. *Hint*: 11.18.d.

c. We may view $\mathbb{R}$ as a linear space over the scalar field $\mathbb{Q}$; a basis for this linear space is called a **Hamel basis**. (Some mathematicians apply that term more widely, as noted in 11.17.) Using such a basis, show that there exists a function $f : \mathbb{R} \to \mathbb{R}$ that is additive — that is, satisfying $f(s+t) = f(s) + f(t)$ — but not continuous. *Remark*: Compare this with 24.42.

d. If V is any linear space, its linear dual separates points of V.

e. Any complex linear space can be represented as the complexification of some real linear space (as defined in 11.11).

f. Let S be a linear subspace of a linear space X. Then S has an additive complement T — that is, X has a linear subspace T satisfying

$$S + T = X \qquad \text{and} \qquad S \cap T = \{0\},$$

or, equivalently, satisfying the condition that

each $x \in X$ can be written in one and only one way as $s + t$ with $s \in S$ and $t \in T$.

It may be instructive to contrast this with 8.16.

g. Let S be a linear subspace of a linear space X. Then S is the range of a linear projection — i.e., there exists a linear map $f : X \to S$ that has range S and satisfies $f(s) = s$ for each $s \in S$.

11.31. Theorem (Löwig, 1934). Let V be an $\mathbb{F}$-linear space. Then any two vector bases for V over $\mathbb{F}$ have the same cardinality. (That cardinality can therefore be called the **dimension** of the linear space.)

Proof. This proof is taken from Hall [1958]. Let S and T be vector bases for V. Each $s \in S$ can be expressed uniquely (except for the order of summation) in the form $s = a_1 t_1 + \cdots + a_n t_n$ for some positive integer n, some nonzero scalars $a_1, a_2, \ldots, a_n$, and some vectors $t_1, t_2, \ldots, t_n \in T$. Let $F(s)$ be the finite set $\{t_1, t_2, \ldots, t_n\}$ obtained in this fashion.

If $s_1, s_2, \ldots, s_k$ are distinct elements of S, then $F(s_1) \cup F(s_2) \cup \ldots \cup F(s_k)$ contains at least k elements, for otherwise $s_1, s_2, \ldots, s_k$ would be linearly dependent (by 11.25). By M. Hall's Marriage Theorem 6.37(ii), there exist points $t(s) \in F(s)$ such that the mapping $s \mapsto t(s)$ is injective; thus $\mathrm{card}(S) \le \mathrm{card}(T)$. Similarly $\mathrm{card}(T) \le \mathrm{card}(S)$; now apply the Schröder-Bernstein Theorem 2.19.

Dimension of the Linear Dual (Optional)

11.32. *Assumptions, notations, and remarks.* The results below make use of the fact (proved in 10.44.f) that $\mathrm{card}(\mathbb{R}) = \mathrm{card}(\mathbb{C}) = \mathrm{card}(2^{\mathbb{N}})$. Also, the results below assume the Axiom of Choice; these results should be contrasted with 27.47.a.

Throughout the discussion below, let $\mathbb{F}$ be the scalar field; assume $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. Let X be a linear space over $\mathbb{F}$, and let X^* be its linear dual — i.e., the set of all linear maps from X into $\mathbb{F}$.

11.33. *Observation.* If $\dim(X) = n < \infty$, then $\dim(X^*) = n$ also, and $\mathrm{card}(X) = \mathrm{card}(X^*) = \mathrm{card}(\mathbb{F})$. *Hint*: 11.22.

11.34. *Proposition.* $\mathrm{card}(X) = \max\{\mathrm{card}(\mathbb{F}), \dim(X)\}$.

Hints: Let B be any vector basis for X; then $\mathrm{card}(\mathbb{F} \times B) = \max\{\mathrm{card}(\mathbb{F}), \dim(X)\}$ by (AC13) in 6.22. Let $\bigsqcup_{b\in B} \mathbb{F}$ be the external direct sum (defined in 11.6.i) of B copies of $\mathbb{F}$. Use 11.18.e, 6.22, and 11.29 to explain

$$\begin{aligned}\mathrm{card}(X) \;=\; \mathrm{card}\left(\bigsqcup_{b\in B} \mathbb{F}\right) \;&\le\; \mathrm{card}\left(\bigcup_{n=1}^{\infty} (\mathbb{F} \times B)^n\right) \\ &=\; \mathrm{card}(\mathbb{F} \times B) \;\le\; \mathrm{card}(X \times X) \;=\; \mathrm{card}(X).\end{aligned}$$

Then use the Schröder-Bernstein Theorem.

11.35. Lemma. If X is infinite-dimensional, then $\dim(X^*) \ge \mathrm{card}(2^{\mathbb{N}})$.

Hints: Let $\{e_0, e_1, e_2, \ldots\}$ be any linearly independent sequence in X. For each real number r, by 11.30.b there exists some $f_r \in X^*$ satisfying $f_r(e_n) = r^n$ for $n = 0, 1, 2, \ldots$. Now apply 11.15, to show that the f_r's are linearly independent members of X^*.

11.36. Theorem. If X is infinite-dimensional, then $\dim(X^*) > \dim(X)$.

Proof. By the preceding results, we have $\dim(X^*) \ge \dim(2^{\mathbb{N}})$, hence $\dim(X^*) = \mathrm{card}(X^*)$. Let B be a vector basis for X; then B is an infinite set; hence $\mathrm{card}(\mathbb{N} \times B) = \mathrm{card}(B)$. Since the X^* is isomorphic to $\mathbb{F}^B$, we have $\mathrm{card}(X^*) = \mathrm{card}(\mathbb{F}^B) = \mathrm{card}((2^{\mathbb{N}})^B) = \mathrm{card}(2^{\mathbb{N}\times B}) = \mathrm{card}(2^B) > \mathrm{card}(B) = \dim(X)$.

Preview of Measure and Integration

11.37. *Definitions.* Let X be an additive monoid. (In most cases of interest X is either $[0,+\infty]$ or a vector space.) Let $\mathcal{S}$ be a collection of subsets of a set Ω with $\varnothing \in \mathcal{S}$, and let $\tau : \mathcal{S} \to X$ be some mapping satisfying $\tau(\varnothing) = 0$. We say that τ is

> **finitely additive** if $\tau(\bigcup_{j=1}^{n} S_j) = \sum_{j=1}^{n} \tau(S_j)$ whenever $S_1, S_2, \ldots, S_n$ are finitely many disjoint members of $\mathcal{S}$ whose union is also a member of $\mathcal{S}$;
>
> **countably additive** (or **σ-additive**) if X is equipped with a metric (or other notion of convergence) and $\tau(\bigcup_{j=1}^{\infty} S_j) = \sum_{j=1}^{\infty} \tau(S_j)$ whenever $S_1, S_2, S_3, \ldots$ is a sequence of disjoint members of $\mathcal{S}$ whose union is also a member of $\mathcal{S}$.

The expression $\sum_{j=1}^{\infty} \tau(S_j)$ is defined as in 10.39.

Of course, every countably additive mapping is also finitely additive, since we may take S_3, S_4, S_5, ... all equal to $\varnothing$. We emphasize that "finitely additive" means "at least finitely additive, and perhaps countably additive;" it does not mean "finitely additive but not countably additive."

Aside from the requirements $\varnothing \in \mathcal{S} \subseteq \mathcal{P}(\Omega)$, the collection $\mathcal{S}$ in the definition above is arbitrary. We now impose some additional restrictions. By a **charge** we shall mean a finitely additive mapping from an *algebra* of sets into an additive monoid. By a **measure** we shall mean a countably additive mapping from a *σ-algebra* of sets into an additive monoid equipped with some convergence structure.

Cautions: The terminology varies considerably throughout the literature. Some mathematicians apply the term "measures" to what we have called charges, or to countably additive charges, or to positive measures (defined below), etc.

Unfortunately, the phrase "μ is a charge (or measure) on W" has two different meanings in the literature: It may mean W is the (σ-)algebra $\mathcal{S}$ on which μ is defined, or it may mean that W is the underlying set Ω on which $\mathcal{S}$ is defined. One must determine from context just which meaning is intended.

11.38. *Remarks on the choice of the codomain X.* In most applications of charges, the monoid X usually is either $[0,+\infty]$ or some vector space; then μ may be called a **positive charge** or a **vector charge**, respectively. Though a wide variety of vector spaces are used in this fashion in spectral theory, in more elementary applications the vector spaces most often used for the monoid X are the one-dimensional vector spaces $\mathbb{R}$ and $\mathbb{C}$. The resulting charge or measure is then called a **real-valued charge or measure** or a **complex charge or measure**, respectively. We shall study positive charges and measures in 21.9 and thereafter; real-valued charges and measures in 11.47 and thereafter; and other vector charges and measures in 29.3 and thereafter.

Positive charges and vector charges differ only slightly in their definition, but more substantially in their use. We are mainly interested in positive charges when they are in fact measures; moreover, it is commonplace to fix one particular positive measure μ and then use it for many different purposes. In contrast, vector charges are sometimes of interest without countable additivity or σ-algebras, but they are of interest mainly in large collections — i.e., we may study the relationships between many different vector charges, which are members of a "space of charges" as in 11.47. An important part of the theory of vector measures ν is the question of just when they can be represented in the form

$$\nu(S) \quad = \quad \int_S f(\omega)\, d\mu(\omega)$$

for some vector-valued function f and some positive measure μ; see 29.20 and 29.21.

11.39. *Remarks on the choice of the domain* $\mathcal{S}$. In most of our elementary examples of charges or measures later in this book, the collection of sets $\mathcal{S}$ is actually equal to $\mathcal{P}(\Omega) = \{\text{subsets of } \Omega\}$. However, our most important measure is Lebesgue measure, which is not so elementary and which is not defined on a σ-algebra of the form $\mathcal{P}(\Omega)$; in 21.22 we prove it cannot be extended in a natural way to $\mathcal{P}(\Omega)$.

In many cases of interest, Ω is a topological space, and $\mathcal{S}$ is either the Borel σ-algebra or some σ-algebra containing the Borel σ-algebra. Recall that the **Borel σ-algebra** is the σ-algebra on Ω generated by the topology — i.e., the smallest σ-algebra containing all the open sets; the members of that σ-algebra are called **Borel sets**.

A **measurable space** is a pair $(\Omega, \mathcal{S})$ consisting of a set Ω and a σ-algebra $\mathcal{S}$ of subsets of Ω; a **measure space** is a triple $(\Omega, \mathcal{S}, \mu)$ in which μ is a positive measure on $\mathcal{S}$. (It might be more descriptive to call $(\Omega, \mathcal{S}, \mu)$ a "positive measure space," but we shall not be concerned with a measure "space" in which μ is a vector measure.)

Thus, a *measurable space* is a space that is capable of being equipped with a measure; a *measure space* is a space that has been equipped with a positive measure. These terms should not be confused with each other, or with a *space of measures* — i.e., a collection of measures equipped with some structure that makes the collection into a vector space, a topological space, or some other sort of "space," as in 11.48.

11.40. Several kinds of integrals will be introduced in this book; still more integrals can be found in the wider literature. When necessary, we shall specify what kind of integral is being used. Fortunately, the several integrals generally agree in those cases where they are all defined. For instance, $\int_0^1 t^2\, dt$ makes sense as a Riemann integral or as a Lebesgue integral, but with either interpretation the expression has the value of $1/3$.

We now informally sketch some of the main features shared by most types of integrals. Precise definitions will be given later.

In general, an integral $\int_S f\, d\mu$ depends on a set S, a function f (called the **integrand**), and a charge μ. In some of our studies of integrals, we may hold one or two of the arguments S, f, μ fixed. When an argument is held fixed and/or its value is understood, then it may be supressed from the notation; thus

$$\int_S f\, d\mu \qquad \text{may be written as} \qquad \int_S f \qquad \text{or} \qquad \int f\, d\mu \qquad \text{or} \qquad \int f.$$

Usually, when S is omitted from the notation, then S is understood to be equal to Ω. When Ω is a subset of $\mathbb{R}^n$ and μ is Lebesgue measure, then $d\mu(\omega)$ may be written simply as $d\omega$.

The integral $\int_S f\, d\mu$ may be written in greater detail as $\int_S f(\omega)\, d\mu(\omega)$. Here ω is a dummy variable, or placeholder. It is sometimes helpful in clarifying just what is the argument of f, particularly if the function f is complicated. The integral is not altered in value if we replace ω with some other letter, or omit it altogether. Thus:

$$\int_S f(\omega)\,d\mu(\omega) \quad = \quad \int_S f(\lambda)\,d\mu(\lambda) \quad = \quad \int_S f(\cdot)\,d\mu(\cdot) \quad = \quad \int_S f\,d\mu.$$

11.41. Using (σ-)algebras and charges, we shall consider integrals $\int f\,d\mu$ of three main types in this book:

(i) μ is a vector charge taking values in a complete normed vector space, and f is a scalar-valued function taking values in the scalar field of that vector space. Then $\int f d\mu$ takes values in the vector space.

We shall call this a **Bartle integral** (though the terminology varies in the literature); this type of integral is introduced in 29.30. The mapping $(f,\mu) \mapsto \int f\,d\mu$ is bilinear — i.e., linear in each variable when the other variable is held fixed. For f and μ held fixed, the mapping $S \mapsto \int_S f\,d\mu$ is finitely additive; i.e., it is a vector charge.

This is algebraically the simplest type of integral we shall consider. We modify this concept in a couple of ways, indicated below, to allow $+\infty$ in our computations.

(ii) μ is a positive measure (and thus may take the value $+\infty$), f is a function taking values in some complete normed vector space, and some restriction is placed on $\|f(\cdot)\|$ so that it is "not too big." Then $\int f d\mu$ takes values in the vector space.

We shall call this a **Bochner integral**; it is introduced in 23.16. It is a linear function of f, for fixed μ. For fixed f, the mapping $\mu \mapsto \int f\,d\mu$ is like the "upper half" of a linear map: It preserves sums and multiplication by positive constants. For f and μ fixed, the mapping $S \mapsto \int_S f\,d\mu$ is countably additive — i.e., it is a vector measure. A central result for Bochner integrals is Lebesgue's Dominated Convergence Theorem, 22.29.

(iii) μ and f both take values in $[0,+\infty]$, and $\int f d\mu$ does, too.

We shall call this a **positive integral**; it is introduced in 21.36. It behaves like the "positive quadrant" of a bilinear mapping: The maps $f \mapsto \int f\,d\mu$ and $\mu \mapsto \int f\,d\mu$ both preserve sums and multiplication by positive constants. For f and μ fixed, the mapping $S \mapsto \int_S f\,d\mu$ is countably additive — i.e., it is a positive measure. A central result for positive integrals is Lebesgue's Monotone Convergence Theorem, 21.38(ii).

We emphasize that for integrals of this type, $\int f\,d\mu$ may take the value $+\infty$. When $\int f d\mu$ exists and is *finite*, we say that f is **integrable**.

Other types of integrals over charges are possible, of course. For instance, for any vector spaces X, Y, Z, we could integrate an X-valued function f with respect to a Y-valued measure μ, using some bilinear map $\langle\ ,\ \rangle : X \times Y \to Z$; then $\int f\,d\mu$ takes place in Z. However, such integrals will not be studied in this book.

A few other integrals will be defined in other fashions, not in terms of charges and algebras. The **Riemann integral** $\int_a^b f(t)\,dt$ is reviewed in Chapter 24; in that chapter we also introduce the **Henstock integral** $\int_a^b f(t)\,dt$ and the **Henstock-Stieltjes integral** $\int_a^b f(t)\,d\varphi(t)$, and show how these integrals are related to the Lebesgue integral. Here f and φ are functions defined on an interval $[a,b]$.

11.42. *Integration of simple functions.* Let $\mathcal{A}$ be an algebra of subsets of a set Ω. A function $f : \Omega \to X$ is called a **simple function** if

> the range of f is a finite subset of X, and $f^{-1}(x) \in \mathcal{A}$ for each $x \in X$ (or equivalently, for each $x \in \operatorname{Ran}(f)$).

Equivalently, a simple function is one that can be written in the form

$$f(\cdot) \quad = \quad \sum_{j=1}^{n} 1_{S_j}(\cdot)x_j \qquad (*)$$

where n is a positive integer, the x_j's are members of X, and $1_{S_j}(\cdot)$ is the characteristic function of some set $S_j \in \mathcal{A}$. (The representation $(*)$ is not unique, since we do not require the x_j's to be nonzero or distinct and we do not require the S_j's to be disjoint.)

If X is a vector space, then it is easy to verify that the simple functions form a linear subspace of X^{Ω}. If $X = [0, +\infty]$, the set of simple functions is not a linear space, but at least it acts like the "upper half" of a linear space: It is closed under addition and under multiplication by nonnegative constants.

Now let μ be a charge defined on $\mathcal{A}$, taking values in some monoid K, and let $f : \Omega \to X$ be a simple function. When it makes sense, we define

$$\int_{\Omega} f\,d\mu \quad = \quad \sum_{x} \mu\left(f^{-1}(x)\right)x.$$

The summation on the right is over all $x \in X$ or, equivalently, (since $\mu(\varnothing) = 0$) the summation is over all $x \in \operatorname{Ran}(f)$. Thus, the summation involves only finitely many terms. Equivalently, if f is represented by $(*)$, then

$$\int_{\Omega} f\,d\mu \quad = \quad \sum_{j=1}^{n} \mu(S_j)\,x_j.$$

For these summations to make sense, we must also make certain restrictions: We must have some notion of how to multiply x times $\mu\left(f^{-1}(x)\right)$ and how to add up the resulting products. This requirement is met by any simple function, in cases 11.41(i) and 11.41(iii).

In case 11.41(ii), the requirement is met by any simple function f that satisfies this additional hypothesis:

$$\mu\Big(\{\omega \in \Omega : f(\omega) \neq 0\}\Big) \quad < \quad \infty.$$

In this case we say that f is an **integrable simple function**. If we use representation $(*)$, then we must choose the S_j's so that no nonzero vector x_j is associated with a set S_j that has infinite measure. (That is accomplished, for instance, if we require that f be an integrable simple function and the S_j's be disjoint.)

11.43. Simple functions should not be confused with step functions, though the two notions are closely related. A **step function** is a mapping $f : [a, b] \to X$, from some subinterval of $\mathbb{R}$ into some vector space, with the property that there exists some partition $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ such that f is constant on each subinterval (t_{j-1}, t_j). (Different partitions may be used for different step functions.)

Step functions are a special case of simple functions, as follows: Let $\mathcal{A}$ be the collection of all finite unions of subintervals of $[a, b]$. (We interpret "subintervals" so that singletons and the empty set belong to $\mathcal{A}$.) Then $\mathcal{A}$ is an algebra of sets, and the resulting X-valued simple functions (defined as in 11.42) are precisely the step functions.

Ordered Vector Spaces

11.44. *Remarks.* We shall only consider ordered vector spaces using $\mathbb{R}$ for the scalar field. It is possible to develop a theory of ordered vector spaces using other scalar fields — see, for instance, Schaefer [1971] — but such a theory is more complicated and less natural and intuitively appealing; it is not recommended for beginners.

Definitions. An **ordered vector space** is a real vector space X equipped with a partial ordering $\preccurlyeq$ such that

(i) $x \preccurlyeq y \iff x + u \preccurlyeq y + u$ (i.e., X is an ordered group); and

(ii) If $x \succcurlyeq 0$ in X and $r \geq 0$ in $\mathbb{R}$, then $rx \succcurlyeq 0$ in X.

We say X is a **Riesz space**, or **vector lattice**, if in addition

(iii) $(X, \preccurlyeq)$ is a lattice — i.e., each finite nonempty subset of X has a supremum and an infimum.

Finally, X is a **lattice algebra** (or **algebra lattice**) if X is also an algebra (in the classical sense, as in 11.3) whose vector multiplication satisfies

(iv) $x, y \succcurlyeq 0 \Rightarrow xy \succcurlyeq 0$.

If X is a Riesz space, then a **Riesz subspace** is a subset S that is closed under the vector operations and the lattice operations — that is,

$$s, t \in S,\ c \in \mathbb{R} \qquad \Rightarrow \qquad s + t,\ cs,\ s \vee t,\ s \wedge t\ \in S.$$

Clearly, such a set is itself a Riesz space, when equipped with the restriction of the operations of X.

11.45. *Example: real-valued functions.* Let Λ be any set. Then the product

$$\mathbb{R}^\Lambda \quad = \quad \{\text{functions from } \Lambda \text{ into } \mathbb{R}\}$$

is a Dedekind complete lattice algebra, when given the product ordering — that is, when ordered by

$$x \preccurlyeq y \qquad \text{if} \qquad x(\lambda) \leq y(\lambda) \quad \text{for every} \quad \lambda \in \Lambda.$$

The vector and lattice operations are defined pointwise:

$$(x+y)(\lambda) \;=\; x(\lambda)+y(\lambda), \qquad (x\cdot y)(\lambda) \;=\; x(\lambda)\cdot y(\lambda),$$

$$(x\vee y)(\lambda) \;=\; \max\{x(\lambda),y(\lambda)\}, \qquad (x\wedge y)(\lambda) \;=\; \min\{x(\lambda),y(\lambda)\}.$$

More generally, for any set $S \subseteq \mathbb{R}^\Lambda$ that is bounded above or below by some real-valued function, we have

$$[\sup(S)]\,(\lambda) \;=\; \sup\{s(\lambda)\ :\ s\in S\}, \qquad [\inf(S)]\,(\lambda) \;=\; \inf\{s(\lambda)\ :\ s\in S\}.$$

When this ordering is used, many mathematicians write $x \leq y$ instead of $x \preccurlyeq y$. However, in this book we shall often write $\preccurlyeq$ for such an ordering, to help beginners avoid inadvertently attributing familiar properties (e.g., a chain ordering) to a familiar symbol.

11.46. *Further examples: subspaces of* $\mathbb{R}^\Lambda$. The pointwise formulas given for $x\vee y$, $x\wedge y$, $\sup(S)$, $\inf(S)$ in the previous paragraph remain valid in many important subsets of $\mathbb{R}^\Lambda$; some of these are listed below.

a. The set $B(\Lambda)=\{$bounded functions from Λ into $\mathbb{R}\}$ is a Dedekind complete lattice subalgebra of $\mathbb{R}^\Lambda$.

b. The space $C[a,b]=\{$continuous functions from $[a,b]$ into $\mathbb{R}\}$ is a lattice subalgebra of $\mathbb{R}^{[a,b]}$, for any real numbers a,b with $a<b$.

$C[a,b]$ is *not* Dedekind complete. *Example.* Show that the sequence of functions $f_n(t) = \sqrt[n]{\max\{0,t\}}$ is bounded above in $C[-1,1]$ but does not have a least upper bound in $C[-1,1]$.

c. The space $C^1[a,b] = \{$continuously differentiable functions from $[a,b]$ into $\mathbb{R}\}$ is a subalgebra of $\mathbb{R}^{[a,b]}$ — i.e., it is closed under addition and both multiplications. Also, it is an ordered vector space.

$C^1[a,b]$ is not a lattice. *Example.* Let $x(t)=t$ and $y(t)=-t$. Show that the set $\{x,y\}$ has an upper bound in $C^1[-1,1]$, but not a least upper bound.

d. If we use $\mathbb{R}$ for the scalar field, then many of the Banach spaces used in the theory of measure and integration are vector lattices. They are not subspaces of $\mathbb{R}^\Lambda$; rather, they are subspaces of a quotient space $\mathbb{R}^\Lambda/\mathfrak{J}$ for some ideal $\mathfrak{J}$. Examples will be developed in later chapters.

11.47. The space of bounded real charges. Let Ω be a set, let $\mathcal{A}$ be an algebra of subsets of Ω, and let

$$ba(\mathcal{A},\mathbb{R}) \quad = \quad \{\text{bounded, real-valued charges on } \mathcal{A}\}.$$

(Here, "ba" stands for "bounded additive.") Then $ba(\mathcal{A},\mathbb{R})$ is a linear subspace of $B(\mathcal{A},\mathbb{R}) = \{$bounded functions from $\mathcal{A}$ into $\mathbb{R}\}$, which is in turn a linear subspace of $\mathbb{R}^{\mathcal{A}} = \{$functions from $\mathcal{A}$ into $\mathbb{R}\}$.

Let $ba(\mathcal{A},\mathbb{R})$ be equipped with the restriction of the product ordering — that is,

$$\mu \preccurlyeq \nu \qquad \text{means that} \qquad \mu(A)\leq\nu(A) \quad \text{for every } A\in\mathcal{A}.$$

Then $ba(\mathcal{A},\mathbb{R})$ is a Dedekind complete vector lattice, with lattice operations as follows:

$$\begin{aligned}(\mu \vee \nu)(A) &= \sup\{\mu(B)+\nu(A\setminus B)\ :\ B\in\mathcal{A},\ B\subseteq A\},\\ (\mu \wedge \nu)(A) &= \inf\{\mu(B)+\nu(A\setminus B)\ :\ B\in\mathcal{A},\ B\subseteq A\}.\end{aligned}$$

Although $ba(\mathcal{A},\mathbb{R})$ is a linear subspace of $\mathbb{R}^{\mathcal{A}}$ as linear spaces, it is not a sublattice. The lattice operations $\vee$ and $\wedge$ shown in the preceding paragraph are *not* simply the restrictions of the lattice operations of $\mathbb{R}^{\mathcal{A}}$. Indeed, $ba(\mathcal{A},\mathbb{R})$, considered as a subset of $\mathbb{R}^{\mathcal{A}}$, is not closed under that space's lattice operations; an elementary example of this is given in the exercise in 21.11.c.

Since $ba(\mathcal{A},\mathbb{R})$ is a vector lattice, each charge μ has a positive part, a negative part, and an absolute value, as defined in 8.39. In the present context those are

$$\begin{aligned}\mu^+(A) &= \sup\{\ \mu(S)\ :\ S\in\mathcal{A},\ S\subseteq A\},\\ \mu^-(A) &= \sup\{\ -\mu(S)\ :\ S\in\mathcal{A},\ S\subseteq A\},\\ /\mu/(A) &= \sup\{\ \mu(S)-\mu(A\setminus S)\ :\ S\in\mathcal{A},\ S\subseteq A\},\end{aligned}$$

respectively. The three functions μ^+, μ^-, and $/\mu/$ are called the **positive variation**, the **negative variation**, and the **variation** (or *total variation*) of μ, respectively; they are members of $ba(\mathcal{A},\mathbb{R})$. The variation of μ may also be written as $\mathrm{Var}(\mu)$. We emphasize that *any bounded real charge has finite variation*; this fact will be important in 29.6.d and 29.6.h.

The lattice $ba(\mathcal{A},\mathbb{R})$ is Dedekind complete. If M is a nonempty subset of $ba(\mathcal{A},\mathbb{R})$, bounded above or below by some member of $ba(\mathcal{A},\mathbb{R})$, then we have

$$[\sup(M)](A) = \sup \sum_{j=1}^{n} \mu_j(S_j) \qquad \text{or} \qquad [\inf(M)](A) = \inf \sum_{j=1}^{n} \mu_j(S_j),$$

respectively, where the sup or inf is over all positive integers n, all finite sets of charges $\{\mu_1,\mu_2,\ldots,\mu_n\}\subseteq M$, and partitions $A=S_1\cup S_2\cup\cdots\cup S_n$ where the S_j's are disjoint elements of $\mathcal{A}$. If $(M,\preccurlyeq)$ or $(M,\succcurlyeq)$ is a directed set, then we obtain these simpler formulas, respectively:

$$[\sup(M)](A) = \sup_{\mu\in M}\mu(A) \qquad \text{or} \qquad [\inf(M)](A) = \inf_{\mu\in M}\mu(A).$$

(*Hints*: Since $ba(\mathcal{A},\mathbb{R})$ is a lattice, we easily reduce the proof to the case where M is a directed set; then use the fact that a setwise limit of charges is a charge.)

The space $ba(\mathcal{A},\mathbb{R})$ is studied in greater depth by Bhaskara Rao and Bhaskara Rao [1983].

11.48. The space of bounded, countably additive real charges. Let Ω be a set, let $\mathcal{A}$ be an algebra of subsets of Ω, and let

$$ca(\mathcal{A},\mathbb{R}) = \{\text{bounded, countably additive, real-valued charges on } \mathcal{A}\}.$$

(We emphasize that $\mathcal{A}$ is not assumed to be a σ-algebra, so the members of $ca(\mathcal{A}, \mathbb{R})$ are not necessarily measures.) Then $ca(\mathcal{A}, \mathbb{R})$ is a sublattice of $ba(\mathcal{A}, \mathbb{R})$ — that is, $ca(\mathcal{A}, \mathbb{R})$ is closed under the binary operations $\vee$ and $\wedge$ of $ba(\mathcal{A}, \mathbb{R})$. (*Exercise.*)

Moreover, $ca(\mathcal{A}, \mathbb{R})$ is Dedekind complete. If M is a subset of $ca(\mathcal{A}, \mathbb{R})$ that is bounded above or below by some member of $ca(\mathcal{A}, \mathbb{R})$, then

$$[\sup(M)](A) = \sup \sum_{j=1}^{\infty} \mu_j(S_j) \qquad \text{or} \qquad [\inf(M)](A) = \inf \sum_{j=1}^{\infty} \mu_j(S_j)$$

respectively, where the sup or inf is over all countable collections $\{\mu_1, \mu_2, \mu_3, \ldots\} \subseteq M$, and partitions $A = S_1 \cup S_2 \cup S_3 \cup \cdots$, where the S_j's are disjoint elements of $\mathcal{A}$. (*Exercise.* Verify.)

11.49. *Notes on the boundedness of charges.*

a. In 29.3 we shall prove that any real-valued measure (i.e., countably additive, on a σ-algebra) is bounded. In fact, any measure taking values in a Banach space is bounded.

b. A finitely additive charge on an algebra of sets need not be bounded. For example:

Let $\mathcal{A} = \{S \subseteq \mathbb{N} : S \text{ is either finite or cofinite}\}$; this is an algebra (but not a σ-algebra) of subsets of $\mathbb{N}$. Define $\lambda : \mathcal{A} \to \mathbb{Z}$ by

$$\lambda(S) = \begin{cases} \operatorname{card}(S) & \text{if } S \text{ is finite} \\ -\operatorname{card}(\mathbb{N} \setminus S) & \text{if } S \text{ is cofinite.} \end{cases}$$

Verify that λ is a real-valued charge that is unbounded.

c. Are there any real-valued charges on a σ-algebra that are not countably additive? Well, yes and no. Such objects exist, but explicitly constructible examples of such objects do not exist. This is discussed further in 29.37.

11.50. *Some basic properties of Riesz spaces.* If X is a Riesz space, then X is a lattice group, so it has all the properties of lattice groups listed earlier in this chapter. It also has the following properties:

a. $r(x \vee y) = (rx) \vee (ry)$ and $r(x \wedge y) = (rx) \wedge (ry)$ for all $x, y \in X$ and any real number $r > 0$. Hence also $/rx/ = r/x/$ and $(rx)^+ = r(x^+)$.

b. $x \vee y = \frac{1}{2}\Big(x + y + /x - y/\Big)$ and $x \wedge y = \frac{1}{2}\Big(x + y - /x - y/\Big)$.

c. $x \vee (-x) = /x/ \succcurlyeq 0$.

d. $-y \preccurlyeq y \iff y \succcurlyeq 0$.

11.51. Proposition. Let X be a Riesz space. Then X has the same ideals, whether we view X as a Riesz space or (by "forgetting" how to multiply by scalars) we view X as a lattice group. Thus, an ideal in a Riesz space X is an additive subgroup satisfying any of the conditions in 9.27.

Proof. Since the category of lattice groups has fewer fundamental operations, it has at least as many ideals — i.e., every Riesz space ideal is a lattice group ideal. Conversely, suppose

$S \subseteq X$ is a lattice group ideal; we must show that it is a Riesz space ideal. We shall use the fact that S is solid (established in 9.27(B)). To show that S satisfies definition 9.25(B) for Riesz spaces, it suffices to show that

$$c \in \mathbb{R}, \qquad s \in S \qquad \Rightarrow \qquad cs \in S.$$

Since S is an additive subgroup, it suffices to prove this implication in the case where $c > 0$. Since $0 \preccurlyeq s^+ \preccurlyeq /s/$ and $0 \preccurlyeq s^- \preccurlyeq /s/$, we have $s^+, s^- \in S$. Since S is closed under addition, $ms^+, ms^- \in S$ for any positive integer m. Let m be some integer greater than c. Since X is a Riesz space, we obtain

$$0 \preccurlyeq cs^+ \preccurlyeq ms^+ \qquad \text{and} \qquad 0 \preccurlyeq cs^- \preccurlyeq ms^-,$$

and therefore $cs^+, cs^- \in S$. Now use the Jordan decomposition:

$$cs \quad = \quad c \cdot (s^+ - s^-) \quad = \quad (c \cdot s^+) - (c \cdot s^-).$$

Since S is an additive group, it follows that $cs \in S$.

Positive Operators

11.52. *Definitions.* Let X and Y be lattices (not necessarily groups or vector spaces). A mapping $f : X \to Y$ is

> a **lattice homomorphism** if it satisfies $f(x_1 \vee x_2) = f(x_1) \vee f(x_2)$ and $f(x_1 \wedge x_2) = f(x_1) \wedge f(x_2)$.
>
> **increasing** (or **isotone**) if $x_1 \preccurlyeq x_2 \ \Rightarrow\ f(x_1) \preccurlyeq f(x_2)$.
>
> **order bounded** if the image of any order interval is contained in an order interval — i.e., for any $x_1, x_2 \in X$ there exist $y_1, y_2 \in Y$ such that
>
> $$f\left(\{x \in X \ : \ x_1 \preccurlyeq x \preccurlyeq x_2\}\right) \quad \subseteq \quad \{y \in Y \ : \ y_1 \preccurlyeq y \preccurlyeq y_2\}.$$

It is clear that f is a lattice homomorphism $\Rightarrow$ f is increasing $\Rightarrow$ f is order bounded. Any of these three types of functions can be used as the morphisms for a category, with lattices for the objects.

Note that a linear operator between Riesz spaces (or more generally, an additive mapping between ordered groups) is increasing if and only if it is a **positive operator** — i.e., if and only if it satisfies $x \succcurlyeq 0 \ \Rightarrow\ f(x) \succcurlyeq 0$.

11.53. *Proposition.* Let X and Y be Riesz spaces; assume that Y is Archimedean (defined as in 10.3). Let $f : X \to Y$ be an additive, increasing map — that is,

$$f(x_1 + x_2) = f(x_1) + f(x_2), \qquad \text{and} \qquad x_1 \preccurlyeq x_2 \ \Rightarrow\ f(x_1) \preccurlyeq f(x_2).$$

Then f is $\mathbb{R}$-linear.

Corollary. Every lattice group homomorphism from a Riesz space into an Archimedean Riesz space is actually a Riesz space homomorphism.

Proof of proposition. It suffices to show that $f(rx) = rf(x)$ for every real number r and every vector $x \in X$. By additivity and the Jordan Decomposition, it suffices to prove that equation when $r \geq 0$ and $x \succcurlyeq 0$.

By additivity, it is easy to see that $f(qx) = qf(x)$ for *rational* numbers q. Since f is order-preserving, we can conclude that

$$x \in X^+, \quad 0 < q_1 \leq r \leq q_2, \quad q_1, q_2 \in \mathbb{Q} \qquad \Rightarrow \qquad q_1 f(x) \preccurlyeq f(rx) \preccurlyeq q_2 f(x).$$

Now, for any integer $m \in \mathbb{N}$, we can find rational numbers $q_1, q_2 > 0$ such that $-\frac{1}{m} < q_1 - r < 0 < q_2 - r < \frac{1}{m}$. It follows that

$$-\frac{1}{m} f(x) \;\preccurlyeq\; (q_1 - r)f(x) \;\preccurlyeq\; f(rx) - rf(x) \;\preccurlyeq\; (q_2 - r)f(x) \;\preccurlyeq\; \frac{1}{m} f(x).$$

Let $\gamma = f(rx) - rf(x)$; it follows that the subgroup $\mathbb{Z}\gamma = \{m\gamma : m \in \mathbb{Z}\}$ is bounded above by $f(x)$. Since Y is Archimedean, it follows that $\gamma = 0$.

11.54. *A pathological example.* In the preceding theorem, we cannot omit the assumption that Y be Archimedean. To see this, let $\mathbb{H}$ be the hyperreal line (see 10.18). We shall prove the existence of a mapping $f : \mathbb{R} \to \mathbb{H}$ that is a homomorphism for lattice groups but is not $\mathbb{R}$-linear.

First represent $\mathbb{R}$ as an internal direct sum, $\mathbb{R} = X \oplus Y$, where X and Y are some additive subgroups of $\mathbb{R}$ other than $\{0\}$ and $\mathbb{R}$ itself. (This can be accomplished using 11.30.a, since $\mathbb{R}$ may be viewed as a linear space over the scalar field $\mathbb{Q}$.) Let ε be a nonzero infinitesimal in $\mathbb{H}$. Define $f : \mathbb{R} \to \mathbb{H}$ by taking $f(x + y) = x + (1 + \varepsilon)y$ for all $x \in X$ and $y \in Y$. Then f is clearly additive. It is not linear, for if x, y are nonzero real numbers with $x \in X$ and $y \in Y$ then $yf(x) = yx \neq (1 + \varepsilon)xy = xf(y)$. It suffices to show that f is order-preserving. Suppose $x_1 + y_1 < x_2 + y_2$, where $x_1, x_2 \in X$ and $y_1, y_2 \in Y$. Then $x_2 + y_2 - x_1 - y_1$ is a positive real number and $(y_1 - y_2)\varepsilon$ is an infinitesimal. Hence $(y_1 - y_2)\varepsilon < x_2 + y_2 - x_1 - y_1$. That is, $f(x_1 + y_1) < f(x_2 + y_2)$.

(This example disproves an erroneous assertion of Birkhoff [1967, page 349].)

11.55. *Proposition (Kantorovič).* Let X, Y be Riesz spaces, and assume Y is Archimedean. Let $f : X_+ \to Y_+$ be any function. Then f extends to a positive operator $F : X \to Y$ if and only if f is additive — i.e., if and only if $f(x_1 + x_2) = f(x_1) + f(x_2)$ for all $x_1, x_2 \in X_+$. If that condition is satisfied, then the extension F is uniquely determined: It satisfies

$$F(x) \quad = \quad f(x^+) - f(x^-). \tag{$**$}$$

Proof. This proof follows the presentation of Aliprantis and Burkinshaw [1985]. Obviously, if f extends to a positive linear operator, then f must be additive and the extension F must satisfy the formula $(**)$. Conversely, assume f is additive and define $F : X \to Y$ by $(**)$; we must show that F is linear. The proof will be in several steps:

a. If $x = u - v$ with $u, v \in X_+$, then $F(x) = f(u) - f(v)$. *Hint*: $x^+ - x^- = x = u - v$, hence $x^+ + v = u + x^-$; now use our assumption that f is additive on X_+.

b. $F(x_1 + x_2) = F(x_1) + F(x_2)$; that is, F is additive on X. *Hint*: Apply the preceding result with $u = x_1^+ + x_2^+$ and $v = x_1^- + x_2^-$.

c. F is an increasing function on X. *Hint*: If $x \succcurlyeq 0 \Rightarrow F(x) = f(x) \succcurlyeq 0$.

Finally, apply 11.53 to complete the proof.

11.56. *Observations.* Let X and Y be Riesz spaces. Then:

a. $\mathcal{L}_b(X, Y) = \{$order bounded linear maps from X into $Y\}$ is a linear subspace of $\mathcal{L}(X, Y) = \{$linear maps from X into $Y\}$.

b. Let $f, g \in \mathcal{L}(X, Y)$. Then the following conditions are equivalent:

(A) $f - g$ is an increasing operator — that is, $x_1 \preccurlyeq x_2 \Rightarrow f(x_1) - g(x_1) \preccurlyeq f(x_2) - g(x_2)$.

(B) $f - g$ is a positive operator — that is, $x \succcurlyeq 0 \Rightarrow f(x) - g(x) \succcurlyeq 0$.

(C) $f(x) \preccurlyeq g(x)$ for all $x \in X_+$. In other words, the restriction of f to X_+ is larger than or equal to the restriction of g to X_+, where functions on X_+ are ordered by the pointwise ordering — i.e., where $\mathbb{R}^{X_+}$ is equipped with the product ordering.

When either (hence both) of these conditions holds, we shall write $f \preccurlyeq g$. This ordering makes $\mathcal{L}(X, Y)$ and $\mathcal{L}_b(X, Y)$ into ordered vector spaces.

11.57. Theorem (Riesz-Kantorovič). Let X and Y be Riesz spaces, and suppose Y is Dedekind complete. Then the linear space

$$\mathcal{L}_b(X, Y) \quad = \quad \{\text{order bounded linear operators from } X \text{ into } Y\}$$

is equal to the set of all linear operators that can be written as the difference of two positive operators. Furthermore, $\mathcal{L}_b(X, Y)$ is a Dedekind complete Riesz space when ordered as in 11.56.b. For any $f \in \mathcal{L}_b(X, Y)$, the positive part is given by this formula:

$$f^+(x) \quad = \quad \sup\{f(u) \ : \ u \in [0, x]\} \qquad \text{when } x \in X_+.$$

Other lattice operations are as follows, for $x \in X_+$:

$$\begin{aligned}
(f \vee g)(x) \quad &= \quad \sup\{f(u) + g(v) \ : \ u, v \in X_+ \text{ and } u + v = x\}, \\
(f \wedge g)(x) \quad &= \quad \inf\{f(u) + g(v) \ : \ u, v \in X_+ \text{ and } u + v = x\}, \\
/f/(x) \quad &= \quad \sup\{f(u) \ : \ u \in [-x, x]\} \quad = \quad \sup\{/f(u)/ \ : \ u \in [-x, x]\}.
\end{aligned}$$

When Φ is a nonempty subset of $\mathcal{L}_b(X, Y)$ that is directed and that is bounded above by some member of $\mathcal{L}_b(X, Y)$, then

$$(\sup \Phi)(x) \quad = \quad \sup_{f \in \Phi} f(x) \qquad \text{for each } x \in X_+.$$

Caution: A formula above shows the relation between $/f/(x)$ and $/f(x)/$. In general they are *not* the same; do not confuse them. In the expression $/f/(x)$, we take the absolute value of the vector f in the lattice $\mathcal{L}_b(X,Y)$; it is a function from X into Y that can be evaluated at x. On the other hand, $f(x)$ is a vector in the lattice Y, and so we can take its absolute value in that lattice to obtain $/f(x)/ \in Y$.

Proof of theorem. Our proof is based on the presentation of Fremlin [1974]). It is easy to show that any positive operator is order bounded; hence any difference of two positive operators is order bounded.

Conversely, suppose $f : X \to Y$ is order bounded. Define a function $g : X_+ \to Y_+$ by $g(x) = \sup\{f(u) : u \in [0,x]\}$; that supremum exists because Y is assumed to be Dedekind complete. We note that g is additive on X_+:

$$\begin{aligned} g(x_1) + g(x_2) &= \sup\{f(u_1) + f(u_2) \ : \ u_1 \in [0,x_1],\ u_2 \in [0,x_2]\} \\ &= \sup\{f(v) : v \in [0,x_1] + [0,x_2]\} \\ &\overset{(!)}{=} \sup\{f(v) : v \in [0,x_1+x_2]\} \quad = \quad g(x_1+x_2) \end{aligned}$$

where equation (!) is by the Riesz Decomposition Property (noted in 8.38). By 11.55, therefore, g extends to a positive linear map from X into Y, which we shall also denote by g. Then $g(x) \succcurlyeq f(x)$ for all $x \in X_+$, so $g - f$ is also a positive linear map. Thus $f = g - (g - f)$ is the difference of two positive linear maps.

It is an easy exercise to verify that the function g constructed above is actually equal to the supremum of the set $\{0, f\}$, in the ordered vector space $\mathcal{L}_b(X,Y)$. Thus $0 \vee f$ exists for each $f \in \mathcal{L}_b(X,Y)$, and therefore that ordered vector space is a vector lattice, by the observations in 8.38.

To show $\mathcal{L}_b(X,Y)$ is Dedekind complete, suppose $\Phi \subseteq \mathcal{L}_b(X,Y)$ is a nonempty set bounded above by some $\beta \in \mathcal{L}_b(X,Y)$; we shall show that $\sup\Phi$ exists in $\mathcal{L}_b(X,Y)$. We may replace Φ by the collection of sups of nonempty finite subsets of Φ; the existence and value of Φ are not thereby affected. Thus we may assume Φ is directed; we shall show that, on X_+, $\sup\Phi$ is then equal to the pointwise supremum of the members of Φ. Fix any $\varphi_0 \in \Phi$. We may replace each $\varphi \in \Phi$ with the function $\varphi - \varphi_0$; this does not affect the existence of $\sup\Phi$, and it replaces the value of $\sup\Phi$ with $\sup\Phi - \varphi_0$; thus we may assume that $0 \in \Phi$. Since Y is Dedekind complete, $h(x) = \sup_{f\in\Phi} f(x)$ exists for each $x \in X_+$. Since $0 \in \Phi$, we have $h(x) \succcurlyeq 0$ for each $x \in X_+$. By 8.32, the function $h : X_+ \to X_+$ is additive. By 11.55, h extends to a positive linear operator from X into Y; clearly that operator is the sup of Φ in $\mathcal{L}_b(X,Y)$.

11.58. *Definition and corollary.* Let X be a Riesz space. Then the linear space

$$\mathcal{L}_b(X,\mathbb{R}) \quad = \quad \{\text{order bounded linear functionals on } X\}$$

is called the **order dual** of X. It is equal to the set of all linear functionals that can be written as the difference of two positive linear functionals. It is a Dedekind complete Riesz space when equipped with this ordering: $f \succcurlyeq g$ if $x \succcurlyeq 0 \ \Rightarrow \ f(x) \geq g(x)$. It also satisfies this formula:

$$/f/(x) \quad = \quad \sup\{|f(u)| \ : \ u \in [-x,x]\}, \qquad \text{if } x \in X_+.$$

This definition is a special case of the notion of "dual" introduced in 9.55. It is investigated further in books on vector lattices; we shall not study it further in this book.

Orthogonality in Riesz Spaces (Optional)

11.59. *Definitions.* Let X be a Riesz space (or more generally, a lattice group). In this context, two elements x, y are **orthogonal** to each other, denoted $x \perp y$, if $/x/ \wedge /y/ = 0$. For any set $S \subseteq X$, the **orthogonal complement** of S is the set

$$S^{\perp} \quad = \quad \{x \in X \ : \ x \perp s \text{ for all } s \in S\}.$$

This definition is a special case of 4.12, with

$$\Gamma \quad = \quad \{(x,y) \ : \ x \perp y\} \quad = \quad \{(x,y) \ : \ /x/ \wedge /y/ = 0\},$$

and so the conclusions of 4.12 are applicable. Thus, $x \perp x \iff x = 0$, and

$$S \subseteq S^{\perp\perp}, \qquad S^{\perp} = S^{\perp\perp\perp}, \qquad \text{and} \quad S_1 \subseteq S_2 \ \Rightarrow \ S_2^{\perp} \subseteq S_1^{\perp},$$

for sets $S, S_1, S_2 \subseteq X$. Also, if $S = T^{\perp}$ and $T = S^{\perp}$, then

$$\{0\} \ = \ S \cap T \ = \ S^{\perp} \cap T^{\perp} \ = \ (S \cup T)^{\perp}.$$

11.60. *Example.* We consider $\mathbb{R}^{\Lambda}$ as in 11.45. Verify that $x \perp y$ if and only if $xy = 0$, where xy is the function defined pointwise — i.e., $(xy)(\lambda) = [x(\lambda)][y(\lambda)]$ for all $\lambda \in \Lambda$. Also prove that two sets $S_1, S_2 \subseteq \mathbb{R}^{\Lambda}$ are orthogonal complements of each other if and only if they are sets of the form

$$S_j \quad = \quad \left\{ x \in \mathbb{R}^{\Lambda} \ : \ x\Big|_{\Lambda_j} = 0 \right\},$$

where Λ_1 and Λ_2 form a partition of Λ.

11.61. *Definition.* Let X be a Riesz space or, more generally, a lattice group. A **band** in X, also known as a **normal sublattice**, is an ideal (as defined in 9.27 and 11.51) that is sup-closed in X (as defined in 4.4.b).

Riesz Theorem on Orthogonal Decompositions. Suppose that X is Riesz space (or, more generally, a lattice group). Assume X is Dedekind complete. Then a subset $S \subseteq X$ is an orthogonal complement of some subset of X if and only if S is a band. Furthermore, if S and T are orthogonal complements of each other, then they form a direct sum decomposition: $X = S \oplus T$, as defined in 8.13. The projections $\pi_S : X \to S$ and $\pi_T : X \to T$ are homomorphisms of lattice groups (or of Riesz spaces, if X is a Riesz space). The projection onto S is given by the formula

$$\pi_S(x) \quad = \quad \sup\{/s/ \wedge x \ : \ s \in S\} \qquad \text{if } x \succcurlyeq 0,$$

and $\pi_S(x) = \pi_S(x^+) - \pi_S(x^-)$ in general. The projection onto T is given by analogous formulas.

Remark. Compare this theorem with 22.52.

Proof of theorem (following Bhaskara Rao and Bhaskara Rao [1983]). First suppose that S is an orthogonal complement. Then S is an ideal; that is a straightforward exercise. To show that S is sup-closed, let M be a nonempty subset of S, and suppose $\mu = \sup(M)$ exists in X. Show that $/\mu/ \preccurlyeq \sup\{/m/ : m \in M\}$, and hence $\mu \in S$. (These arguments actually do not require that X be Dedekind complete.) Now assume X is a Dedekind complete lattice group, and S is a band in X. Most of our proof will be concerned with showing that

(* * *) If $T = S^\perp$ and $x \in X^+$, then $x = s_x + t_x$ for some $s_x \in S$ and $t_x \in T$.

The set $M = \{/s/ \wedge x : s \in S\}$ is bounded above; since X is Dedekind complete, $s_x = \sup(M)$ exists in X. Since S is a sup-closed ideal, we have $M \subseteq S$ and also $s_x \in S$. The elements of M are nonnegative; hence $s_x \succcurlyeq 0$ also. Let $t_x = x - s_x$; next we shall show that t_x lies in $T = S^\perp$. Let any $\sigma \in S$ be given; we are to show that $/\sigma/ \wedge /t_x/ = 0$. Since M is bounded above by x, we have $s_x \preccurlyeq x$; therefore $t_x \succcurlyeq 0$ and $/t_x/ = t_x$. Let $u = \sigma^{ab} \wedge t_x$; then $u \succcurlyeq 0$ and it suffices to show that $u \preccurlyeq 0$. Since $0 \preccurlyeq u \preccurlyeq /\sigma/$ and $\sigma \in S$ and S is a sup-closed ideal, it follows that $u \in S$; hence also $u + s_x \in S$. Then

$$0 \quad \preccurlyeq \quad u + s_x \quad = \quad (/\sigma/ \wedge t_x) + (x - t_x) \quad \preccurlyeq \quad x,$$

and so

$$u + s_x \quad = \quad (u + s_x) \wedge /u + s_x/ \wedge x \quad \in \quad M,$$

whence $u + s_x \preccurlyeq \sup(M) = s_x$, and thus $u \preccurlyeq 0$. This completes our proof of (* * *). Next we prove the conclusion of (* * *) with the hypothesis weakened: We shall permit x to be any element of X, not necessarily nonnegative. Applying the Jordan Decomposition, we have $x = p - n$, where $p, n \in X^+$. Then p, n have Riesz decompositions

$$p \quad = \quad s_p + t_p \quad \in \quad S + T \qquad \text{and} \qquad n \quad = \quad s_n + t_n \quad \in \quad S + T.$$

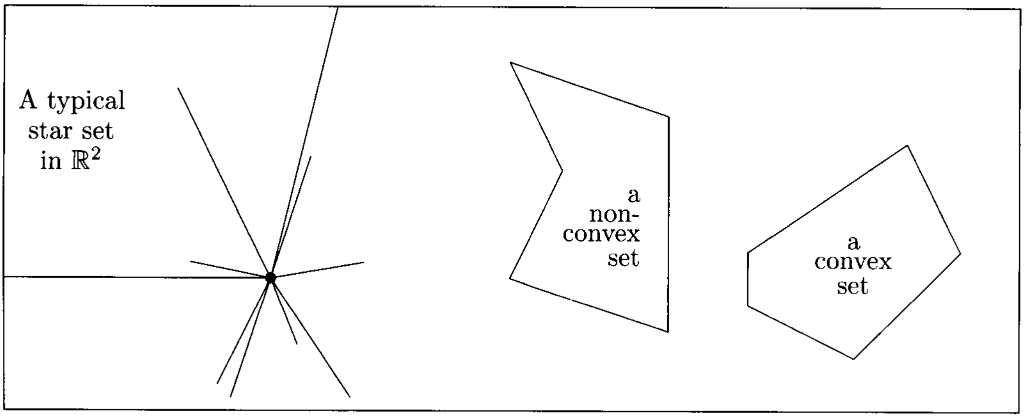
We obtain $x = s_x + t_x$, with $s_x = s_p - s_n \in S$ and $t_x = t_p - t_n \in T$. To show every sup-closed ideal is an orthogonal complement, let S be an sup-closed ideal. Clearly $S \subseteq S^{\perp\perp}$. For the reverse inclusion, let $x \in S^{\perp\perp}$ have decomposition $x = s + t \in S + T$. Then $t \in T$, but also $t = x - s \in S^{\perp\perp} = T^\perp$. Hence $t = 0$, and $x = s \in S$.

Whenever S and T are orthogonal complements, they satisfy $S \cap T = \{0\}$; see 8.13. In the present context, we have also shown that $S + T = X$. Hence $S \oplus T = X$ (see 8.11), and the projections π_S, π_T are uniquely determined group homomorphisms. The arguments of the preceding paragraphs show that π_S must satisfy the formula stated in the theorem. Note that if $u \succcurlyeq 0$ then $0 \preccurlyeq \pi_S(u) \preccurlyeq u$. For any $x \in X$, both x^+ and x^- are nonnegative, so $0 \preccurlyeq \pi_S(x^+) \preccurlyeq x^+$ and $0 \preccurlyeq \pi_S(x^-) \preccurlyeq x^-$. Since $x^+ \wedge x^- = 0$, it follows that $\pi_S(x^+) \wedge \pi_S(x^-) = 0$. From the Jordan Decomposition $x = x^+ - x^-$, we obtain $\pi_S(x) = \pi_S(x^+) - \pi_S(x^-)$, which is therefore the Jordan Decomposition of $\pi_S(x)$. Hence $[\pi_S(x)]^+ = \pi_S(x^+)$. By 8.45, π_S is a homomorphism of lattice groups. If X is a Riesz space, then π_S is a homomorphism of Riesz spaces, by 11.53. The same conclusions can be drawn for π_T.

Chapter 12

Convexity

12.1. *Preview.* The diagram below shows examples of a star set, a nonconvex set, and a convex set, all of which will be defined soon. The distinction between convex and nonconvex may be easier to understand after 12.5.i.



12.2. *Notational convention.* Throughout the remainder of this book (except where noted otherwise), the scalar field of a linear space will always be either $\mathbb{R}$ or $\mathbb{C}$. Usually the scalar field will be denoted by $\mathbb{F}$, and we shall not specify which field is intended; this intentional ambiguity will permit us to treat both the real and complex cases simultaneously. However, we shall make free use of certain properties and structures enjoyed by $\mathbb{R}$ and $\mathbb{C}$ that are not shared by all other fields — e.g., the real part, imaginary part, complex conjugate, and absolute value (see 10.31), and the completeness of the metric determined by that absolute value (see Chapter 19).

Convex Sets

12.3. *Definitions.* Several types of sets will now be introduced together; they have similar definitions and basic properties. Let X be a linear space with scalar field $\mathbb{F}$ (equal to $\mathbb{R}$ or

$\mathbb{C}$), and let $S \subseteq X$. We say that the set S is

a **linear subspace** of X if $s, t \in S$ and $\lambda, \mu \in \mathbb{F}$ imply $\lambda s + \mu t \in S$;

convex if $s, t \in S$ and $\lambda \in (0,1)$ imply $\lambda s + (1-\lambda)t \in S$;

affine if $s, t \in S$ and $\lambda \in \mathbb{F}$ imply $\lambda s + (1-\lambda)t \in S$;

symmetric if $s \in S \Rightarrow -s \in S$.

Also, a nonempty set $S \subseteq X$ is said to be

balanced (or *circled*) if, whenever $s \in S$ and α is a scalar with $|\alpha| \leq 1$, then $\alpha s \in S$;

absolutely convex if, whenever $s, t \in S$ and α, β are scalars with $|\alpha| + |\beta| \leq 1$, then $\alpha s + \beta t \in S$;

a **star set** if, whenever $s \in S$ and $\lambda \in [0,1)$, then $\lambda s \in S$.

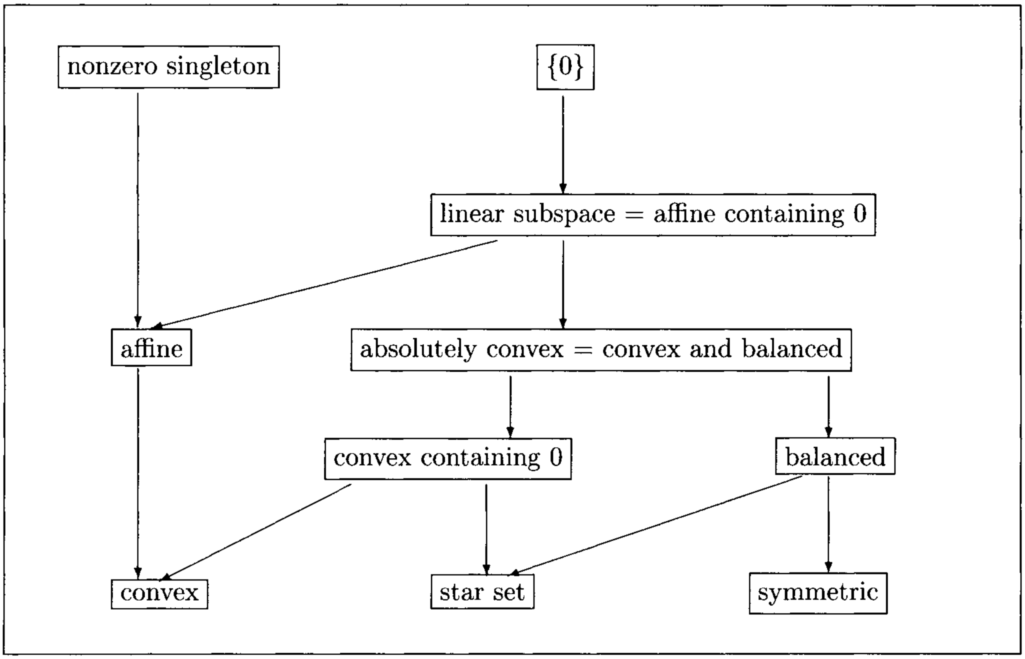
Caution: This definition of "star set" is well suited for our purposes, but it differs slightly from the definitions of "star body," "star-like set," etc., used elsewhere in the literature.

Though these different classes of sets ultimately must be studied separately, they do share a few basic properties: They are classes of sets that are closed under certain fundamental operations, and thus they are Moore collections (as in 4.6). For instance, a set $S \subseteq X$ is a linear subspace of X if and only if S is closed under all the binary operations $b_{\lambda,\mu} : X \times X \to X$ defined by $b_{\lambda,\mu}(x, y) = \lambda x + \mu y$, for all choices of λ, μ in the scalar field. Likewise, S is a convex set if and only if S is closed under all the binary operations $b_{\lambda,1-\lambda}$ for $\lambda \in [0,1]$. The other classes of sets can be characterized similarly — using not only binary operations, but also unary operations ($s \mapsto \lambda s$ for balanced sets and star sets, $s \mapsto -s$ for symmetric sets) and the nullary operation 0 (for balanced sets, absolutely convex sets, and star sets).

Since these classes are Moore collections, they are closed under intersection. Thus, any intersection of convex sets is a convex set, etc. In fact, all the fundamental operations involved are finitary, and so the resulting classes of sets are algebraic closure systems, in the sense of 4.8.

Since these classes of sets are Moore collections, they yield Moore closures (see 4.3) — in fact, they yield algebraic closures (see 4.8). However, in this context it is not customary to use the term "closure." Instead we use different terms for the different kinds of closures: The smallest linear subspace containing a set T is the **(linear) span** of T. The smallest convex set containing a set T is the **convex hull** of T. Analogously we define the **affine hull** of T, the **symmetric hull** of T, the **balanced hull** of T, the **absolutely convex hull** of T, and the **star hull** of T. Notations for these hulls vary throughout the literature. In this book the convex hull of T and balanced hull of T will be denoted by $\mathrm{co}(T)$ and $\mathrm{bal}(T)$, respectively.

12.4. *Some relations between convexity and its relatives.* These relationships are summarized in the following chart.



a. A set is absolutely convex if and only if it is convex and balanced.

b. Every balanced set is a symmetric star set.

c. Every convex set that contains 0 is a star set.

d. Every affine set is convex.

e. A subset of X is a linear subspace of X if and only if it is affine and contains 0. Thus any linear subspace of X (in particular, X itself) is convex, affine, symmetric, balanced, absolutely convex, and a star set.

f. If $x \in X \setminus \{0\}$, then the singleton $\{x\}$ is an affine set, but it is not balanced.

Moreover, suppose that the scalar field $\mathbb{F}$ is $\mathbb{R}$. Then:

g. A set is balanced if and only if it is a symmetric star set.

h. A set is absolutely convex if and only if it is nonempty, symmetric, and convex.

12.5. *Further elementary properties.* Let X be an $\mathbb{F}$-linear space. Then:

a. Any union of symmetric sets or balanced sets or star sets is, respectively, a symmetric or balanced or star set.

b. Suppose that $\mathcal{F}$ is a nonempty collection of subsets of X that is directed by inclusion — i.e., such that for each $F_1, F_2 \in \mathcal{F}$ there exists some $F \in \mathcal{F}$ such that $F_1 \cup F_2 \subseteq F$. If every member of $\mathcal{F}$ is convex or affine or absolutely convex, then the union of the members of $\mathcal{F}$ also has that property, respectively. *Hint:* 4.8(B).

c. The convex hull of a set T is equal to the set of all **convex combinations** of members of T — i.e., all vectors of the form

$$x \quad = \quad c_1 t_1 + c_2 t_2 + \cdots + c_n t_n$$

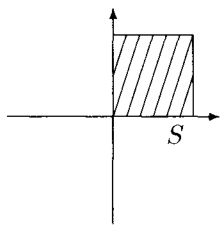
where n is a positive integer, the t_j's are members of T, and the c_j's are positive numbers whose sum is 1.

d. The convex hull of a set T is the union of the convex hulls of the finite subsets of T.

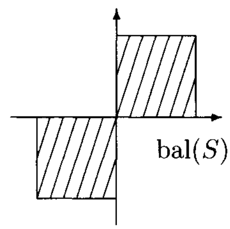
e. The convex hull of a balanced set is balanced.

f. The absolutely convex hull of any set $S \subseteq X$ is equal to $\mathrm{co}(\mathrm{bal}(S))$.

g. The balanced hull of a convex is set is *not* necessarily convex. See the example in the following diagram.



Example. Let the scalar field be $\mathbb{R}$, let the vector space be $\mathbb{R}^2$, and let $S = [0,1] \times [0,1]$. Then S is convex. However, $T = \mathrm{bal}(S) = [0,1] \times [0,1] \ \cup \ [-1,0] \times [-1,0]$ is balanced but not convex.



h. If $x, y \in X$, then the **straight line** through x and y is the set $\{\alpha x + (1-\alpha)y \ : \ \alpha \in \mathbb{R}\}$. It is the affine hull of the set $\{x, y\}$, if the scalar field is $\mathbb{R}$. In a real vector space, a set $S \subseteq X$ is affine if and only if it contains the straight line through each pair of its members.

i. If $x, y \in X$, then the **straight line segment** from x to y is the set $\{\alpha x + (1-\alpha)y \ : \ \alpha \in [0,1]\}$. It is the convex hull of the set $\{x, y\}$. The points x and y are its **endpoints**. A set $S \subseteq X$ is convex if and only if it contains the straight line segment connecting each pair of its members (regardless of whether the scalar field is $\mathbb{R}$ or $\mathbb{C}$).

j. A subset of $\mathbb{R}$ is convex if and only if it is an interval.

k. Let X be a real linear space, and let $C \subseteq X$. Then there exists an ordering $\preccurlyeq$ on X that makes X into an ordered vector space with nonnegative cone X_+ equal to C if and only if C satisfies these conditions: (i) C is convex, (ii) $C \cap (-C) = \{0\}$, and (iii) if $x \in C$ and $r > 0$ then $rx \in C$.

12.6. *Exercises: arithmetic operations on convex sets.*

a. For each λ in some index set Λ, suppose that C_λ is a convex subset of some linear space X_λ. Then $\prod_{\lambda \in \Lambda} C_\lambda$ is a convex subset of the linear space $\prod_{\lambda \in \Lambda} X_\lambda$.

b. Let $f : X \to Y$ be a linear map. If $S \subseteq X$ is a convex set, then so is $f(S) \subseteq Y$. If $T \subseteq Y$ is a convex set, then so is $f^{-1}(T) \subseteq X$.

In particular, if $S \subseteq X$ is a convex set, then $cS = \{cs : s \in S\}$ is convex for any scalar c, and $x_0 + S = \{x_0 + s : s \in S\}$ is convex for any vector $x_0 \in X$.

c. If S and T are convex subsets of X, then

$$\mathrm{co}(S \cup T) \quad = \quad \bigcup_{0 \le \alpha \le 1} \Big[\alpha S + (1-\alpha)T\Big].$$

d. For any sets $A_1, A_2, \ldots, A_k \subseteq X$, the convex hull of the sum is the sum of the convex hulls. That is, $\mathrm{co}\left(\sum_{i=1}^k A_i\right) \;=\; \sum_{i=1}^k \mathrm{co}(A_i)$.

e. If C is a convex set, then $C + C = 2C$. That is, $\{x + y : x, y \in C\} = \{2u : u \in C\}$. Equivalently, $\frac{1}{2}C + \frac{1}{2}C = C$.

12.7. (*Optional.*) As we noted in 12.3, it is possible to consider convex sets as algebraic systems, with fundamental operations given by the binary operations

$$c_r(x, y) \;=\; rx + (1-r)y \qquad\qquad \text{for} \;\; r \in (0, 1).$$

One may be tempted to try to view convex sets as an equational variety and thus apply to them all the theory of equational varieties.

However, convex sets do *not* form a variety, for they are not closed under the taking of homomorphic images that respect the fundamental operations. It can be proved (see Romanowska and Smith [1985]) that the smallest variety containing all convex sets is the variety of *barycentric algebras*. These are the algebraic systems that have fundamental operations given by some binary operations c_r for $r \in (0, 1)$, where the binary operations satisfy these identities:

$$c_r(x, x) \;=\; x \qquad \text{and} \qquad c_r(x, y) \;=\; c_{1-r}(y, x) \qquad\qquad \text{when } 0 < r < 1;$$

$$c_{t/(s+1)}\left(c_{s/t}(x, y), z\right) \quad = \quad c_{s/(s+1)}\left(x, c_{t-s}(y, z)\right) \qquad \text{when } 0 < s < t < s + 1.$$

The convex sets are the barycentric algebras that can be embedded in vector spaces; not all barycentric algebras can be so embedded.

The following example, from [Romanowska and Smith], shows that the class of convex sets is not closed under the taking of homomorphic images that respect the fundamental operations. Let n be an integer greater than 1. Let $\Omega = \{e_1, e_2, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$ — that is, let $e_j = (0, 0, \ldots, 1, \ldots, 0, 0)$ be the vector with 1 in the jth place and 0s elsewhere. Let Δ be the convex hull of the set Ω; it is a convex subset of $\mathbb{R}^n$ (called the **standard simplex**). We shall also consider the set $\mathcal{P}(\Omega) = \{\text{subsets of } \Omega\}$ as an algebraic system, with binary operations defined by

$$c_r(A, B) \;=\; A \cup B \qquad\qquad \text{for} \;\; r \in (0, 1).$$

(We emphasize that all the c_r's, for different values of r, are the *same* binary operation.) The set $\mathcal{P}(\Omega)$ cannot possibly be isomorphic to a convex subset of a real vector space, for any convex set that contains more than one point must contain infinitely many points. However, the mapping $f : \Delta \to \mathcal{P}(\Omega)$ defined by

$$f\left(\sum_{j=1}^{n} r_j e_j\right) \;=\; \{e_j \in \Omega \;:\; r_j > 0\}$$

is a homomorphism — i.e., it preserves the fundamental operations of the algebraic systems. Thus $\mathcal{P}(\Omega)$ is a homomorphic image of a convex set. Therefore it preserves any identities that could be used to define the variety of convex sets — but it is not a convex set. Thus convex sets do not form a variety. In fact, $\mathcal{P}(\Omega)$ is a barycentric algebra.

12.8. *Definition.* Let X be a linear space, with scalar field $\mathbb{R}$ or $\mathbb{C}$. A set $S \subseteq X$ is **absorbing** (or *radial*) if for each $x \in X$ we have $cx \in S$ for all scalars c sufficiently small (i.e., for all scalars c satisfying $|c| \leq r$, where r is some positive number that may depend on x and S).

Show that the absorbing sets form a proper filter on X; thus they are sets that are "large" in the sense of 5.3.

Absorbing sets will be important in the theory of Minkowski functionals (see 12.29.c and 12.29.g) and topological vector spaces (see 26.26, 27.9.e, and 27.20).

Combinatorial Convexity in Finite Dimensions (Optional)

12.9. Radon's Affineness Lemma. Let $x_0, x_1, \ldots, x_k$ be vectors in $\mathbb{R}^n$, for some positive integers k and n with $k > n$. Then there exist real numbers $p_0, p_1, \ldots, p_k$, not all zero, such that $\sum_{j=0}^{k} p_j = 0$ and $\sum_{j=0}^{k} p_j x_j = 0$.

Hint: First show that the vectors $x_1 - x_0, x_2 - x_0, \ldots, x_k - x_0$ are linearly dependent — see 11.25.

12.10. Carathéodory's Theorem. Let $S \subseteq \mathbb{R}^n$. Then every point in $\mathrm{co}(S)$ can be expressed as a convex combination of $n+1$ or fewer elements of S.

Proof. The proof is in several steps.

(i) Let T_k be the set of all convex combinations of k or fewer elements of S. It suffices to show that if $k > n$ then $T_{k+1} \subseteq T_k$. (Why?)

(ii) Let $x \in T_{k+1}$. Then $x = \alpha_0 x_0 + \cdots + \alpha_k x_k$ for some $x_0, x_1, \ldots, x_k \in S$ and $\alpha_0, \alpha_1, \ldots, \alpha_k \in (0, 1]$ with $\alpha_0 + \cdots + \alpha_k = 1$. (Explain.)

(iii) Choose real numbers $p_0, p_1, \ldots, p_k$ as in Radon's Lemma. For $j = 0, 1, \ldots, k$ and any real number r, let $\beta_j(r) = \alpha_j - r p_j$. Show that $x = \sum_{j=0}^{k} \beta_j(r) x_j$ and $1 = \sum_{j=0}^{k} \beta_j(r)$.

(iv) By a suitable choice of r, show that $x \in T_k$.

12.11. Radon's Intersection Theorem. Let S be a subset of $\mathbb{R}^n$ consisting of at least

$n + 2$ points. Then S can be partitioned into disjoint subsets Q and R such that $\operatorname{co}(Q)$ meets $\operatorname{co}(R)$.

Hints: Let $S \supseteq \{x_0, x_1, \ldots, x_{n+1}\}$. Choose real numbers $p_0, p_1, \ldots, p_{n+1}$ as in Radon's Lemma. By relabeling and reordering, we may assume

$$p_0, p_1, \ldots, p_r > 0 \qquad \text{and} \qquad p_{r+1}, p_{r+2}, \ldots, p_{n+1} \leq 0,$$

where $0 \leq r < n + 1$ (explain). Now let

$$Q \supseteq \{x_0, x_1, \ldots, x_r\} \qquad \text{and} \qquad R \supseteq \{x_{r+1}, x_{r+2}, \ldots, x_{n+1}\}.$$

Then what?

12.12. Helly's Intersection Theorem. Let $S_0, S_1, \ldots, S_k$ be convex subsets of $\mathbb{R}^n$, where k and n are positive integers and $k > n$. Suppose that each $n + 1$ of these sets have nonempty intersection. Then $\bigcap_{i=0}^{k} S_i$ is nonempty.

Hints: By induction on k, we may assume that the intersection of any k of the S_j's is nonempty (explain). For each $j = 0, 1, \ldots, k$, pick some $x_j \in \bigcap_{i \neq j} S_i$. Apply Radon's Intersection Theorem to the points x_j. (How?)

12.13. The following result is interesting enough to deserve mention, though its proof is too difficult to include here:

> **Shapley-Folkman Theorem.** Suppose $x \in \sum_{j=1}^{m} \operatorname{co}(A_j)$, in $\mathbb{R}^n$. Then x can be expressed as $x = \sum_{j=1}^{m} x_j$, where each $x_j \in \operatorname{co}(A_j)$ and where $\{j : x_j \notin A_j\}$ has cardinality at most n.

Taking m much larger than n, this shows that the sum of a large number of arbitrary sets is "almost convex." Proofs can be found in the appendices of Arrow and Hahn [1971] and Starr [1969]. Actually, those proofs assume the sets A_j are compact, but the problem can easily be reduced to that case by using Carathéodory's Theorem and its consequences; see 26.23.g.

Other matters related to the theorems of Radon, Helly, and Carathéodory are considered by Danzer, Grünbaum, and Klee [1963]. Additional material on convexity, especially in finite dimensions, can be found in Roberts and Varberg [1973], Rockafellar [1970], and Stoer and Witzgall [1970].

Convex Functions

12.14. *Remarks.* For the definitions below, we consider functions f taking values in $[-\infty, +\infty]$. The definitions can be simplified slightly when f is known to be *real-valued* — i.e., when $-\infty, +\infty \notin \operatorname{Range}(f)$ — and certainly that restricted case still covers most of the applications. For these reasons, some mathematicians define "convex" only for real-valued functions. However, the greater generality of extended real-valued functions is occasionally useful, because $[-\infty, +\infty]$ is order complete — i.e., we can always take sups and infs in $[-\infty, +\infty]$.

Arithmetic in $[-\infty, +\infty]$ is defined as in 1.17. Note that a sum of finitely many terms, $r_1 + r_2 + \cdots + r_n$, is defined if and only if $-\infty$ and $+\infty$ are not both among $r_1, r_2, \ldots, r_n$.

12.15. *Definition.* Let C be a convex subset of a linear space X, and let $f : C \to [-\infty, +\infty]$ be some function. Then the following conditions are equivalent; if they are satisfied we say f is a **convex** function.

(A) The set $\{(x, r) \in C \times \mathbb{R} : f(x) \le r\}$ is a convex subset of $C \times \mathbb{R}$. (This set is called the **epigraph** of f.)

(B) The set $\{(x, r) \in C \times \mathbb{R} : f(x) < r\}$ is a convex subset of $C \times \mathbb{R}$.

(C) Whenever $x_0, x_1 \in C$ and $0 < \lambda < 1$, then

$$f\left((1-\lambda)x_0 + \lambda x_1\right) \quad \le \quad (1-\lambda)f(x_0) + \lambda f(x_1)$$

whenever the right side is defined (see remark in 12.14).

(D) Whenever n is a positive integer and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are positive numbers summing to 1 and $x_1, x_2, \ldots, x_n \in C$, then

$$f\left(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n\right) \le \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n)$$

whenever the right side is defined (see remark in 12.14).

(E) Whenever n is a positive integer and $\mu_1, \mu_2, \ldots, \mu_n$ are positive numbers and $x_1, x_2, \ldots, x_n \in C$, then

$$f\left(\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_n x_n\right) \le \frac{\mu_1 f(x_1) + \mu_2 f(x_2) + \cdots + \mu_n f(x_n)}{\mu_1 + \mu_2 + \cdots + \mu_n}$$

whenever the right side is defined.

If f is *real-valued* — i.e., if $\pm\infty \notin \operatorname{Range}(f)$ — then the following conditions are also equivalent.

(F) For each $v \in X$ and $\xi \in C$, the function

$$p \quad \mapsto \quad h_{\xi.v}(p) \quad = \quad \frac{f(\xi + pv) - f(\xi)}{p}$$

is increasing on the interval $\{p \in \mathbb{R} : p > 0,\ \xi + pv \in C\}$.

(G) For each $v \in X$ and $\xi \in C$, the function $p \mapsto h_{\xi,v}(p) = [f(\xi + pv) - f(\xi)]/p$ is increasing on the set where it is defined — i.e., on the set $\{p \in \mathbb{R} \setminus \{0\} : \xi + pv \in C\}$.

Hints: The equivalence of (A), (B), (C) follows by considering various cases, according to whether each number involved is $+\infty$, $-\infty$, or a finite real number. Obviously (D) implies (C) as a special case; conversely, (D) follows from (C) by induction. Condition (E) is just a reformulation of (D).

Now suppose f is real-valued. To prove (C) $\Rightarrow$ (F), show $h_{\xi,v}(\lambda p) \leq h_{\xi,v}(p)$ for $0 < \lambda < 1$ by taking $x_0 = \xi$ and $x_1 = \xi + pv$. To prove (F) $\Rightarrow$ (C), take $\xi = x_0$ and $v = x_1 - x_0$; use the fact that $h_{\xi,v}(\lambda) \leq h_{\xi,v}(1)$. Obviously (G) implies (F). To prove that (C) and (F) together imply (G), note that $h_{\xi,-v}(-p) = -h_{\xi,v}(p)$; also, the inequality $h_{\xi,v}(-p) \leq h_{\xi}, v(p)$ for $p > 0$ follows from the convexity of f.

12.16. *Further definitions.* A function $g : C \to [-\infty, +\infty]$ is **concave** if $-g$ is convex. A function $h : C \to [-\infty, +\infty]$ is **affine** if it is both concave and convex. An equivalent condition for h to be affine is that

> whenever $x_0, x_1 \in C$ and $0 < \lambda < 1$, then
> $$h\Big((1-\lambda)x_0 + \lambda x_1\Big) \quad = \quad (1-\lambda)h(x_0) + \lambda h(x_1)$$
> whenever the right side is defined — i.e., whenever we do not have one of $h(x_0), h(x_1)$ equal to $-\infty$ and the other equal to $+\infty$.

12.17. *Some elementary properties of convex functions.* Let X be a vector space, let C be a convex subset of X, and let $f : C \to [-\infty, +\infty]$ be some function. Then:

a. f is convex if and only if the restriction $f\big|_L$ is a convex function for each line segment L whose endpoints are elements of C — equivalently, if and only if for each $x_0, x_1 \in C$, the function $\lambda \mapsto f\left((1-\lambda)x_0 + \lambda x_1\right)$ is a convex function from the interval $[0,1]$ into $[-\infty, +\infty]$.

b. We say f is **quasiconvex** if the set $\{x \in C : f(x) \leq r\}$ is a convex set for each $r \in [-\infty, +\infty]$. Show that

(i) Every convex function is quasiconvex.

(ii) Every increasing function from $\mathbb{R}$ into $[-\infty, +\infty]$ is quasiconvex.

(iii) (*Example.*) The function $f(x) = x^3$ is increasing on $\mathbb{R}$, hence quasiconvex, but it is not convex. (*Hint*: Use 12.19(E).)

c. We say f is **strictly convex** if it has this property: Whenever x and y are two distinct points in C and $0 < \lambda < 1$, then $f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$.

Show that if C is an open interval in the real line, then any convex function from C into $\mathbb{R}$ is either affine or strictly convex.

d. If f is a real-valued function defined on a linear space, then f is affine if and only if $f - f(0)$ is linear.

Caution: In some contexts the term "linear" is used for affine maps as well. Especially, a "**piecewise-linear**" map is a map that is defined separately on various parts of its domain and is affine on each of those parts. This terminology is especially common in numerical analysis.

e. (*Optional.*) If f is real-valued and convex and its domain C is the convex hull of a finite set, then f is bounded.

Hints: Say $C = \mathrm{co}\{x_1, x_2, \ldots, x_n\}$. First show $\sup_{x \in C} f(x) \le \max_j f(x_j)$. Then let $u = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$. For each $y \in C$, show there is some corresponding $z \in C$ satisfying $u = \frac{1}{n}y + \frac{n-1}{n}z$. Use this to obtain a lower bound on $f(y)$.

f. (*Optional.*) Let V be a real linear space. Let Φ be the collection of graphs of functions f that have the property that they can be extended to convex functions from convex subsets of V into $\mathbb{R}$. Then Φ has finite character (see 3.46).

12.18. *Remarks.* Let C be a convex subset of a vector space X, and let $f : C \to [-\infty, +\infty]$ be a convex function. Then the set

$$\{x \in C \ : \ f(x) < +\infty\}$$

is a convex subset of C, sometimes called the **effective domain** of f. Most interesting behavior of convex functions occurs in the effective domain, and we can replace f with its restriction to this set without seriously affecting most results about convex functions. Conversely, any convex function f defined on any convex set $S \subseteq X$ can be extended to a convex function on any larger convex set C, by taking $f(x) = +\infty$ whenever $x \in C \setminus S$. The extension and the original function have the same effective domain.

Here is a simple special case: Let f be the constant function 0 on some convex set S. Let C be any larger convex set. Then f can be extended to the convex function $I_S : C \to \{0, +\infty\}$ defined by

$$I_S(x) \quad = \quad \begin{cases} 0 & \text{when } x \in S \\ +\infty & \text{when } x \in C \setminus S. \end{cases}$$

Then I_S is a convex function, sometimes called the **indicator function** of C. Note that its definition depends on not only S but also C, though the choice of C is not reflected by our notation I_S. (The indicator function should not be confused with the *characteristic function* $1_S : C \to \{0, 1\}$, defined in 2.2.b.)

12.19. *Derivatives and convexity.* (These results assume some familiarity with college calculus.) Let $C \subseteq \mathbb{R}$ be an interval, and assume $f : C \to \mathbb{R}$ is continuously differentiable. Show that the following are equivalent:

(A) f is convex,

(B) $(y - x)f'(y) \ge f(y) - f(x)$ for all $x, y \in C$,

(C) $(y - x)[f'(y) - f'(x)] \ge 0$ for all $x, y \in C$,

(D) $f'(x)$ is an increasing function of x on C.

If f is twice continuously differentiable, then this condition is also equivalent:

(E) $f'' \ge 0$ on C.

Remark. In 25.25 we shall determine precisely how much differentiability a convex function must possess.

12.20. *Corollaries.* We now note some specific applications of the preceding results.

a. Show that $t \mapsto e^t$ is convex on $\mathbb{R}$. Then use that fact to show that if $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, then $\sqrt[p]{t}\,\sqrt[q]{u} \le \frac{t}{p} + \frac{u}{q}$ for $t, u \ge 0$.

b. The function $x \mapsto x^p$, defined on $[0, +\infty)$, is convex if $1 \le p < \infty$ and concave if $0 < p \le 1$.

c. Show that $\tan : [0, \pi/2) \to [0, +\infty)$ is convex. Taking limits, define $\tan(\pi/2) = +\infty$, and show that $\tan : [0, \pi/2] \to [0, +\infty]$ is convex.

12.21. Combining convex functions. Let C be a convex subset of a linear space.

a. *Sums*: Let f and g be convex functions defined on C, both taking values in $(-\infty, +\infty]$ or both taking values in $[-\infty, +\infty)$. Then $f + g$ is convex.

b. *Products*: Let $f, g : C \to [0, +\infty)$ be convex functions. Assume also that

$$0 \quad \le \quad \Big[f(x) - f(y)\Big]\Big[g(x) - g(y)\Big] \qquad \text{for all } x, y \in C.$$

(This condition is satisfied, for instance, if the linear space is $\mathbb{R}$, and f and g are both increasing or both decreasing.) Show that the product function $x \mapsto f(x)g(x)$ is also convex.

c. *Compositions*: Let $J \subseteq \mathbb{R}$ be an interval — i.e., a convex subset of $\mathbb{R}$. Let $f : C \to J$ and $g : J \to [-\infty, +\infty]$ both be convex, and assume g is increasing. Show that the composition $g \circ f : C \to [-\infty, +\infty]$ is convex. As a particular example, show that $x \mapsto \exp(\tan(x))$ is convex on $[0, \pi/2)$.

d. *Pointwise suprema*: Let $\{f_\lambda : \lambda \in \Lambda\}$ be a nonempty family of convex functions from C into $[-\infty, +\infty]$. For each $x \in C$ let $\sigma(x) = \sup_{\lambda \in \Lambda} f_\lambda(x)$. Then σ is convex. (Dually, the pointwise infimum of concave functions is concave.)

Hint: Rather than bother with separate cases according to whether $f_\lambda(x)$ is $+\infty$, $-\infty$, or a member of $\mathbb{R}$, just note that the epigraph of σ is the intersection of the epigraphs of the f_λ's.

Example. Using this result (or arguing directly), prove that the mapping $x \mapsto -\min\{1, x\}$ is convex on $\mathbb{R}$.

Remarks. A converse of this result is given by (HB4) in 12.31. Compare also the supremum results in 15.23 and 16.16(D).

12.22. *The infimum of convex functions.* Let $\{f_\lambda : \lambda \in \Lambda\}$ be a nonempty family of convex functions from C into $[-\infty, +\infty]$.

a. In general, the pointwise infimum $p(x) = \inf_{\lambda \in \Lambda} f_\lambda(x)$ is not convex.

For instance, let C be the real line, and let $\{f_\lambda : \lambda \in \Lambda\}$ consist of just the two functions x and $-x$. Then the pointwise infimum is $-|x|$, which is not convex.

b. Define $\iota : C \to [-\infty, +\infty]$ by taking $\iota(x) \;=\; \inf \sum_{j=1}^{n} c_j f_{\lambda_j}(x_j)$, where the infimum is over all choices of n, c_j, λ_j, x_j such that

> n is a positive integer, the c_j's are positive numbers summing to 1, the λ_j's are members of Λ, and the x_j's are members of C satisfying $x = \sum_{j=1}^{n} c_j x_j$

and such that $+\infty$ and $-\infty$ are not both among the values $f_{\lambda_1}(x_1)$, $f_{\lambda_2}(x_2)$, $\ldots$, $f_{\lambda_n}(x_n)$.

Then ι is convex, and in fact ι is the largest convex function that satisfies $\iota \leq f_\lambda$ for all λ. We may refer to it as the **convex infimum** of the f_λ's.

Thus, the convex functions from C into $[-\infty, +\infty]$ form a complete lattice.

Of course, in some cases, the convex infimum may simply be the constant $-\infty$. That is the case, for instance, when $C = \mathbb{R}$ and the collection of functions consists of just $\{-x, x\}$.

c. Suppose $\{f_\lambda : \lambda \in \Lambda\}$ is directed downward — i.e., suppose that for each finite set $\Lambda_0 \subseteq \Lambda$ there is some $\mu \in \Lambda$ such that $f_\mu \leq \min\{f_\lambda : \lambda \in \Lambda_0\}$. Then the pointwise infimum p is equal to the convex infimum ι.

Norms, Balanced Functionals, and Other Special Functions

12.23. *Definition and exercise.* Let X be a real or complex linear space; let the scalar field be denoted by $\mathbb{F}$. Let $\rho : X \to [0, +\infty)$ be some mapping. Show that the following conditions are equivalent. If one, hence all of them, is satisfied, we shall say that ρ is a **balanced** function.

(A) $|c| \leq 1 \;\Rightarrow\; \rho(cx) \leq \rho(x)$ for scalars c and vectors x.

(B) $|c_1| \leq |c_2| \;\Rightarrow\; \rho(c_1 x) \leq \rho(c_2 x)$ for scalars c_1, c_2 and vectors x.

(C) For each number $b \in [0, +\infty]$, the set $\{x \in X : \rho(x) \leq b\}$ is balanced (in the sense of 12.3).

Show that any balanced function also satisfies $\rho(cx) = \rho(|c|x)$. In particular, if $|c| = 1$ then $\rho(cx) = \rho(x)$.

12.24. *Definitions.* Let X be a real or complex vector space.

a. A function $g : X \to [-\infty, +\infty]$ is **positively homogeneous** if it satisfies

$$g(tx) \;=\; tg(x) \qquad \text{for all } t \in [0, +\infty) \text{ and } x \in X.$$

Here we follow the convention that $0 \cdot \infty = 0$. Thus, a positively homogeneous function g may have ∞ in its range, but it must satisfy $g(0) = 0$.

A function $g : X \to [-\infty, +\infty)$ is **homogeneous** if it satisfies

$$g(tx) \;=\; |t| g(x) \qquad \text{for all scalars } t \text{ and all } x \in X.$$

(Here g is not permitted to take an infinite value.)

Exercise. Let $g : X \to [0, +\infty)$. Then g is homogeneous if and only if g is both balanced and positively homogeneous.

b. Let C be a subset of X that is closed under addition. Suppose $\beta : C \to [-\infty, +\infty]$ does not have both $-\infty$ and $+\infty$ in its range. If $\beta(x+y) \le \beta(x) + \beta(y)$ for all $x, y \in C$, we say β is **subadditive** (at least, in the context of vector spaces; the term "subadditive" has another meaning in measure theory — see 29.29.b).

c. A function $f : X \to [-\infty, +\infty]$ is **sublinear** if it is both subadditive and positively homogeneous. *Exercise.* Such a function is convex.

A **seminorm** is a function $f : X \to [0, +\infty)$ that is subadditive and homogeneous. Note that any such function is sublinear, hence convex. A **norm** is a seminorm f that also satisfies $x \ne 0 \;\Rightarrow\; f(x) > 0$. Seminorms and norms will be studied in greater detail in Chapter 22 and thereafter.

Remarks. We shall use sublinearity very seldom in this book; most of our functions will satisfy either stronger hypotheses (e.g., linear or seminorm) or weaker hypotheses (e.g., convexity). An exception is the proof in 28.37, which uses a sublinear functional. See also the remarks in 12.31.

12.25. *Elementary properties and examples.*

a. Any norm is a seminorm; any seminorm is sublinear; any linear function is sublinear; any sublinear function is convex.

b. If p is subadditive, then $p(x) \le p(y) + p(x-y)$ and $p(y) \le p(x) + p(y-x)$, hence

$$-p(y-x) \quad \le \quad p(x) - p(y) \quad \le \quad p(x-y).$$

c. The map $f \mapsto f^+ = \max\{f, 0\}$ is sublinear on $\mathbb{R}^X$, for any set X.

d. If $\beta : [0, +\infty) \to [0, +\infty)$ is concave and $\beta(0) = 0$, then β is subadditive.

Hint: Use the decomposition $x = (1-\lambda)0 + \lambda(x+y)$ to show $\beta(x) \ge \dfrac{x}{x+y}\beta(x+y)$. Similarly, $\beta(y) \ge \dfrac{y}{x+y}\beta(x+y)$. Now add these two results.

e. In particular, some subadditive functions are the functions

$$\arctan(s), \quad \frac{s}{1+s}, \quad \min\{1, s\}, \quad \tanh(s), \quad \text{and} \quad s^p \quad \text{for } p \in (0,1].$$

(Some beginners may be unfamiliar with $\tanh(s)$, which is the function $(e^s - e^{-s})/(e^s + e^{-s})$.) All of these functions except the last are also *bounded*; that fact will be significant in 18.14.

f. If $\rho : X \to [0, +\infty)$ is balanced and subadditive, then

$$\rho(cx) \quad \le \quad \big([\![\,|c|\,]\!] + 1\big)\rho(x) \quad \le \quad \big(|c| + 1\big)\rho(x)$$

for all vectors x and scalars c, where $[\![s]\!]$ is the greatest integer less than or equal to s.

12.26. Pryce's Sublinearity Lemma. Let X be a linear space, let $Y \subseteq X$ be a convex set, and let $\xi \in X$. Let $\rho : X \to \mathbb{R}$ be a sublinear functional. Let $a, b, c \in (0, +\infty)$, with

$$\rho(\xi) + ac \quad < \quad \inf_{y \in Y} \rho(\xi + ay)$$

Then there is a point $\eta \in Y$ such that

$$\rho(\xi + a\eta) + bc \quad < \quad \inf_{y \in Y} \rho(\xi + a\eta + by).$$

(This rather technical result will not be needed until 28.37.)

Proof (Pryce [1966]). The hypothesis can be restated as:

$$-\rho(\xi) \quad = \quad ac - \inf_{y \in Y} \rho(\xi + ay) + \delta \qquad\qquad \text{for some } \delta > 0.$$

Consider any $y_0, y_1 \in Y$ and let $\widehat{y} = \frac{ay_0 + by_1}{a+b}$. Then $\widehat{y} \in Y$ and

$$\xi + ay_0 + by_1 \quad = \quad \xi + (a+b)\widehat{y} \quad = \quad (1 + \frac{b}{a})(\xi + a\widehat{y}) - \frac{b}{a}\xi.$$

By sublinearity of ρ,

$$\rho(\xi + ay_0 + by_1) \quad \geq \quad (1 + \frac{b}{a})\rho(\xi + a\widehat{y}) - \frac{b}{a}\rho(\xi).$$

Hence for any fixed $\eta \in Y$, we have

$$\begin{aligned}
&\inf_{y \in Y} \rho(\xi + a\eta + by) \\
\geq \quad & (1 + \frac{b}{a}) \inf \left\{ \rho(\xi + a\widehat{y}) \ : \ \widehat{y} = \frac{a\eta + by}{a+b}, \ y \in Y \right\} - \frac{b}{a}\rho(\xi) \\
\geq \quad & (1 + \frac{b}{a}) \inf_{y \in Y} \rho(\xi + ay) - \frac{b}{a}\rho(\xi) \\
= \quad & (1 + \frac{b}{a}) \inf_{y \in Y} \rho(\xi + ay) + \frac{b}{a}\left(ac - \inf_{y \in Y} \rho(\xi + ay)\right) + \frac{b}{a}\delta \\
= \quad & \left[bc + \inf_{y \in Y} \rho(\xi + ay) + \frac{b}{a}\delta \right].
\end{aligned}$$

The last expression, in square brackets [ ], is greater than $bc + \rho(\xi + a\eta)$ if we choose $\eta \in Y$ appropriately.

Minkowski Functionals

12.27. *Definitions.* Let S be a star set in a vector space X. Note that if x is a point in the vector space (not necessarily a member of S), then

$$\begin{aligned}
&\{r \in [0, +\infty) \ : \ rx \in S\} && \text{is a subinterval of } [0, +\infty) \text{ that includes } 0, \text{ so} \\
&\{k \in (0, +\infty] \ : \ \frac{1}{k}x \in S\} && \text{is a subinterval of } (0, +\infty] \text{ that includes } \infty.
\end{aligned}$$

Thus the number

$$\mu_S(x) \quad = \quad \inf\left\{k \in (0,+\infty] \;:\; k^{-1}x \in S\right\}$$

is well defined (though it may be ∞). The function $\mu_S : X \to [0,+\infty]$ is the **Minkowski functional** of the set S.

It is easy to see that the mapping $S \mapsto \mu_S$ is direction-reversing — i.e., if S and T are star sets and $S \subseteq T$, then $\mu_S \geq \mu_T$. The largest star set, X, has the smallest Minkowski functional: μ_X is just the constant function 0. The smallest star set, $\{0\}$, has the largest Minkowski functional; it is easily seen to be

$$\mu_{\{0\}}(x) \quad = \quad \begin{cases} 0 & \text{if } x = 0 \\ \infty & \text{if } x \neq 0. \end{cases}$$

In our applications, we will use Minkowski functionals μ_S mainly when S is a *convex* set — a case investigated in 12.29.e — but the most basic properties of Minkowski functionals do not involve convexity.

12.28. *Proposition.* Let X be a vector space. Let $g : X \to [0,+\infty]$ be some function, and let S be a nonempty subset of X. Then the following two conditions are equivalent:

(A) S is a star set and g is the Minkowski functional of S.

(B) g is a positively homogeneous function and

$$\{x \in X \;:\; g(x) < 1\} \quad \subseteq \quad S \quad \subseteq \quad \{x \in X \;:\; g(x) \leq 1\}.$$

Hints for (B) $\Rightarrow$ (A): To show S is a star set, observe that

$$x \in S,\ \lambda \in [0,1) \quad \Rightarrow \quad g(\lambda x) = \lambda g(x) < 1 \quad \Rightarrow \quad \lambda x \in S.$$

To show $\mu_S \leq g$, observe that

$$g(x) < r \;\Rightarrow\; g\left(\frac{1}{r}x\right) < 1 \;\Rightarrow\; \frac{1}{r}x \in S \;\Rightarrow\; \mu_S(x) \leq r.$$

To show $g \leq \mu_S$, similarly,

$$\mu_S(x) < r \;\Rightarrow\; \frac{1}{r}x \in S \;\Rightarrow\; g\left(\frac{1}{r}x\right) \leq 1 \;\Rightarrow\; g(x) \leq r.$$

12.29. *Corollaries and further properties.*

a. A function on a vector space is positively homogeneous if and only if it is the Minkowski functional of some star set.

b. The Minkowski functional of any balanced set is a balanced function.

c. Let S be a balanced star set. Then μ_S is finite-valued (i.e., does not take the value $+\infty$) if and only if S is absorbing (as defined in 12.8).

d. If g is a positively homogeneous function, then both the sets $\{x \in X \;:\; g(x) < 1\}$ and $\{x \in X \;:\; g(x) \leq 1\}$ are star sets with Minkowski functional equal to g.

e. If S is a convex star set, then μ_S is a convex function.

Hints: If $0 < \lambda < 1$ and $\alpha > \mu_S(x)$ and $\beta > \mu_S(y)$, then $\alpha^{-1}x$ and $\beta^{-1}y$ belong to S, hence

$$\frac{\lambda x + (1-\lambda)y}{\lambda\alpha + (1-\lambda)\beta} \quad = \quad \frac{\lambda\alpha}{\lambda\alpha + (1-\lambda)\beta}\,\frac{x}{\alpha} + \frac{(1-\lambda)\beta}{\lambda\alpha + (1-\lambda)\beta}\,\frac{y}{\beta} \quad \in \quad S.$$

f. The converse of that last result is false; a star set S may be nonconvex and still have μ_S convex.

For instance, let $X = \mathbb{R}^2$ and define $g(x,y) = \max\{|x|,|y|\}$; this function is convex. Thus the sets $A = \{(x,y) \in \mathbb{R}^2 \ : \ g(x,y) < 1\}$ and $B = \{(x,y) \in \mathbb{R}^2 \ : \ g(x,y) \leq 1\}$ are convex. The function g is the Minkowski functional of both A and B, and also of any set between A and B, but $A \subseteq S \subseteq B$ does not imply S is convex.

g. Lemma on the construction of seminorms. Let $S \subseteq X$ be a convex, balanced set (not necessarily absorbing), and let μ_S be its Minkowski functional. Then the linear span of S is the set $X_0 = \bigcup_{n=1}^{\infty} nS = \{x \in X : \mu_S(x) < \infty\}$, and μ_S is a seminorm on that set. If S is also absorbing, then $X_0 = X$; that is, μ_S is a seminorm on X.

Remark. Minkowski functionals will be used in 22.11, 22.25, 22.28, and 26.29.

Hahn-Banach Theorems

12.30. *Introduction.* The literature contains many closely related theorems, any one of which may be referred to as "the Hahn-Banach Theorem." These theorems are useful in different ways in different parts of analysis. We shall prove about 20 of these theorems, in this and later chapters — see 12.31, 23.18, 23.19, 26.56, 28.4, 28.14.a, and 29.32. Still other forms of the Hahn-Banach Theorem are given by Buskes [1993], Gluschankof and Tilli [1987], Holmes [1975], Luxemburg [1969], Thierfelder [1991], Tuy [1972], and Zowe [1978].

We shall view the Hahn-Banach Theorems as weak forms of the Axiom of Choice. The various equivalent forms will be denoted by (HB1), (HB2), (HB3), etc.; collectively we shall refer to them all as HB. (It is surprising that some seemingly weaker forms of HB commonly presented as "corollaries" in the literature — such as (HB1), (HB4), (HB9) — are in fact equivalent to HB in their set-theoretic strength.) We shall keep track of effective proofs because the Hahn-Banach Theorem is nonconstructive: It implies the existence of certain pathological objects for which we have no explicit examples. However, many analysts prefer to view the Axiom of Choice (AC) as simply being "true" and will therefore view the Hahn-Banach Theorem in the same fashion; these readers can skip some of the converse proofs.

Considered as a set-theoretic principle, the Hahn-Banach Theorem is weaker than the Ultrafilter Principle, which is in turn weaker than the Axiom of Choice. (In 17.6 we shall prove UF $\Rightarrow$ HB.) In fact, the Hahn-Banach Theorem is *strictly* weaker than the Ultrafilter Principle; that fact was established by Pincus [1972], but its proof is beyond the scope of this book. A survey comparing the relative strengths the Hahn-Banach Theorem and other weak forms of Choice is given by Pincus [1974]. Some theorems that appear similar to the

Hahn-Banach Theorem are in fact equivalent to the Axiom of Choice (see Lembcke [1979]) or the Ultrafilter Principle (see Buskes and van Rooij [1992]).

We begin with vector space versions of the Hahn-Banach Theorem which do not involve any topology. In later chapters we shall present other versions in normed vector spaces, topological vector spaces, and Boolean algebras.

12.31. Real-valued, Nontopological Hahn-Banach Theorems. Following are our most basic versions of the Hahn-Banach Theorem; we shall show that they are equivalent to one another and are consequences of the Axiom of Choice. However, the proofs will be postponed until 12.36, 12.37, and 12.38, where we present proofs of more general results.

Many of the Hahn-Banach Theorems can be extended to complex vector spaces via the Bohnenblust-Sobczyk Correspondence (11.12): If X is a complex vector space on which λ is a linear functional, then X can also be viewed as a real vector space on which $\mathrm{Re}\,\lambda$ is a linear functional. We shall omit the details of that argument; for simplicity we shall generally only consider real vector spaces.

For brevity we combine certain theorems. Theorems (HB2) assumes p is convex where (HB3) assumes p is sublinear; otherwise those two theorems are identical. Theorems (HB4) and (HB5) differ in the same fashion. Of course, in each case the sublinear version is just a weakened form of the convex version, since any sublinear function is convex. The sufficiency of convexity was noted at least as early as Nakano [1959], but it seems not to be widely known that the assumption of sublinearity can be replaced by the weaker hypothesis of convexity. Most of the literature assumes sublinearity — a notable exception being the excellent textbook of Reed and Simon [1972]. We will use convexity instead of sublinearity throughout this book.

It is interesting to compare (HB4) with 16.16(D) and (HB17) (see 28.4). It is also interesting to compare (HB6) with Dowker's Sandwich Theorem 16.30.

Banach limits, introduced below, will be discussed further in 12.33. Note that any member of $B(\Delta)$ is a bounded net of real numbers, hence it has a limsup.

> **(HB1) Existence of Banach Limits.** Let $(\Delta, \preccurlyeq)$ be a directed set, and let $B(\Delta) = \{\text{bounded functions from } \Delta \text{ into } \mathbb{R}\}$. Then there exists a real-valued *Banach limit* for $(\Delta, \preccurlyeq)$ — that is, a linear map $\mathrm{LIM} : B(\Delta) \to \mathbb{R}$ that satisfies $\mathrm{LIM}(f) \le \limsup_{\delta\in\Delta} f(\delta)$ for each $f \in B(\Delta)$.
>
> **(HB2) Convex Extension Theorem** and **(HB3) Sublinear Extension Theorem.** Suppose X is a real vector space, X_0 is a linear subspace, $\lambda : X_0 \to \mathbb{R}$ is a linear map, $p : X \to \mathbb{R}$ is a convex (or sublinear) function, and $\lambda \le p$ on X_0. Then λ can be extended to a linear map $\Lambda : X \to \mathbb{R}$ that satisfies $\Lambda \le p$ on X.
>
> **(HB4) Convex Support Theorem** and **(HB5) Sublinear Support Theorem.** Any convex (or sublinear) function from a real vector space into $\mathbb{R}$ is the pointwise maximum of the affine functions that lie below it. That is, if $p : X \to \mathbb{R}$ is convex (respectively, sublinear), then for each $x_0 \in X$ there exists some affine function $f : X \to \mathbb{R}$ that satisfies $f(x) \le p(x)$ for all $x \in X$ and

$f(x_0) = p(x_0)$.

(HB6) Sandwich Theorem. Let C be a convex subset of a real vector space. Suppose that $e : C \to \mathbb{R}$ is a concave function, $g : C \to \mathbb{R}$ is a convex function, and $e \leq g$ everywhere on C. Then there exists an affine function $f : C \to \mathbb{R}$ satisfying $e \leq f \leq g$.

Proofs will be given in 12.36, 12.37, and 12.38.

Convex Operators

12.32. Before proving that the principles in 12.31 follow from each other and from the Axiom of Choice, we shall generalize slightly. This will require more definitions:

Definitions. Let C be a convex subset of a vector space and let $(Z, \preccurlyeq)$ be an ordered vector space (not necessarily a vector lattice). A mapping $p : C \to Z$ is

sublinear	if $p(x+y) \preccurlyeq p(x) + p(y)$ and $p(tx) = tp(x)$ for all $t \in [0, +\infty)$;
convex	if $p(tx + (1-t)y) \preccurlyeq tp(x) + (1-t)p(y)$ for all $t \in [0,1]$;
concave	if $p(tx + (1-t)y) \succcurlyeq tp(x) + (1-t)p(y)$ for all $t \in [0,1]$;
affine	if $p(tx + (1-t)y) = tp(x) + (1-t)p(y)$ for all $t \in [0,1]$;

in each case the condition is to hold for all $x, y \in C$. Note that any sublinear function is convex.

Remarks. The theory of convex operators obviously includes the theory of convex real-valued functions. It also includes the theory of affine operators between (unordered) linear spaces, for if $f : X \to Y$ is any affine mapping, we can make it into a convex operator by equipping Y with the trivial ordering described in 8.37. We shall study convex operators further in Chapters 26 and 27.

Another way to unite the theories of linear operators and convex functionals is via Ioffe's "fans;" these are convex-set-valued functions. An introduction to the subject and further references are given by Ioffe [1982].

12.33. The notion of generalized limits can be traced back at least as far as Banach [1932]. However, the precise definition of "Banach limit" varies slightly from one paper to another in the literature. In the case of $Z = \mathbb{R}$, our definition of "Banach limit" agrees with the definition given by Yosida [1964].

Let $(\Delta, \sqsubseteq)$ be a directed set, and let $(Z, \preccurlyeq)$ be a Dedekind complete, ordered vector space. (The reader should keep in mind the special case of $Z = \mathbb{R}$; that is the simplest and most important case.) Let

$$B(\Delta, Z) \quad = \quad \{\text{bounded functions from } \Delta \text{ into } Z\};$$

in this context a function f is **bounded** if its range has an upper bound and a lower bound. Note that $B(\Delta, Z)$ is itself an ordered vector space.

A member of $B(\Delta, Z)$ may be viewed as a bounded net based on Δ, taking values in Z. Since Z is Dedekind complete, for any $f \in B(\Delta, Z)$ the objects

$$\liminf(f) \;=\; \sup_{\alpha\in\Delta}\,\inf_{\beta\sqsupseteq\alpha} f(\beta), \qquad\qquad \limsup(f) \;=\; \inf_{\alpha\in\Delta}\,\sup_{\beta\sqsupseteq\alpha} f(\beta)$$

both exist in Z, and $\liminf(f) \preccurlyeq \limsup(f)$. We say that the net f **converges** if and only if $\liminf(f)$ and $\limsup(f)$ are equal, in which case their common value is the **limit** of f, denoted $\lim(f)$. We may sometimes refer to this as the **order limit**. We may also call it the **ordinary limit**, to contrast it with the generalized limit developed in the next few paragraphs.

Note that (*exercise*) the order limit is a positive linear operator, from a linear subspace of $B(\Delta, Z)$ into Z. That linear subspace — i.e., the space of all convergent nets — generally is not all of $B(\Delta, Z)$, since some bounded nets are not convergent in the sense of ordinary limits.

Let $\mathrm{LIM} : B(\Delta) \to Z$ be a linear map. At the end of this section we shall show that the following three conditions are equivalent:

(A) $\mathrm{LIM}(f) \preccurlyeq \limsup_{\delta\in\Delta} f(\delta)$ for each $f \in B(\Delta, Z)$.

(B) $\liminf_{\delta\in\Delta} f(\delta) \preccurlyeq \mathrm{LIM}(f) \preccurlyeq \limsup_{\delta\in D} f(\delta)$ for each $f \in B(\Delta, Z)$.

(C) LIM is a positive operator that extends the ordinary limit to all of $B(\Delta, Z)$. That is, $f \succcurlyeq 0 \;\Rightarrow\; \mathrm{LIM}(f) \succcurlyeq 0$, and $\mathrm{LIM}(f) = \lim_{\delta\in\Delta} f(\delta)$ whenever the right side of that equation exists.

If one, hence all, of these three conditions are satisfied, we say LIM is a ***Z*-valued Banach limit** for the directed set $(\Delta, \sqsubseteq)$ If, moreover, $(\Delta, \sqsubseteq)$ is the ordered set $(\mathbb{N}, \leq)$ — that is, the positive integers with their usual ordering — then we shall call LIM a **sequential Banach limit**.

In the next few sections we shall prove the existence of Banach limits. The Banach limit is an extension of the ordinary limit; it gives us a way of saying that, in a generalized sense, *every* bounded net "converges." In particular, when $Z = \mathbb{R}$, it says that every bounded net of real numbers "converges" to a real number; a real-valued sequential Banach limit is a way of saying that every bounded sequence of real numbers converges to a real number. We emphasize that there may be many different Z-valued Banach limits for a directed set $(\Delta, \sqsubseteq)$. They agree on those nets that converge in the usual sense, but they may give different generalized limits for those nets that do not converge in the usual sense.

Banach limits could be contrasted with limsups and liminfs (discussed as "generalized limits" in 7.46). Banach limits have slightly better algebraic properties — they are given by a linear map — but they do not preserve topological properties quite as well as the limsup and liminf do.

Proof of equivalence: First, (A) $\iff$ (B) since $\limsup f = -\liminf(-f)$ and $\mathrm{LIM}(-f) = -\mathrm{LIM}(f)$ by the linearity of LIM. Clearly, condition (A) implies that LIM is a positive operator, and condition (B) implies that LIM agrees with lim wherever the latter exists.

It remains only to prove (C) $\Rightarrow$ (A). For each $\varepsilon \in \Delta$, let $u(\varepsilon) = \sup_{\delta\sqsupseteq\varepsilon} f(\delta)$. Then $u(\varepsilon) \succcurlyeq f(\varepsilon)$; thus $u - f \succcurlyeq 0$, so $\mathrm{LIM}(u - f) \succcurlyeq 0$. Also, the net $(u(\varepsilon) : \varepsilon \in \Delta)$ decreases to $\limsup_{\delta\in\Delta} f(\delta)$. Thus $\mathrm{LIM}(f) \preccurlyeq \mathrm{LIM}(u) = \lim_{\varepsilon\in\Delta} u(\varepsilon) = \limsup_{\delta\in\Delta} f(\delta)$.

12.34. Vector-valued Hahn-Banach Theorems. We now generalize the theorems of 12.31. We shall show that the following principles are equivalent to each other and that they are all consequences of the Axiom of Choice. (It is not yet known whether they are equivalent to the Axiom of Choice or are strictly weaker.)

Hypothesis. Let Z be a Dedekind complete, ordered vector space.

(VHB1) Existence of Banach Limits. If $(\Delta, \preccurlyeq)$ is any directed set, then there exists a Z-valued Banach limit for $(\Delta, \preccurlyeq)$ (as defined in 12.32).

(VHB2) Convex Extension Theorem and **(VHB3) Sublinear Extension Theorem.** Suppose X is a real vector space, X_0 is a linear subspace, $\lambda : X_0 \to Z$ is a linear map, $p : X \to Z$ is a convex (or sublinear) function, and $\lambda \preccurlyeq p$ on X_0. Then λ can be extended to a linear map $\Lambda : X \to Z$ that satisfies $\Lambda \preccurlyeq p$ on X.

(VHB4) Convex Support Theorem and **(VHB5) Sublinear Support Theorem.** Any convex (or sublinear) function from a real vector space into Z is the pointwise maximum of the affine functions that lie below it. That is, if $p : X \to Z$ is convex (respectively, sublinear), then for each $x_0 \in X$ there exists some affine function $f : X \to Z$ that satisfies $f(x) \preccurlyeq p(x)$ for all $x \in X$ and $f(x_0) = p(x_0)$.

(VHB6) Sandwich Theorem. Let C be a convex subset of a real vector space. Suppose that $e : C \to Z$ is a concave function, $g : C \to Z$ is a convex function, and $e \preccurlyeq g$ everywhere on C. Then there exists an affine function $f : C \to Z$ satisfying $e \preccurlyeq f \preccurlyeq g$.

Proofs will be given in the next few sections.

12.35. Finite Extension Lemma (FEL). Suppose X is a real vector space and Z is a Dedekind complete, ordered vector space. Suppose $X_0 \subseteq X$ is a linear subspace, $\lambda_0 : X_0 \to Z$ is linear, $p : X \to Z$ is convex, and $\lambda_0 \preccurlyeq p$ on X_0. Also assume that

$$X = \operatorname{span}(X_0 \cup S) \qquad \text{for some finite set } S \subseteq X. \tag{$*$}$$

Then λ_0 can be extended to a linear map $\lambda : X \to Z$ satisfying $\lambda \preccurlyeq p$ on X.

Remarks. We shall use FEL in proving (VHB2). Note that FEL differs from (VHB2) only in the addition of the hypothesis $(*)$.

FEL does not require the Axiom of Choice or any of its weaker relatives; FEL can be proved using just ZF. FEL or a similar result was already known to Banach; it appears explicitly in Luxemburg [1969].

Proof of FEL. The proof is by induction on the cardinality of S; thus we may assume S contains just one element ξ. Using the linearity of λ_0, the convexity of p, and the fact that

$\lambda_0 \preccurlyeq p$ on X_0, we can verify that

$$\sup_{w \in X_0,\ s<0} \frac{p(w+s\xi)-\lambda_0(w)}{s} \quad \preccurlyeq \quad \inf_{v \in X_0,\ r>0} \frac{p(v+r\xi)-\lambda_0(v)}{r}$$

(the details of the verification are left as an exercise). Now let $\lambda(\xi)$ be any member of Z lying between those two values. The function λ, being linear, must be defined by

$$\lambda(x+r\xi) \quad = \quad \lambda_0(x)+r\lambda(\xi) \qquad \text{for all } x \in X_0,\ r \in \mathbb{R}.$$

From our choice of $\lambda(\xi)$ it follows that $\lambda \preccurlyeq p$ on X (again, the details of the verification are left as an exercise). This completes the proof.

12.36. *Proof of* AC $\Rightarrow$ (VHB2). We shall give two different proofs.

Version (i): This is the more traditional proof. Consider all linear maps $\Lambda : W \to Z$, where W is a linear subspace of X that includes X_0 and Λ is an extension of λ that satisfies $\Lambda \preccurlyeq p$ on W. Partially order such Λ's by inclusion of their graphs. By Zorn's Lemma, there is a maximal member of this partially ordered set — i.e., an extension $\Lambda : W \to Z$ that cannot be extended farther. If $W \subsetneqq X$, choose any $\xi \in X \setminus W$. By the Finite Extension Lemma 12.35, Λ can be extended to all of $\operatorname{span}(W \cup \{\xi\})$ — contradicting the maximality of W. Thus $W = X$, completing the proof.

Version (ii): Some mathematicians may find the Finite Character Principle more intuitively appealing than Zorn's Lemma, so we offer an alternative proof. Consider all functions $\Lambda : W \to Z$ where W is a subset of X that includes X_0, Λ is an extension of λ that satisfies $\Lambda \preccurlyeq p$ on W, and Λ is a function that can be extended to a linear function (see 11.10). Partially order such Λ's by inclusion of their graphs. Use the Finite Character Principle ((AC5), in 6.20) to show that there is a maximal Λ in this collection.

12.37. *Most of the equivalence proofs.* In this section we prove most of the equivalences stated in 12.34; the one remaining argument is much longer and will be given separately in 12.38.

Proof of (VHB2) $\Rightarrow$ (VHB3). Obvious.

Proof of (VHB2) $\Rightarrow$ (VHB4). Define $q(x) = p(x+x_0) - p(x_0)$. Then q is also convex, and $q(0) = 0$. By (VHB2) (with $X_0 = \{0\}$), there exists some linear function $g : X \to Z$ that satisfies $g \preccurlyeq q$ everywhere on X. Now let $f(x) = g(x - x_0) + p(x_0)$; this completes the proof.

Proof of (VHB3) $\Rightarrow$ (VHB1). Let $X = B(\Delta, Z)$, and let $X_0 = \{\text{convergent nets}\}$. Let $p(x) = \limsup_{\delta\in\Delta} x(\delta)$; this functional is easily verified to be sublinear. For $x \in X_0$, let $\lambda(x) = \lim x$; then $\lambda \preccurlyeq p$ on X_0.

Proof of (VHB4) $\Rightarrow$ (VHB5). Obvious.

Proof of (VHB5) $\Rightarrow$ (VHB1). The mapping $\Phi : f \mapsto \limsup_{\delta\in\Delta} f(\delta)$ is a sublinear mapping from $B(\Delta, Z)$ into Z, which vanishes on the zero element of the linear space $B(\Delta, Z)$. Let $\mathrm{LIM} : B(\Delta, Z) \to Z$ be an affine mapping satisfying $\mathrm{LIM} \preccurlyeq \Phi$ and $\mathrm{LIM}(0) = \Phi(0)$.

Proof of (VHB1) $\Rightarrow$ (VHB2). Let $f_0 : X_0 \to Z$ and p be given. For each finite set $S \subseteq X$, let Φ_S be the set of all functions $g : X \to Z$ that have the following properties: g is an extension of f_0, $-p(-x) \preccurlyeq g(x) \preccurlyeq p(x)$ for all $x \in X$, and the restriction of g to $\operatorname{span}(X_0 \cup S)$ is linear. By Banach's Finite Extension Lemma (in 12.35), each Φ_S is nonempty. Since $\Phi_S \cap \Phi_T = \Phi_{S \cup T}$, the family of sets Φ_S has the finite intersection property. Now let

$$\Delta \quad = \quad \{(g, S) \quad : \quad S \text{ is a finite subset of } X \text{ and } g \in \Phi_S\},$$

and let Δ be ordered by: $(g_1, S_1) \preccurlyeq (g_2, S_2)$ if $S_1 \subseteq S_2$; then Δ is a directed set. Define $\mathrm{LIM} : B(\Delta, Z) \to Z$ as in (VHB1).

For each $x \in X$, define a function $\psi_x : \Delta \to Z$ by taking $\psi_x(g, S) = g(x)$. Then ψ_x is bounded, since $-p(-x) \preccurlyeq \psi_x(\delta) \preccurlyeq p(x)$ for all $\delta \in \Delta$. Define a function $f : X \to Z$ by taking $f(x) = \mathrm{LIM}(\psi_x)$. Observe that $f(x) = \mathrm{LIM}(\psi_x) \preccurlyeq \mathrm{LIM}(p(x)) = p(x)$, since $p(x)$ does not depend on δ. Also observe that if $x \in X_0$, then $\psi_x(\delta) = f_0(x)$, so $f(x) = \mathrm{LIM}(f_0(x)) = f_0(x)$.

It remains to show that f is linear. Fix any $x, y \in X$ and $\alpha, \beta \in Z$. Then for all $\delta = (g, S)$ sufficiently large in Δ, we have $x, y \in S$, so g is linear on the span of $\{x, y\}$. Thus $g(\alpha x + \beta y) - \alpha g(x) - \beta g(y) = 0$. Therefore $\psi_{\alpha x + \beta y}(\delta) - \alpha\psi_x(\delta) - \beta\psi_y(\delta) = 0$ for all δ sufficiently large. Since LIM is a linear operator, we have

$$\begin{aligned} f(\alpha x + \beta y) - \alpha f(x) - \beta f(y) \quad &= \quad \mathrm{LIM}(\psi_{\alpha x + \beta y}) - \alpha \mathrm{LIM}(\psi_x) - \beta \mathrm{LIM}(\psi_y) \\ &= \quad \mathrm{LIM}(\psi_{\alpha x + \beta y} - \alpha\psi_x - \beta\psi_y) \quad = \quad 0. \end{aligned}$$

(The proof above is a reformulation of an argument of Luxemburg [1969]. We use nets where Luxemburg used reduced powers of Z and the terminology of nonstandard analysis; our condition (VHB1) is essentially a translation of Luxemburg's Theorem 6.1.)

Proof of (VHB6) $\Rightarrow$ (VHB1). Let $e = \liminf$ and $g = \limsup$.

12.38. *Proof of* (VHB1) $\Rightarrow$ (VHB6). This proof takes several ingredients from Neumann [1994]. However, Neumann was not concerned with weak forms of Choice, and so he used Zorn's Lemma to find a minimal element of the set $\mathcal{M}$ investigated below. In our present investigation of weak forms of Choice, we are not permitted to use Zorn's Lemma — an equivalent of AC — and so it is not clear that $\mathcal{M}$ has a minimal element. Instead, in our proof (heretofore unpublished) we shall use infinitely many decreasing sequences in $\mathcal{M}$.

Let $\iota = \inf\{g(x) - e(x) : x \in C\}$; then $\iota \in Z_+$. By replacing g with $g - \iota$, we may assume $\iota = 0$. Let $\mathcal{M}$ be the set of those convex functions $f : C \to Z$ that satisfy $e \preccurlyeq f \preccurlyeq g$ on C. The set $\mathcal{M}$ is nonempty, since $g \in \mathcal{M}$.

The proof will be in several steps. We first show that

(a) From any $f \in \mathcal{M}$ and $\alpha, \beta \in (0, 1)$ with $\alpha + \beta = 1$, we can canonically construct a function $p \in \mathcal{M}$ that satisfies $p(x) \preccurlyeq f(x)$ and $p(\alpha x + \beta y) \succcurlyeq \alpha p(x) + \beta e(y)$ for all $x, y \in C$.

(The term "canonically" here refers to the fact that no arbitrary choices are needed; the function p is constructed from α, β, and f by a uniquely specified algorithm. That fact is important, since we shall apply this construction infinitely many times later in this proof.)

To prove (a), we define a decreasing sequence $f_0 \succcurlyeq f_1 \succcurlyeq f_2 \succcurlyeq f_3 \succcurlyeq \cdots$ in $\mathcal{M}$ as follows: Let $f_0 = f$. Now assume some $f_n \in \mathcal{M}$ is given. Since e is concave and $e \preccurlyeq f_n$, we have $e(x) \preccurlyeq [f_n(\alpha x + \beta y) - \beta e(y)]/\alpha$ for all $x, y \in C$. Hence we may define a function $f_{n+1} : C \to Z$ by

$$f_{n+1}(x) \quad = \quad \inf\left\{\frac{f_n(\alpha x + \beta y) - \beta e(y)}{\alpha} \quad : \quad y \in C\right\}. \tag{1}$$

Then $e \preccurlyeq f_{n+1}$. From the convexity of f_n and the concavity of e, we find that f_{n+1} is convex, too (easy exercise). From $f_n \preccurlyeq g$ and the convexity of f_n, we obtain

$$\frac{f_n(\alpha x + \beta y) - \beta e(y)}{\alpha} \preccurlyeq f_n(x) + \frac{\beta}{\alpha}[f_n(y) - e(y)] \preccurlyeq f_n(x) + \frac{\beta}{\alpha}[g(y) - e(y)]$$

for all $x, y \in C$. Hold x fixed and take the infimum over all y; this yields $f_{n+1}(x) \preccurlyeq f_n(x)$ since $\inf\{g(y) - e(y) : y \in C\} = 0$. Thus $f_{n+1} \in \mathcal{M}$, too. This completes the recursive construction of the sequence $\{f_n\}$.

Now define $p(x) = \inf_{n \in \mathbb{N}} f_n(x)$. The pointwise infimum of a chain of convex functions is convex, so p is convex. Hence $p \in \mathcal{M}$. By taking the infimum over all n in both sides of (1), we obtain

$$p(x) \quad = \quad \inf\left\{\frac{p(\alpha x + \beta y) - \beta e(y)}{\alpha} \ : \ y \in C\right\}.$$

This completes the proof of (a).

Next we shall show that

(b) From any function $f \in \mathcal{M}$ and any finite sequence $\big((\alpha_1, \xi_1), (\alpha_2, \xi_2), \ldots, (\alpha_J, \xi_J)\big)$ in $(0,1) \times C$, we can canonically construct a function $q \in \mathcal{M}$ that satisfies $q(y) \preccurlyeq f(y)$ and $q(\alpha_j \xi_j + (1 - \alpha_j)y) = \alpha_j q(\xi_j) + (1 - \alpha_j)q(y)$ for all $y \in C$ and $1 \le j \le J$.

Extend the given finite sequence to an infinite sequence $((\alpha_j, \xi_j) : j \in \mathbb{N})$ in $(0,1) \times C$ by making it periodic in j with period J. Let $\beta_j = 1 - \alpha_j$. Construct a sequence

$$f_1 \succcurlyeq p_1 \succcurlyeq f_2 \succcurlyeq p_2 \succcurlyeq f_3 \succcurlyeq p_3 \succcurlyeq \cdots \qquad \text{in } \mathcal{M} \tag{2}$$

as follows: Let $f_1 = f$. Given f_n, choose $p_n \in \mathcal{M}$ as in statement (a) — that is, with $p_n \preccurlyeq f_n$ and with

$$p_n(\alpha_n x + \beta_n y) \succcurlyeq \alpha_n p_n(x) + \beta_n e(y) \quad \text{for all } x, y \in C. \tag{3}$$

Then we can define a convex function $f_{n+1} : C \to Z$ by

$$f_{n+1}(y) \quad = \quad \frac{p_n(\alpha_n \xi_n + \beta_n y) - \alpha_n p_n(\xi_n)}{\beta_n} \qquad \text{for } y \in C. \tag{4}$$

From (3) we may deduce that $e \preccurlyeq f_{n+1}$ on C. On the other hand, $f_{n+1} \preccurlyeq p_n$ follows from the convexity of p_n. Thus $f_{n+1} \in \mathcal{M}$, completing the recursive construction of the sequence (2). Now let q be the pointwise infimum of the functions in that sequence. Then

q is the order limit of the f_n's, as well as the order limit of the p_n's. The order limit is a linear operator, so it preserves linear combinations. Fix some particular j, and consider the subsequence of equations obtained from (4) by taking $n = j, j+J, j+2J, j+3J, \ldots$. Taking order limits, we obtain

$$q(y) \quad = \quad \frac{q(\alpha_j \xi_j + \beta_j y) - \alpha_j q(\xi_j)}{\beta_j} \qquad \text{for } y \in C.$$

This proves (b).

Finally we proceed to our main construction. Let Δ be the set of all ordered pairs (Q, q) where Q is a finite subset of $(0,1) \times C$ and q is a member of $\mathcal{M}$ that satisfies

$$q(\alpha\xi + (1-\alpha)y) = \alpha q(\xi) + (1-\alpha)q(y) \quad \text{for all } y \in C \text{ and } (\alpha, \xi) \in Q.$$

For such an ordered pair (Q, q) and for each $x \in C$, let $\Psi_x(Q, q) = q(x)$. In this fashion we define a function $\Psi_x : \Delta \to Z$. This function has bounded range: $e(x) \preccurlyeq \Psi_x(Q, q) \preccurlyeq g(x)$ for all $(Q, q) \in \Delta$. Thus $\Psi_x \in B(\Delta, Z)$.

Say that $(Q, q) \sqsubseteq (R, r)$ whenever $Q \subseteq R$; from statement (b) it follows that $\sqsubseteq$ is a directed ordering of Δ. By (VHB1) there exists a Z-valued Banach limit $\mathrm{LIM} : B(\Delta, Z) \to Z$. Define $f(x) = \mathrm{LIM}(\Psi_x)$. Then $e(x) \preccurlyeq f(x) \preccurlyeq g(x)$.

To show f is affine, fix any $x, y \in C$ and $\alpha \in (0,1)$. For all (Q, q) sufficiently large in Δ, we have $(\alpha, x) \in Q$. Then $q(\alpha x + (1-\alpha)y) = \alpha q(x) + (1-\alpha)q(y)$. That is,

$$\Psi_{\alpha x + (1-\alpha)y}(Q, q) \quad = \quad \alpha \Psi_x(Q, q) + (1-\alpha)\Psi_y(Q, q)$$

for all (Q, q) sufficiently large. Apply the Banach limit on both sides; we obtain $f(\alpha x + (1-\alpha)y) = \alpha f(x) + (1-\alpha)f(y)$. This completes the proof.

Chapter 13

Boolean Algebras

Boolean Lattices

13.1. *Definition.* Let $(X, \preccurlyeq)$ be a lattice that has a first element and a last element — denoted 0 and 1, respectively. If $x \in X$, then a **complement** of x is an element y that satisfies

$$x \wedge y = 0 \qquad \text{and} \qquad x \vee y = 1.$$

The lattice X is **complemented** if each of its elements has at least one complement.

Recall from 4.23 that a lattice is **distributive** if its binary operations $\vee, \wedge$ distribute over each other — i.e., if

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \qquad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

for all $x, y, z \in X$. *Exercise.* If X is a distributive lattice with smallest and largest elements 0 and 1, then each member of X has at most one complement.

A **Boolean lattice** is a complemented lattice that is also distributive. (Some mathematicians add the further requirement that $0 \neq 1$, but we shall not impose that restriction; see the remarks in 13.4.a and 13.13.) A Boolean lattice is **complete** if its ordering is complete — i.e., if every subset $S \subseteq X$ has a supremum (written $\bigvee S$) and an infimum (written $\bigwedge S$).

By the exercise above, if X is a Boolean lattice, then each $x \in X$ has exactly one complement, which we shall denote by $\complement x$. It is convenient to also define the **symmetric difference** of two elements x and y:

$$x \bigtriangleup y = (x \wedge \complement y) \vee (\complement x \wedge y).$$

A Boolean lattice is essentially the same thing as a *Boolean algebra*, and the two terms may be used interchangeably. However, the term "Boolean algebra" emphasizes the universal algebra viewpoint, as discussed in 13.7. Boolean rings are introduced in 13.13; although Boolean rings and Boolean lattices are not the same, there is a natural correspondence between them, and for that reason the terms "Boolean rings" and "Boolean lattices" are occasionally used interchangeably.

Much of this chapter is based on Halmos [1963], Monk [1989], and Sikorski [1964].

13.2. *Exercise* (*optional*). Let X be a distributive lattice with smallest and largest elements 0 and 1. Then the set S of all complemented elements of X is a sublattice of X; it is a Boolean lattice if we restrict the lattice operations of X to S. If X contains more than one element, then no element of X can be its own complement.

13.3. The basic example: algebras of sets. If $\mathcal{S}$ is an algebra of subsets of a set Ω (as defined in 5.25), then $\mathcal{S}$ is a Boolean lattice, when ordered by $\subseteq$. In this case we have a conversion of symbols as described in the table below. Of course, in the algebra of sets, the complement of a set S is the set $\complement S = \{x \in \Omega : x \notin S\}$. We emphasize that in Boolean lattices, $\complement$ may have other meanings.

Algebra of subsets of Ω	$\subseteq$	$\varnothing$	Ω	$\cup$	$\cap$	$\complement$	$\triangle$
Boolean lattice	$\preccurlyeq$	0	1	$\vee$	$\wedge$	$\complement$	$\triangle$

Boolean lattices are not really much more general than algebras of sets. In fact, the *Stone Representation Theorem*, proved later in this chapter, states that every Boolean algebra is isomorphic to some algebra of sets. However, that isomorphism is sometimes an inconvenient representation; the proof of the Stone Representation Theorem involves arbitrary choices and intangibles.

Aside from the conceptual difficulty of intangibles, the chief difference between Boolean lattices and algebras of sets is one of viewpoint: When considering algebras of sets, we are permitted to consider the *points* that make up those sets. In contrast, the members of a Boolean lattice are considered as urelements, not necessarily containing any "points." (Compare the remarks in 5.21 about "pointless" open sets.)

13.4. Further examples of Boolean lattices.

a. If $0 = 1$ in a Boolean lattice X, then X contains just the single element 0. We shall call $\{0\}$ the **degenerate Boolean lattice**; it is the smallest Boolean lattice. It is isomorphic to $\mathcal{P}(\varnothing) = \{\varnothing\}$.

Any Boolean lattice with $0 \neq 1$ will be called **nondegenerate**.

To call $\{0\}$ a Boolean lattice reflects a recent trend among algebraists. The older literature imposed the restriction that $0 \neq 1$ in any Boolean lattice (and that additional restriction is still imposed by some mathematicians today). That restriction only excludes one Boolean lattice, and so that restriction has little effect on the ultimate results of the theory if one is careful to keep track of the degenerate case. However, that restriction complicates the notation and the development of the theory, because with that restriction Boolean algebras and Boolean rings (discussed later in this chapter) do not form equational varieties. We emphasize that

in this book, $\{0\}$ *is a Boolean lattice*

and so Boolean algebras and Boolean rings do form equational varieties.

b. The next smallest Boolean lattice is the set $2 = \{0, 1\}$ consisting of two elements, ordered by $0 \prec 1$. This example, though quite elementary, is extremely important; it will be used in 13.19. The set 2 is isomorphic to $\mathcal{P}(S)$ if S is any singleton.

c. Some algebras of sets, such as $\mathcal{P}(\Omega)$, are complete Boolean lattices.

Others are not complete. For instance, let X be the algebra of all finite or cofinite subsets of $\mathbb{Z}$. (This is a special case of 5.26.f.) Show that $S = \{\text{finite subsets of the set of even integers}\}$ is a subset of X that does not have a least upper bound in X.

Hints: Any upper bound for S is a set B that contains all the even integers; hence it is not finite; hence it is cofinite; hence it contains all but finitely many odd integers. If r is an odd integer belonging to B, then $B \setminus \{r\}$ is a slightly smaller upper bound for S.

d. The lattice M_3 given in 4.18 is complemented but not distributive; thus it is *not* a Boolean lattice.

e. Let S be a topological space. Recall that a subset of S is **clopen** if it is both open and closed. The collection $\text{clop}(S) = \{\text{clopen subsets of } S\}$ is an algebra of sets, and thus a Boolean lattice. It will play an important role in 17.44.

f. Let X and $\perp$ be as in 4.12, and let $\mathcal{C} = \{S \subseteq X \ : \ S^{\perp\perp} = S\}$ — that is, let $\mathcal{C}$ be the collection of closed subsets of X. Show that $(\mathcal{C}, \subseteq)$ is a complete Boolean lattice, with

$$\bigwedge_{t\in T} A_t = \bigcap_{t\in T} A_t, \qquad \bigvee_{t\in T} A_t = \mathrm{cl}\left(\bigcup_{t\in T} A_t\right), \qquad \complement A = A^{\perp}.$$

Hint: To prove $(\mathcal{C}, \subseteq)$ is distributive, observe that for any sets $P, Q \subseteq X$ we have $\mathrm{cl}(P \cap Q) = \mathrm{cl}(P) \cap \mathrm{cl}(Q)$, as noted in 4.12. Apply this with $P = A \cup B$ and $Q = A \cup C$ to show that

$$A \vee (B \wedge C) \quad = \quad (A \vee B) \wedge (A \vee C) \qquad \text{for any } A, B, C \in \mathcal{C}.$$

g. Let $\mathcal{L}$ be some language, and let $\mathbb{F}$ be the set of all formulas that can be formed in that language. Assume that the language is equipped with some suitable collection of axioms and rules of inference, which reflect ordinary methods of reasoning. (A precise specification of such "suitable" axioms and rules will be given in 14.32.) Call two formulas $\mathcal{A}$ and $\mathcal{B}$ "equivalent" provided each implies the other via the given axioms and rules of inference. This turns out to be an equivalence relation on $\mathbb{F}$. The resulting quotient algebra is a Boolean lattice, with the binary operations $\wedge$, $\vee$, and $\complement$ corresponding to the logical notions "and," "or," and "not" respectively. See 14.33.

h. Let Ω be a topological space; define closure and interior as in 5.16. A set $S \subseteq \Omega$ is **regular open** if $S = \mathrm{int}(\mathrm{cl}(S))$. Clearly, any clopen set is regular open; a topological space may have other regular open sets as well. The regular open sets may be described as those open sets that have no "cracks" or "pinholes." The collection $RO(\Omega) = \{\text{regular open subsets of } \Omega\}$, ordered by $\subseteq$, forms a complete Boolean lattice, with Boolean lattice operations given by

$$\bigvee_{i\in I} S_i = \mathrm{int}\left(\mathrm{cl}\left(\bigcup_{i\in I} S_i\right)\right), \qquad \bigwedge_{i\in I} S_i = \mathrm{int}\left(\bigcap_{i\in I} S_i\right), \qquad \complement S = \Omega \backslash \mathrm{cl}(S)$$

for $S, S_i \in RO(\Omega)$. Although $RO(\Omega)$ is a subcollection of $\mathcal{P}(\Omega)$ and its ordering is the restriction of the ordering of $\mathcal{P}(\Omega)$, the Boolean lattice operations of $RO(\Omega)$ generally are *not* just the restrictions of the Boolean lattice operations of $\mathcal{P}(\Omega)$. The Boolean lattice $RO(\Omega)$ may appear rather complicated, but it arises naturally in certain applications: It turns out to be the *smallest* complete Boolean lattice that fits certain constructions. It plays an important role in the theory of forcing. See 14.53 and Bell [1985].

Another Boolean lattice that is of particular interest to analysts is given in 21.9.

13.5. *A few computations.* If X is a Boolean lattice, show that

a. $\complement\complement x = x$. (Thus, $\complement$ is an involution of X, in the sense of 2.4.)

b. $x \preccurlyeq y \iff \complement x \succcurlyeq \complement y$.

c. $x = y \iff x \triangle y = 0 \iff x \wedge \complement y = \complement x \wedge y = 0$.

d. De Morgan's Laws. $\complement(x \vee y) = (\complement x) \wedge (\complement y)$ and $\complement(x \wedge y) = (\complement x) \vee (\complement y)$.

e. $b \succcurlyeq \complement b \iff b = 1$, and $b \preccurlyeq \complement b \iff b = 0$. Hence no Boolean lattice with more than two elements is also a chain.

13.6. *The duality principle.* Whenever $B = (X, \preccurlyeq, 0, 1, \complement, \wedge, \vee)$ is a Boolean lattice, then $B^{op} = (X, \succcurlyeq, 1, 0, \complement, \vee, \wedge)$ is another Boolean lattice — i.e., we obtain a new Boolean lattice if we keep the same set X and the same complementation operation, but swap 0 and 1, and swap meets and joins. (In fact, the mapping $x \mapsto \complement x$ is an isomorphism, in the sense of 13.7, from B onto B^{op}.)

Any statement about Boolean lattices has a *dual* statement that follows as a consequence by this swapping. For instance, the two De Morgan's Laws in 13.5.d are dual to each other. When two statements are dual to each other in this fashion, for brevity we may state just one of them.

Boolean Homomorphisms and Subalgebras

13.7. *Definitions.* We may view Boolean lattices as an equational variety, in the sense of 8.50. The **fundamental operations** are $\vee, \wedge, \complement, 0, 1$. A Boolean lattice satisfies the axioms of a lattice (that is, L1–L3 in 4.20), together with these axioms:

$$x \wedge 0 = x \wedge \complement x = 0, \qquad x \vee 1 = x \vee \complement x = 1.$$

A Boolean lattice, viewed as an algebraic system in this fashion, is usually called a **Boolean algebra**. We may occasionally revert to the term "Boolean lattice" to emphasize the ordering structure. We emphasize that, in this book, the singleton $\{0\}$ is a Boolean algebra (albeit a degenerate one); see the remarks in 13.4.a.

A **Boolean homomorphism** is a homomorphism in this variety — i.e., a mapping that preserves the fundamental operations. Thus, a *Boolean homomorphism* means a mapping

$f : X \to Y$ from one Boolean algebra into another that satisfies

$$f(x_1 \vee x_2) = f(x_1) \vee f(x_2), \qquad f(x_1 \wedge x_2) = f(x_1) \wedge f(x_2),$$
$$f(\complement x) = \complement f(x), \qquad f(0) = 0, \qquad f(1) = 1,$$

for all $x, x_1, x_2 \in X$. We may call this a *Boolean algebra homomorphism* for emphasis or clarification. *Exercise.* It suffices to show that f preserves $\vee$ and $\complement$; the other conditions then follow as consequences. *Hint*: $0 = 0 \wedge \complement 0$.

13.8. *Definition and a concrete example.* A **two-valued homomorphism** on a Boolean algebra X is a Boolean homomorphism from X into the Boolean algebra $2 = \{0, 1\}$.

If X is an algebra of subsets of some set Ω, and $\omega_0 \in \Omega$, then one two-valued homomorphism on X is the **probability concentrated at ω_0**:

$$\mu(S) = \begin{cases} 1 & \text{if } \omega_0 \in S \\ 0 & \text{if } \omega_0 \notin S \end{cases} \qquad \text{for } S \in X.$$

Exercise. If Y is a nondegenerate Boolean algebra, then there does not exist any homomorphism from the degenerate Boolean algebra $\{0\}$ into Y. In particular, there is no two-valued homomorphism on the degenerate Boolean algebra $\{0\}$.

13.9. *More definitions.* If X is a Boolean algebra, then a **Boolean subalgebra** of X is a subobject of X in the variety of Boolean algebras — i.e., it is a set $S \subseteq X$ that is closed under the fundamental operations. Thus, a *Boolean subalgebra* of X is a nonempty set $S \subseteq X$ that satisfies

$$x_1, x_2 \in S \quad \Rightarrow \quad x_1 \vee x_2,\ x_1 \wedge x_2,\ \complement x_1,\ 0,\ 1 \ \in\ S.$$

Note that S itself is then a Boolean algebra, when equipped with the restrictions of the operations of X.

We can apply to Boolean subalgebras all the conclusions of Chapter 4 about Moore closed sets and all the conclusions of Chapter 9 about subalgebras in an equational variety. Thus, X is a Boolean subalgebra of itself; the intersection of any collection of Boolean subalgebras is a Boolean subalgebra; any homomorphic image of a Boolean subalgebra is a Boolean subalgebra; etc. The Boolean subalgebra S **generated** by a set $G \subseteq X$ is the smallest Boolean subalgebra that includes G; it is equal to the intersection of all the Boolean subalgebras that include G; the set G is then called a **generating set**, or a **set of generators**, for the Boolean subalgebra S. In the special case where $X = \mathcal{P}(\Omega)$ for some set Ω and G is a collection of subsets of Ω, we find that S is the algebra of sets generated by G (see 5.26.e).

13.10. Normal Form Theorem. Let X be a Boolean algebra, and let $G \subseteq X$. Then the Boolean subalgebra S generated by G can be described more concretely in three stages, as follows: Let

$$\begin{aligned} G_{\complement} &= \{x \in X \ :\ x \in G \text{ or } \complement x \in G\}, \\ G_{\complement\wedge} &= \{x \in X \ :\ x \text{ is the inf of finitely many members of } G_{\complement}\}, \\ G_{\complement\wedge\vee} &= \{x \in X \ :\ x \text{ is the sup of finitely many members of } G_{\complement\wedge}\}. \end{aligned}$$

(We have $1 \in G_{\complement\wedge}$ and $0 \in G_{\complement\wedge\vee}$ since the inf of no members of X is 1 and the sup of no members of X is 0.) Then $G_{\complement\wedge\vee} = S$. In other words, S consists of the elements of X that can be written in the form

$$s \quad = \quad \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} \complement^{n(i,j)}\, g_{i,j} \tag{$*$}$$

where $I, J_1, J_2, \ldots, J_I$ are nonnegative integers, the $n(i,j)$'s are nonnegative integers (or more simply, 0s and 1s), and the $g_{i,j}$'s are elements of G. An expression such as the right side of $(*)$ is said to be in **normal form**.

In particular, the subalgebra generated by a finite set is also finite.

Hints: Obviously $G_{\complement\wedge\vee}$ is closed under finite sups. Use the Distributive Law and De Morgan's Laws to show that $G_{\complement\wedge\vee}$ is also closed under finite infs and under complementation.

Remarks. An important special case is that in which $X = \mathcal{P}(\Omega)$ for some set Ω. Then $G, G_{\complement}, G_{\complement\wedge}, G_{\complement\wedge\vee}$ are collections of subsets of Ω, and "inf" and "sup" mean "intersection" and "union," respectively. The theorem above shows that the algebra of sets $\mathcal{S}$ generated by a given collection of sets $\mathcal{G}$ can be obtained by a three-stage construction.

An analogous three-stage construction does *not* work for σ-algebras: If we start with a collection $\mathcal{G}$ of subsets of Ω, and then close it under complementation, then under countable intersection, then under countable union, the resulting collection $\mathcal{A}_{\delta\sigma}$ is not necessarily equal to the σ-algebra generated by $\mathcal{G}$. Indeed, $\mathcal{A}_{\delta\sigma}$ is contained in the σ-algebra generated by $\mathcal{G}$, but $\mathcal{A}_{\delta\sigma}$ is not necessarily closed under complementation or countable intersection. An example is given by $\Omega = 2^{\mathbb{N}}$, with $\mathcal{A}$ equal to the collection of all sets of the form $P \times \prod_{j=m+1}^{\infty}\{0,1\}$, where m is any positive integer and P is any subset of $\prod_{j=1}^{m}\{0,1\}$. We omit the lengthy computation that shows that the resulting collection $\mathcal{A}_{\delta\sigma}$ is not closed under complementation or countable intersection.

It is easy to see where the analogy between algebras and σ-algebras breaks down: the product of finitely many finite sets is finite, but a product $\prod_{\gamma\in C} A_\gamma$ (as in 1.38) of countably many countable sets is not necessarily countable — see 2.20.k and 2.20.l.

13.11. Sikorski's extension criterion. Let G be a subset of a Boolean algebra X, which generates a Boolean subalgebra $S \subseteq X$. Let Y be another Boolean algebra, and let $f : G \to Y$ be some mapping. Then f can be extended to a Boolean homomorphism $F : S \to Y$ if and only if f satisfies the following condition:

$$\complement^{n_1} g_1 \wedge \complement^{n_2} g_2 \wedge \cdots \wedge \complement^{n_k} g_k \quad = \quad 0$$

implies

$$\complement^{n_1} f(g_1) \wedge \complement^{n_2} f(g_2) \wedge \cdots \wedge \complement^{n_k} f(g_k) \quad = \quad 0$$

for every nonnegative integer k and every choice of

$$g_1, g_2, \ldots, g_k \in G \qquad \text{and} \qquad n_1, n_2, \ldots, n_k \in \{0,1\}.$$

Moreover, if this condition is satisfied, then the extension $F : S \to Y$ is uniquely determined.

Remark. This theorem is similar in nature to 11.10, though a bit more complicated.

Proof of theorem. The uniqueness of F is clear: If an extending homomorphism $F : S \to Y$ exists, then it must satisfy

$$F(s) = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} \complement^{n(i,j)} f(g_{i,j}) \qquad \text{whenever} \qquad s = \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} \complement^{n(i,j)} g_{i,j}\,. \tag{1}$$

Every $s \in S$ can be expressed as a combination of $g_{i,j}$'s in G as above, by 13.10; hence there is at most one homomorphism $F : S \to Y$ that extends f. It is not immediately clear that equation (1) determines a *function*, however. Some $s \in S$ may be representable in normal form in terms of $g_{i,j}$'s in more than one way, and so we must verify that the resulting value of $F(s)$ does not depend on the particular representation of s. After we establish that, we shall show that the function F defined by (1) is indeed a homomorphism.

To show that (1) actually does define a function, suppose that

$$\bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} \complement^{n(i,j)} g_{i,j} \quad = \quad s \quad = \quad \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} \complement^{m(k,l)} h_{k,l}$$

for some $g_{i,j}$'s and $h_{k,l}$'s in G, and let

$$\varphi_1 \quad = \quad \bigvee_{i=1}^{I} \bigwedge_{j=1}^{J_i} \complement^{n(i,j)} f(g_{i,j}) \qquad \text{and} \qquad \varphi_2 \quad = \quad \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} \complement^{m(k,l)} f(h_{k,l}).$$

We must show that φ_1 and φ_2 are equal to each other.

To show this, first observe — by De Morgan's Law and the Distributive Law — that

$$\complement s \quad = \quad \bigvee_{(j_i)} \bigwedge_{i=1}^{I} \complement\complement^{n(i,j_i)} g_{i,j_i} \tag{2}$$

where the join $\bigvee_{(j_i)}$ is over all sequences $(j_i) \in \prod_{i=1}^{I}\{1, 2, \ldots, J_i\}$ — i.e., all sequences $(j_1, j_2, \ldots, j_I)$ that satisfy $1 \le j_i \le J_i$ for each i.

Now, from the two representations for s, together with De Morgan's Law and the Distributive Law, we see that

$$0 \quad = \quad s \wedge \complement s \quad = \quad \bigvee_{(j_i)} \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} \bigwedge_{i=1}^{I} \left[\complement\complement^{n(i,j_i)} g_{i,j_i} \;\wedge\; \complement^{m(k,l)} h_{k,l}\right].$$

The right side of this equation is an expression of the type in the hypothesis of Sikorski's criterion. Hence, by assumption,

$$\bigvee_{(j_i)} \bigvee_{k=1}^{K} \bigwedge_{l=1}^{L_k} \bigwedge_{i=1}^{I} \left[\complement\complement^{n(i,j_i)} f(g_{i,j_i}) \;\wedge\; \complement^{m(k,l)} f(h_{k,l})\right] \quad = \quad 0.$$

Unwinding our computations, we find that $(\complement\varphi_1) \wedge \varphi_2 = 0$. Similarly, $\varphi_1 \wedge (\complement\varphi_2) = 0$. Therefore $\varphi_1 = \varphi_2$.

Thus F is well-defined. It remains to show that F is a Boolean homomorphism. Suppose that s and $F(s)$ are represented as in (1). We claim that

$$\complement F(s) \quad = \quad \bigvee_{(j_i)} \bigwedge_{i=1}^{I} \complement \complement^{n(i,j_i)} f(g_{i.j_i}) \quad = \quad F(\complement s).$$

Indeed, the first equation follows from our representation of $F(s)$ in (1), by an argument analogous to our proof of (2). The second equation follows from our representation of $\complement s$ in (2), using the definition of F. Thus F preserves $\complement$. That F also preserves $\vee$ is obvious from our definition of F. Now it follows from an exercise in 13.7 that F is a homomorphism.

13.12. Tarski-Scott-Luxemburg Epimorphism Theorem. Every Boolean algebra is the homomorphic image of some algebra of sets. That is, if X is any Boolean algebra, then there exists some $\mathfrak{H}$ that is an algebra of sets and some surjective Boolean homomorphism $f : \mathfrak{H} \to X$.

Proof. We may specify such a surjective homomorphism as follows:

Temporarily forget the Boolean algebra structure of X, and just view X as a set. Let $\mathcal{A} = \mathcal{P}(X)$. For each $x \in X$, let

$$\mathcal{F}_x \quad = \quad \{A \subseteq X \ : \ x \in A\} \quad = \quad \{A \in \mathcal{A} \ : \ x \in A\} \quad \subseteq \quad \mathcal{A}.$$

(This is just the ultrafilter on X, fixed at x, as in 5.5.c.) Thus $\mathfrak{G} = \{\mathcal{F}_x : x \in X\}$ is a collection of subsets of $\mathcal{A}$; let $\mathfrak{H}$ be the algebra of subsets of $\mathcal{A}$ generated by $\mathfrak{G}$.

Now recall the Boolean algebra structure of X. We shall show that the mapping $\mathcal{F}_x \mapsto x$ satisfies Sikorski's criterion (13.11); hence it extends uniquely to a Boolean algebra homomorphism from $\mathfrak{H}$ onto X. Let

$$x_1, x_2, \ldots, x_{m-1}, x_m \qquad \text{and} \qquad y_1, y_2, \ldots, y_{n-1}, y_n$$

be any elements of X, such that

$$(\mathcal{F}_{x_1} \cap \mathcal{F}_{x_2} \cap \cdots \cap \mathcal{F}_{x_{m-1}} \cap \mathcal{F}_{x_m}) \ \cap \ (\complement\mathcal{F}_{y_1} \cap \complement\mathcal{F}_{y_2} \cap \cdots \cap \complement\mathcal{F}_{y_{n-1}} \cap \complement\mathcal{F}_{y_n}) \ = \ \varnothing. \qquad (a)$$

(We permit m or n to be 0, with the understanding that the intersection of no subsets of $\mathcal{A}$ is just $\mathcal{A}$.) It suffices to show that

$$(x_1 \wedge x_2 \wedge \cdots \wedge x_{m-1} \wedge x_m) \ \wedge \ (\complement y_1 \wedge \complement y_2 \cdots \wedge \complement y_{n-1} \wedge \complement y_n) \ = \ 0. \qquad (b)$$

The set on the left side of (a) can be rewritten as

$$\Big\{S \subseteq X \quad : \quad \{x_1, \ldots, x_m\} \ \subseteq \ S \ \subseteq \ X \setminus \{y_1, \ldots, y_n\}\Big\}.$$

The fact that this set is empty implies that

$$\{x_1, x_2, \ldots, x_m\} \quad \text{is not a subset of} \quad X \setminus \{y_1, y_2, \ldots, y_n\},$$

and therefore $x_i = y_j$ for some i and j. But then $x_i \wedge \complement y_j = 0$, implying (*b*).

Remarks and alternate proof. This theorem could be taken as a corollary of the Stone Representation Theorem (UF6) in 13.22. However, the Stone Representation Theorem is an equivalent of the Ultrafilter Principle — i.e., it is a weak form of the Axiom of Choice. In contrast, the present theorem does not require any arbitrary choices. We shall use that fact in our proof of (HB12) $\Rightarrow$ (HB13), in 23.19.

This theorem was announced by Tarski [1954], who credited it to Scott. It was subsequently used by Luxemburg [1969].

Readers with a greater background in algebra may prefer the following proof. Let X_b denote the given Boolean algebra X, and let X_s denote the underlying set of X — i.e., without its Boolean structure. Let A be the free Boolean algebra that has the set X_s for its set of generators. By the definition of a free Boolean algebra (not covered in this book), the identity map $i : X_s \to X_b$ extends uniquely to a homomorphism $I : A \to X_b$, which is therefore surjective. The construction of free algebras is a well-known construction and can be found in many algebra books. By an argument similar to the proof of 13.11, it can be shown (without using the Axiom of Choice or any of its weaker offspring) that the free Boolean algebra with generators X_s is isomorphic to a subalgebra of $\mathcal{P}(\mathcal{P}(X_s))$; thus A is isomorphic to an algebra of sets.

Boolean Rings

13.13. *Definitions and remarks.* A **Boolean ring** is a ring X with unit, in which every element is idempotent — i.e., in which $x^2 = x$ for every $x \in X$. The collection of all Boolean rings is an equational variety, as defined in 8.50, and so our results on algebraic systems are applicable. The fundamental operations are the ring operations: $\cdot$, $+$, $-$, 0, 1. Boolean rings are a full subcategory of the category of rings with unit.

We emphasize that, by our definition, the degenerate ring $\{0\}$ is a Boolean algebra. Some of the older literature imposes the further restriction that $0 \neq 1$ as part of the definition of a Boolean ring; see the remarks in 13.4.a. In this book, $\{0\}$ *is* a Boolean ring, and so Boolean rings do form an equational variety.

Exercise. Let X be a Boolean ring. Show that

$$-x \;=\; x \qquad \text{and} \qquad xy \;=\; yx$$

for all $x, y \in X$. *Hints*: $(x+x)^2 = (x+x)$ and $(x+y)^2 = (x+y)$.

13.14. Any Boolean ring $(X, +, -, \cdot, 0, 1)$ can be made into a Boolean lattice with the same underlying set, $(X, \wedge, \vee, \complement, 0, 1)$, by the definitions

$$p \wedge q = pq, \qquad p \vee q = p + q + pq, \qquad \complement p = 1 + p.$$

The unary operations 0 and 1 are left unchanged. Conversely, we can make any Boolean lattice into a Boolean ring by defining

$$pq \;=\; p \wedge q, \qquad p + q \;=\; p \,\triangle\, q, \qquad -p \;=\; p.$$

These two transformations are inverses to each other; they yield a bijection between Boolean rings and Boolean lattices. (The relevant verifications are left as a tedious but straightforward exercise.)

In the mathematical literature, the phrases "Boolean lattice," "Boolean ring," and "Boolean algebra" are sometimes used interchangeably. The same set X may be viewed as a Boolean lattice or a Boolean ring. However, some caution must be exercised. It should be noted that

> if X contains more than one element, then the ring and lattice structures are *not* compatible with each other in the sense of ordered groups (discussed in 8.30).

Proof. We noted in 10.2 that an ordered group containing more than one element cannot have a lowest or highest element, but any Boolean lattice has both.

13.15. *Proposition (optional).* Let $p(x_1, x_2, \ldots, x_n)$ and $q(x_1, x_2, \ldots, x_n)$ be terms for the variety of Boolean algebras (as in 8.50) — that is, assume p and q are functions of the variables $x_1, x_2, \ldots, x_n$, expressed by formulas using only those variables and the fundamental operations $0, 1, \vee, \wedge, \complement$ or $0, 1, \cdot, +$. Then

> the equation $p(x_1, x_2, \ldots, x_n) = q(x_1, x_2, \ldots, x_n)$ is satisfied by every choice of $x_1, x_2, \ldots, x_n$ in every Boolean algebra X

if and only if

> the equation $p(x_1, x_2, \ldots, x_n) = q(x_1, x_2, \ldots, x_n)$ is satisfied by every choice of $x_1, x_2, \ldots, x_n$ in the Boolean algebra $2 = \{0, 1\}$.

In other words, the identities that are true for all Boolean algebras are the same as the identities that are true for 2.

Remarks. We omit the proof, since this result is not needed later in this book. A proof is given by Johnstone [1987] and other books.

13.16. Some features associated with Boolean rings are equivalent to features associated with Boolean lattices:

a. If X is a Boolean ring, then a **Boolean subring** of X is a nonempty subset $S \subseteq X$ that is closed under the fundamental operations of rings with unit — i.e., that satisfies

$$x_1, x_2 \in S \quad \Rightarrow \quad x_1 x_2,\ x_1 + x_2,\ -x_1,\ 0,\ 1\ \in\ S.$$

It is then a Boolean ring in its own right, when equipped with the restrictions of the operations of X.

Let X be a Boolean algebra, and let $S \subseteq X$. Show that S is a Boolean subring (i.e., closed under the ring-with-unit operations) if and only if S is a Boolean sublattice (i.e., closed under the Boolean lattice operations). Hereafter, we shall use the terms "Boolean sublattice," "Boolean subring," and "Boolean subalgebra" interchangeably.

b. Recall that a ring-with-unit homomorphism is a mapping $f : X \to Y$, from one ring with unit to another, that preserves the fundamental operations of rings with units — i.e., that satisfies

$$f(x_1x_2) = f(x_1)f(x_2), \qquad f(x_1 + x_2) = f(x_1) + f(x_2),$$

$$f(-x_1) = -f(x_1), \qquad f(0) = 0, \qquad f(1) = 1$$

for all $x_1, x_2 \in X$.

Let $f : X \to Y$ be a mapping from one Boolean algebra into another. Show that f is a homomorphism of Boolean lattices (as defined in 13.7) if and only if f is a homomorphism of rings with unit (as in the preceding paragraph). Hereafter, either type of homomorphism may be described more briefly as a **Boolean homomorphism**.

13.17. *More definitions.* Let X be a Boolean ring.

a. By an **ideal** in X we shall mean a set $S \subseteq X$ that is an ideal in the sense of 9.25, in the category of rings with unit or in the category of Boolean rings. (We obtain the same ideals either way; see 9.26.g.) Clearly the set X itself is an ideal; any other ideal is called a **proper ideal**.

An equivalent definition can be given in terms of the lattice structure. *Exercise.* Let X be a Boolean algebra, and let $S \subseteq X$. Show that S is an ideal in X if and only if S is nonempty and

(i) $s, t \in S \;\Rightarrow\; s \vee t \in S$, and

(ii) $s \in S,\ x \in X \;\Rightarrow\; x \wedge s \in S$.

b. Dual to the notion of "ideal" is that of "filter." Let X be a Boolean algebra, and let $T \subseteq X$. We shall say that T is a **filter** if T is nonempty and

(i) $s, t \in T \;\Rightarrow\; s \wedge t \in T$, and

(ii) $t \in T,\ x \in X \;\Rightarrow\; x \vee t \in T$.

Equivalently, a filter is a set of the form $f^{-1}(1)$, where f is any Boolean homomorphism. Clearly the set X itself is a filter; any other filter is called a **proper filter**. (Note that the improper filter is of the form $f^{-1}(1)$ because we may take $f : X \to \{0\}$; recall that $0 = 1$ in the degenerate Boolean algebra $\{0\}$.)

c. Let $T = \{\complement s : s \in S\}$; then T is a filter if and only if S is a ideal. We say S and T are *dual* to each other.

d. If $X = \mathcal{P}(\Omega)$ for some set Ω, with ordering given by $\subseteq$, then an ideal or filter in X as defined above is the same thing as an ideal of sets or a filter of sets (as defined in 5.2 and 5.1).

13.18. Let I be a proper ideal in a Boolean algebra X, and let $F = \{\complement x : x \in I\}$; thus F is a proper filter. Show that the following conditions are equivalent. When any of them is satisfied, we say I is a **prime ideal** and F is a **Boolean ultrafilter**.

(A) I is a maximal ideal — i.e., a proper ideal that is not included in any other proper ideal. (Equivalently, F is a maximal filter.)

(B) For each $x \in X$, exactly one of $x, \complement x$ is an element of I (and hence the other is a member of F). That is, the filter $F = \{x \in X : \complement x \in I\}$ is also equal to $\{x \in X : x \notin I\}$.

(C) Whenever $x \wedge y \in I$, then at least one of x, y is an element of I. Equivalently: if $x \vee y \in F$, then at least one of x, y belongs to F.

(D) The quotient Boolean algebra X/I is isomorphic to $\{0, 1\}$.

(E) The characteristic function of F is a two-valued homomorphism on X (defined as in 13.8).

If $X = \mathcal{P}(\Omega)$ for some set Ω, then another equivalent condition is

(F) F is an ultrafilter of sets on Ω (in the sense of 5.8).

Hint for (A) $\Rightarrow$ (B): Show that if I is an ideal in X containing neither z nor $\complement z$, then $\{x \vee y : x \preccurlyeq z \text{ and } y \in I\}$ is a strictly larger ideal that contains z.

13.19. The **dual** of a Boolean algebra A is defined to be the set

$$\begin{aligned} A^* &= \{\text{two-valued homomorphisms on } A\} \\ &= \{\text{characteristic functions of Boolean ultrafilters in } A\}, \end{aligned}$$

which is a subset of $2^A = \{\text{mappings from } A \text{ into } \{0,1\}\}$. From the examples in 13.8 we see that A^* is empty when A is the degenerate algebra $\{0\}$, but A^* is nonempty when $A = \mathcal{P}(\Omega)$ for some nonempty set Ω. In (UF8) we shall prove that A^* is nonempty whenever A is any nondegenerate Boolean algebra. In 17.44 we shall introduce a natural topology on A^*.

Remarks. Since we may identify sets with their characteristic functions, we could say that a Boolean algebra A has dual $A^* = \{\text{Boolean ultrafilters in } A\} \subseteq \mathcal{P}(A)$. However, we prefer to view A^* as a subset of 2^A because this makes the topology on A^* (introduced in 17.44) more obvious and also makes more obvious the analogy between Boolean duality and the other kinds of dualities described in 9.55.

Caution: Some mathematicians instead define A^* to be the set of all prime ideals in A. Switching to this definition would require no changes of substance; we would simply have to replace each argument with a dual argument. Of course, switching back and forth between the two conventions is a tedious process, so we find it more convenient to stick with just one of the conventions. We prefer to use ultrafilters rather than ideals here, because this makes our Boolean duality more like the other kinds of duality discussed in 9.55. In this book we shall always take A^* to be the set of (characteristic functions of) Boolean ultrafilters.

13.20. *Proposition.* Any finite, nondegenerate Boolean algebra X has a nonempty dual. In fact, if $x_0 \in X \setminus \{0\}$, then there exists at least one $f \in X^*$ with $f(x_0) = 1$.

Remark. This proposition does not require the Axiom of Choice or any of its weakened forms. This proposition will be used, together with a weak form of Choice, to prove (UF8) in 13.22, which removes the restriction to finite X's.

Proof of proposition. Since $S_0 = \{u \in X : 0 \prec u \preccurlyeq x_0\}$ is a nonempty, finite poset, it has a minimal element. Let u_0 be any minimal element of S_0. (Different choices of u_0 may yield different functions f, but in the present argument we are only concerned with proving the existence of *at least one* such f; we do not need a canonical, particular f.) Observe that $y \wedge u_0$ is either u_0 or 0, for each $y \in X$. Use that fact to show that the set $T = \{y \in X : y \succcurlyeq u_0\}$ is a Boolean ultrafilter in X. Hence its characteristic function

$$f(y) \quad = \quad \begin{cases} 1 & \text{if } y \in T \\ 0 & \text{if } y \in X \setminus T \end{cases}$$

is a member of X^* with $f(x_0) = 1$.

13.21. Lemma on Stone's Epimorphism. Let X be any Boolean algebra. Assume that X^* is nonempty. Then there exists a Boolean homomorphism from X onto an algebra $\mathfrak{S}$ of subsets of X^*.

In fact, one such homomorphism may be defined as follows: For each $x \in X$, let $S(x) = \{f \in X^* : f(x) = 1\}$. Let $\mathfrak{S}$ be the range of this mapping. Verify that

$$S(x \vee y) = S(x) \cup S(y), \qquad S(x \wedge y) = S(x) \cap S(y),$$

$$S(\complement x) = X^* \setminus S(x), \qquad S(0) = \varnothing, \qquad S(1) = X^*.$$

These equations show that $\mathfrak{S}$ is an algebra of subsets of X^* and that the mapping $x \mapsto S(x)$ is a homomorphism from X onto $\mathfrak{S}$.

These observations do not require the Axiom of Choice or any of its weakened forms. However, we can draw further conclusions about Stone's epimorphism if we assume some weakened form of the Axiom of Choice; see (UF6) in 13.22.

Boolean Equivalents of UF

13.22. We shall show that the several principles listed below are equivalent to the Ultrafilter Principle. In Chapter 6 we proved (UF1) $\Rightarrow$ (UF2) (and in Chapters 7 and 9 we proved that (UF1) $\Leftrightarrow$ (UF3) $\Leftrightarrow$ (UF4)); now we shall complete the cycle by proving that (UF2) $\Rightarrow$ (UF5) $\Rightarrow$ (UF6) $\Rightarrow$ (UF7) $\Rightarrow$ (UF8) $\Rightarrow$ (UF9) $\Rightarrow$ (UF10) $\Rightarrow$ (UF1).

Remark. Although (UF1) is probably the version of the Ultrafilter Principle most useful for analysts, the variants listed below are as well known in logic and algebra. The mathematician who is searching through the literature for equivalents of UF and related material would do well to look under not only "ultrafilter," but also "prime ideal" and "Boolean."

> **(UF5) Boolean Separation Theorem.** The dual of a Boolean algebra separates its points. That is, if X is a Boolean algebra and x_0, x_1 are distinct members of X, then there exists $f \in X^*$ with $f(x_0) \neq f(x_1)$. Or, by translation, we may restate this as: If X is a Boolean algebra and $x_0 \in X \setminus \{0\}$, then there exists $f \in X^*$ with $f(x_0) = 1$.

(UF6) Stone Representation Theorem (explicit version). If X is a nondegenerate Boolean algebra, then X^* is nonempty and the Stone mapping (described in 13.21) is an isomorphism from X onto an algebra of sets.

(UF7) Stone Representation Theorem (simple version). Every Boolean algebra is isomorphic to some algebra of sets.

(UF8) Boolean Prime Ideal Existence Theorem. If X is a nondegenerate Boolean algebra, then X has a prime ideal. (Equivalently, X has a Boolean ultrafilter; X^* is nonempty; there exists a *two-valued probability* on X, in the terminology of 23.19.b.)

(UF9) Boolean Prime Ideal Extension Theorem. Let X be a Boolean algebra. Then every proper ideal in X is included in a prime ideal. (Equivalently, every proper filter in X is included in a Boolean ultrafilter.)

(UF10) Boolean Ultrafilter Extension Theorem. Let X be a Boolean algebra, and let S be a nonempty subset of X. Then S is included in a Boolean ultrafilter if and only if S has this *finite meet property*:

$$s_1 \wedge s_2 \wedge \cdots \wedge s_n \neq 0 \quad \text{for each finite set} \quad \{s_1, s_2, \ldots, s_n\} \subseteq S.$$

(Equivalently, S is included in a prime ideal if and only if S has the analogous finite join property.)

Proof of (UF2) $\Rightarrow$ (UF5). Let Φ be the collection of all functions f from subsets of X into $\{0,1\}$ that have the following property:

> f can be extended to a Boolean homomorphism from some subalgebra of X into $\{0,1\}$, where that subalgebra includes the point x_0 and where that homomorphism maps x_0 to 1.

Then Φ can be described also as the set of all functions f from subsets of X into $\{0,1\}$, such that

> $\{(x_0, 1)\} \cup \mathrm{Graph}(f)$ is the graph of a function that satisfies Sikorski's extension criterion (13.11).

It is easy to verify that Φ has finite character, in the sense of (UF2)(iii). Also, Φ satisfies (UF2)(i) trivially, since the set $\{0,1\}$ is finite. To verify (UF2)(ii), let S be any finite subset of X. Then the Boolean subalgebra generated by $S \cup \{x_0\}$ is finite, and so we can apply 13.20 to it. Thus, (UF2) is applicable; this completes the proof. This argument is a modification of one by Rice [1968].

Proof of (UF5) $\Rightarrow$ (UF6). X^* is nonempty, so Stone's mapping $S : X \to \mathfrak{S}$ in 13.21 is well-defined. Also, from (UF5) we see that $x \in X \backslash \{0\} \Rightarrow S(x) \neq \varnothing$. Thus the ring-with-unit homomorphism $S : X \to \mathfrak{S}$ has kernel equal to $\{0\}$, so S is injective. Therefore S is an isomorphism from X onto $\mathfrak{S}$.

Proof of (UF6) $\Rightarrow$ (UF7). Obvious.

Proof of (UF7) $\Rightarrow$ (UF8). Immediate from the example in 13.8.

Proof of (UF8) $\Rightarrow$ (UF9). Let I be a given proper ideal in X. Let $\pi : X \to X/I$ be the quotient map, onto the quotient Boolean algebra. By (UF8), X/I has a prime ideal P. Verify that $\pi^{-1}(P)$ is a prime ideal in X that includes I.

Proof of (UF9) $\Rightarrow$ (UF10). The "only if" part is obvious and does not require (UF9); any subset of an ultrafilter (or more generally, any subset of a proper filter) has the finite meet property. For the "if" part, conversely, suppose S has the finite meet property. Then the set

$$\Big\{x \in X \ : \ x \succcurlyeq s_1 \wedge \cdots \wedge s_n \ \text{ for some finite set } \ \{s_1, \ldots, s_n\} \subseteq S\Big\}$$

is a proper filter containing S. (In fact, it is the smallest such filter; it is the filter *generated* by S.) By (UF9), this filter is contained in some ultrafilter.

Proof of (UF10) $\Rightarrow$ (UF1). Immediate from 13.18(F).

Heyting Algebras

13.23. In this subchapter we shall consider two types of algebraic systems that are slightly more general than Boolean algebras. They have most of the properties of Boolean algebras, but not quite all. In particular, they lack some of the symmetry or duality of Boolean algebras; thus they might be thought of as "one-sided Boolean algebras." That relatively pseudocomplemented lattices are more general than Heyting algebras and Heyting algebras are more general than Boolean algebras can be seen from the examples in 13.28.

13.24. *Definition.* Let X be a lattice, and let $a, b \in X$. Then the set $\{x \in X \ : \ a \wedge x \preccurlyeq b\}$ is nonempty — for instance, b is a member. The **pseudocomplement of a relative to b** is the element of X denoted by $a \Rightarrow b$ and defined by this formula:

$$(a \Rightarrow b) \quad = \quad \max\{x \in X \ : \ a \wedge x \preccurlyeq b\}$$

if such a maximum exists. We shall also refer to $\Rightarrow$ as the **Heyting implication**.

We say that X is a **relatively pseudocomplemented lattice** if the Heyting implication is a binary operation — that is, $a \Rightarrow b$ exists in X for all $a, b \in X$, and thus the Heyting implication is a mapping from $X \times X$ into X.

13.25. *Basic properties.* Let X be a relatively pseudocomplemented lattice. Then:

a. $x \preccurlyeq (a \Rightarrow b)$ if and only if $a \wedge x \preccurlyeq b$.

b. $x \preccurlyeq (a \Rightarrow (b \Rightarrow c))$ if and only if $a \wedge b \wedge x \preccurlyeq c$.

c. Interchange of Hypotheses. $(a \Rightarrow (b \Rightarrow c)) = (b \Rightarrow (a \Rightarrow c))$.

d. $(a \wedge (a \Rightarrow b)) \quad \preccurlyeq \quad b \quad \preccurlyeq \quad (a \Rightarrow b)$.

e. If $a \preccurlyeq b$, then $(c \Rightarrow a) \preccurlyeq (c \Rightarrow b)$ and $(a \Rightarrow c) \succcurlyeq (b \Rightarrow c)$. *Hint:*

$$\begin{aligned} \{x \in X : c \wedge x \preccurlyeq a\} &\subseteq \{x \in X : c \wedge x \preccurlyeq b\}, \\ \{x \in X : a \wedge x \preccurlyeq c\} &\supseteq \{x \in X : b \wedge x \preccurlyeq c\}. \end{aligned}$$

f. $(a \Rightarrow b) \preccurlyeq ((b \Rightarrow c) \Rightarrow (a \Rightarrow c))$.

g. X has a largest element, hereafter denoted by 1. It is equal to $(a \Rightarrow a)$, for any $a \in X$.

h. $(1 \Rightarrow a) = a$ and $(a \Rightarrow 1) = 1$ for any $a \in X$.

i. $a \preccurlyeq b$ if and only if $(a \Rightarrow b) = 1$.

j. X is a distributive lattice. In fact, it satisfies one of the infinite distributive laws:

> If $a \in X$ and $S \subseteq X$ and $\sigma = \sup(S)$ exists, then $\sup\{a \wedge s : s \in S\}$ exists and equals $a \wedge \sigma$.

Proof. Let $T = \{a \wedge s : s \in S\}$. In any lattice, if $\sigma = \sup(S)$ exists, then $a \wedge \sigma$ is an upper bound for T (easy exercise). To show that in a relatively pseudocomplemented lattice, $a \wedge \sigma$ is the *least* upper bound, we shall show that $a \wedge \sigma \preccurlyeq \tau$, where τ is any given upper bound for T. By assumption, $a \wedge s \preccurlyeq \tau$ for each $s \in S$. Then each $s \in S$ satisfies $s \preccurlyeq (a \Rightarrow \tau)$, by definition of $\Rightarrow$. Thus $\sigma \preccurlyeq (a \Rightarrow \tau)$. Using the definition of $\Rightarrow$ again, we have $a \wedge \sigma \preccurlyeq \tau$, as required. (This proof can be found in Rasiowa and Sikorski [1963] and other books.)

k. It can be shown that relatively pseudocomplemented lattices form an equational variety. We omit the proof; it can be found in Rasiowa [1974].

13.26. A **Heyting algebra** (also known as a **Brouwerian lattice** or a **pseudo-Boolean algebra**) is a relatively pseudocomplemented lattice with the further property that

X has a smallest element, denoted hereafter by 0.

In a Heyting algebra X, we also define a unary operation $\complement : X \to X$ by

$$\complement a \quad = \quad (a \Rightarrow 0) \quad = \quad \max\{x \in X \ : \ a \wedge x = 0\}.$$

The operation $\complement$ is called the **pseudocomplement**.

13.27. *Basic properties.* Let X be a Heyting algebra. Prove the following properties. (Several of these are just specializations of results of 13.25, obtained by setting one of the variables to 0.)

a. $\complement 0 = 1$, $\complement 1 = 0$, $(0 \Rightarrow b) = 1$.

b. If $a \preccurlyeq b$, then $\complement b \preccurlyeq \complement a$.

c. Contrapositive Law. $\big(a \Rightarrow (\complement b)\big) = \big(b \Rightarrow (\complement a)\big)$.

d. Double Negation Law. $a \preccurlyeq \complement\complement a$. *Hint:* Use the Contrapositive Law with $b = \complement a$, and use 13.25.i.

e. $(a \Rightarrow b) \preccurlyeq \big((\complement b) \Rightarrow (\complement a)\big)$.

f. $a \wedge (\complement a) = 0$.

g. $\big(a \Rightarrow (\complement a)\big) = (\complement a)$.

h. **Brouwer's Triple Negation Law.** $\complement\complement\complement a = \complement a$.

Hints: $(\complement a) \preccurlyeq \complement\complement(\complement a)$ by applying the Double Negation Law to $\complement a$. Also, apply $\complement$ to both sides of the Double Negation Law; by 13.27.b this yields $\complement(\complement\complement a) \preccurlyeq \complement(a)$.

i. $\big((\complement a) \wedge (\complement b)\big) = \complement(a \vee b)$.

j. $\big((\complement a) \vee (\complement b)\big) \preccurlyeq \complement(a \wedge b)$.

k. $\big((\complement a) \vee b\big) \preccurlyeq (a \Rightarrow b)$.

l. $\complement\complement\left(b \vee (\complement b)\right) = 1$. *Hint*: Use 13.27.f with $a = \complement b$; also use 13.27.i.

m. Although we omit the proof, it can be shown that Heyting algebras form an equational variety. See Rasiowa [1974].

n. Every Boolean lattice is a Heyting algebra.

13.28. *Topological examples.* Let X be a set, and let $\mathcal{D}$ be a collection of subsets of X that is closed under finite intersections and arbitrary unions. Assume that the partially ordered set $(\mathcal{D}, \subseteq)$ is a lattice.

Then $(\mathcal{D}, \subseteq)$ is a relatively pseudocomplemented lattice. To see this, let S, T be any two given members of $\mathcal{D}$. Let $\mathcal{K}_{S,T} = \{G \in \mathcal{D} : S \cap G \subseteq T\}$. Then the union of the members of $\mathcal{K}_{S,T}$ is itself a member of $\mathcal{K}_{S,T}$, and thus the largest member of $\mathcal{K}_{S,T}$; hence it satisfies the requirements for a relative pseudocomplement.

We note two particular instances of this when X is a topological space; these examples are from Rasiowa and Sikorski [1963]:

a. The lattice of open sets, discussed in 5.21, is a Heyting algebra, since it also has a smallest member — namely, the empty set. (Although the proof is too long to present here, it can be shown that, conversely, any Heyting algebra is lattice isomorphic to the lattice of open sets of some topological space; a proof of this is given by Rasiowa and Sikorski [1963, page 128].) In 5.21 we verified directly that the lattice of open sets satisfies one of the infinite distributive laws; that fact also follows from 13.25.j. The lattice of open sets may or may not be a Boolean algebra; in 5.21 we gave an example in which the lattice of open sets does not satisfy the other infinite distributive law and thus is not a Boolean algebra.

b. The open dense subsets of a topological space X form a lattice, with binary lattice operations $\cup$, $\cap$ (see 15.13.c). In fact, it is a relatively pseudocomplemented lattice, by the argument given at the beginning of this section. It may or may not have a smallest member, and thus it may or may not be a Heyting algebra. For instance, (*exercise*) it does not have a smallest member if X is the real line with its usual topology.

13.29. *Which Heyting algebras are Boolean?* Suppose X is a Heyting algebra. Then the following conditions are equivalent:

(A) $a \vee (\complement a) = 1$ for all $a \in X$.

(B) $\complement\complement a \preccurlyeq a$ for all $a \in X$.

(C) $(a \Rightarrow b) \preccurlyeq \big((\complement a) \vee b\big)$ for all $a, b \in X$.

(D) $\big((\complement a) \Rightarrow (\complement b)\big) \preccurlyeq (b \Rightarrow a)$ for all $a, b \in X$.

(E) $\big((\complement a) \ \Rightarrow \ b\big) \preccurlyeq \big((\complement b) \ \Rightarrow \ a\big)$ for all $a, b \in X$.

(F) X is a Boolean algebra.

Proof. If X is a Boolean algebra, then it is easy to verify that all the other conditions listed above are satisfied. Conversely:

If (A) holds, then $\complement$ is a complementation operation (as in 13.1), not just pseudocomplementation; since any Heyting algebra is a distributive lattice, (F) follows. For (B) implies (A), note that $\complement\complement\big(a \vee (\complement a)\big) = 1$ and $a \preccurlyeq \complement\complement a$ in any Heyting algebra. For (C) implies (A), let $b = a$. For (D) implies (B), let $b = \complement\complement a$ and simplify. For (E) implies (B), let $b = \complement a$.

Chapter 14

Logic and Intangibles

14.1. *Introduction.* Contrary to the assumption of many nonmathematicians, the study of formal logic does *not* make us more "logical" in the usual sense of that word — i.e., the study of logic does not make us more precise or unemotional. Formal logic is not merely a more accurate or more detailed version of ordinary mathematics. Rather, it is a whole other subject, with its own methods and its own theorems, which are of a rather different nature than the theorems of other branches of mathematics.

Because many of logic's most important applications are in set theory, those two subjects are often presented together, and they may be confused in the minds of some beginners. However, logic and set theory are really different subjects. It is possible to do some interesting things in set theory without any formal logic (see Chapter 6). Conversely, logic can be applied to other theories besides set theory — e.g., real analysis, ring theory, etc. We have already seen examples of this in 8.51 and 13.15.

14.2. *Chapter overview.* This chapter provides a brief introduction to formal logic. Our presentation is mostly conventional, but we follow the unconventional approach of Rasiowa and Sikorski [1963] in our definition of "free variables" and "bound variables;" this is discussed in 14.20.

We shall cover the basics of logic, up to and including a proof of the Completeness and Compactness Principles, which show that the syntactic and semantic views of consistency are equivalent. An easy corollary of the Compactness Principle is the existence of nonstandard models of arithmetic and analysis in 14.63; this is one way to introduce the subject of nonstandard analysis. The Completeness and Compactness Principles are also interesting to us because they are equivalent to the Ultrafilter Principle, a weak form of Choice studied extensively in other chapters of this book.

After the Completeness and Compactness Principles we shall state a few more advanced results, with references in lieu of proofs. Our main goal is to develop some understanding of the notion of "consistency," so that we can understand Shelah's alternative to conventional set theory,

$$\mathrm{Con}(\mathrm{ZF}) \qquad \Rightarrow \qquad \mathrm{Con}(\mathrm{ZF} + \mathrm{DC} + \mathrm{BP}).$$

This result was proved by Shelah [1984], but the proof is too long and too advanced to be included in this book. Our goal is only to understand the statement of Shelah's result and some of its applications. At the end of this chapter, we shall use Shelah's consistency result to explain intangibles — i.e., objects that "exist" in conventional mathematics but that lack "examples." For a first reading, some may choose to skip ahead to the end of this chapter and just read the summary of consistency results and the explanation of intangibles; the rest of this chapter will not be needed elsewhere in the book.

Some Informal Examples of Models

14.3. In logic we separate a language from its meanings. An **interpretation** of a language is a way of assigning meanings to its symbols. Formulas are not true or false in any absolute sense; they are only true or false when we give a particular interpretation to the language. For instance, the axioms of ZF set theory are usually regarded as true, but they become false if we interpret "set" and "member" in the peculiar fashion indicated in 1.48. In the view of some logicians — especially, formalists — mathematical objects such as sets do not really "exist;" all that really "exists" is the language we use to discuss sets and the reasoning we can perform in that language. When we change the language or its interpretation, then the nature of sets or other mathematical objects changes. Bertrand Russell took such a viewpoint when he said

> Mathematics is the only science where one never knows what one is talking about nor whether what is said is true.

If we cannot establish absolute truth, the next best thing is **syntactic consistency** — i.e., knowing that our axioms do not lead by logical deduction to a contradiction. By the Completeness Theorem (proved in 14.57), syntactic consistency is equivalent to **semantic consistency** — i.e., knowing that our collection of axioms has at least one model. A **model** of a collection of formulas is an interpretation that makes those formulas true; it is a sort of "example" for that collection of formulas.

An interpretation of a language may be highly unconventional, unwieldly, and not at all intuitive. It may be constructed just for a brief, one-time use — e.g., to prove the consistency of a given collection of axioms. After a model has been used to establish consistency of some axioms, in some cases we may choose to discard the model and think solely in terms of the axioms, because they are conceptually simpler. (A good example of this is in 14.4.) Application-oriented mathematicians may choose to skip the modeling step altogether, and begin with the axioms, trusting that other mathematicians have already justified those axioms with a model.

All of the terms introduced above — interpretation, consistent, model, etc. — will be given more specialized and precise meanings later in this chapter. But first, to introduce the basic ideas, in the present subchapter we shall present some informal examples of models, where "model" has the broad and slightly imprecise meaning indicated above. Most of these examples are mere sketches, intended only to indicate the flavor of the ideas. The omitted details are considerable, and are not intended as exercises; the reader who wishes to fill in the details should consult the references in the bibliography.

14.4. *Models of the reals.* The axioms for an ordered field were given in 10.7; those plus

Dedekind completeness make up the axioms for the real number system. Many analysis books simply "define" $\mathbb{R}$ to be a Dedekind complete, ordered field. But how do we know that that list of axioms makes sense? We must show (or trust other mathematicians who say they have shown) that

(i) there is such a field, and

(ii) any two such fields are isomorphic.

Proof of (ii) is given in 10.15.e. Proofs of (i) by different constructions in terms of the rationals are given in 10.15.d, 10.45, and 19.33.c. Any one of these constructions is a model of the axioms of $\mathbb{R}$, and it therefore demonstrates the consistency of the axioms of $\mathbb{R}$. However, the constructions — involving Dedekind cuts, equivalence classes of Cauchy sequences, etc. — are rather complicated and generally have little to do with our intended applications of the reals. The axioms for $\mathbb{R}$ are usually much simpler conceptually and more convenient for applications. Thus, after we have demonstrated consistency we may discard the model and think of the real numbers in terms of their axioms: The real number system *is* a complete ordered field.

14.5. *A non-Euclidean geometry modeled in Euclidean geometry.* During the 18th and 19th centuries, mathematicians became concerned about Euclid's **Parallel Postulate**, which says, in one formulation:

> through a given point p not on a given line L, there passes exactly one line that lies in the same plane as L but does not meet L.

The other postulates of geometry are concerned with objects of finite size, such as triangles. In contrast, the Parallel Postulate is concerned with behavior of points that are very far away — perhaps infinitely far away — and so the Parallel Postulate is less self-evident. Some mathematicians attempted to remove any doubts by *proving* this axiom as a consequence of the other axioms of Euclidean geometry. In these attempts, one approach was to replace the Parallel Postulate with some sort of alternative that negates the Parallel Postulate, and then try to derive a contradiction. Some alternatives to the Parallel Postulate did indeed lead to clear contradictions, but other alternatives merely led to very peculiar conclusions.

The peculiar conclusions made up new, non-Euclidean geometries. For instance, a paper of Riemann (1854) developed a geometry, now called **double elliptic geometry** or **Riemannian geometry**, in which any two lines meet in two points, and the sum of the angles of a triangle is greater than 180 degrees. At first these geometries were not seen to have anything to do with the "real" world; they were merely viewed as imaginary mathematical constructs.

However, in 1868 Eugenio Beltrami observed that the axioms of two-dimensional double elliptic geometry are satisfied by the surface of an ordinary sphere of Euclidean geometry, if we interpret "line" to mean "great circle" (i.e., a circle whose diameter is the diameter of the sphere). Therefore, if a contradiction arises in our reasoning about double elliptic geometry, then by the same argument with a different interpretation of the words we can obtain a contradiction in Euclidean geometry. Thus, we have a model that establishes relative consistency: *If* the axioms of Euclidean geometry are noncontradictory *and* the theorems we have proved about the sphere in Euclidean geometry are correct, *then* the axioms of double elliptic geometry are also consistent.

Even if we find these bizarre geometries distasteful and prefer to concern ourselves only with Euclidean geometry, Beltrami's reasoning leads to this important conclusion:

> The Parallel Postulate of Euclidean geometry is not implied by the other axioms of Euclidean geometry.

As Hirsch [1995] has put it so aptly, before the 19th century Euclidean geometry was "not merely an axiomatic study, but our best scientific description of physical space." In retrospect, we can now see that double elliptic geometry is every bit as "realistic" as Euclidean geometry. Ants on a very large sphere might think they were on a plane, if they thought at all. Indeed, many humans thought that way until Colombus sailed. In much the same fashion, our three-dimensional space may be very slightly curved in a fourth direction, but the curvature may be so small that we have not yet detected it. Perhaps a space ship that travels far enough in a seemingly straight line will eventually return to its home planet. We can only be certain of what is near at hand.

For a more detailed discussion of the history of these ideas, see Kline [1980]. A similar approach to the Parallel Postulate, using the interior of a circle in the Euclidean plane, was developed by Cayley; it is discussed by Young [1911/1955].

14.6. *Specifying a universe.* Here is one way to construct **models of set theory**: Let $\mathcal{M}$ be some given class of sets. Hereafter, interpret the term "set" to mean "member of $\mathcal{M}$." Thus, the phrases "for each set" and "for some set" will be interpreted as "for each member of $\mathcal{M}$" and "for some member of $\mathcal{M}$." Then statements in the language of set theory can be interpreted in terms of $\mathcal{M}$.

For instance, the definition of equality of sets (given in 1.47) says that if A and B are two sets, then

$$A = B \text{ holds if and only if for each set } T\text{, we have } T \in A \leftrightarrow T \in B.$$

This condition is satisfied when "set" and "member" have their usual meanings, but not when those terms have certain unconventional meanings, as in 1.48. Are they satisfied in the model $\mathcal{M}$? Yes, for some choices of $\mathcal{M}$; no, for others. In the model $\mathcal{M}$, the definition of equality has this interpretation: Let $A, B \in \mathcal{M}$; then

$$A = B \text{ if and only if for each } T \in \mathcal{M}\text{, we have } T \in A \leftrightarrow T \in B.$$

It is easy to see that "equality" of sets in the model $\mathcal{M}$ coincides with the restriction of ordinary equality to the collection $\mathcal{M}$ if and only if $\mathcal{M}$ has this property:

$$\text{whenever } A, B \in \mathcal{M} \text{ and } \mathcal{M} \cap A = \mathcal{M} \cap B\text{, then } A = B. \tag{!}$$

Easy exercise (from Doets [1983]). If $\mathcal{M}$ is a transitive set (defined as in 5.42), then $\mathcal{M}$ satisfies the condition (!) above. A converse to this exercise is **Mostowski's Collapsing Lemma**, which states that any model that satisfies (!) is isomorphic to a transitive model. We shall not prove this lemma; it can be found in books on axiomatic set theory.

If $\mathcal{M}$ includes some, but not all, members of von Neumann's universe V (described in 5.53), then sets $A, B \in \mathcal{M}$ may have different properties when viewed in $\mathcal{M}$ or in V. For instance, since $\mathcal{M}$ has fewer sets than V, it also has fewer functions and fewer bijective functions. It is quite possible that there exists a bijective function in V between A and B, but there does not exist such a function in $\mathcal{M}$. Thus $\text{card}(A) = \text{card}(B)$ in V, but $\text{card}(A) \neq \text{card}(B)$ in $\mathcal{M}$. When we go from the smaller universe $\mathcal{M}$ to the larger universe V, some distinct cardinalities coalesce; this phenomenon is called **cardinal collapse**.

14.7. *Gödel's universe.* A subclass of V was used for an important model of set theory by Gödel around 1939. He interpreted "set" to mean "member of L," using the universe L of sets that are "constructible relative to the ordinals," as described in 5.54. That universe is (perhaps) smaller than the usual universe V. With this interpretation, he was able to show that the axioms of ZF set theory plus AC (the Axiom of Choice) plus GCH (the Generalized Continuum Hypothesis) are all *true.* He constructed his model L inside the conventional universe V, and his use of V assumed the consistency of the ZF axioms. Thus, he concluded that

if ZF is consistent, then ZF + AC + GCH is consistent.

Though Gödel's proof involved constructible sets, this conclusion does not mention constructible sets, and is not restricted to any particular meaning for "sets."

(In 1963 Cohen showed, by other methods, that $\neg$CH is also consistent with set theory; see 14.8. Thus the Continuum Hypothesis and the Generalized Continuum Hypothesis are *independent* of the axioms of conventional set theory.)

Gödel's construction also shows that if ZF is consistent, then so is ZF + AC + GCH + $(V = L)$. The axiom $V = L$ is called the **Axiom of Constructibility**, which says that *all* sets are constructible relative to the ordinals. Thus, we cannot be sure that Gödel's constructible universe L really is smaller than von Neumann's universe V; we do not obtain a contradiction if we assume that those two universes are the same. On the other hand, it has been proved by other methods that $V \neq L$ is also consistent with set theory — see for instance Bell [1985] — so the Constructibility Axiom is independent of the axioms of conventional set theory.

14.8. *Modeling the reals with random variables.* The Continuum Hypothesis (CH) can be formulated as a statement about subsets of the reals: It says that no set S satisfies $\text{card}(\mathbb{N}) < \text{card}(S) < \text{card}(\mathbb{R})$. By now the literature contains several different variants of Cohen's proof that CH is independent of ZF + AC. One of the simplest to outline is the following:

Let (Ω, Σ, μ) be a probability space, where the set Ω has a very high cardinality. Let $\mathcal{R}$ be the space of all equivalence classes of real-valued random variables. For a suitable choice of (Ω, Σ, μ), it is possible to show that

(i) For a suitably formulated axiomatization of $\mathbb{R}$, every axiom of $\mathbb{R}$ is satisfied with probability 1 by $\mathcal{R}$.

(ii) If statements that are true with probability 1 are used to generate new statements, via the rules of logic, then the new statements are also true with probability 1.

(iii) The Continuum Hypothesis, interpreted as a statement about $\mathcal{R}$, is not true with probability 1.

Here $\mathbb{R}$ is modeled by $\mathcal{R}$, and "truth" is replaced by "truth with probability 1." The axioms of set theory and of $\mathbb{R}$, though interpreted in a peculiar fashion, remain unchanged in superficial appearance, and the rules of logic remain unchanged insofar as they deal with strings of symbols. *Therefore*, regardless of what kind of "truth" and "sets" and "real numbers" we use, the axioms of sets and of $\mathbb{R}$ cannot be used, via the rules of logic, to deduce the Continuum Hypothesis.

The explanation sketched above is only the merest outline; the omitted details are numerous and lengthy. Some of them are given by Manin [1977].

14.9. *A topos model for constructivists.* We mentioned in 6.6 that the Trichotomy Law for real numbers is not constructively provable. We now sketch part of a demonstration of that unprovability, due to Scedrov. Our treatment is modified from Bridges and Richman [1987].

By a "Scedrov-real number" we shall mean a continuous function from $[0,1]$ into $\mathbb{R}$. Let f be such a function, let P be a statement about a real number, and let $x \in [0,1]$; then we say P is **true for f at x** if P is a true statement about the real number $f(y)$ for all y in some neighborhood of x in $[0,1]$. The collection of all such points x is the **truth value** of P for f; it is an open set. A statement is **true** if its truth value is the entire interval $[0,1]$. It can be demonstrated (though we shall omit the details here) that the Scedrov-real numbers are a model of the real numbers *with constructivist rules of inference.*

Now let $f(x) = x$ and $g(x) = 0$, for all $x \in [0,1]$. Then the truth value of the statement $f \leq g$ is the empty set (since the interior of a singleton is empty), while the truth value of $f > g$ is the interval $(0,1]$. Hence the truth value of the statement "$f \leq g$ or $f > g$" is the interval $(0,1]$, and thus that statement is not true in this model.

14.10. *A finite model.* The following example (from Nagel and Newman [1958]) is a bit contrived, but it illustrates a point well. We consider a mathematical system consisting of two classes of objects, K and L, which must satisfy these axioms:

1. Any two members of K are contained in just one member of L.
2. No member of K is contained in more than two members of L.
3. The members of K are not all contained in a single member of L,
4. Any two members of L intersect in just one member of K.
5. No member of L contains more than two members of K.

The consistency of this axiom system can be established by the following model:

(*) Let T be a triangle. Let L be the set of edges of T, and let K be the set of vertices of T.

We can verify that this model satisfies the preceding five axioms; thus they cannot be contradictory. We emphasize that those five axioms might also have *other* interpretations;

we are not restricted to (*) as the only possible interpretation. However, we do have at least one model, given in (*). This is sufficient to prove that the five axioms by themselves cannot lead to a contradiction.

We have used Euclidean geometry to make (*) easy to visualize, but perhaps we do not feel certain about the reliability of Euclidean geometry. The use of geometry is not essential for our present axiom system. We can reformulate (*) without mentioning triangles:

(**) Let a, b, c be distinct objects. Let $K = \{a, b, c\}$ and $L = \big\{\{a, b\},\ \{b, c\},\ \{c, a\}\big\}$.

The model (**) has only finitely many parts; thus, it leaves very little room for doubt. The importance of such models is discussed in 14.70.

Languages and Truths

14.11. A **language** is a collection $\mathcal{L}$ of symbols, together with rules of grammar that govern the ways in which those symbols may be put together into strings of symbols called "formulas." For instance, one of the most important languages we shall study is the *language of set theory.* This language includes symbols such as $\in, \subseteq, \cap$, etc. Its grammatical rules tell us that $A \cap B \in C$ is a formula, but

$$A \in\in B, \qquad A \cap \cup = B, \qquad A\cap =$$

are not formulas.

In formal logic we separate a language from its meanings. For instance, in a formal language, "$1 + 2$" and "3" are different, unrelated strings of meaningless symbols. When we *interpret* that language in its usual fashion, then the strings "$1 + 2$" and "3" represent the same object. Although ultimately we shall be concerned with attaching meanings to the symbols in the language $\mathcal{L}$, at the outset it is best to disregard such meanings — even the meanings of familiar symbols such as $\in, \subseteq, \cap, +, =$. Conceptually, a good place to start is the monoid of meaningless strings of symbols, described in 8.4.g.

14.12. *Formal versus informal systems.* Ordinary mathematicians (i.e., those other than logicians) may study sentences, theorems, and proofs about (for instance) rings or differential equations. However, logicians study sentences, theorems, and proofs about sentences, theorems, and proofs. The ordinary mathematician uses a language that describes rings or differential equations; the logician uses a language that describes languages. The logician is related to the mathematician much as a linguist is related to a novelist.

As Rosser [1939] pointed out, in works of logic we may commonly identify at least two distinct systems of reasoning:

a. The *inner* system of reasoning is the *subject* of the work. It is a sort of microcosm of reasoning. It may be less powerful than the reasoning that we use in "ordinary" mathematics, but it is delimited more precisely. Just as a theorem about rings must have precise hypotheses ("Let G be a commutative ring ..."), so too a theorem in logic must have precise hypotheses ("Let $\mathcal{L}$ be a language with infinitely many free variable symbols ..."). The inner language is also called the **object language**. Formulas such as $(\forall_\xi\ P(\xi, x)) \sqcup Q\,(x, f(y,z))$ will occur in the object languages studied in this chapter.

b. The *outer* system of reasoning is ordinary reasoning. It is conducted in the language of ordinary discourse, also sometimes known as the **metalanguage** — a natural language such as English or Japanese, modified slightly to suit the specialized needs of mathematicians. In logic, as in algebra or analysis, the outer language usually does not have to be formal — we can communicate effectively without first discussing in detail how we will communicate.

When the inner system is mathematics, the outer system is often called "**metamathematics**," which translates roughly to "beyond mathematics" or "above mathematics" or "about mathematics." For instance, the Soundness Principle 14.55(iv) and the Gödel-Mal'cev Completeness Principle in 14.57 are results *about* formal systems; thus they are "metatheorems" which reside in the outer system.

The inner and outer systems do not necessarily have the same truths; one of these systems may be stronger than the other. For instance, we must assume ZF plus the Ultrafilter Principle (UF) in our outer system when we want to prove the Gödel-Mal'cev Completeness Principle. That principle can be applied to inner systems that are weaker (such as ZF) or stronger (ZF + AC) or perhaps not even directly comparable.

Here is another example: Let "Con" denote consistency. Then "Con(ZF)" and "Con(ZF + AC + GCH)" are two statements about the consistency of certain axiom systems in formal set theory. Thus they are metamathematical statements, where the mathematics in this case is set theory. Then "Con(ZF) $\Rightarrow$ Con(ZF + AC + GCH)" is a metatheorem — or, if we prefer, a *metametatheorem*, as discussed below.

c. Beginners may find it helpful to view the Ultrafilter Principle and the Completeness Principle as "true;" then they will only need to deal with the two levels of reasoning described above. However, more advanced readers can consider a third level: Throughout many chapters of this book we study equivalents of AC and of UF, viewing them as principles which "might be true" or "might be false," depending on what kind of universe we decide to live in. The implication (UF8) $\Rightarrow$ (UF11), proved in 14.57, is a **metametatheorem** — it is a theorem *about* metatheorems such as the Completeness Principle. It is even more "outer" than the "outer system" — but to avoid confusion, hereafter we shall not discuss such results in this fashion.

There is some resemblance between the logician's inner and outer systems — both systems include "sentences," "implications," "theorems," and "proofs." This may cause some confusion for beginners. No such confusion arises in other subjects, such as ring theory or differential equations — e.g., a theorem about rings generally does not look like a ring.

14.13. The beginner is cautioned to carefully maintain in his or her mind the distinction between inner and outer systems. Throughout this chapter we shall use notations that support that distinction. First, however, we shall give an example of the kind of difficulties that arise when the distinction is not maintained carefully:

Berry's Paradox. Call a positive integer *succinct* if it can be described in sentences of the English language using less than 1000 characters (where a character means a letter, a space, or a punctuation symbol). There are only finitely many different characters, and so it is clear that there are only finitely many succinct numbers. Let n_0 be the first positive integer that is not succinct. We have described n_0 in this paragraph, which is shorter than 1000 characters. So n_0 is succinct after all, a contradiction.

Explanation of the flaw in the reasoning. The first sentence suggests that we are to use English for the formal language of our inner system. However, English is a very fluid language, which changes even while it is being used. Everyday, nonmathematical English is permitted to talk about itself, and this kind of self-referencing can lead to paradoxes, but they are not taken seriously because, after all, English is not mathematics.

In attempting to make mathematically precise sense out of Berry's Paradox, we must use some frozen, unchanging "version" of English for the formal language of our inner system. We assume that some particular version of English has been selected and is understood by all parties participating in this endeavor. The term "English" hereafter is understood to refer to this frozen, formal language. This object language cannot discuss itself, and thus cannot discuss what is a "sentence of the English language." The notion of a "sentence" (as it is used here) is a metamathematical concept — i.e., a concept *about* the language, rather than a concept expressed *in* the language.

Now we can give a more precise definition of a "succinct" positive integer: It is a positive integer that can be described *in the object language* in fewer than 1000 characters. Likewise, n_0 is the first positive integer that cannot be described *in the object language* in fewer than 1000 characters. These definitions are mathematically precise, quite brief, and not at all fallacious, but they are formulated in the metalanguage, not in the object language. Our definition of n_0 is formulated in the metalanguage; we have not given a definition of n_0 in the object language. We certainly have not given a definition of n_0 that is 1000 characters or fewer in the object language. We cannot conclude that n_0 is succinct, so no contradiction is reached.

14.14. When $\mathcal{F}$ and $\mathcal{G}$ are formulas in our *object language*, then $\mathcal{F} \to \mathcal{G}$ is also a formula in that language. It is most often read as "$\mathcal{F}$ implies $\mathcal{G}$."

We now consider two types of implications in our *metalanguage*, which will be investigated in greater detail later in this chapter. Let Σ be a set of formulas, and let $\mathcal{F}$ be a formula.

a. A **derivation** of $\mathcal{F}$ from Σ is a finite sequence of formulas $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ such that $\mathcal{E}_n = \mathcal{F}$ and each $\mathcal{E}_j$ is either (i) an axiom, (ii) a member of Σ, or (iii) obtained from previous members of the sequence by rules of inference. When such a sequence exists, then we say $\mathcal{F}$ is a **syntactic consequence** of Σ; this is abbreviated as $\Sigma \vdash \mathcal{F}$.

When Σ is the empty set, the derivation is called a **proof**, and the consequence $\mathcal{F}$ is called a **syntactic theorem**; we may write $\varnothing \vdash \mathcal{F}$ or, more briefly, $\vdash \mathcal{F}$. Observe that this notation does not reflect the choice of the axioms, which are nevertheless available for use in the derivation. The statement " $\Sigma \vdash \mathcal{F}$ " is equivalent to this statement, which may be preferred by some readers: "any set of axioms that makes all the members of Σ into syntactic theorems, also makes $\mathcal{F}$ into a syntactic theorem."

A set of formulas Σ is **syntactically inconsistent** if some formula and its negation are both syntactic consequences of Σ; otherwise the set of formulas is **syntactically consistent**. The study of syntactic consequences is sometimes called **proof theory**.

b. We write $\Sigma \vDash \mathcal{F}$ to say that $\mathcal{F}$ is a **semantic consequence** of Σ. This means that every model of Σ is also a model of $\mathcal{F}$ — i.e., that every interpretation of the language that makes Σ true (and makes all of our unmentioned axioms true), also makes $\mathcal{F}$ true. The axioms, if any, are understood from the context and are not mentioned in this notation.

When Σ is the empty set, the consequence $\mathcal{F}$ is called a **semantic theorem**. We may write $\varnothing \vDash \mathcal{F}$ or, more briefly, $\vDash \mathcal{F}$. This means that any interpretation that makes the axioms true also makes $\mathcal{F}$ true.

A set of formulas Σ is **semantically inconsistent** if it has no models or **semantically consistent** if it has at least one model. The study of semantic consequences is sometimes called **model theory**.

On the surface, proof theory and model theory seem rather different. A fundamental and nontrivial result of first-order logic is that proof theory and model theory are *equivalent*, in this sense: they actually yield the same notions of "theorem" and "consistency," provided that our system of reasoning is sound. This equivalence will be established in 14.57 and 14.59; thereafter, we can simply refer to a "theorem" and to "consistency." However, to establish the equivalence, we must first develop the syntactic and semantic views separately.

A few cautionary remarks.

(i) Terminology varies in the literature. Some mathematicians prefer either the viewpoint of proof theory or the viewpoint of model theory, and so they define the terms "**theorem**," "**valid formula**," "**true formula**," or "**tautology**" to be synonomous with what we have called a "syntactic theorem" or with what we have called a "semantic theorem." Which term is applied to which type of theorem varies from one paper or book to another. That may confuse beginners, but it does not affect the ultimate results since the two kinds of theorems will eventually be shown equivalent.

(ii) The symbol $\vDash$ has another meaning, which will not be used in this chapter. We mention it to prevent confusion when the reader runs across it in some other book. If $\mathcal{M}$ is some *particular* model in which a formula $\mathcal{F}$ is true, some mathematicians may write $\mathcal{M} \vDash \mathcal{F}$. This is read as: $\mathcal{M}$ is a model of $\mathcal{F}$, or $\mathcal{M}$ **satisfies** $\mathcal{F}$, or $\mathcal{F}$ **holds** in $\mathcal{M}$.

(iii) The symbols $\vdash$ (syntactic implication) and $\vDash$ (semantic implication) should not be confused with these similar symbols:

$$\top \quad (\text{truth}), \qquad\qquad \bot \quad (\text{falsity}), \qquad\qquad \Vdash \quad (\text{forcing}),$$

which are used in some books (but not this one). Forcing is discussed briefly in 14.53.

(iv) Although $\mathcal{F} \vdash \mathcal{G}$ and $\mathcal{F} \models \mathcal{G}$ will ultimately be proved equivalent to each other, they are not equivalent to $\mathcal{F} \to \mathcal{G}$. In fact, no direct comparison is possible, since $\mathcal{F} \to \mathcal{G}$ is a statement in the object language. We can modify that statement slightly if we wish to make comparisons: Each of the four expressions

$$\mathcal{F} \vdash \mathcal{G}, \qquad \mathcal{F} \models \mathcal{G}, \qquad \vdash (\mathcal{F} \to \mathcal{G}), \qquad \models (\mathcal{F} \to \mathcal{G})$$

is a statement in the metalanguage. The first two are equivalent to each other, and the last two are equivalent to each other. In propositional logic, all four statements are equivalent to each other. In predicate logic, the last two statements are *slightly stronger* than the first two statements; this will be discussed further in 14.38 and thereafter.

14.15. The kind of logic used most often in the literature is **first-order logic**; it is also known as **predicate logic** or the **predicate calculus**.

To be precise, we may subdivide a theory into these ingredients:

A **first-order language** includes an alphabet of symbols — punctuation symbols, symbols for individuals, symbols for operations, and quantifiers — and grammatical rules for forming those symbols into formulas. All of the symbols are understood as meaningless characters of a meaningless alphabet; they will only take on meaning when we consider an *interpretation* or *quasi-interpretation* (as in 14.47). The specification of the symbols and rules is understood to include a specification of the arities of the operation symbols, as explained in 14.18 below. First-order language is discussed in further detail in the next subchapter.

A **first-order logic** includes the language, plus rules of inference and logical axioms. It may also be viewed as including the resulting theorems — i.e., the syntactic and semantic consequences of those rules and axioms. The rules of inference and logical axioms are discussed in the subchapter which begins with 14.25, and the resulting theorems are discussed in the subchapters after that.

A **first-order theory** includes the logic, plus extra-logical axioms. It may also be viewed as including the resulting theorems. Some examples of extra-logical axioms are given in 14.27.

Ingredients of First-Order Language

We shall now list the ingredients. Some readers may wish to glance ahead to 14.24, where we consider *propositional logic*, a special case that has fewer ingredients.

14.16. *Punctuation symbols.* These are *parentheses* for grouping — i.e., to avoid ambiguity — and *commas* for delimiting items in a list. It is possible to give precise rules for the use of parentheses and commas, but we shall omit the details.

14.17. *Symbols for individuals.* (These are omitted in propositional logic; see 14.24.) In predicate logic, there are three types of symbols for individuals:

- **individual constant symbols**, denoted in the discussions below by

$$a, b, c, \ldots \qquad \text{or} \qquad a_1, a_2, a_3, \ldots \qquad \text{or} \qquad a, a', a'', \ldots$$

  (Actually, constant symbols may be dispensed with as a separate class of symbols, since they may be viewed as function symbols of arity 0; see 14.18 below.)

- **individual free variable symbols**, denoted in the discussions below by

$$x, y, z, \ldots \qquad \text{or} \qquad v_1, v_2, v_3, \ldots \qquad \text{or} \qquad v, v', v'', \ldots$$

- **individual bound variable symbols**, denoted in the discussions below by

$$\xi, \eta, \zeta, \ldots \qquad \text{or} \qquad \xi_1, \xi_2, \xi_3, \ldots \qquad \text{or} \qquad \xi, \xi', \xi'', \ldots$$

In most texts on logic, the free and bound variables are taken from the same set of symbols; in some texts either the constant symbols or the variable symbols are omitted altogether. However, we prefer to use three separate sets of symbols; this is discussed further in 14.20.

The sets of constants and variables are countably infinite in most applications, but these sets could be larger or smaller. Practical, everyday mathematics uses only countably many symbols — for instance, although there are uncountably many real numbers, we have no way of actually writing down distinct representations for most of those numbers. It is not even humanly possible to write down a countably infinite collection of symbols; that would require more time than any mere mortal has. Nevertheless, for some theoretical purposes it can be useful to conceptualize and investigate a language with uncountably many symbols — e.g., we could say "let $\mathcal{L}$ be a language that includes a constant symbol c_r for each real number r." We can talk about the c_r's in the abstract, even if we can't write them all down concretely.

Hereafter, we shall assume that

> *the set of free variables and the set of bound variables are both empty* (*as in propositional logic*) *or both infinite* (*as in ordinary mathematics*).

The case of finitely many variables turns out to be technically different and difficult, and will not be considered in this chapter. That case is manageable for elementary results but becomes difficult starting in 14.41; for simplicity of exposition we shall exclude that case from the outset. One difficulty with that case can be explained roughly as follows: Reasoning in formal logic (or in other parts of mathematics, for that matter) involves substitutions, usually replacing all occurrences of one free variable with copies of some term whose free variables are not already in use. A single computation may involve many substitutions and thus many free variables. It will involve only a finite number of free variables, but in general we do not know in advance how large or small that finite number will be — the number may vary from one computation to another, and in general there is no finite upper bound for the number of variables needed for a computation. If our language has only finitely many free variables — i.e., if some particular finite number is specified in advance — then we may run out of variables before some computations are completed. On the other hand, if we have infinitely many free variables, we can complete any computation and still have plenty of free variables left over.

14.18. *Symbols for operations.* Each operation symbol has an **arity**, or **rank** — i.e., an associated nonnegative integer that specifies how many arguments each of these symbols should be followed by. For instance, if f has arity 4, then we may form expressions such as $f(w,x,y,z)$. The precise rules for forming such expressions are given in 14.22 and 14.23.

It is convenient to write $x+y$ instead of $+(x,y)$. The abstract discussions of expressions $f(x,y)$ will apply to expressions $x+y$ with obvious modifications. Analogous modifications also apply for other commonly used binary operation symbols, such as $\cdot, \times, \wedge, \vee$, etc.

We have three types of operation symbols, listed below. Although interpretations are not part of the formal language, a preview of typical interpretations may make the formal language easier to understand, so we include a few examples of interpretations here:

(i) **Function symbols**, here denoted f, g, h, etc. (These are omitted in propositional logic; see 14.24.)

Examples in arithmetic or analysis. We might use the function symbols $+, -, \cdot, /$, all with arity 2, and the function symbols cos and - with arity 1. A function symbol with arity 0 is a symbol that gets interpreted as a constant — e.g., the symbol "3" or "$\sqrt{5}$."

In ordinary mathematics, the character "$-$" represents both the binary operator of subtraction and the unary operator of additive inverse, but those are actually different operators, and for purposes of logic it would be best to represent them with different characters such as "$-$" and "-." Interestingly, these two operators are represented by different keys on some recent handheld electronic calculators, a source of confusion for mathematicians who grew up using one character for the two operations.

Examples in set theory. We might use the function symbols $\cap, \cup$ with arity 2, and $\complement$ with arity 1; we might use $\varnothing$ for a symbol with arity 0.

Examples in group theory. A character such as $\circ$ or $\square$ might be used as a function symbol of arity 2.

(ii) **Relation (or predicate) symbols**, here denoted P, Q, R, etc. (In propositional logic, these occur only with arity 0, and are then called *primitive proposition symbols*; see 14.24 and 14.18(ii).)

Examples. Some common relations of arity 2 are $<, >, \leq, \geq, =, \neq, \in, \notin$. Many other meanings are possible for relation symbols; for instance, in arithmetic, $R(x,y)$ might have the interpretation "x is a divisor of y." An example of a relation with arity 1 is "x is a prime number." A relation with arity 0 is just a statement that does not mention any variables.

Remark. It is actually possible to dispense with function symbols, by viewing each function of arity n as a relation with arity $n+1$. For instance, the equation $z = x+y$ determines a function $z = f(x,y)$ with arity 2, but it also determines a relation $R(x,y,z)$ with arity 3.

(iii) **Logical connective symbols**. The precise choice of logical connective symbols may vary slightly from one exposition to another. The ones we shall use are:

$\neg$	(arity 1)	not, negation
$\sqcup$	(arity 2)	or, disjunction
$\sqcap$	(arity 2)	and, conjunction
$\rightarrow$	(arity 2)	implies, implication

Some mathematicians use additional connectives — e.g., the connective $\leftrightarrow$ (iff) or the connective $|$ (the Sheffer stroke). Also, some mathematicians prefer to define some of the connectives in terms of others — e.g., the connective $\sqcup$ may be defined by the equation $\mathcal{A} \sqcup \mathcal{B} = (\neg\mathcal{A}) \rightarrow \mathcal{B}$. However, we prefer to begin with unrelated symbols and then find relationships as a consequence of axioms.

The notations vary slightly. For instance, among some mathematicians,

"not"	may be written instead as	$\sim$
"or"	may be written instead as	$\vee$ or $\cup$
"and"	may be written instead as	$\wedge$ or $\cap$ or &
"implies"	may be written instead as	$\Rightarrow$ or $\supset$

We have chosen our notation in this book so that different symbols are used in logics ($\sqcup, \sqcap$), in lattices ($\vee, \wedge$), and in algebras of sets ($\cup, \cap$). This may reduce some confusion when two of these different kinds of structures must interact — see especially 14.27.d, 14.32, and 14.38.

It should be understood that meanings are not yet attached to the symbols — not even to familiar symbols such as $\neg, +, \mathsf{C}, \circ, =, \in, \rightarrow$. We may call $\neg$ the "negation" or call $+$ the "plus sign" to make them easier to read aloud and to lend some intuition about what this is all leading up to, but we do not yet associate these symbols with their usual meanings or any other meanings. Meanings will be attached later, when we consider interpretations in 14.47. In the formal language, these symbols are merely viewed as meaningless symbols, with arities assigned to grammatically govern the joining of these meaningless symbols into meaningless strings. In fact, the symbols $\neg$ and $\sqcup$ have slightly different meanings in intuitionist and classical logic, both of which are introduced in the following pages.

In the formal theory, our meaningless symbols may also be accompanied by some axioms. The logical connective symbols are governed by the logical axioms (see 14.25); function symbols may be governed by extra-logical axioms as in 14.27.c; relation symbols may be governed by extra-logical axioms as in 14.27.a. No other meaning is attached to any of these symbols in the formal theory.

14.19. *Quantifiers.* There are two kinds of quantifiers:

$\forall_\xi$, the **universal quantifier**, usually read "for each ξ."

$\exists_\xi$, the **existential quantifier**, usually read "there exists ξ such that."

Here ξ is a bound variable; any other bound variable may be used in the same fashion. (In propositional logic there are no variables and thus no quantifiers; see 14.24) We caution that the symbols $\forall$ and $\exists$ occasionally have meanings slightly different from "for each" and "there exists;" see 14.47.j. Until we study their interpretations in 14.47.j, the symbols $\forall$ and $\exists$ should be viewed as not having any meaning at all; they are simply meaningless symbols whose use is governed by grammatical rules and inference rules listed in 14.23(iii) and 14.26.

The quantifier $\forall$ is commonly read as "for all" in the mathematical literature, but we prefer to read it as "for each." In common English, "for all" suggests that the objects are perhaps being treated *all in the same fashion.* The customary mathematical meaning of $\forall$ is closer to "for each," which emphasizes that the objects under consideration can all be treated separately, one by one, perhaps with each treated differently.

We emphasize that in a first-order language, a quantifier is understood to act only on an *individual variable.* Thus, it is possible to say "for each individual ξ," but grammatically it is not permitted to say "for each formula $\mathcal{F}$" or "for each class S of individuals." Those expressions are permitted in *higher-order languages* — i.e., languages of second or third order, etc.; we shall not investigate such languages in this book.

14.20. *Discussion of bindings.* Some popular expositions of logic are Mendelson [1964] and Hamilton [1978]; those textbooks have been used widely and their treatment can now be considered "conventional" or "customary." Our own treatment will follow Rasiowa and Sikorski [1963], which is unconventional in some minor respects. For instance, the Rasiowa-Sikorski treatment uses fewer definitions of symbols and more axioms governing the use of undefined symbols.

A more important difference is in the use of bound and free variables:

- In conventional treatments such as Mendelson [1964] or Hamilton [1978], the rule for incorporating quantifiers into formulas is trivial: If $\mathcal{A}$ is any formula and x is any variable, then $\forall x\mathcal{A}$ and $\exists x\mathcal{A}$ are formulas, regardless of how x is already being used in the formula $\mathcal{A}$ or elsewhere. The same symbols are used for free variables and bound variables; one defines whether a variable symbol x is bound or free according to how and where it appears in a formula. The definitions of bound and free variables and the rules for substitution are (in this author's opinion) rather complicated and nonintuitive. The definitions involve the "scope" of a quantifier and the rather convoluted notion of "a term t that is free for the variable x in the formula $\mathcal{F}$."

- In Rasiowa and Sikorski [1963], the rule for defining free and bound variables is trivial: Before we even begin to think about how to make formulas, we agree in advance which symbols $(x, y, z, \ldots)$ will be free variables and which symbols $(\xi, \eta, \zeta, \ldots)$ will be bound variables; two disjoint sets of symbols are used. The rules for incorporating those variables into formulas (described in 14.22 and 14.23) and for making substitutions (described in 14.26) are not trivial, but they are not particularly complicated.

To motivate either approach, we shall now discuss bindings in general.

In ordinary mathematics (i.e., outside of formal logic), bindings are generated by certain operators such as $\int$, $\sum$, $\prod$. For instance, the equation

$$f(x) \quad = \quad \int_0^x \xi^2 \, d\xi$$

makes sense whenever x is a real number. In this equation, x is a *free* variable, and ξ is a *bound* or *dummy* variable (also sometimes known as an *apparent* variable). The function f is a function of x; it does not really involve ξ. In some sense, ξ is not really a "variable" at all — it is just a "placeholder," and the place can be held just as well by nearly any other letter. All of the expressions

$$\int_0^x \omega^2 \, d\omega, \qquad \int_0^x \xi^2 \, d\xi, \qquad \int_0^x \eta^2 \, d\eta, \qquad \int_0^x \zeta^2 \, d\zeta$$

represent exactly the same function $f(x)$. In fact, that function can also be represented without any dummy variables: $f(x) = x^3/3$. However, dummy variables are unavoidable for certain other functions; for instance, it is well known (but not easy to prove) that the function $g(x) = \int_0^x \exp\left(\xi^2\right) d\xi$ *cannot* be represented in terms of the classical elementary functions (algebraic expressions, trigonometric functions, exponentiation, logarithms, and compositions of such).

In the paragraph above, we have followed the typographical convention of Rasiowa and Sikorski, using different sorts of letters for free variables (x, y, z, etc.) and bound variables (ξ, η, ζ, etc.). But in the wider literature, that convention generally is not observed, and any letter can be used for either type of variable. For instance, the function f described in the preceding paragraph could be defined as easily by the equation $f(\xi) = \int_0^\xi x^2 \, dx$. In this respect, the Mendelson/Hamilton approach follows the convention of "ordinary" mathematics (i.e., mathematics outside of logic). However, in other respects the Mendelson/Hamilton approach differs from the conventions of ordinary mathematics, as we shall now describe.

In the equation $f(x) = \int_0^x \xi^2 \, d\xi$, we can replace ξ by *nearly* any other letter. There is one exception: We should not replace ξ with x itself. Polite mathematicians prefer not to write $f(x) = \int_0^x x^2 \, dx$ since that equation uses the same letter x for two different purposes — as a free variable and a bound variable. Admittedly, that type of expression can be found in some physics or engineering books — it is interpreted to mean the same thing as $f(x) = \int_0^x \xi^2 \, d\xi$ — but mathematicians frown upon such constructions. Likewise, an expression such as

$$g(x, y) \quad = \quad (x + y)^2 + \int_0^y \exp(x^2) \, dx$$

will make any well-bred mathematician uncomfortable, but we know that what is probably meant is

$$g(x, y) \quad = \quad (x + y)^2 + \int_0^y \exp(\xi^2) \, d\xi.$$

Analogous beasts appear in conventional logic books, but with little or no stigma attached. In the formula

$$Q(x, y) \quad \sqcup \quad \Big(\forall x \, \big(R(x, z)\big)\Big), \tag{$*$}$$

the variable x has one free occurrence and two bound occurrences. (The "x" immediately after the $\forall$ is one of the bound occurrences.) Such a distasteful formula is not absolutely necessary for proofs, since $\forall x\ (R(x,z))$ is in most respects equivalent to $\forall w\ (R(w,z))$, as explained in 14.42. Thus we can replace $(*)$ with the formula $Q(x,y) \sqcup (\forall w\ (R(w,z)))$, which does not mix free and bound occurrences of one symbol.

An integral formula such as $g(u) = \int_0^1 \int_0^1 x^3 u\, dx\, dx$ has no clear meaning in ordinary mathematics. Nevertheless, analogous formulas appear often in logic; one such formula is

$$\exists x\ \Big(\forall x\ (P(x,u))\Big), \tag{$**$}$$

which has one variable bound twice. Such a formula may seem unnatural, since it has no analogue outside of formal logic. Again, such beasts are not really necessary: Since $\forall x\ P(x,u)$ is in most respects equivalent to $\forall w\ P(w,u)$ (see 14.42), it may be helpful to view $(**)$ as having the same meaning as $\exists x\ (\forall w\ (P(w,u)))$, an expression with no double bindings. (An analogous interpretation would make $\int_0^1 \int_0^1 x^3 u\, dx\, dx$ equal to $\int_0^1 \int_0^1 w^3 u\, dw\, dx$.)

Thus, in any explanation of logic, it is necessary to either

(i) prohibit nonintuitive expressions such as $(*)$ and $(**)$, or

(ii) provide rules for dealing with such expressions.

Conventional books such as Mendelson [1964] and Hamilton [1978] have followed option (ii), but the rules are necessarily rather complicated. Rasiowa and Sikorski [1963] have taken option (i), and so shall we in this book. Since the nonintuitive expressions can always be replaced by more acceptable ones anyway, the difference between options (i) and (ii) has only a superficial or cosmetic effect; it has no effect on deeper results discussed later in this chapter, such as the Completeness Principle.

A word of caution: Even the Rasiowa-Sikorski approach is not entirely trivial. Among other things, it permits expressions such as $(\exists_\xi\ P(\xi)) \sqcup (\forall_\xi\ Q(\xi))$. The ξ's in the first half of this expression are unrelated to the ξ's in the second half of this expression. Some confusion might be avoided if we replace this formula with the equivalent formula $(\exists_\xi\ P(\xi)) \sqcup (\forall_\eta\ Q(\eta))$.

Actually, variables could be dispensed with altogether; books on **combinatory logic** such as Hindley, Lercher and Seldin [1972] show that everything can be expressed in terms of functions. That approach will not be followed in this book, however.

14.21. *Substitution notation.* Throughout the discussions in the next few pages, we shall frequently use this notation:

> Let x be a free variable symbol. Let $\mathcal{A}(x)$ be a finite string of symbols in which x may occur 0 or more times, and in which other free or bound variables may occur 0 or more times. Let σ be any finite string of symbols. Then $\mathcal{A}(\sigma)$ will denote the string of symbols obtained from $\mathcal{A}(x)$ by replacing each occurrence of x (if there are any) with a copy of the string σ.

Of course, if x does not appear in the string $\mathcal{A}(x)$, then $\mathcal{A}(\sigma)$ is identical to $\mathcal{A}(x)$.

14.22. *Grammatical rules, part 1.* In a first-order language, **terms** are certain finite strings of symbols formed recursively by these two rules:

(i) Any constant symbol or free variable symbol is a term.

(ii) If f is an n-ary function symbol and $t_1, t_2, \ldots, t_n$ are terms, then the expression $f(t_1, t_2, \ldots, t_n)$ is a term.

There are no other terms besides those formed via these rules.

Condition (ii) is not self-referential — i.e., it does not involve any circular reasoning that leads to a contradiction. Indeed, we may classify strings of symbols according to their length (i.e., how many symbols appear in a string) or their depth (i.e., how many times we have functions nested within functions). Then the construction in condition (ii) always forms longer terms from shorter ones or forms deeper terms from shallower ones. To prove a statement about all the terms of a language, it is often possible to proceed by induction on the length or depth of the terms.

Observation: By our definition, no bound variables appear in terms.

Example. Consider a language in which 2, 5, 6 are among the constant symbols, x and y are among the variable symbols, $\surd$ and cos are function symbols of arity 1, and $+, -, \cdot$ are function symbols of arity 2. Then the string of symbols $(5 \cdot x) + (\surd((6 \cdot y) - (\cos(2))))$ is a term. When it is given its usual interpretation involving real numbers, then that string of symbols is a real-valued function of two real variables, written more commonly as $5x + \sqrt{6y - \cos 2}$.

Remark. In the discussions below, terms will generally be represented by the letters $t, t_1, t_2, \ldots$ and $s, s_1, s_2, \ldots$, etc. However, it should be understood that these letters are not actually symbols making up a part of our formal language (the inner system). Rather, these letters are **metavariables** — i.e., they are part of the metalanguage; they are informal conventions adopted for our discussion in the outer system. The precise expression "let t be a term" is an abbreviation for the imprecise and unwieldly expression "let us consider any term, such as $f\big(x_1, x_2, g(c_1, c_2), h(x_3, c_3, c_4)\big)$."

14.23. *Grammatical rules, part 2.* Certain finite strings of symbols are known as **formulas** (or, in some books, **well-formed formulas**, or **wff**'s). The definitions are recursive:

(i) An **atomic formula** (or *atom*) is an expression of the form $P(t_1, t_2, \ldots, t_n)$, where P is an n-ary relation symbol and $t_1, t_2, \ldots, t_n$ are terms. It is a formula.

We permit $n = 0$. Thus a primitive proposition symbol (i.e., relation symbol of arity 0) is an atomic formula. (In propositional logic, this is the *only* type of atomic formula, since the only relation symbols we have in propositional logic are those of arity 0; see 14.24.)

(ii) If $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ are formulas and $\mathfrak{F}$ is an n-ary logical connective symbol, then $\mathfrak{F}(\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n)$ is a formula. Since we will only use a few connectives, this rule can be restated as: If $\mathcal{A}_1$ and $\mathcal{A}_2$ are formulas, then

$$(\neg \mathcal{A}_1), \qquad (\mathcal{A}_1 \sqcup \mathcal{A}_2), \qquad (\mathcal{A}_1 \sqcap \mathcal{A}_2), \qquad (\mathcal{A}_1 \rightarrow \mathcal{A}_2)$$

are formulas. We may omit the parentheses when no confusion is likely.

(iii) Suppose $\mathcal{A}(x)$ is a formula in which the bound variable ξ does not occur. Apply the substitution notation of 14.21. Then $\forall_\xi\, \mathcal{A}(\xi)$ and $\exists_\xi\, \mathcal{A}(\xi)$ are also formulas. (Of course, no formulas can be formed in this fashion if the sets of variable symbols are empty, as in propositional logic.)

There are no other formulas besides those formed recursively using the rules above.

Remarks. The beginner might be concerned that 14.23(ii) seems self-referential and thus might permit circular reasoning. However, there is no need for worry — no circularity is possible here. Each formula formed as in 14.23(ii) is longer (in number of symbols used) than the formulas from which it was formed. A statement about the set of all formulas can be proved by induction on the lengths of the formulas; this is a common method of proof.

In the discussions below, formulas will generally be represented by the letters $\mathcal{A}, \mathcal{B}, \mathcal{C}$, etc. However, it should be understood that these letters are not actually symbols making up a part of our formal language (the inner system). Rather, they are *metavariables*, adopted for our informal discussion in the outer system. The precise expression "let $\mathcal{F}$ be a formula" is an abbreviation for the imprecise and unwieldly expression "let us consider any formula, such as $(\neg P(x, f(y,z))) \sqcup (Q(x,z) \sqcap ((R \sqcap (S(x,y,z,g(z))))))$."

14.24. An important special case of predicate logic is **propositional logic** (or **propositional calculus**, also known as **sentential logic** or **sentential calculus**). Historically, it developed before other kinds of logic. It is simpler than predicate logic, in that it has fewer ingredients. A typical formula in propositional logic is $(P \to (P \sqcap (\neg P))) \to (\neg P)$; a typical formula in predicate logic is $(\forall_\xi\ P(\xi, x)) \sqcup Q\,(x, f(y,z))$.

In propositional logic, there are no symbols for individuals — i.e., no constant individuals, no bound variables, and no free variables — and there are no function symbols. Consequently, the only relation symbols have arity 0, and there are no quantifiers and no terms.

Assumptions in First-Order Logic

In addition to its language (i.e., alphabet and grammatical rules, described above), a logical theory also involves certain assumptions. These are listed below.

14.25. *Logical axioms.* The literature contains many different axiomatizations of logic. We shall follow the development of Rasiowa and Sikorski [1963].

Our first nine axioms determine what is known as **positive logic**.

(i) $(\mathcal{A} \to \mathcal{B}) \to \big((\mathcal{B} \to \mathcal{C}) \to (\mathcal{A} \to \mathcal{C})\big)$. This is called the **Syllogism Law**.

(ii) $\mathcal{A} \to (\mathcal{A} \sqcup \mathcal{B})$.

(iii) $\mathcal{B} \to (\mathcal{A} \sqcup \mathcal{B})$.

(iv) $(\mathcal{A} \to \mathcal{C}) \to \Big((\mathcal{B} \to \mathcal{C}) \to ((\mathcal{A} \sqcup \mathcal{B}) \to \mathcal{C})\Big)$.

(v) $(\mathcal{A} \sqcap \mathcal{B}) \to \mathcal{A}$.

(vi) $(\mathcal{A} \sqcap \mathcal{B}) \to \mathcal{B}$.

(vii) $(\mathcal{C} \to \mathcal{A}) \to \Big((\mathcal{C} \to \mathcal{B}) \to \big(\mathcal{C} \to (\mathcal{A} \sqcap \mathcal{B})\big)\Big)$.

(viii) $\big(\mathcal{A} \to (\mathcal{B} \to \mathcal{C})\big) \to \big((\mathcal{A} \sqcap \mathcal{B}) \to \mathcal{C}\big)$. This is the **Importation Law**.

(ix) $\big((\mathcal{A} \sqcap \mathcal{B}) \to \mathcal{C}\big) \to \big(\mathcal{A} \to (\mathcal{B} \to \mathcal{C})\big)$. This is the **Exportation Law**.

The nine axioms above, plus the next two axioms below, determine what is commonly known as **intuitionist logic**. It was developed largely by Heyting and corresponds closely to intuitionist or constructivist thinking.

(x) $\big(\mathcal{A} \sqcap (\neg\mathcal{A})\big) \to \mathcal{B}$. This is the **Duns Scotus Law**.

(xi) $\Big(\mathcal{A} \to \big(\mathcal{A} \sqcap (\neg\mathcal{A})\big)\Big) \to (\neg\mathcal{A})$.

Finally, the eleven axioms above plus the twelfth axiom below determine what is known as **classical logic**, which is close to the way of thinking of most mathematicians.

(xii) $\mathcal{A} \sqcup (\neg\mathcal{A})$. This is the **Law of the Excluded Middle**, or *tertium non datur.*

The twelve rules listed above are actually axiom *schemes* — each of them represents infinitely many axioms. For instance, Axiom Scheme (ii) yields the axiom $P \to (P \sqcup Q)$, but it also yields the axiom $(P(x) \sqcap R(f(a,y))) \to ((P(x) \sqcap R(f(a,y))) \sqcup (Q \sqcap (\neg S)))$ by using different formulas for $\mathcal{A}$ and $\mathcal{B}$. (Recall that $\mathcal{A}$ and $\mathcal{B}$ only belong to the metalanguage. They are informal shorthand abbreviations for expressions such as $P(x) \sqcap R(f(a,y))$, which belong to the object language.)

14.26. The **rules of inference** of our logical system are rules by which, from a given set of formulas, we may **deduce** (or **infer**) another formula. Here, "deduce" and "infer" merely mean "obtain." We are not necessarily obtaining "true" formulas — we are merely collecting "obtainable" formulas, and the rules of inference tell us which formulas are obtainable. Admittedly, the rules of inference are most often applied to formulas that are in some sense "true," but this is not always the case. For instance, in a proof by contradiction, we may *assume* the negation of the desired conclusion, and then use the rules of inference to try to infer various consequences of that and other assumptions, until a contradiction is reached, thereby proving the desired conclusion.

The rules of inference and logical axioms vary slightly from one exposition to another. Indeed, what one book calls a rule of inference is what another book may call a logical axiom. Our own rules, listed below, follow those of Rasiowa and Sikorski [1963]. These rules will be "justified" in 14.55(ii).

(R1) **Modus ponens**, also known as the **rule of detachment**. Suppose $\mathcal{A}$ and $\mathcal{B}$ are formulas. Then from $\mathcal{A}$ and $(\mathcal{A} \to \mathcal{B})$ we can infer $\mathcal{B}$.

Our first rule, *modus ponens*, is present in all versions of logic. The remaining rules below involve variables, and so they can be skipped in considering any logic that does not involve variables (such as propositional logic).

(R2) **Rule of substitution.** Let $x_1, x_2, \ldots, x_n$ be distinct free individual variables, and let $t_1, t_2, \ldots, t_n$ be (not necessarily distinct) terms. Let $\mathcal{A}(x_1, x_2, \ldots, x_n)$ be a formula, in which each of the free individual variables $x_1, x_2, \ldots, x_n$ occurs 0 or more times. Let $\mathcal{A}(t_1, t_2, \ldots, t_n)$ be the formula obtained from $\mathcal{A}(x_1, x_2, \ldots, x_n)$ by simultaneously replacing all occurrences of the x_j's with copies of the corresponding t_j's. Then from $\mathcal{A}(x_1, x_2, \ldots, x_n)$ we can infer $\mathcal{A}(t_1, t_2, \ldots, t_n)$.

In the four rules below, let $\mathcal{A}(x)$ be a formula in which the bound variable ξ does not occur; we follow the substitution notation of 14.21. Also, let $\mathcal{B}$ be any formula. Then:

(R3) *Introduction of existential quantifiers.* Suppose $\mathcal{B}$ contains no occurrence of x. Then from $\mathcal{A}(x) \to \mathcal{B}$ we can infer $(\exists_\xi\ \mathcal{A}(\xi)) \to \mathcal{B}$.

(R4) *Introduction of universal quantifiers.* Suppose $\mathcal{B}$ contains no occurrence of x. Then from $\mathcal{B} \to \mathcal{A}(x)$ we can infer $\mathcal{B} \to (\forall_\xi\ \mathcal{A}(\xi))$.

(R5) *Elimination of existential quantifiers.* (We make no assumption about whether x appears in $\mathcal{B}$.) From $(\exists_\xi\ \mathcal{A}(\xi)) \to \mathcal{B}$ we can infer $\mathcal{A}(x) \to \mathcal{B}$.

(R6) *Elimination of universal quantifiers.* (We make no assumption about whether x appears in $\mathcal{B}$.) From $\mathcal{B} \to (\forall_\xi\ \mathcal{A}(\xi))$ we can infer $\mathcal{B} \to \mathcal{A}(x)$.

The rules of inference form the basis for our syntactic implications, and in fact the rules of inference are our most basic *examples* of syntactic implications. The rule of *modus ponens* says that for certain formulas $\mathcal{F}, \mathcal{G}, \mathcal{H}$ we have $\mathcal{F}, \mathcal{G} \vdash \mathcal{H}$; the other five rules of inference are of the form $\mathcal{F} \vdash \mathcal{G}$.

The rules of inference listed above will be *assumed* — i.e., taken as *hypotheses* in our reasoning about reasoning. Some auxiliary rules of inference will be proved as *consequences* of modus ponens and the logical axioms in 14.30.

14.27. *Examples of extra-logical axioms.* Besides the logical axioms shared by essentially all first-order theories determining our reasoning methods, a particular first-order theory may have additional, specialized axioms, determining the mathematical objects that we wish to study with that reasoning process. We refer to these as *extra-logical* or *nonlogical* axioms. Below are some examples. It should be understood that these examples are not part of the general explanation of predicate logic developed in this chapter — i.e., we shall not assume these axioms later in this chapter.

a. In many first-order systems, a special role is played by a relation symbol of rank two, called *equality* or *equals.* It is equipped with several axioms; the precise list of axioms varies from one exposition to another. Most often this symbol is denoted by " $=$ " — though in some expositions it may instead be denoted by " $\approx$ " or " $\equiv$ " or some other symbol, to emphasize that it is a symbol with precisely specified properties under formal study rather than just ordinary informal equality. Here is a typical set of **axioms for equality**:

(i) $x = x$.

(ii) $(x = y) \to (y = x)$.

(iii) $\big((x = y) \sqcap (y = z)\big) \to (x = z)$.

(iv) Let s_1, s_2, t be terms, and let y be a free variable. For $j = 1, 2$, let t_j be the term obtained from t by replacing each occurrence of the variable y with the term s_j. Then $(s_1 = s_2) \to (t_1 = t_2)$ is an axiom.

(v) With notation as in 14.21, if s_1, s_2 are terms, then $(s_1 = s_2) \to (\mathcal{A}(s_1) \to \mathcal{A}(s_2))$ is an axiom.

The first three of these axioms say that equality is an "equivalence relation," in a sense similar to that in 3.8 and 3.10. The last two axioms say that "equals can be substituted for equals." Actually, there is some redundancy in our formulation; our axioms (ii) and (iii) actually follow from axioms (i), (iv), and (v). (See Hamilton [1978], for instance.)

A logical system that includes such axioms is generally called **predicate logic with equality**. In this book we shall consider the axioms of equality to be extra-logical axioms, since they do not occur in all first-order systems. However, we caution that some mathematicians are concerned solely with first-order systems with equality, and some of these mathematicians find it convenient to designate the axioms of equality as "logical axioms" — which means that those axioms may sometimes get used without being mentioned.

b. For the **theory of preordered sets**, two of the binary relation symbols are $=$ and $\preccurlyeq$. Axioms used are the axioms for equality (described above) plus these axioms for the ordering:

$$\begin{array}{rl} \text{(reflexive)} & (x \preccurlyeq x), \\ \text{(transitive)} & \big((x \preccurlyeq y) \sqcap (y \preccurlyeq z)\big) \to (x \preccurlyeq z). \end{array}$$

For the **theory of partially ordered sets**, we add this axiom:

$$\text{(antisymmetric)} \qquad \big((x \preccurlyeq y) \sqcap (y \preccurlyeq x)\big) \to (x = y).$$

All of the axioms above are of first order – i.e., they deal only with individual members of a preordered or partially ordered set D, not with subsets of that set. In contrast, the Dedekind completeness of a poset D is a statement requiring a higher-order language. Recall that D is Dedekind complete if each nonempty subset that is bounded above has a least upper bound. In symbols, that condition is

$$\text{for each } S \subseteq D,\ ((\neg\ (S = \varnothing)) \sqcap (\exists x\ \forall y\ (y \in S \to y \preccurlyeq x))) \to (\exists u\ \forall x\ ((u \preccurlyeq x) \leftrightarrow (\forall y\ (y \in S \to y \preccurlyeq x))))$$

where $\mathcal{A} \leftrightarrow \mathcal{B}$ is an abbreviation for $(\mathcal{A} \to \mathcal{B}) \wedge (\mathcal{B} \to \mathcal{A})$. The condition begins with "for each set S that is a subset of D;" thus it involves a quantifier that ranges over *subsets* of D. There are other, equivalent ways to formulate the condition of Dedekind completeness of D, but none of them can be expressed in first-order language over D. Contrast this with 14.27.d, below.

c. For the **theory of monoids**, one of the binary relations is $=$, one of the binary functions is $\circ$, and some nullary function (i.e., function of arity 0) is denoted by i. Axioms used are the axioms of equality (described above) plus these axioms:

$$\begin{array}{rl} \text{(associative)} & (x \circ y) \circ z = x \circ (y \circ z), \\ \text{(right identity)} & x \circ i = x, \\ \text{(left identity)} & i \circ x = x. \end{array}$$

Additional algebraic axioms can be used to determine the theory of other types of algebraic systems — e.g., groups, rings, etc.

d. In the **language of set theory**, the individual elements a, b, c, x, y, z, etc., that we discuss are intended to represent sets. In conventional (i.e., atomless) set theory, the only undefined constant is $\varnothing$; all other constants are defined in terms of it. Thus, 0 is an abbreviation for $\varnothing$, 1 is an abbreviation for $\{\varnothing\}$, 2 is an abbreviation for $\{\varnothing, \{\varnothing\}\}$, etc., as in 5.44.

A basic binary relation is $\in$ (membership). Other relations can be defined in terms of membership. For instance, $u \subseteq v$ means $(x \in u) \to (x \in v)$, and $u = v$ means $(u \subseteq v) \sqcap (v \subseteq u)$.

The most commonly used axioms of set theory are the ZF axioms listed in 1.47. To make ZF into a first-order theory we must view the Axiom of Comprehension and the Axiom of Replacement not as single axioms, but as axiom *schemes*. Each of these two schemes represents infinitely many different axioms. We have one Axiom of Comprehension for each property P that can be formulated in the first-order language, and one Axiom of Replacement for each function f that can be formulated in the first-order language. (See also the reinterpretation of these axioms indicated in 14.67.)

The language of set theory is extremely powerful — it is more expressive than any of the other languages mentioned above. As we remarked in 1.46, all familiar objects of mathematics can be expressed in this language. The integers can be built up using the Axiom of Infinity; the rational numbers can be built up using equivalence classes of pairs of integers; the real numbers can be built up using Dedekind cuts of rationals. The language of set theory is sufficiently expressive for us to assert that a certain poset X is Dedekind complete: We can describe the ordering as a subset of $X \times X$, and the subsets of X are members of $\mathcal{P}(X)$. Contrast this with 14.27.b, above.

Some Syntactic Results (Propositional Logic)

14.28. *Remark.* We now consider some consequences of the logical axioms and inference rules listed in 14.25 and 14.26. We begin with some results that do not mention variables or constants; these results will not require any rules of inference except *modus ponens*; these results will apply equally well to propositional logic or predicate logic. In 14.39 we shall begin to consider results that do involve variables and constants.

14.29. *Some basic syntactic theorems of positive logic.*

(i) $\mathcal{A} \to \mathcal{A}$.

(ii) $\mathcal{A} \to (\mathcal{B} \to \mathcal{A})$.

(iii) $\mathcal{B} \to (\mathcal{A} \to \mathcal{A})$.

(iv) $\mathcal{A} \to ((\mathcal{A} \to \mathcal{A}) \to \mathcal{A})$.

(v) $\big((\mathcal{A} \to \mathcal{B}) \sqcap \mathcal{A}\big) \to \mathcal{B}$.

(vi) $\mathcal{A} \to (\mathcal{B} \to (\mathcal{A} \sqcap \mathcal{B}))$.

These results will be proved for all formulas $\mathcal{A}$ and $\mathcal{B}$,using the axioms of positive logic (i.e., Axioms (i) through (ix) of 14.25). These results will be used later in proofs.

Proof. The formula

$$\left(\Big(((\mathcal{A} \sqcap \mathcal{A}) \to \mathcal{A}) \sqcap \mathcal{A}\Big) \to \mathcal{A}\right) \to \Big(((\mathcal{A} \sqcap \mathcal{A}) \to \mathcal{A}) \to (\mathcal{A} \to \mathcal{A})\Big)$$

is an instance of Axiom (ix), and $\Big(((\mathcal{A} \sqcap \mathcal{A}) \to \mathcal{A}) \sqcap \mathcal{A}\Big) \to \mathcal{A}$ is an instance of Axiom (vi). Combine these, by *modus ponens*, to prove the formula $\big((\mathcal{A} \sqcap \mathcal{A}) \to \mathcal{A}\big) \to (\mathcal{A} \to \mathcal{A})$. Next, $(\mathcal{A} \sqcap \mathcal{A}) \to \mathcal{A}$ is another instance of Axiom (vi); combine that with the preceding formula to prove Theorem (i).

Theorem (ii) is immediate from Axioms (v) and (ix) via *modus ponens* (with the substitution $\mathcal{C} = \mathcal{A}$).

The formula $(\mathcal{A} \to \mathcal{A}) \to (\mathcal{B} \to (\mathcal{A} \to \mathcal{A}))$ is an instance of Theorem (ii). Combine it with Theorem (i), by *modus ponens*, to prove Theorem (iii).

Theorem (iv) is just an instance of Theorem (ii).

The formula $((\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{B})) \to \Big(((\mathcal{A} \to \mathcal{B}) \sqcap \mathcal{A}) \to \mathcal{B}\Big)$ is an instance of Axiom (viii). An instance of Theorem (i) is $(\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to \mathcal{B})$. Combine those, by *modus ponens*, to prove Theorem (v).

The formula $\big((\mathcal{A} \sqcap \mathcal{B}) \to (\mathcal{A} \sqcap \mathcal{B})\big) \to \big(\mathcal{A} \to (\mathcal{B} \to (\mathcal{A} \sqcap \mathcal{B}))\big)$ is an instance of Axiom (ix), and the formula $(\mathcal{A} \sqcap \mathcal{B}) \to (\mathcal{A} \sqcap \mathcal{B})$ is an instance of Theorem (i). Combine these to prove Theorem (vi).

14.30. *Additional rules of inference.* We shall use the axioms of positive logic to prove:

(i) If $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are some formulas such that $\mathcal{A} \to \mathcal{B}$ and $\mathcal{B} \to \mathcal{C}$ are syntactic theorems, then $\mathcal{A} \to \mathcal{C}$ is also a syntactic theorem.

(ii) $\mathcal{A} \sqcap \mathcal{B}$ is a syntactic theorem if and only if both $\mathcal{A}$ and $\mathcal{B}$ are syntactic theorems.

(iii) $(\mathcal{A} \to \mathcal{A}) \to \mathcal{A}$ is a syntactic theorem if and only if $\mathcal{A}$ is a syntactic theorem.

(iv) If $\mathcal{A} \to (\mathcal{B} \to \mathcal{C})$ and $\mathcal{A} \to \mathcal{B}$ are syntactic theorems, then $\mathcal{A} \to \mathcal{C}$ is a syntactic theorem.

Proofs. We shall use not only Axioms (i) through (ix), but also some of the Theorems of the previous section.

Rule (i) follows easily from Theorem (i) and *modus ponens.*

For Rule (ii), observe that if $\mathcal{A} \sqcap \mathcal{B}$ is a syntactic theorem, then $\mathcal{A}$ and $\mathcal{B}$ are syntactic theorems by Axioms (v) and (vi) via *modus ponens.* Conversely, if $\mathcal{A}$ and $\mathcal{B}$ are syntactic theorems, then $\mathcal{A} \sqcap \mathcal{B}$ follows from Theorem (vi) and two applications of *modus ponens.*

To prove Rule (iii): If $\mathcal{A}$ is a syntactic theorem, then from Theorem (iv) by *modus ponens* we know $(\mathcal{A} \to \mathcal{A}) \to \mathcal{A}$ is a syntactic theorem. Conversely, if $(\mathcal{A} \to \mathcal{A}) \to \mathcal{A}$ is a syntactic theorem, then from Theorem (i) by *modus ponens* we can conclude that $\mathcal{A}$ is also a syntactic theorem.

To prove Rule (iv): Note that $(\mathcal{A} \to (\mathcal{B} \to \mathcal{C})) \to ((\mathcal{A} \to \mathcal{B}) \to (\mathcal{A} \to ((\mathcal{B} \to \mathcal{C}) \sqcap \mathcal{B})))$ is an instance of Axiom (vii). Combine it with the two given syntactic theorems, via *modus ponens*; thus we obtain $\mathcal{A} \to (((\mathcal{B} \to \mathcal{C}) \sqcap \mathcal{B}))$ as a syntactic theorem. On the other hand, an instance of 14.29(v) gives us $((\mathcal{B} \to \mathcal{C}) \sqcap \mathcal{B}) \to \mathcal{C}$. Combine these two results, using 14.30(i); thus $\mathcal{A} \to \mathcal{C}$ is a syntactic theorem.

14.31. A set Σ of formulas is **syntactically inconsistent** if we can use Σ to deduce both $\mathcal{A}$ and $\neg\mathcal{A}$, for some formula $\mathcal{A}$. Note that we can then use Σ to deduce *any* formula; that is clear from the Duns Scotus Law (Axiom (x) in 14.25).

If Σ is not syntactically inconsistent, then it is **syntactically consistent**.

A derivation is understood to involve only finitely many steps, and so it can only involve finitely many of the axioms. Therefore,

> a collection of formulas is syntactically consistent if and only if each finite subset of that collection is syntactically consistent.

In other words, syntactic consistency of sets of formulas is a property with *finite character*, in the sense of 3.46.

14.32. *Definition of the ordering of the language.* Let $\mathbb{F}$ be the set of all formulas. We use the positive logic axioms from 14.25, plus whatever additional axioms we may choose. We now define two binary relations on $\mathbb{F}$ as follows: For formulas $\mathcal{A}$ and $\mathcal{B}$,

> $\mathcal{A} \preccurlyeq \mathcal{B}$ will mean that the formula $\mathcal{A} \to \mathcal{B}$ is a syntactic theorem;
>
> $\mathcal{A} \approx \mathcal{B}$ will mean that both $\mathcal{A} \to \mathcal{B}$ and $\mathcal{B} \to \mathcal{A}$ are syntactic theorems.

It should be emphasized that the relations $\preccurlyeq$ and $\approx$ are part of our metalanguage, not part of our object language (see 14.12.b). Thus, the expressions $\mathcal{A} \preccurlyeq \mathcal{B}$ and $\mathcal{A} \approx \mathcal{B}$ are not "formulas."

It follows from Theorem (i) of 14.29 and Rule (i) of 14.30 that

$$\preccurlyeq \text{ is a preorder, and } \approx \text{ is an equivalence relation, on } \mathbb{F}.$$

Let $\mathbb{L} = (\mathbb{F}/\approx)$ be the set of equivalence classes. The set $\mathbb{L}$, equipped with the operations discussed below, is commonly known as the **Lindenbaum algebra**.

The preorder $\preccurlyeq$ on $\mathbb{F}$ determines a partial order on $\mathbb{L}$, which we shall also denote by $\preccurlyeq$. Let $[\ ] : \mathbb{F} \to \mathbb{L}$ be the quotient map; that is, $[\mathcal{A}]$ is the equivalence class containing the formula $\mathcal{A}$. Thus

$$[\mathcal{A}] \preccurlyeq [\mathcal{B}] \qquad \text{if and only if} \qquad (\mathcal{A} \to \mathcal{B}) \text{ is a syntactic theorem.}$$

We shall use equality $(=)$ in its usual fashion as a relation between equivalence classes. Thus, $[\mathcal{A}] = [\mathcal{B}]$ means that $\mathcal{A}$ and $\mathcal{B}$ belong to the same equivalence class — i.e., it means that $\mathcal{A} \approx \mathcal{B}$.

It should be emphasized that any axioms whatsoever (or no axioms at all) may be used in addition to Axioms (i)–(ix). Different choices of additional axioms yield different relations $\preccurlyeq, \approx$ and thus yield different Lindenbaum algebras. Throughout most of our discussions in this and the next chapter, we assume that some particular choice is made regarding the additional axioms, and thus we may speak of *the* Lindenbaum algebra.

14.33. Theorem. The Lindenbaum algebra $(\mathbb{L}, \preccurlyeq)$ defined above is, in fact, a relatively pseudocomplemented lattice (as defined in 13.24), with operations given by

$$\begin{array}{rclcrcl} [\mathcal{A}] \vee [\mathcal{B}] & = & [\mathcal{A} \sqcup \mathcal{B}], & \qquad & ([\mathcal{A}] \Rightarrow [\mathcal{B}]) & = & [\mathcal{A} \to \mathcal{B}], \\ [\mathcal{A}] \wedge [\mathcal{B}] & = & [\mathcal{A} \sqcap \mathcal{B}], & & 1 & = & [\mathcal{A} \to \mathcal{A}] \end{array}$$

for any formulas $\mathcal{A}, \mathcal{B}$. A formula $\mathcal{A}$ is a syntactic theorem if and only if it satisfies $[\mathcal{A}] = 1$.

If we assume the axioms of intuitionist logic, then $\mathbb{L}$ is a Heyting algebra, with

$$0 = [\mathcal{A} \sqcap (\neg\mathcal{A})], \qquad \complement[\mathcal{A}] = [\neg\mathcal{A}]$$

for any formula $\mathcal{A}$. The Heyting algebra is degenerate (i.e., satisfying $0 = 1$) if and only if our set of axioms is syntactically inconsistent (i.e., there is some formula such that $\mathcal{A} \sqcap (\neg\mathcal{A})$ is a syntactic theorem), in which case every formula is provable.

If we assume the axioms of classical logic, then $\mathbb{L}$ is a Boolean algebra. Its greatest member 1 is also equal to $[\mathcal{A} \sqcup (\neg\mathcal{A})]$ for any formula $\mathcal{A}$.

The Boolean algebra $\mathbb{L}$ is more than just $\{0, 1\}$ if and only if at least one formula $\mathcal{F}$ is neither provable nor disprovable from the axioms.

Proof of theorem. (This argument follows the exposition of Rasiowa and Sikorski [1963].)

We first show that any two-element subset of $(\mathbb{L}, \preccurlyeq)$ has a supremum. Let any two elements of $\mathbb{L}$ be given; then those two elements can be represented as $[\mathcal{A}]$ and $[\mathcal{B}]$ for some formulas $\mathcal{A}$ and $\mathcal{B}$ (which are not uniquely determined by the given elements of $\mathbb{L}$). From Axioms (ii) and (iii) we see that $[\mathcal{A} \sqcup \mathcal{B}]$ is an upper bound for the set $\{[\mathcal{A}], [\mathcal{B}]\}$ in the poset $(\mathbb{L}, \preccurlyeq)$. Is it least among the upper bounds? Let any other upper bound for $\{[\mathcal{A}], [\mathcal{B}]\}$ be given; say that upper bound can be represented by $[\mathcal{C}]$ for some formula $\mathcal{C}$. Then $\mathcal{A} \to \mathcal{C}$ and $\mathcal{B} \to \mathcal{C}$ are syntactic theorems, by our definition of $\preccurlyeq$. By Axiom (iv) via *modus ponens*, it follows that $(\mathcal{A} \vee \mathcal{B}) \to \mathcal{C}$ is also a syntactic theorem; thus $[\mathcal{A} \sqcup \mathcal{B}] \preccurlyeq [\mathcal{C}]$. Thus $[\mathcal{A} \sqcup \mathcal{B}]$ is indeed least among the upper bounds of $\{[\mathcal{A}], [\mathcal{B}]\}$.

An analogous argument works for lower bounds, using Axioms (v), (vi), and (vii). Thus $\mathbb{L}$ is a lattice, with $[\mathcal{A}] \vee [\mathcal{B}] = [\mathcal{A} \sqcup \mathcal{B}]$ and $[\mathcal{A}] \wedge [\mathcal{B}] = [\mathcal{A} \sqcap \mathcal{B}]$.

Next we shall show that $([\mathcal{A}] \Rightarrow [\mathcal{B}]) = [\mathcal{A} \to \mathcal{B}]$ defines a relative pseudocomplementation operator. Let any formulas $\mathcal{A}, \mathcal{B}$ be given; we are to show that $[\mathcal{A} \to \mathcal{B}]$ is the largest λ in $\mathbb{L}$ that satisfies $\lambda \wedge [\mathcal{A}] \preccurlyeq [\mathcal{B}]$. By Theorem (v) of 14.29 we know that $[\mathcal{A} \to \mathcal{B}]$ is one of the λ's with that property. Is it the largest? Let any formula $\mathcal{D}$ satisfying $[\mathcal{D}] \wedge [\mathcal{A}] \preccurlyeq [\mathcal{B}]$ be given. Then $(\mathcal{D} \sqcap \mathcal{A}) \to \mathcal{B}$ is a syntactic theorem. The formula

$$\big((\mathcal{D} \sqcap \mathcal{A}) \to \mathcal{B}\big) \to \big(\mathcal{D} \to (\mathcal{A} \to \mathcal{B})\big)$$

is also a syntactic theorem, as it is an instance of Axiom (ix). From those two syntactic theorems via *modus ponens* we deduce the syntactic theorem $\mathcal{D} \to (\mathcal{A} \to \mathcal{B})$. In other words, $\mathcal{D} \preccurlyeq (\mathcal{A} \to \mathcal{B})$, so $[\mathcal{A} \to \mathcal{B}]$ is indeed largest.

That proves our claim about relative pseudocomplements. As in any relatively pseudocomplemented lattice, we now know that the largest element of $\mathbb{L}$ is $1 = [\mathcal{A} \rightarrow \mathcal{A}]$, for any formula $\mathcal{A}$. By 14.30(iii) we know that $\mathcal{A}$ is a syntactic theorem if and only if $1 \preccurlyeq [\mathcal{A}]$; that is, if and only if $1 = [\mathcal{A}]$.

Now suppose our axioms include the axioms of intuitionist logic. By Axiom (x) we see that $[\mathcal{A} \sqcap (\neg\mathcal{A})] \preccurlyeq [\mathcal{B}]$ for any formulas $\mathcal{A}, \mathcal{B}$. Thus $\mathbb{L}$ has a smallest element, given by the rule $0 = [\mathcal{A} \sqcap (\neg\mathcal{A})]$ for any formula $\mathcal{A}$. Hence $\mathbb{L}$ is a Heyting algebra. The conclusion about inconsistency and degeneracy is now obvious.

We have $0 = [\mathcal{A}] \wedge [\neg\mathcal{A}]$. By 13.25.a it follows that $[\neg\mathcal{A}] \preccurlyeq ([\mathcal{A}] \Rightarrow 0)$. On the other hand, we also have $([\mathcal{A}] \Rightarrow 0) = [\mathcal{A} \rightarrow (\mathcal{A} \sqcap (\neg\mathcal{A}))] \preccurlyeq [\neg\mathcal{A}]$ by Axiom (xi). Thus the pseudocomplement $([\mathcal{A}] \Rightarrow 0)$ is equal to $[\neg\mathcal{A}]$.

If we also assume Axiom (xii), then $[\mathcal{A}] \vee (\complement[\mathcal{A}]) = [\mathcal{A} \sqcup (\neg\mathcal{A})] = 1$, so $\mathbb{L}$ is a Boolean algebra.

14.34. *Further consequences in intuitionistic logic.* In any intuitionist logic, all the formulas given by the following schemes are syntactic theorems. This follows from the fact that the formulas correspond to identities that are satisfied in any Heyting algebra.

- **a. Interchange of Hypotheses.** $(\mathcal{A} \rightarrow (\mathcal{B} \rightarrow \mathcal{C})) \rightarrow (\mathcal{B} \rightarrow (\mathcal{A} \rightarrow \mathcal{C}))$.
- **b. Contrapositive Law.** $(\mathcal{A} \rightarrow (\neg\mathcal{B})) \rightarrow (\mathcal{B} \rightarrow (\neg\mathcal{A}))$.
- **c. Double Negation Law.** $\mathcal{A} \rightarrow (\neg\neg\mathcal{A})$.
- **d.** $(\mathcal{A} \rightarrow (\neg\mathcal{A})) \rightarrow (\neg\mathcal{A})$.
- **e.** $(\mathcal{A} \rightarrow \mathcal{B}) \rightarrow ((\neg\mathcal{B}) \rightarrow (\neg\mathcal{A}))$.
- **f. Brouwer's Triple Negation Law.** $(\neg\neg\neg\mathcal{A}) \rightarrow (\neg\mathcal{A})$ and $(\neg\mathcal{A}) \rightarrow (\neg\neg\neg\mathcal{A})$.
- **g.** $((\neg\mathcal{A}) \sqcap (\neg\mathcal{B})) \rightarrow (\neg(\mathcal{A} \sqcup \mathcal{B}))$ and $(\neg(\mathcal{A} \sqcup \mathcal{B})) \rightarrow ((\neg\mathcal{A}) \sqcap (\neg\mathcal{B}))$.
- **h.** $((\neg\mathcal{A}) \sqcup (\neg\mathcal{B})) \rightarrow (\neg(\mathcal{A} \sqcap \mathcal{B}))$.
- **i.** $((\neg\mathcal{A}) \sqcup \mathcal{B}) \rightarrow (\mathcal{A} \rightarrow \mathcal{B})$.

In classical logic we have those formulas, plus the ones listed in the section below.

14.35. *Some nonconstructive techniques of reasoning.* In the setting of intuitionist logic, the following formula schemes are undecidable, in the sense that they can neither be proved nor disproved syntactically, using just the intuitionist logical axioms. Moreover, they are equivalent to each other, in the sense that any one of them can be deduced from any of the others. They are all derivable in classical logic; adding any one of them to intuitionist logic yields classical logic.

Some of the formulas below could be viewed as symbolic representations of the principle of proof by contradiction.

- (A) **Law of the Excluded Middle:** $\mathcal{A} \sqcup (\neg\mathcal{A})$
- (B) **Converse of the Double Negation Law:** $(\neg\neg\mathcal{A}) \rightarrow \mathcal{A}$
- (C) $(\mathcal{A} \rightarrow \mathcal{B}) \rightarrow ((\neg\mathcal{A}) \sqcup \mathcal{B})$
- (D) $((\neg\mathcal{A}) \rightarrow (\neg\mathcal{B})) \rightarrow (\mathcal{B} \rightarrow \mathcal{A})$

(E) $((\neg\mathcal{A}) \to \mathcal{B}) \to ((\neg\mathcal{B}) \to \mathcal{A})$

for all formulas $\mathcal{A}, \mathcal{B}$.

Proof. The equivalence of these conditions is just a restatement of the result of 13.29. To say that these conditions cannot be disproved in intuitionist logic is just to say that the axioms of classical logic are syntactically consistent; that will be established in the next chapter. To show that these conditions cannot be disproved in intuitionist logic, let $(H, 0, 1, \vee, \wedge, \Rightarrow, \complement)$ be some particular Heyting algebra that is not a Boolean algebra. (An example of such is mentioned in 13.28.a.) Say the members of H are $0, 1, a, b, c$, etc. Form a propositional calculus that has one primitive propositional symbol for each member of H; say the primitive propositional symbols are denoted P_0, P_1, P_a, P_b, P_c, etc. Now *interpret* each formula in the propositional logic as the corresponding member of the Heyting algebra. For example, one instance of the Duns Scotus Law is the formula $\big(P_c \sqcap (\neg P_c)\big) \to P_a$; interpret this as the member of the Heyting algebra represented by $\big(c \wedge (\complement c)\big) \Rightarrow a$. That expression simplifies to 1, in any Heyting algebra. It is now a tedious but straightforward matter to verify that (i) each of the eleven logical axiom schemes of intuitionist logic is represented by 1 in the Heyting algebra; and (ii) if $\mathcal{F}$ and $\mathcal{F} \to \mathcal{G}$ are formulas represented by 1 in the Heyting algebra, then $\mathcal{G}$ is also represented by 1 in the Heyting algebra — i.e., *modus ponens* preserves "truth;" but (iii) the Law of the Excluded Middle is *not* represented by 1 in this particular Heyting algebra. This completes the proof.

Remarks. In effect, we have used H as a "quasimodel" for our propositional calculus. Models and quasimodels will be explored in greater detail later in this chapter. However, for brevity we shall only consider classical logics, and so all our formal models and quasimodels will be Boolean-valued. Thus the argument given in the preceding paragraph does not quite fit the formal framework developed later in this chapter.

14.36. *Discussion of intuitionist logic.* The axiom system of classical logic is slightly stronger than that of intuitionist logic. Hence, the set of syntactic theorems in classical logic is slightly larger than that of intuitionist logic.

The connective $\sqcup$ has rather different meanings in classical logic and intuitionist logic. In classical logic, if P is a primitive proposition symbol about which nothing in particular is assumed, then neither P nor $\neg P$ is a theorem; nevertheless $P \sqcup (\neg P)$ is a theorem — this is just the Law of the Excluded Middle. In contrast,

> in intuitionist propositional logic, if $\mathcal{A}$ and $\mathcal{B}$ are some formulas such that $\mathcal{A} \sqcup \mathcal{B}$ is a syntactic theorem, then at least one of $\mathcal{A}$ or $\mathcal{B}$ is a syntactic theorem.

(The proof is too difficult to give here; a topological proof is given by Rasiowa and Sikorski [1963, page 394]. An analogous result for predicate logic can be found on page 430 of that book.) This result may surprise many readers, because it is so different from what we are familiar with in classical logic. It may also puzzle some readers, because it seems to give a *stronger* conclusion in intuitionist logic than in classical logic — even though intuitionist logic is the weaker logic. But read carefully! Since intuitionist logic has fewer axioms and fewer syntactic theorems than classical logic, the hypothesis that "$\mathcal{A} \sqcup \mathcal{B}$ is a syntactic theorem of intuitionist logic" is *stronger* than the hypothesis that "$\mathcal{A} \sqcup \mathcal{B}$ is a syntactic theorem of classical logic."

Heyting developed his algebraic approach to intuitionist logic in a paper in 1930. In 1932, Kolmogorov published some related results, including this intuitive (i.e., real-world) interpretation of Heyting's formalism:

Let us use letters such as $\mathcal{A}, \mathcal{B}, \mathcal{C}$, etc., to denote problems that are to be solved. Interpret connectives as follows:

> $\mathcal{A} \sqcap \mathcal{B}$ means the problem "to solve both $\mathcal{A}$ and $\mathcal{B}$,"
>
> $\mathcal{A} \sqcup \mathcal{B}$ means the problem "to solve at least one of $\mathcal{A}$ or $\mathcal{B}$,"
>
> $\mathcal{A} \rightarrow \mathcal{B}$ means the problem "to show how any solution of $\mathcal{A}$ would yield a solution of $\mathcal{B}$,"
>
> $\neg\mathcal{A}$ means the problem "to show how any solution of $\mathcal{A}$ would yield a contradiction."

Then the properties of Kolmogorov's system of problem-solving coincides with the properties of Heyting's formal intuitionist propositional calculus. (For further references and discussion, see Kneebone [1963].) The Law of the Excluded Middle (which we introduced in 6.4) is not taken as an axiom in the intuitionist system of Heyting or Kolmogorov — i.e., although for some particular problems $\mathcal{A}$ we may be able to solve at least one of $\mathcal{A}$ or $\neg\mathcal{A}$, we do not have a general method for doing that. The systems of Heyting and Kolmogorov reflect a somewhat constructive viewpoint, but we shall not try to make that description precise, for there are many different schools of constructivism.

Some Syntactic Results (Predicate Logic)

14.37. *Remark.* Our preceding syntactic results did not make any direct use of variables; they would apply equally well to propositional logic or predicate logic. We now turn to syntactic results that do involve variables. Most of these results are only relevant to predicate logic. A few of them are also relevant to propositional logic, but take a simplified form in that case; see 14.40.b for instance.

14.38. We begin by considering the relation between these two kinds of implications:

$$\text{(i)} \quad \Sigma \vdash (\mathcal{F} \rightarrow \mathcal{G}), \qquad\qquad \text{(ii)} \quad \Sigma \cup \mathcal{F} \vdash \mathcal{G}.$$

Here $\mathcal{F}$ and $\mathcal{G}$ are any formulas, and Σ is any set of formulas.

It is easy to see that (i) $\Rightarrow$ (ii) — i.e., that whenever $\mathcal{F}$ and $\mathcal{G}$ are some particular formulas that satisfy (i), then they also satisfy (ii). (*Proof.* Assume $\Sigma \vdash (\mathcal{F} \rightarrow \mathcal{G})$, and assume we are given the set of formulas $\Sigma \cup \mathcal{F}$. Since we are given Σ, we may deduce $\mathcal{F} \rightarrow \mathcal{G}$. Since we are also given $\mathcal{F}$, by *modus ponens* we may deduce $\mathcal{G}$.)

Under certain additional assumptions we can show that (ii) $\Rightarrow$ (i), and therefore (i), (ii) are equivalent; that is the subject of 14.39 and 14.40. However, in general (i.e., without additional assumptions), (ii) does *not* imply (i); that is shown by an example in 14.60.

14.39. The Deduction Principle. Let $\mathcal{F}$ and $\mathcal{G}$ be formulas, and let Σ be a set of formulas. Suppose that $\Sigma \cup \{\mathcal{F}\} \vdash \mathcal{G}$; that is, there exists a derivation of $\mathcal{G}$ from $\Sigma \cup \{\mathcal{F}\}$. Suppose, moreover, that the derivation can be chosen so that

> whenever any of the inference rules (R2), (R3), (R4) is used, then the free variables $x, x_1, x_2, x_3, \ldots$ being replaced are symbols that do not appear in $\mathcal{F}$.

Then $\Sigma \vdash (\mathcal{F} \to \mathcal{G})$.

Proof. We view Σ as a collection of extra-logical axioms. Let $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ be the given derivation. Thus, $\mathcal{E}_n = \mathcal{G}$, and each $\mathcal{E}_j$ is either a logical or extra-logical axiom, or $\mathcal{F}$, or a consequence of previous $\mathcal{E}_i$'s by the rules of inference. It suffices to prove, by induction on $k = 1, 2, \ldots, n$, that $\Sigma \vdash (\mathcal{F} \to \mathcal{E}_k)$. We prove this by considering cases according to the method by which $\mathcal{E}_k$ enters the given derivation.

If $\mathcal{E}_k$ is an axiom, then from $\mathcal{E}_k \to (\mathcal{F} \to \mathcal{E}_k)$ (in 14.29(ii)) and $\mathcal{E}_k$ we may deduce $\mathcal{F} \to \mathcal{E}_k$. If $\mathcal{E}_k$ is equal to $\mathcal{F}$, then $\mathcal{F} \to \mathcal{E}_k$ is the formula $\mathcal{F} \to \mathcal{F}$, which was proved in 14.29(i). Thus, in these cases we have $\Sigma \vdash (\mathcal{F} \to \mathcal{E}_k)$, without even referring to the induction hypothesis.

Next, consider the case in which $\mathcal{E}_k$ follows from previous formulas $\mathcal{E}_i$ and $\mathcal{E}_j$ via *modus ponens*. Then (with i and j switched if necessary) we may assume $\mathcal{E}_i$ is the formula $\mathcal{E}_j \to \mathcal{E}_k$. By our induction hypothesis we have $\Sigma \vdash (\mathcal{F} \to (\mathcal{E}_j \to \mathcal{E}_k))$ and $\Sigma \vdash (\mathcal{F} \to \mathcal{E}_j)$. By 14.30(iv) it follows that $\Sigma \vdash (\mathcal{F} \to \mathcal{E}_k)$.

Next, consider the case in which $\mathcal{E}_k$ follows from some previous formula $\mathcal{E}_j$ by the Rule of Substitution (R2) — i.e., by replacing some or all of the free variables with specified terms. Since none of those free variables appear in $\mathcal{F}$, the same substitution leaves $\mathcal{F}$ unaffected. Thus $\mathcal{F} \to \mathcal{E}_k$ follows from $\mathcal{F} \to \mathcal{E}_j$ by the Rule of Substitution.

Next, consider the cases in which $\mathcal{E}_k$ follows from some previous formula $\mathcal{E}_j$ by one of the remaining inference rules (R3), (R4), (R5), (R6). In these cases, $\mathcal{E}_j$ is a formula $\mathcal{C} \to \mathcal{D}$ of a certain type, and from it we can deduce $\mathcal{E}_k$, a formula $\mathcal{C}' \to \mathcal{D}'$ of a certain type. It suffices to show, by different reasonings in these four cases, that

$$\text{from } \mathcal{F} \to (\mathcal{C} \to \mathcal{D}), \text{ we can deduce } \mathcal{F} \to (\mathcal{C}' \to \mathcal{D}'). \tag{a}$$

For some of these cases it is helpful to use 14.34.a; thus (a) is established if we can just show that

$$\text{from } \mathcal{C} \to (\mathcal{F} \to \mathcal{D}), \text{ we can deduce } \mathcal{C}' \to (\mathcal{F} \to \mathcal{D}'). \tag{b}$$

For other cases, it is helpful to use Axioms (viii) and (ix) in 14.25; thus (a) is established if we can just show that

$$\text{from } (\mathcal{C} \sqcap \mathcal{F}) \to \mathcal{D}, \text{ we can deduce } (\mathcal{C}' \sqcap \mathcal{F}) \to \mathcal{D}'. \tag{c}$$

Applications of (R3) are of this form: $\mathcal{A}(x)$ contains no occurrence of ξ, $\mathcal{B}$ contains no occurrence of x, and from $\mathcal{A}(x) \to \mathcal{B}$ we infer $(\exists_\xi\, \mathcal{A}(\xi)) \to \mathcal{B}$. By assumption, x does not occur in $\mathcal{F}$, hence it does not occur in $(\mathcal{F} \to \mathcal{B})$, and so from $\mathcal{A}(x) \to (\mathcal{F} \to \mathcal{B})$ we infer $(\exists_\xi\, \mathcal{A}(\xi)) \to (\mathcal{F} \to \mathcal{B})$, by the same rule of inference. That is just (b).

Applications of (R4) are of this form: $\mathcal{A}(x)$ contains no occurrence of ξ, $\mathcal{B}$ contains no occurrences of x, and from $\mathcal{B} \to \mathcal{A}(x)$ we infer $\mathcal{B} \to (\forall_\xi\, \mathcal{A}(\xi))$. By assumption, $\mathcal{F}$ contains no occurrences of x, hence $\mathcal{B} \sqcap \mathcal{F}$ contains no occurrences of x, hence by the same inference rule from $(\mathcal{B} \sqcap \mathcal{F}) \to \mathcal{A}(x)$ we infer $(\mathcal{B} \sqcap \mathcal{F}) \to (\forall_\xi\ \mathcal{A}(\xi))$. This is just (*c*).

Applications of (R5) are of this form: $\mathcal{A}(x)$ contains no occurrence of ξ and from $(\exists_\xi\ \mathcal{A}(\xi)) \to \mathcal{B}$ we infer $\mathcal{A}(x) \to \mathcal{B}$. By the same inference rule, from $(\exists_\xi\ \mathcal{A}(\xi)) \to (\mathcal{F} \to \mathcal{B})$ we infer $\mathcal{A}(x) \to (\mathcal{F} \to \mathcal{B})$. This is (*b*).

Applications of (R6) are of this form: $\mathcal{A}(x)$ contains no occurrence of ξ, and from $\mathcal{B} \to (\forall_\xi\ \mathcal{A}(\xi))$ we infer $\mathcal{B} \to \mathcal{A}(x)$. By the same rule, from $(\mathcal{B} \sqcap \mathcal{F}) \to (\forall_\xi\ \mathcal{A}(\xi))$ we infer $(\mathcal{B} \sqcap \mathcal{F}) \to \mathcal{A}(x)$. This is (*c*).

This completes the proof.

14.40. *Corollaries.* Refer to 14.38. In special circumstances we obtain simplified versions of the Deduction Principle:

a. Deduction Principle for Closed Formulas. Let $\mathcal{F}$ and $\mathcal{G}$ be formulas, and let Σ be a set of formulas. Assume $\mathcal{F}$ has no free variables. Then $\Sigma \cup \{\mathcal{F}\} \vdash \mathcal{G}$ if and only if $\Sigma \vdash (\mathcal{F} \to \mathcal{G})$.

b. Deduction Principle for Propositional Logic. In propositional logic, let $\mathcal{F}$ and $\mathcal{G}$ be formulas, and let Σ be a set of formulas. Then $\Sigma \cup \{\mathcal{F}\} \vdash \mathcal{G}$ if and only if $\Sigma \vdash (\mathcal{F} \to \mathcal{G})$.

(In particular, taking $\Sigma = \varnothing$, we see that $\mathcal{F} \vdash \mathcal{G}$ if and only if $\vdash (\mathcal{F} \to \mathcal{G})$. Thus, "$\to$" means just what one would expect it to mean, at least in propositional logic.)

c. A weak form of proof by contradiction. Suppose Σ is a set of formulas and $\mathcal{A}$ is a formula with no free variables. If $\Sigma \cup \{\mathcal{A}\}$ is syntactically inconsistent, then $\Sigma \vdash (\neg\mathcal{A})$. (Moreover, this result is valid both in classical logic and in intuitionist logic.)

Proof of c. By assumption, $\Sigma \cup \{\mathcal{A}\} \vdash (\mathcal{B} \sqcap (\neg\mathcal{B}))$ for some formula $\mathcal{B}$. Now, the formula $(\mathcal{B} \sqcap (\neg\mathcal{B})) \to (\neg\mathcal{A})$ is an instance of the Duns Scotus Law, Axiom (x) from 14.25. Hence by *modus ponens* we have $\Sigma \cup \{\mathcal{A}\} \vdash (\neg\mathcal{A})$. Since $\mathcal{A}$ has no free variables, we have $\Sigma \vdash (\mathcal{A} \to (\neg\mathcal{A}))$ by the Deduction Principle 14.40.a. However, by 14.34.e we have the syntactic theorem $(\mathcal{A} \to (\neg\mathcal{A})) \to (\neg\mathcal{A})$, based solely on the logical axioms. Hence by *modus ponens* we have $\Sigma \vdash (\neg\mathcal{A})$.

14.41. Theorem characterizing quantifiers as sup and inf. Let ξ be a bound variable that does not appear in the formula $\mathcal{A}(x)$; we follow the substitution notation of 14.21. Let T be the set of all free variables (— an infinite set, by our assumption in 14.17), or more generally let T be any set of terms with $T \supseteq \{x_1, x_2, x_3, \ldots\}$ where the x_j's are distinct free variables. Then, in the Lindenbaum algebra $(\mathbb{L}, \preccurlyeq)$, we have

$$[\exists_\xi\ \mathcal{A}(\xi)] = \sup_{t \in T}\, [\mathcal{A}(t)], \qquad [\forall_\xi\ \mathcal{A}(\xi)] = \inf_{t \in T}\, [\mathcal{A}(t)].$$

Proof. By 14.29(i) we know $(\exists_\xi \mathcal{A}(\xi)) \to (\exists_\xi \mathcal{A}(\xi))$ is a syntactic theorem. From that syntactic theorem and inference rule (R5) we can infer that $\mathcal{A}(x) \to (\exists_\xi\ \mathcal{A}(\xi))$ is also a syntactic theorem. Let t be any term in T. Note that the variable x does not appear in the formula $(\exists_\xi\ \mathcal{A}(\xi))$. When we replace each x with t in the formula $\mathcal{A}(x) \to (\exists_\xi\ \mathcal{A}(\xi))$, the result is the formula $\mathcal{A}(t) \to (\exists_\xi\ \mathcal{A}(\xi))$, which is therefore is a syntactic theorem by (R2).

Thus $[\mathcal{A}(t)] \preccurlyeq [\exists_\xi \mathcal{A}(\xi)]$ in the poset $(\mathbb{L}, \preccurlyeq)$. That is, $[\exists_\xi \; \mathcal{A}(\xi)]$ is an upper bound for the set $\{[\mathcal{A}(t)] : t \in T\}$.

Is it the least upper bound? Let $\mathcal{B}$ be a formula such that $[\mathcal{B}]$ is also an upper bound for $\{[\mathcal{A}(t)] : t \in T\}$; we are to show that $[\exists_\xi \mathcal{A}(\xi)] \preccurlyeq [\mathcal{B}]$. For each $t \in T$, our assumption about $\mathcal{B}$ says that $\mathcal{A}(t) \to \mathcal{B}$ is a syntactic theorem. Since T contains infinitely many free variables, and $\mathcal{B}$ and $\mathcal{A}(x)$ are strings of only finitely many symbols, we have $\mathcal{A}(z) \to \mathcal{B}$ for some free variable z that belongs to T but does not appear in $\mathcal{B}$ or in $\mathcal{A}(x)$. Then ξ does not appear in $\mathcal{A}(z)$, and $\mathcal{A}(\xi)$ is the same as the string of symbols obtained from $\mathcal{A}(z)$ by replacing each z with ξ. By (R3), the formula $(\exists_\xi \mathcal{A}(\xi)) \to \mathcal{B}$ is a syntactic theorem. That is, $[\exists_\xi \mathcal{A}(\xi)] \preccurlyeq [\mathcal{B}]$, as required.

A dual argument proves $[\forall_\xi \; \mathcal{A}(\xi)] = \inf_{t \in T} [\mathcal{A}(t)]$.

14.42. *Corollary.* Suppose that neither of the bound variables ξ, η appears in the formula $\mathcal{A}(x)$. Then

$$[\forall_\xi \; \mathcal{A}(\xi)] = [\forall_\eta \; \mathcal{A}(\eta)], \qquad [\exists_\xi \; \mathcal{A}(\xi)] = [\exists_\eta \; \mathcal{A}(\eta)].$$

(Thus, the validity of a syntactic theorem does not depend on our choice of the bound variables.) Also,

$$[\neg (\forall_\xi \; \mathcal{A}(\xi))] = [\exists_\xi \; (\neg \mathcal{A}(\xi))],$$
$$[(\forall_\xi \; \mathcal{A}(\xi)) \sqcup \mathcal{B}] = [\forall_\xi \; (\mathcal{A}(\xi) \sqcup \mathcal{B})],$$
$$[(\exists_\xi \; \mathcal{A}(\xi)) \sqcup \mathcal{B}] = [\exists_\xi \; (\mathcal{A}(\xi) \sqcup \mathcal{B})],$$

etc. By repeated use of these formulas, quantifiers can be moved to the beginning of any formula. Thus, any formula is equivalent to a formula in **prenex normal form** — i.e., a formula consisting of a string of quantifiers, followed by a string of symbols other than quantifiers.

14.43. Rule of Generalization. Let ξ be a bound variable that does not appear in the formula $\mathcal{A}(x)$; we follow the substitution notation of 14.21. Then $\mathcal{A}(x)$ is a syntactic theorem if and only if $\forall_\xi \; \mathcal{A}(\xi)$ is a syntactic theorem.

Proof. By 14.29(i) we know $(\forall_\xi \mathcal{A}(\xi)) \to (\forall_\xi \mathcal{A}(\xi))$ is a syntactic theorem; by inference rule (R6) it follows that $(\forall_\xi \mathcal{A}(\xi)) \to \mathcal{A}(x)$ is also a syntactic theorem. If $\forall_\xi \mathcal{A}(\xi)$ is a syntactic theorem, then *modus ponens* tells us $\mathcal{A}(x)$ is also a syntactic theorem.

Conversely, assume $\mathcal{A}(x)$ is a syntactic theorem. Let T be the set of all terms. For each $t \in T$, we know by (R2) that $\mathcal{A}(t)$ is a syntactic theorem; that is, $[\mathcal{A}(t)] = 1$. Hence $[\forall_\xi \mathcal{A}(\xi)] = \inf_{t \in T} [\mathcal{A}(t)] = 1$, so $\forall_\xi \mathcal{A}(\xi)$ is a syntactic theorem.

14.44. *Definitions.* A **sentence**, or **closed formula**, is a formula with no free variables. (Note that "formulas" and "sentences" are the same thing in propositional logic, since that logic has no free variables.)

Let $\mathcal{A}(x_1, x_2, \ldots, x_n)$ be some given formula whose free variables are the distinct symbols $x_1, x_2, \ldots, x_n$, and no others. Let $\xi_1, \xi_2, \ldots, \xi_n$ be any n distinct bound variables that do not appear in the given formula. Then the **closure** of the given formula is the formula $\forall_{\xi_1} \; \forall_{\xi_2} \; \cdots \; \forall_{\xi_n} \; \mathcal{A}(\xi_1, \xi_2, \ldots, \xi_n)$; that is, it is the formula obtained by replacing all occurrences of each x_i with the corresponding ξ_i and binding all the ξ_i's with universal quantifiers. (Observe that the closure is a closed formula.)

By applying the Rule of Generalization n times, we obtain:

Proposition. Let $\mathcal{F}$ be any formula. Then $\mathcal{F}$ is a syntactic theorem if and only if its closure is a syntactic theorem.

14.45. *Lemma.* Let $(\mathfrak{L}, \Sigma)$ be a first-order theory (with language $\mathfrak{L}$ and axioms Σ) that is syntactically consistent. Assume $\mathfrak{L}$ has infinitely many free variables. Let c be a constant symbol that is not already in use in the language $\mathfrak{L}$; let $\mathfrak{L}' = \mathfrak{L} \cup \{c\}$ be the larger language obtained by adding that one symbol. Then $(\mathfrak{L}', \Sigma)$ is also syntactically consistent.

Proof. Suppose not — i.e., suppose that $\Sigma \vdash \mathcal{F} \sqcap (\neg\mathcal{F})$ for some formula $\mathcal{F}$ when we use the enlarged language $\mathfrak{L}'$. Consider any derivation of the formula $\mathcal{F} \sqcap (\neg\mathcal{F})$ from the axioms of Σ. The derivation involves only finitely many steps, hence only finitely many symbols. Select some free variable z that does not appear in the derivation, and replace all occurrences of c with z throughout the derivation. This yields a derivation $\Sigma \vdash \mathcal{G} \sqcap (\neg\mathcal{G})$ in the language $\mathfrak{L}$, for some formula $\mathcal{G}$ — contradicting the assumed consistency of $(\mathfrak{L}, \Sigma)$.

14.46. *Technical lemma.* (This lemma will be used in the proof of 14.57.)

Let $(\mathfrak{L}, \Sigma)$ be a first-order theory (with language $\mathfrak{L}$ and axioms Σ) that is syntactically consistent. Assume that the language $\mathfrak{L}$ contains infinitely many free variables, but all the axioms in Σ are closed formulas. Let $(\mathfrak{L}', \Sigma')$ be a first-order theory with larger language $\mathfrak{L}'$ and larger axiom collection Σ', formed by adding more constant symbols and more axioms according to these rules:

(i) Whenever t is a term in the language $\mathfrak{L}$ that involves no free variables (i.e., it involves only constant symbols and function symbols), and $\forall_\xi\ \mathcal{A}(\xi)$ is an axiom in Σ, then add the axiom $\mathcal{A}(t)$ as one of the ingredients of Σ'.

(ii) For each axiom in Σ of the form $\exists_\xi\ \mathcal{A}(\xi)$, add to the language some constant symbol c not already present in $\mathfrak{L}$, and add the axiom $\mathcal{A}(c)$.

Then the new system $(\mathfrak{L}', \Sigma')$ is also syntactically consistent.

Remarks. A constant c that satisfies $\mathcal{A}(c)$ is sometimes called a **witness** for the axiom $\exists_\xi\ \mathcal{A}(\xi)$. Thus, in part (ii) we add a new symbol c to serve as a witness. That new symbol can be chosen canonically — i.e., this does not require the Axiom of Choice — for instance, we could use the symbol $\boxed{\exists_\xi\ \mathcal{A}(\xi)}$, where the box and all the marks inside it are parts of the symbol.

Proof of lemma. Suppose that $(\mathfrak{L}', \Sigma')$ is not syntactically consistent — i.e., adding axioms and constants by methods (i) and (ii) yields a contradiction $\mathcal{F} \sqcap (\neg\mathcal{F})$. Then finitely many of those axioms and symbols yield a contradiction, so the contradiction results from extending $(\mathfrak{L}, \Sigma)$ finitely many times using steps of type (i) or (ii) above. Take those finitely many steps, one at a time, and stop at the point where the system becomes inconsistent. Say we have the systems

$$(\mathfrak{L}_0, \Sigma_0) \text{ consistent} \qquad \text{and} \qquad (\mathfrak{L}_1, \Sigma_1) \text{ inconsistent},$$

separated by one step.

The step that causes the inconsistency cannot be of type (i), since steps of that type do not enlarge the language, and they enlarge the axiom system only by adding formulas that were already syntactic theorems (by the Rule of Generalization 14.43 and the Rule of Substitution 14.26(R2)). Thus, the step causing inconsistency is of type (ii). We have

$$\mathfrak{L}_1 = \mathfrak{L}_0 \cup \{c\} \qquad \text{and} \qquad \Sigma_1 = \Sigma_0 \cup \{\mathcal{A}(c)\},$$

where $\mathcal{A}$ is some formula such that

$$\exists_\xi\ \mathcal{A}(\xi) \quad \text{belongs to } \Sigma \text{ and hence also belongs to } \Sigma_0. \tag{$**$}$$

By 14.45 we know that $(\mathfrak{L}_1, \Sigma_0)$ is consistent. We shall use the language $\mathfrak{L}_1$ throughout the remainder of this proof.

Since $\Sigma_0 \cup \{\mathcal{A}(c)\}$ is inconsistent and $\mathcal{A}(c)$ has no free variables, 14.40.c tells us that $\Sigma_0 \vdash (\neg\mathcal{A}(c))$. Let some derivation of $(\neg\mathcal{A}(c))$ from Σ_0 be specified, and let z be a free variable that does not appear in that derivation. Replacing c with z throughout that derivation yields a derivation of the formula $\neg\mathcal{A}(z)$. By the Rule of Generalization 14.43, $\forall_\xi\ (\neg\mathcal{A}(\xi))$ is a syntactic theorem of Σ_0. By 14.41, that theorem can be restated as $\neg(\exists_\xi\ \mathcal{A}(\xi))$. This contradicts $(**)$ and completes the proof.

The Semantic View

14.47. We shall now discuss *interpretations* and *quasi-interpretations* of a language. In the discussion below, $|\ |$ will represent several different but related mappings.

By a **quasi-interpretation** of a first-order language $\mathfrak{L}$ we shall mean[1] a triple $(B, D, |\ |)$, with these ingredients:

a. A Boolean algebra B is specified, which must be nondegenerate (i.e., satisfying $0 \neq 1$). We may refer to B as the set of **truth values**. We require that the Boolean algebra B satisfy a certain *completeness condition*, which is slightly complicated; it will be described in 14.47.j($\natural$). The condition is satisfied if B happens to be a complete Boolean algebra (i.e., if every subset of B has a supremum and an infimum), but the condition may also be satisfied by certain other Boolean algebras. In the simplest cases we may take $B = \{0, 1\}$, and then the quasi-interpretation is called an **interpretation**. However, other choices of B are also of interest, and so we consider the more general theory of quasi-interpretations. The element 1 (the greatest element of B) may be called **true** or **truth** or **tautology**. The element 0 (the least element of B) may be called **false** or **falsehood** or **contradiction**.

b. A collection D of objects is specified; it is sometimes called the **domain**. It may be a set or a proper class. Members of D are called **individuals**. Typical examples: For analysis, D could be the set of real numbers. For set theory, D could be the class

[1] See also the alternate terminologies in 14.50.

of all sets or some smaller collection of sets such as the collection $\mathcal{M}$ in 14.6. For propositional logic, we generally take D to be empty.

In some discussions, we may refer to D itself as the "interpretation" or "quasi-interpretation" but it is understood that the interpretation or quasi-interpretation also involves the Boolean algebra B and the mappings $|\ |$ discussed below. A **countable** quasi-interpretation is a quasi-interpretation whose domain D is a countable set.

c. A mapping $|\ |$ is given from the set of individual constant symbols of the language $\mathcal{L}$, into D. Thus, each constant symbol c in $\mathcal{L}$ is understood to represent some individual in D, which will be denoted in this discussion by $|c|$. We say that c is the **name** of the individual $|c|$. In ordinary mathematics we do not distinguish between an individual object and the symbol that is its name, but that distinction is important in logic.

We emphasize that this mapping $|\ |$ is not necessarily injective — i.e., one quasi-interpretation may give several names to the same individual. In other words, several constant symbols in $\mathcal{L}$ could conceivably all be names for the same number or set or other mathematical object in D.

Also, we emphasize that this mapping $|\ |$ is not necessarily surjective — i.e., not every individual in D necessarily has a name. (For theoretical purposes it is sometimes useful to work with $\mathcal{L}(D)$, the language obtained by adding to $\mathcal{L}$ another constant symbol for each individual in D — i.e., a language that gives a name to every member of D — but that language will not be studied here.)

d. For each n-ary function symbol f in the formal language, the mapping $|\ |$ must specify some n-ary function $|f|$ on D — that is, a mapping $|f| : D^n \to D$. Note this distinction: f is a *meaningless letter*, while $|f|$ is a *function*. Typical examples with $n = 2$: If D is the real number system, then $|f|$ might be addition or multiplication. If D is the class of all sets, then $|f|$ might be the union operation.

We emphasize that $|f|(d_1, d_2, \ldots, d_n)$ is defined to be some member of D whenever $d_1, d_2, \ldots, d_n$ are *any* individuals; they need not be individuals with names. Thus, the function $|f|$ goes beyond merely specifying meanings for expressions that can be formulated in the formal language $\mathcal{L}$.

e. For each n-ary relation symbol R in the formal language, we specify some corresponding function $|R| : D^n \to B$.

Note that if $B = \{0, 1\}$, then $|R|$ may be viewed as (the characteristic function of) some subset of D^n; thus it is an n-ary relation on D. In particular, if $n = 2$, then $|R|$ is (the characteristic function of the graph of) a binary relation on D, such as $\leq$ or $\subseteq$. Here is a typical example: When the language of analysis is interpreted in its usual fashion, then D is the set of real numbers and $<$ is a binary relation symbol that is interpreted as a function $|<|$ from $\mathbb{R}^2$ into {"true," "false"}. Some of its values are $|<|(3, 5) =$ "true" and $|<|(3, 3) =$ "false."

Note that in propositional logic, each relation symbol R has arity 0, and so its quasi-interpretation $|R|$ is just a constant member of B.

Continuation of the quasi-interpretation. The ingredients indicated above can be specified arbitrarily (except that B must satisfy the completeness condition 14.47.j(♮)). Once those ingredients have been specified, we extend the mappings $|\ |$ according to the following rules, which are not at all arbitrary.

f. Each *free variable symbol* x is interpreted as the identity map from D into D. Thus, it acts as a variable whose possible values are the members of D.

g. Any *term* in the formal language is interpreted as a function from D^m into D for some nonnegative integer m; here m is the number of distinct free variables appearing in the expression term. Indeed, we have already interpreted each constant symbol as a mapping from D^0 into D (i.e., as a member of D — see 14.47.c) and each free variable symbol as a mapping from D^1 into D (see 14.47.f). Recall from 14.22 that other terms are defined recursively. If

$$t_1, t_2, \ldots, t_n \text{ are terms interpreted to have values } |t_1|, |t_2|, \ldots, |t_n| \text{ in } D,$$

and f is an n-ary function symbol with interpretation $|f| : D^n \to D$, then

$$f(t_1, t_2, \ldots, t_n) \text{ is interpreted to have value } |f|\Big(|t_1|, |t_2|, \ldots, |t_n|\Big) \text{ in } D$$

— it is the composition of the function $|f| : D^n \to D$ and the functions $|t_j|$.

For instance, suppose a_1, a_2, a_3 are constant symbols, f_1, f_2 are function symbols with arity 3, and g is a function symbol with arity 4; assume some interpretation is given for each of these symbols. Let x_1, x_2, x_3, x_4 be any free variable symbols. Then

$$g\left(f_1(x_1, a_1, a_2),\ f_2(x_2, a_3, x_3),\ x_3,\ f_1(x_1, x_4, x_2)\right)$$

is interpreted as a function from D^4 into D, since it yields particular values in D when particular values are substituted for x_1, x_2, x_3, x_4.

For more concrete examples we turn to arithmetic. If "3," "5," and "+" have their usual interpretations, then the term "$3 + 5$" will be interpreted to have value 8. The term "$x + 5$" will be interpreted as a function from D into D (where D is $\mathbb{N}$ or $\mathbb{Z}$ or whatever); this function takes numerical values when particular numbers are substituted for x.

h. An *atomic formula* $P(t_1, t_2, \ldots, t_n)$ (defined as in 14.23(i)) is interpreted by composing the function $|P|$ (introduced in 14.47.e) with the functions $|t_1|, |t_2|, \ldots, |t_n|$. Thus, it is interpreted as a function from D^m into B, where m is the number of distinct free variables appearing in the atomic formula.

For instance, in ordinary arithmetic, the formula $3 + 5 < 10$ is interpreted to have the constant value "true;" the formula $3 + 5 < x$ is interpreted as a mapping from the integers or the real numbers or other domain D, to the set {"true," "false"}.

i. The logical connectives $\neg, \sqcup, \sqcap, \to$ are mapped to the corresponding fundamental operations of the Boolean algebra B. *Formulas* in the formal language are now interpreted recursively, in the obvious fashion. If $\mathcal{A}$ and $\mathcal{B}$ are formulas, then

$$\begin{aligned}
|\neg\mathcal{A}| &= (\complement|\mathcal{A}|),\\
|\mathcal{A} \sqcup \mathcal{B}| &= (|\mathcal{A}| \vee |\mathcal{B}|),\\
|\mathcal{A} \sqcap \mathcal{B}| &= (|\mathcal{A}| \wedge |\mathcal{B}|),\\
|\mathcal{A} \to \mathcal{B}| &= (|\mathcal{A}| \Rightarrow |\mathcal{B}|).
\end{aligned}$$

In each of these equations, the connective on the left side is the formal, logical symbol; the connective on the right side is a unary or binary operation in the Boolean algebra B. In general, a formula with n distinct free variables is interpreted as a mapping from D^n into B.

j. *Quasi-interpretation of quantifiers.* If the language has infinitely many free variable symbols, then quantifiers are interpreted as suprema and infima in the Boolean algebra B, as follows:

For simplicity we explain quantifiers first in the case of a formula involving only one free variable. Suppose x is the only free variable occurring in $\mathcal{A}(x)$, and ξ does not occur in $\mathcal{A}(x)$; we follow the substitution notation of 14.21. Then $x \mapsto |\mathcal{A}(x)|$ is a function from D into B, which takes some truth value whenever x is replaced by some $d \in D$ (whether that d has a name or not). Then we define the quasi-interpretations

$$|\exists_\xi\ \mathcal{A}(\xi)| \;=\; \sup_{d\in D} |\mathcal{A}(d)|\,, \qquad\qquad |\forall_\xi\ \mathcal{A}(\xi)| \;=\; \inf_{d\in D} |\mathcal{A}(d)|\,.$$

The sup and inf are with respect to the ordering of the Boolean lattice B.

More generally, assume ξ is a bound variable, $x_1, x_2, \ldots, x_n$ are distinct free variables, $\mathcal{A}(x_1, x_2, \ldots, x_n)$ is a formula whose only free variables are $x_1, x_2, \ldots, x_n$, and ξ does not appear in $\mathcal{A}(x_1, x_2, \ldots, x_n)$. Then $(x_1, x_2, \ldots, x_n) \mapsto |\mathcal{A}(x_1, x_2, \ldots, x_n)|$ is a function from D^n into B, taking a truth value whenever $x_1, x_2, \ldots, x_n$ are replaced by some $d_1, d_2, \ldots, d_n \in D$ (whether those d_j's have names or not). Now we define the quasi-interpretations

$$\begin{aligned} |\exists_\xi\ \mathcal{A}(\xi, x_2, x_3, \ldots, x_n)| &= \sup_{d\in D} |\mathcal{A}(d, x_2, x_3, \ldots, x_n)|\,, \\ |\forall_\xi\ \mathcal{A}(\xi, x_2, x_3, \ldots, x_n)| &= \inf_{d\in D} |\mathcal{A}(d, x_2, x_3, \ldots, x_n)|\,; \end{aligned}$$

these are functions from D^{n-1} into B.

To make sense of this definition, we require that

($\natural$) all sups and infs of the types indicated above must exist in B.

One simple way to satisfy ($\natural$) is by insisting that B be a complete Boolean algebra — i.e., by requiring that *every* subset of B have a sup and an inf. However, that simple requirement may be too strong in some cases — see 14.56 — so we merely keep it in mind for motivation; ($\natural$) is the one condition we shall actually impose as part of our definition of "quasi-interpretation."

Caution: When B is the two-element Boolean algebra $\{0, 1\}$, then each of those sups or infs indicated above is actually a maximum or a minimum, and so $\forall$ and $\exists$ actually *do* have the meanings "*for each* ... *in* D ..." and "*there exists* ... *in* D *such that* ...." However, when B is a larger Boolean algebra, then the sups and infs need not be maxima or minima. It is quite possible that $\sup_{d\in D} |\mathcal{A}(d, x_2, x_3, \ldots, x_n)|$ is equal to 1, and yet no one of the elements $d \in D$ actually satisfies $|\mathcal{A}(d, x_2, x_3, \ldots, x_n)| = 1$. (Thus, in the terminology of 14.46 and 14.48, an existential formula may be valid and yet not have a witness.) The possible lack of a suitable d is the cause of some of the complications in the pages that follow — e.g., this is why 14.46 is needed as a step in the proof of the Completeness Principle.

14.48. *More terminology.* Let $(B, D, |\ |)$ be a quasi-interpretation of some language $\mathfrak{L}$, and let $\mathcal{F}$ be some formula in that language with n free variables. Then $|\mathcal{F}|$ is a function from D^n into B. It is a constant function (i.e., a member of B) if $n = 0$; it may or may not be a constant function otherwise.) We say that the formula $\mathcal{F}$ is **valid** in the quasi-interpretation if that function $|\mathcal{F}|$ is a constant function that takes only the value 1.

Observation. From the definition given in 14.47.j for the quasi-interpretation of quantifiers and the definition of closures given in 14.44, we see immediately that *in any quasi-interpretation, the closure of any valid formula is valid.*

Example. Consider quasi-interpretations with domain D equal to the set $\mathbb{Z}$ of integers. Let $\mathcal{F}(x)$ be a formula that states (in some appropriate symbolism) that "x is even;" then $\neg(\mathcal{F}(x))$ states that "x is odd." When the language of arithmetic is given its usual interpretation, then the formula $(\mathcal{F}(x)) \sqcup (\neg(\mathcal{F}(x)))$ is valid, since every integer is either even or odd; this is an instance of the Law of the Excluded Middle. However, neither of the formulas $\mathcal{F}(x)$ or $\neg(\mathcal{F}(x))$ is valid, since neither of the statements "every integer is even" or "every integer is odd" is correct.

14.49. *Definition.* Let V be the set of all free variable symbols in a language $\mathfrak{L}$, and let $(B, D, |\ |)$ be a quasi-interpretation of $\mathfrak{L}$. By a **valuation** (or *assignment*) on $(B, D, |\ |)$ we shall mean a map $\psi : V \to D$; thus it is just a member of D^V. We shall denote by $|\ |_\psi$ the effect of combining a quasi-interpretation $|\ |$ and a valuation ψ. Expressions are interpreted with values in B, as in 14.47–14.47.j, but in addition each free variable v is replaced by its valuation $\psi(v) \in D$. Thus, all the functions get evaluated. For any formula $\mathcal{F}$ — even one involving free variables — the result $|\mathcal{F}|_\psi$ is a particular member of the Boolean algebra B, not just a function from D^n into B. We emphasize that $|\mathcal{F}|_\psi$ is not necessarily 0 or 1; it may be some other member of B. Observe that

> a formula $\mathcal{F}$ is valid in $|\ |$, as defined in 14.48, if and only if it satisfies $|\mathcal{F}|_\psi = 1$ for every valuation ψ.

14.50. *More terminology.* Let $(B, D, |\ |)$ be a quasi-interpretation of the language, and let Σ be a collection of formulas in that language. We say that $(B, D, |\ |)$ is a **quasimodel of the theory Σ** if every formula in Σ is valid in $(B, D, |\ |)$.

A two-valued quasimodel of a theory Σ (i.e., a quasimodel in which $B = \{0, 1\}$) will be called a **model** of Σ.

Alternate terminology. Instead of "quasi-interpretation" or "quasimodel," some mathematicians use the terms **Boolean-valued interpretation** or a **Boolean-valued model**. We prefer not to use those terms, for this reason: Following common nonmathematical English usage, those terms would appear to refer to notions that are *less* general than "interpretation" or "model." Our prefix of "quasi-" suggests *greater* generality, and is therefore more descriptive.

Also, Rasiowa and Sikorski [1963] use the term "*realization*" for both quasi-interpretations and quasimodels, but perhaps that is less helpful to the beginner's intuition.

A few mathematicians use the term "model" where we have used the term "valuation." This changes the nature of the theory, but not by very much if (as in some books) we do not distinguish between constants and variables.

14.51. Quasi-interpretations of propositional logic. The ingredients of a quasi-interpretation can be simplified substantially when we work with propositional logic — i.e., when there are no quantifiers, individual constants, individual variables, or relation symbols of arity greater than 0. In this special case we can take the domain D to be empty.

For a theory in propositional logic, a *quasi-interpretation* means an assignment of some truth value for each primitive proposition symbol P; we may denote that assignment by $|P|$. Thus, we only need to specify a mapping $|\ |$ as in 14.47.e, and only for $n = 0$. After that, the quasi-interpretation recursively assigns a true value to compound propositions, as in 14.47.i.

Note that if each primitive proposition is true or false — i.e., if $|P| \in \{0,1\}$ for each primitive proposition symbol P — then in fact $|\mathcal{F}| \in \{0,1\}$ for each formula $\mathcal{F}$; this follows by induction on the lengths or depths of formulas. In this case, the resulting quasi-interpretation is in fact an interpretation.

14.52. *Example.* **Peano arithmetic** uses a constant symbol "u" (the "unit" or "urelement"), a unary function "σ" (the **successor function**), the binary relation "=" with the axioms for equality (listed in 14.27.a), plus these three further axioms:

(i) $\neg\,(\exists_\xi\ (\sigma(\xi) = u))$. That is, u is not the successor of any number.

(ii) $((\sigma(x) = \sigma(y)) \to (x = y))$. That is, σ is injective.

(iii) The **Induction Axiom**. If S is a subset of the domain that satisfies $u \in S$ and also satisfies $((x \in S) \to (\sigma(x) \in S))$, then $S = D$.

A few models of Peano's Axioms are given by:

- $D = \mathbb{N} = \{1, 2, 3, \ldots\},\ \ u = 1,\ \ \sigma(x) = x + 1;$
- $D = \mathbb{N} \sqcup \{0\},\ \ u = 0,\ \ \sigma(x) = x + 1;$
- $D = 2\mathbb{N} = \{2n : n \in \mathbb{N}\},\ \ u = 2,\ \ \sigma(x) = x + 2.$

It is not difficult to manufacture more models of Peano's Axioms. However, all these models are *isomorphic* — i.e., it can be shown that if (D, u, σ) is any model of Peano's Axioms, then there is a unique bijection $b : D \to \mathbb{N}$ such that $b(u) = 1$ and $b(\sigma(x)) = 1 + b(x)$. Thus, Peano's Axioms determine $\mathbb{N}$ uniquely up to isomorphism.

Peano's first two axioms fit into a first-order language, but the last axiom requires a higher-order language, since it quantifies over sets $S \subseteq D$. In a first-order language, we have no precise representation of the notion of "a subset of D."

For most purposes, we can replace axiom (iii) with a *scheme* of infinitely many first-order axioms: For each property $P(x)$ that can be expressed in first-order language, we have an axiom

$$(\mathrm{iii})_P \qquad ((P(u) \sqcap (P(x) \to P(\sigma(x)))) \to (\forall_\xi\ P(\xi)).$$

This axiom scheme is slightly weaker than Peano's Axiom (iii). One way to see that fact is to note that if we only have finitely or countably many symbols in our language, then there are only countably many properties P that can be expressed in the language, but $\mathbb{N}$ has uncountably many subsets S for us to consider. For another demonstration that the first-order axiom scheme $(\mathrm{iii})_P$ is weaker than (iii), see 14.63, where we shall show that no system of first-order properties of $\mathbb{N}$ can uniquely determine $\mathbb{N}$ up to isomorphism.

14.53. A brief introduction to forcing (optional). Cohen's method of forcing is a technique for creating models and quasimodels, particularly of set theory. Our presentation below is based on Bell [1985].

Let B be a complete Boolean algebra. We shall describe classes $V^{(B)}$ and $V^{(\Gamma)}$, which can be used for the domains for quasimodels of set theory, taking truth values in B.

The **Boolean-valued universe** $V^{(B)}$, will be defined recursively, in a fashion somewhat analogous to the construction of the von Neumann universe V in 5.53, but with this difference: When we ask whether $x \in y$ and whether $x = y$, the answers are not necessarily members of the Boolean algebra $2 = \{0, 1\} = \{\text{"no," "yes"}\}$; rather, the answers may be members of the Boolean algebra B. More precisely,

> for each ordinal α, let $V_\alpha^{(B)}$ be the set of all B-valued functions x that have $\mathrm{Dom}(x) \subseteq V_\beta^{(B)}$ for some ordinal $\beta < \alpha$;

then let $V^{(B)}$ be the union of all the $V_\alpha^{(B)}$'s.

Truth values in this quasi-interpretation are defined recursively too. The language of set theory expresses everything — ordered pairs, the integers, functions, etc. — in terms of set membership, so in our formal language we can dispense with function symbols and with most relation symbols; we only need the two relation symbols $\in$ and $=$. A term is a constant or a free variable; an atomic formula is an expression of the form $s \in t$ or $s = t$ where s, t are terms. Truth values of atomic formulas are defined thus:

$$|u \in v| \quad = \quad \sup_{y \in \mathrm{Dom}(v)} \left(v(y) \wedge |u = y| \right),$$

$$|u = v| \;=\; \left(\sup_{x \in \mathrm{Dom}(u)} \left(u(x) \Rightarrow |x \in v| \right) \right) \wedge \left(\sup_{y \in \mathrm{Dom}(v)} \left(v(y) \Rightarrow |y \in u| \right) \right).$$

Other formulas are built from atomic formulas and evaluated in a fashion similar to that in 14.47.i, 14.47.j.

With these evaluations, $V^{(B)}$ is a quasimodel of conventional set theory, ZF + AC (if we assume ZF + AC in the outer system). Making different choices of the complete Boolean algebra B yields different additional properties of $V^{(B)}$, and hence various semantic consistency results. For instance, with a suitable choice of B, $V^{(B)}$ does not satisfy the Continuum Hypothesis; therefore

$$\mathrm{Con}(\mathrm{ZF}) \quad \Rightarrow \quad \mathrm{Con}(\mathrm{ZF} + \mathrm{AC} + \neg\mathrm{CH}).$$

However, all the quasi-interpretations constructed in the fashion above will satisfy ZF + AC. To get negations of AC, we need a more complicated construction, based on automorphisms of B.

An **automorphism** of B is a Boolean isomorphism $g : B \to B$ — i.e., a Boolean homomorphism that is also a permutation of B. The automorphisms of B form a group, $\mathrm{Aut}(B)$, with group operation given by the composition of functions.

An automorphism $g : B \to B$ can be extended naturally to a map $g^{(B)} : V^{(B)} \to V^{(B)}$ recursively by this rule: Whenever $u \in V^{(B)}$ with domain $\mathrm{Dom}(u)$, then $g^{(B)}u$ is the member of $V^{(B)}$ that has $\mathrm{Dom}(g^{(B)}u) = \{g^{(B)}x : x \in \mathrm{Dom}(u)\}$ and is defined on that domain by $(g^{(B)}u)(g^{(B)}x) = g(u(x))$. (That last g is just the original mapping from B into B.) It is not hard to verify that the map $g \mapsto g^{(B)}$ is a group homomorphism; that is, it preserves compositions: $(gh)^{(B)} = g^{(B)}h^{(B)}$.

Let G be a subgroup of $\mathrm{Aut}(B)$. For each $x \in V^{(B)}$, define the **stabilizer group**

$$\mathrm{stab}_G(x) \quad = \quad \{g \in G \ : \ g(x) = x\};$$

it is a subgroup of G.

Now let Γ be a collection of subgroups of $\mathrm{Aut}(B)$. We now recursively define the Boolean-valued universe $V^{(\Gamma)}$, a subclass of $V^{(B)}$, as follows:

> For each ordinal α, let $V_\alpha^{(\Gamma)}$ be the set of all B-valued functions x that have $\mathrm{Dom}(x) \subseteq V_\beta^{(\Gamma)}$ for some ordinal $\beta < \alpha$ and satisfy $\mathrm{stab}_G(x) \in \Gamma$;

then let $V^{(\Gamma)}$ be the union of all the $V_\alpha^{(\Gamma)}$'s.

Truth values can be defined on $V^{(\Gamma)}$ just as they were defined on $V^{(B)}$. Certain choices of G and Γ yield quasimodels of certain set theories. For instance, Bell [1985] shows a quasimodel of this sort in which a set is infinite but Dedekind finite (see 6.27); hence the axiom of Countable Choice is not satisfied. Thus

$$\mathrm{Con(ZF)} \quad \Rightarrow \quad \mathrm{Con(ZF} + \neg\mathrm{CC)}.$$

The omitted details are very large and numerous, and are not intended as an exercise. The interested reader should consult Bell [1985] and other books on forcing.

The main ideas of forcing can be reformulated in syntactic terms. Let P be a suitable subset of $B \setminus \{0\}$. For $p \in P$ and formulas $\mathcal{A}$, let $p \Vdash \mathcal{A}$ be an abbreviation for $p \preccurlyeq |\mathcal{A}|$, where $|\ |$ is the truth-value mapping and $\preccurlyeq$ is the ordering of the Boolean algebra B; then p is called a "forcing condition." The basic properties of the Boolean-valued universe $V^{(\Gamma)}$ can be reformulated as properties of the forcing relation $\Vdash$. In fact, it is possible to study $\Vdash$ without referring to Boolean-valued universes. This approach is more difficult for newcomers to logic and will not be explained here, but it seems to be preferred by logicians — they find it more intuitive than the Boolean-valued approach. This is the approach originally used by Cohen. The approach via Boolean-valued universes is a later reformulation, due largely to Scott and Solovay.

Historical note: The interested reader may search in vain for an important paper of Scott and Solovay, often referenced as "to appear." That work actually did not appear. It is subsumed by Bell [1985], as explained in Scott's foreword in that book.

Soundness, Completeness, and Compactness

14.54. *Observation.* Any first-order language (as described in 14.15–14.23) has at least one interpretation (as described in 14.47–14.50).

Proof. Here is one trivial construction: Let $D = \{0\}$, where "0" is some object — i.e., let D be a singleton. Interpret every constant symbol to have value 0; interpret all the relation symbols to only take the value "true."

14.55. *Proposition.* Every quasi-interpretation of the language $\mathcal{L}$ is also a quasimodel of the logic. That is, if $(B, D, |\ |)$ is a quasi-interpretation of a first-order language $\mathcal{L}$, then

(i) each of the twelve logical axioms listed in 14.25 is valid in $|\ |$; and

(ii) each of the six rules of inference listed in 14.26 is valid in the following sense: Whenever $\mathcal{E}$ and $\mathcal{F}$ are formulas that are valid in $|\ |$ and $\mathcal{G}$ is a formula that can be deduced from $\mathcal{E}$ and $\mathcal{F}$ using one of the rules of inference, then $\mathcal{G}$ is also valid in $|\ |$.

Using the two preceding results plus an induction argument, it follows that

(iii) If Σ is any given set of extra-logical axioms and $\mathcal{A}$ is a formula that can be deduced from Σ and the logical axioms via the rules of inference, then $\mathcal{A}$ is valid in every quasimodel of Σ.

Since every model is a quasimodel, as a corollary we obtain this slightly weaker result:

(iv) **The Soundness Principle.** If Σ is any given set of extra-logical axioms and $\mathcal{A}$ is a formula that can be deduced from Σ and the logical axioms via the rules of inference, then $\mathcal{A}$ is valid in every model of Σ. In other words,

$$\text{if } \Sigma \vdash \mathcal{A}, \text{ then } \Sigma \models \mathcal{A}.$$

In other words, every syntactic theorem is a semantic theorem.

Remark. Some mathematicians use the term "theorem" only for syntactic theorems and call a formula "true" if it is valid in every model. With that terminology, the Soundness Principle takes this more memorable form: *Every theorem is true.*

Proof. We first consider the validity of the twelve logical axioms. We shall demonstrate validity only for Axiom (ii); the other axioms can be verified in a similar fashion and are left as exercises. Let $|\ |$ be some quasi-interpretation of the language $\mathcal{L}$; we wish to prove that $|\mathcal{A} \to (\mathcal{A} \sqcup \mathcal{B})|_\psi = 1$ for every valuation ψ. By 14.47.i, that condition can be restated as $(|\mathcal{A}|_\psi \Rightarrow (|\mathcal{A}|_\psi \vee |\mathcal{B}|_\psi)) = 1$. But $(\alpha \Rightarrow (\alpha \vee \beta)) = 1$ is true for any elements α, β in any Boolean algebra B. This proves the validity of Axiom (ii).

Next, we consider the validity of the inference rules. We shall verify this only for (R3) and (R5); verification of the other inference rules is left as an exercise. We assume ξ is a bound variable that does not occur in the formula $\mathcal{A}(x)$; we follow the substitution notation of 14.21. Let $\mathcal{B}$ be some formula; for (R3) we also assume that x does not occur in $\mathcal{B}$. Let

$|\,|$ be a given quasi-interpretation of the language. The conditions "$\mathcal{A}(x) \to \mathcal{B}$ is valid" and "$(\exists_\xi\, \mathcal{A}(\xi)) \to \mathcal{B}$ is valid" can be restated, respectively, as

$$|\mathcal{A}(x)|_\varphi \preccurlyeq |\mathcal{B}|_\varphi \text{ for every valuation } \varphi, \tag{1}$$

$$\left|\exists_\xi\, \mathcal{A}(\xi)\right|_\psi \preccurlyeq |\mathcal{B}|_\psi \text{ for every valuation } \psi. \tag{2}$$

If ψ is any given valuation, for each $d \in D$ we may define an auxiliary valuation by

$$\psi_d(v) \quad = \quad \begin{cases} \psi(v) & \text{when } v \neq x \\ d & \text{when } v = x. \end{cases}$$

From the definition in 14.47.j we see that $\left|\exists_\xi\, \mathcal{A}(\xi)\right|_\psi = \sup_{d\in D} |\mathcal{A}(x)|_{\psi_d}$. Thus (2) can be restated

$$|\mathcal{A}(x)|_{\psi_d} \preccurlyeq |\mathcal{B}|_\psi \text{ for every } \psi, |\,|, \text{ and } d. \tag{2$'$}$$

To verify (R3), we need to show that (1) implies (2$'$). Since x does not occur in the formula $\mathcal{B}$, we find that $|\mathcal{B}|_{\psi_d} = |\mathcal{B}|_\psi$ for every d; hence $|\mathcal{A}(x)|_{\psi_d} \preccurlyeq |\mathcal{B}|_{\psi_d} = |\mathcal{B}|_\psi$. To verify (R5), we need to show that (2$'$) implies (1); just observe that when $\delta = \psi(x)$, then $\psi_\delta = \psi$.

14.56. *Observation.* Let Σ be a syntactically consistent set of formulas. Then Σ has a quasimodel. In fact, one can be specified as follows:

For domain D use the set of all terms in the language, with the interpretation mapping $|\,|$ defined on terms by the identity mapping. For the Boolean algebra of truth values use the Lindenbaum algebra $\mathbb{L}$, with the interpretation mapping $|\,|$ defined on formulas by the equivalence class mapping $[\,]$ defined in 14.32.

Remark. We do *not* assert that the Lindenbaum algebra is necessarily complete. The fact that it satisfies condition 14.47.j($\natural$) follows from 14.41.

14.57. We shall show that the following two principles (and two more covered in 14.59) are equivalent to the Ultrafilter Principle; we refer especially to other equivalents in 13.22.

(UF11) for Propositional Logic:
(UF12) for Predicate Logic:

Gödel-Mal'cev Completeness Principle (consistency version). If Σ is any set of formulas, then these three conditions are equivalent:

(A) Σ is syntactically consistent — i.e., Σ cannot be used to deduce a contradiction.

(B) Σ is semantically consistent — i.e., Σ has at least one model.

(C) Σ has at least one quasimodel.

As an intermediate step between (UF11) and (UF12), we shall also prove the equivalence of this more complicated principle:

(UF13) Let $\mathcal{L}$ be a language that has no variable symbols and no quantifiers (but may still have constants and functions). Let Σ be a set of formulas in $\mathcal{L}$ that is syntactically consistent in $\mathcal{L}$. Then Σ has at least one model $(\{0,1\}, D, |\ |)$. Furthermore, the model can be chosen so that the mapping $|\ | : \{\text{terms of } \mathcal{L}\} \to D$ is surjective — i.e., so that for each individual $d \in D$ there is at least one term t satisfying $|t| = d$.

Proof. The implication (C) $\Rightarrow$ (A) is proved as follows: Suppose Σ is not syntactically consistent. Then there is some formula $\mathcal{A}$ such that both $\mathcal{A}$ and $\neg\mathcal{A}$ are syntactic theorems. If $|\ |$ is a quasimodel of Σ, then it makes both $\mathcal{A}$ and $\neg\mathcal{A}$ valid — that is, $|\mathcal{A}|$ and $|\neg\mathcal{A}|$ are both equal to the constant function 1. Then $1 = |\neg\mathcal{A}| = \complement|\mathcal{A}| = \complement 1 = 0$. Thus the Boolean algebra B is degenerate, contrary to the requirement in 14.47.a. This shows (C) $\Rightarrow$ (A).

The implication (B) $\Rightarrow$ (C) is trivial, since every model is a quasimodel. It only remains to prove (A) $\Rightarrow$ (B). (That implication by itself is sometimes known as the Completeness Principle.)

Proof of (UF8) $\Rightarrow$ (UF11). As we noted in 14.56, the Lindenbaum algebra $\mathbb{L}$ is a quasimodel of Σ. By (UF8), there exists a Boolean homomorphism from $\mathbb{L}$ into $\{0,1\}$. Use that homomorphism to map the truth values in $\mathbb{L}$ to truth values in $\{0,1\}$. The homomorphism preserves the action of $\vee, \wedge, \complement, \Rightarrow$. We need not concern ourselves with $\exists, \forall$ since we are considering only propositional logic, which has no individuals or quantifiers. Thus the resulting map into $\{0,1\}$ is a model of Σ.

Proof of (UF11) $\Rightarrow$ (UF13). Let $\mathcal{C}$ be the given predicate calculus — i.e., the language $\mathcal{L}$ equipped with the logical axioms and rules of inference, the given extra-logical axioms Σ, and the resulting syntactic theorems. To construct a model, we shall first form a related language $\widehat{\mathcal{L}}$ and propositional calculus $\widehat{\mathcal{C}}$; a model for that propositional calculus will be used to form a model for $\mathcal{C}$.

Form a language $\widehat{\mathcal{L}}$ by taking each atomic formula of $\mathcal{L}$ as a primitive propositional variable symbol of $\widehat{\mathcal{L}}$. Thus, an expression such as $P(f(a,b), g(c))$ will be treated as a single symbol, grammatically on the same level as Q or R. The terms $f(a,b)$ and $g(c)$ and the constants a, b, c play no role in $\widehat{\mathcal{L}}$, except as meaningless marks on paper that serve to make up *parts* of that single symbol. The language $\widehat{\mathcal{L}}$ will have no individual variable symbols, individual constant symbols, or functions. It will have the same logical connective symbols $\neg, \sqcap, \sqcup, \to$ as the language $\mathcal{L}$. Each formula in either of the languages $\mathcal{L}$ or $\widehat{\mathcal{L}}$ can be reinterpreted as a formula in the other language by reading it in a different fashion. For instance, in the original language $\mathcal{L}$, the expression $P \sqcup Q(a, b, g(c))$ consists of seven symbols

$$P \qquad \sqcup \qquad Q \qquad a \qquad b \qquad g \qquad c$$

joined together with commas, parentheses, and juxtapositions; but in the new language $\widehat{\mathcal{L}}$ the expression $P \sqcup Q(a, b, g(c))$ consists of just the three symbols

$$P \qquad\qquad \sqcup \qquad\qquad Q(a, b, g(c))$$

joined together with juxtaposition.

Form a new propositional calculus $\widehat{\mathcal{C}}$ using the language $\widehat{\mathcal{L}}$ and the same set Σ of extra-logical axioms (but read in a different fashion, as noted above). Since $\mathcal{L}$ has no variable symbols or quantifiers, rules of inference (R2) through (R6) are irrelevant; thus both $\mathcal{C}$ and $\widehat{\mathcal{C}}$ have *modus ponens* as their only rule of inference. Therefore, proofs in the two systems are identical in appearance (though read differently). By assumption, $\mathcal{C}$ is syntactically consistent; therefore $\widehat{\mathcal{C}}$ is, too. By (UF11), $\widehat{\mathcal{C}}$ has a model, as described in 14.51. That is, there exists a mapping

$$| \ | \quad : \quad \{\text{formulas of } \widehat{\mathcal{L}}\} \quad \longrightarrow \quad \{\text{"true," "false"}\}$$

defined on the primitive proposition symbols and then defined recursively on other formulas, in such a way that all the axioms of $\widehat{\mathcal{C}}$ become true.

Next, let D be the set of all *terms* that can be formed in the language $\mathcal{L}$, as defined in 14.22 — i.e., expressions such as $f(a,b)$ and $g(c)$. We shall now construct a model for $\mathcal{C}$ whose domain is the set D. To do that, we must describe an interpretation mapping $| \ |$ that can be applied to constant symbols, to function symbols, and to relation symbols, as explained in 14.47.

For terms, the mapping $| \ |$ will just be the identity mapping. In other words, any term t in $\mathcal{L}$ is a string of symbols that is a single element d of D; we interpret $|t| = d$.

Next we shall interpret atomic formulas: If P is an n-ary predicate symbol in the language $\mathcal{L}$, and $t_1, t_2, \ldots, t_n$ are terms, then the atomic formula $P(t_1, t_2, \ldots, t_n)$ in $\mathcal{L}$ will be given the same truth value (1 or 0) that it had when we viewed it as a primitive proposition symbol in the propositional calculus $\widehat{\mathcal{C}}$ introduced a few paragraphs ago.

Finally, we recursively assign truth values for compound propositions, as in 14.47.i. Thus we obtain an interpretation of $\mathcal{L}$, which is in fact a model of Σ.

Proof of (UF13) $\Rightarrow$ (UF12). Let Σ be a syntactically consistent set of axioms; we wish to prove that Σ has a model. By repeated use of the Rule of Generalization 14.43, we may replace all the members of Σ with closed formulas — i.e., formulas with no free variables. Then, replacing members of Σ by equivalent formulas (where equivalence is as in 14.32), by 14.42 we may assume that each axiom in Σ is in prenex normal form — i.e., with all the quantifiers at the beginning of the formula.

Let $\mathcal{T}_0$ represent the given logical system — i.e., the given language and syntactically consistent set of axioms. Form new, syntactically consistent systems $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3, \ldots$ recursively; obtain $\mathcal{T}_{n+1}$ from $\mathcal{T}_n$ by adding new axioms and new constant symbols, using the construction given in 14.46; the axioms added in this fashion are also without free variables.

Let $\mathcal{T}_\infty = \bigcup_{n=0}^{\infty} \mathcal{T}_n$, in the obvious sense — i.e., let $\mathcal{T}_\infty$ be the original system $\mathcal{T}_0$ plus all the additional axioms and constant symbols of the $\mathcal{T}_n$'s. Clearly, $\mathcal{T}_\infty$ is also syntactically consistent, by 14.31. Let $\mathcal{U}$ be the subset of $\mathcal{T}_\infty$ consisting of those statements that do not contain any quantifiers. Now let $\mathcal{M} = (\{0,1\}, D, | \ |)$ be a model for $\mathcal{U}$, of the type described in (UF13) — i.e., with each individual named by at least one term. We shall show that $\mathcal{M}$ is also a model for $\mathcal{T}_\infty$ (and hence also for our original system $\mathcal{T} = \mathcal{T}_0$). It suffices to show, by induction on integers $k \geq 0$, that

> if $\mathcal{F}$ is an axiom of $\mathcal{T}_\infty$ in which k or fewer quantifiers appear, then $\mathcal{F}$ is valid in the interpretation $\mathcal{M}$.

This is clear for $k = 0$, since such axioms are just the axioms of $\mathcal{U}$. Suppose it is true for some k, and let $\mathcal{F}$ be an axiom of $\mathcal{T}_\infty$ involving $k+1$ quantifiers; we shall show that this $\mathcal{F}$ is also valid in $\mathcal{M}$. There are two cases to consider: $\mathcal{F}$ is either of the form $\forall_\xi\ \mathcal{A}(\xi)$ or of the form $\exists_\xi\ \mathcal{A}(\xi)$, for some formula $\mathcal{A}$. For these two cases we refer again to the construction in 14.46.

Case (i). $\mathcal{F}$ is of the form $\forall_\xi\ \mathcal{A}(\xi)$. Thus $\forall_\xi\ \mathcal{A}(\xi)$ is an axiom in $\mathcal{T}_\infty$, hence in $\mathcal{T}_n$ for all integers n sufficiently large — say for all $n \geq j$. If t is any term in the language of $\mathcal{T}_\infty$, then t is a term in $\mathcal{T}_n$ for some $n \geq j$, and so (by our construction of $\mathcal{T}_{n+1}$ from $\mathcal{T}_n$) we know that the statement $\mathcal{A}(t)$ is an axiom of $\mathcal{T}_{n+1}$, hence of $\mathcal{T}_\infty$. The statement $\mathcal{A}(t)$ involves only k quantifiers, hence it is valid in $\mathcal{M}$. Thus, $1 = |\mathcal{A}(t)| = |\mathcal{A}|(|t|) = |\mathcal{A}|(d)$, where $d = |t|$. By our assumption about the model $\mathcal{M}$ in (UF13), the mapping $|\ |$ takes terms onto individuals; thus $|\mathcal{A}|(d) = 1$ for every individual d in the domain D. By the definition in 14.47.j, therefore, $|\forall_\xi\ \mathcal{A}(\xi)| = 1$, so $\mathcal{F}$ is valid.

Case (ii). $\mathcal{F}$ is of the form $\exists_\xi\ \mathcal{A}(\xi)$. Then $\exists_\xi\ \mathcal{A}(\xi)$ is an axiom of $\mathcal{T}_n$ for some n. By our construction of $\mathcal{T}_{n+1}$ from $\mathcal{T}_n$, the axiom has a witness — i.e., there is some constant symbol c in $\mathcal{T}_{n+1}$ such that $\mathcal{A}(c)$ is an axiom of $\mathcal{T}_{n+1}$. Now $\mathcal{A}(c)$ is an axiom of $\mathcal{T}_\infty$ involving only k quantifiers, so it is valid in $\mathcal{M}$. Therefore $\exists_\xi\ \mathcal{A}(\xi)$ is valid in $\mathcal{M}$ by the definition of $|\exists_\xi\ \mathcal{A}(\xi)|$ in 14.47.j.

This completes the proof.

14.58. *Remarks.* The earliest version of the completeness principle was due to Gödel, so it is sometimes known as the **Gödel Completeness Theorem**. It should not be confused with Gödel's *Incompleteness* Theorems, introduced in 14.62 and 14.70. In mathematics, the term "complete" generally means "not missing any parts," or "not having any holes in it" — see 4.14. Predicate logic is complete in some respects, but incomplete in other respects.

The equivalence of the completeness principles with other forms of UF was proved by Rasiowa and Sikorski [1951], Loś [1954], and Henkin [1954]. Our exposition is based on Cohen [1966] and several other works.

With a bit more work (not shown in detail here), our proof of (UF12) can be modified to show a slightly stronger principle:

> If Σ is a syntactically consistent set of formulas, then Σ has a model whose domain D satisfies $\mathrm{card}(D) \leq \max\{\mathrm{card}(\Sigma), \mathrm{card}(\mathbb{N})\}$.

In particular, *any syntactically consistent first-order theory with a countable language has a countable model.* Most languages used in practice are countable (i.e., have only countably many symbols).

As we remarked in 14.6, it is possible to form a model of set theory by replacing von Neumann's universe V with some other class $\mathcal{M}$ of sets, which may be smaller. The class $\mathcal{M}$ need not be a proper class — it may be a set. In fact, it may even be a *countable* set, since set theory can be described with a countable language. Thus arises a situation which, at first, seems paradoxical: Set theory describes and proves the existence of various uncountable objects, and yet some of the sets that can be used as domains for models of set theory are countable! This is known as **Skolem's Paradox**. To understand this, we must distinguish between the "inner" and "outer" systems, described in 14.12. The set D may be countable in the outer system, but not in the inner system — i.e., there may be a bijection between D and $\mathbb{N}$ in the informal, outer system that we use to analyze the model, but there may be no such function in the formal, inner system.

14.59. Here is another form of the Completeness Principle:

> **(UF14) for Propositional Logic:**
> **(UF15) for Predicate Logic:**
>
> **Completeness Principle (theorems version).** Let Σ be a collection of axioms, and let $\mathcal{A}$ be a formula. Then the following are equivalent:
>
> (A) $\mathcal{A}$ is a syntactic theorem. That is, $\Sigma \vdash \mathcal{A}$.
>
> (B) $\mathcal{A}$ is a semantic theorem. That is, $\Sigma \models \mathcal{A}$. That is, in every model of Σ, the formula $\mathcal{A}$ is also valid.
>
> (C) In every quasimodel of Σ, the formula $\mathcal{A}$ is also valid.

Proof that the "consistency versions" (UF11) *and* (UF12) *imply the "theorem versions"* (UF14) *and* (UF15), *respectively.* The implication (A) $\Rightarrow$ (C) was given in 14.55(iii). The implication (C) $\Rightarrow$ (B) is trivial, since every model of Σ is a quasimodel of Σ. It suffices to prove (B) $\Rightarrow$ (A).

Let $\mathcal{K}$ be the closure of $\mathcal{A}$ (defined as in 14.44). In every model of Σ, since $\mathcal{A}$ is valid, $\mathcal{K}$ is also valid, by 14.48. If $\Sigma \cup \{\neg\mathcal{K}\}$ is syntactically consistent, then it has a model by by (UF11) or (UF12), but that model would make $\mathcal{K}$ and $\neg\mathcal{K}$ both valid, a contradiction. Thus $\Sigma \cup \{\neg\mathcal{K}\}$ is syntactically inconsistent. The formula $\neg\mathcal{K}$ has no free variables, so by 14.40.c we obtain $\Sigma \vdash \neg\neg\mathcal{K}$. Since we are using classical logic, that simplifies to $\Sigma \vdash \mathcal{K}$. By the result in 14.44, then, we have $\Sigma \vdash \mathcal{A}$. This completes the proof.

Proof that the "theorem versions" (UF14) *and* (UF15) *imply the "consistency versions"* (UF11) *and* (UF12), *respectively.* The only part that requires proof is (A) $\Rightarrow$ (B). Let $\mathcal{A}$ be any formula. Suppose that Σ has no model. Then it is vacuously true that every model of Σ makes $\mathcal{A} \sqcap (\neg\mathcal{A})$ valid. Thus $\mathcal{A} \sqcap (\neg\mathcal{A})$ is a semantic theorem, and therefore a syntactic theorem. Thus from Σ we can deduce $\mathcal{A} \sqcap (\neg\mathcal{A})$; that is, Σ is syntactically inconsistent.

14.60. *A pathological example.* To complete the discussion in 14.38, we shall now present an example in which

$$\mathcal{F} \vdash \mathcal{G} \qquad \text{but not} \qquad \vdash (\mathcal{F} \to \mathcal{G}).$$

It is clear from 14.40.a that in such an example, $\mathcal{F}$ must have at least one free variable. Actually, what we shall prove is that $\mathcal{F} \models \mathcal{G}$ but not $\vdash (\mathcal{F} \to \mathcal{G})$; the desired conclusion then follows from the Completeness Theorem.

Assume that our language includes (among other things) the constant symbols 0 and 1, the binary relation symbols $=$ and $\neq$, and at least one free variable symbol x. Assume that our axiom system includes at least the usual axioms for equality (these are listed in 14.27.a) and the axiom $0 \neq 1$. Let $\mathcal{G}$ be the formula "$0 = 1$;" thus $\neg\mathcal{G}$ is one of our axioms.

The formulas $x = 1$ and $x \neq 1$ are negations of each other, but neither of these formulas is a valid formula in any interpretation of the language, since each can be falsified by at least one valuation — i.e., by at least one choice of the value of x. Let $\mathcal{F}$ be either one of these two formulas (it doesn't matter which). Then neither $\mathcal{F}$ nor $\neg\mathcal{F}$ has a model, hence neither is a semantic theorem, hence neither is a syntactic theorem.

Since there are no models of $\mathcal{F}$, we can say (vacuously) that every model of $\mathcal{F}$ is also a model of $\mathcal{G}$. That is, $\mathcal{F} \models \mathcal{G}$.

If $\mathcal{F} \to \mathcal{G}$ were a syntactic theorem, then its contrapositive, $(\neg\mathcal{G}) \to (\neg\mathcal{F})$ would also be a theorem. Then $(\neg\mathcal{F})$ would also be a theorem, by *modus ponens*, since $(\neg\mathcal{G})$ is one of our axioms. But we already know that $(\neg\mathcal{F})$ is not a theorem. Thus, $\mathcal{F} \to \mathcal{G}$ is not a syntactic theorem — i.e., we do not have $\vdash (\mathcal{F} \to \mathcal{G})$.

14.61. Following are two more equivalents of UF:

(UF16) for Propositional Logic:
(UF17) for Predicate Logic:

Compactness Principle. If Σ is a set of formulas, every finite subset of which has a model, then Σ has a model.

Remarks. The name "Compactness Principle" stems from some topological considerations described in 17.25. A nonlogicians' variant of the Compactness Principle is given by (UF2) — see the remarks in 6.35.

Proof of (UF11) $\Rightarrow$ (UF16) *and proof of* (UF12) $\Rightarrow$ (UF17). Immediate from the observation about finite character, in 14.31.

Proof of (UF17) $\Rightarrow$ (UF16). Propositional logic is a special case of predicate logic.

Proof of (UF16) $\Rightarrow$ (UF8). Let X be a nondegenerate Boolean algebra; it suffices to show that the dual of X is nonempty — i.e., we are to show the existence of a function $f : X \to \{0, 1\}$ that satisfies

$$f(x \vee y) = f(x) \vee f(y), \qquad f(\complement x) = \complement f(x), \qquad f(1) = 1 \qquad (!)$$

for all $x, y \in X$. Let $\mathcal{B} \leftrightarrow \mathcal{C}$ be an abbreviation for the formula $(\mathcal{B} \to \mathcal{C}) \sqcap (\mathcal{C} \to \mathcal{B})$. Define a propositional calculus that has one primitive proposition symbol P_x for each $x \in X$, and let Σ be the set of formulas

$$P_{x \vee y} \leftrightarrow P_x \sqcup P_y, \qquad P_{\complement x} \leftrightarrow \neg P_x, \qquad P_1 \leftrightarrow T$$

for all $x, y \in X$. Then a model of Σ is the same thing as a function f satisfying (!).

We shall apply (UF16); thus it suffices to show that each finite subset of Σ has at least one model. Any finite subset $\Phi \subseteq \Sigma$ involves only finitely many x's and y's. Let X_0 be the nondegenerate Boolean subalgebra of X generated by those x's and y's. Then the dual of X_0 is nonempty, by 13.20. Thus there exists a Boolean homomorphism $f : X_0 \to \{0, 1\}$. Now

$$\text{interpret } P_x \text{ as } \begin{cases} \text{true} & \text{if } x \notin X_0 \text{ or } f(x) = 1 \\ \text{false} & \text{if } x \in X_0 \text{ and } f(x) = 0. \end{cases}$$

This is a model of Φ.

14.62. Let us emphasize the difference between a model and a quasimodel. A model "answers every question" that can be expressed in the formal language by assigning a truth value of "true" or "false" (1 or 0) to every closed formula. A quasimodel does not give such a definite answer, since its truth values may range through a Boolean algebra. The Lindenbaum algebra, which yields a quasimodel as in 14.56, plays this special role: It tells us which formulas are provable or disprovable, by assigning them the truth values of 1 or 0. Some formulas may be neither provable nor disprovable — the Lindenbaum algebra may have other values besides 1 and 0. We can answer some of the unanswered questions by adding more axioms, but we would have to keep adding more axioms; that will be evident from Gödel's Incompleteness Theorem, described below.

If we really want to have an answer to every question, the Completeness Theorem gives us one way to accomplish that. Any consistent theory has a (not necessarily unique) model and thus a (not necessarily unique) method for assigning the value "true" or "false" to every closed formula.

We can even make each closed formula provable or disprovable, in this rather contrived fashion: Form a model, and then use that model's valid formulas as the axioms for a new theory. The new theory extends the old one, is consistent, and has Lindenbaum algebra equal to $\{0, 1\}$. However, this formulation is not constructive, since the Completeness Theorem is not constructively provable. The resulting axiom system is extremely large and not recursive.

Gödel's First Incompleteness Theorem, published in 1931, says that for sufficiently complicated theories we cannot answer all the questions. Somewhat more precisely:

> Let $\mathcal{T}$ be a formal theory that includes arithmetic, and assume that the axioms of $\mathcal{T}$ can be described in a mechanical fashion (i.e., *recursively* — we shall not give a precise definition of this term). Assume the language of $\mathcal{T}$ includes only countably many symbols. Then:
>
> **Gödel's First Incompleteness Theorem.** If $\mathcal{T}$ is consistent, then there exist formulas that can be formulated in the language of $\mathcal{T}$, but cannot be proved or disproved within the formal system from the axioms of $\mathcal{T}$.

We shall not prove this theorem, but we shall sketch some of the ideas of the proof. The remainder of this section is optional; it will not be needed later in the book.

Let $\mathcal{T}$ be a theory that contains arithmetic; say $\mathcal{L}$ is the language of $\mathcal{T}$. Some properties of numbers can be expressed in $\mathcal{L}$ — for instance, a number x is composite if it satisfies

$$\exists_\xi \, \exists_\eta \qquad (\xi \neq 1) \sqcap (\eta \neq 1) \sqcap (x = \xi\eta).$$

Now assume that $\mathcal{T}$'s language $\mathcal{L}$ has only countably many symbols, and let $\mathcal{S}$ be the set of all finite strings of symbols. Then it is possible to number these strings — i.e., to define a canonical injective mapping $\# : \mathcal{S} \to \mathbb{N}$. Statements about strings can be transformed to statements about numbers — for instance, define a relation $\rhd$ on the positive integers by saying that $m \rhd n$ if

$$m = \#(S) \text{ and } n = \#(T) \text{ for strings } S, T \text{ such that } S \text{ is a proof of } T.$$

The relation $\rhd$ is a purely numerical relation — i.e., it is a relation whose graph is a subset of $\mathbb{N} \times \mathbb{N}$. Although we defined $\rhd$ in terms of the language $\mathcal{L}$ and the correspondence $\#$, it is possible to describe this same relation $\rhd$ in purely numerical terms, without mentioning $\mathcal{L}$ or $\#$.

When $\mathcal{A}$ is a formula, we shall call $\#(\mathcal{A})$ the **Gödel number** of $\mathcal{A}$. Let $\mathbb{G}$ be the set of all Gödel numbers; it is a subset of $\mathbb{N}$. We shall now outline a proof of:

> *Lemma.* Let Q be a property of some natural numbers — i.e., assume that $Q(x)$ is true for some natural numbers x and false for others. Assume, moreover, that the statement "x has the property Q" is expressible in the formal language $\mathcal{L}$. Then there exist a particular number n and a formula $\mathcal{A}$ such that (i) the Gödel number of $\mathcal{A}$ is n, and (ii) $\mathcal{A}$ expresses the statement that "the number n has the property Q."

Sketch of proof of lemma. Let v be some particular free variable, which will not change for the remainder of this discussion. Define a special function $\varphi : \mathbb{G} \times \mathbb{N} \to \mathbb{G}$ as follows: To evaluate $\varphi(m, n)$,

> let S_m be the string of symbols with Gödel number m. The number n can be expressed in the language $\mathcal{L}$; let T_n be the string of symbols that expresses the number n. Let $U_{m,n}$ be the string obtained from S_m by replacing each occurrence of v in it with a copy of the string T_n; finally, let $\varphi(m, n)$ be the Gödel number of $U_{m,n}$.

The equation $z = \varphi(x, x)$ is a statement about numbers. Gödel proved that it can be expressed in the formal language $\mathcal{L}$. The map $x \mapsto \varphi(x, x)$ is an operation somewhat analogous to the operation of *quining*, which was introduced in 1.12.

Now, let $\mathcal{G}$ be the formula that expresses, in the language $\mathcal{L}$, the statement "$\varphi(v, v)$ has the property Q." Say the Gödel number of $\mathcal{G}$ is p. The number p can be expressed in the formal language. Obtain $\mathcal{A}$ from $\mathcal{G}$ by replacing each occurrence of v with the string that expresses p, and let $n = \varphi(p, p)$; this proves the lemma.

Proof of theorem, continued. For property $Q(x)$, Gödel uses a property such as "x is the Gödel number of a formula that is not provable in $\mathcal{T}$." Of course, it is a nontrivial matter to establish that the lemma is applicable to this property. The lemma then yields an effectively constructible formula that says, roughly, "I am not provable." (However, it uses indirect self-referencing as in Quine's Paradox (1.12), rather than direct self-referencing as in Epimenides's Paradox (1.11).) Such a statement cannot be provable and therefore must be true and therefore cannot be disprovable either. In this fashion we obtain Gödel's First Incompleteness Theorem.

Remarks. Longer informal expositions of this subject can be found in Rosser [1939], Nagel and Newman [1958], Hofstadter [1979], and Mac Lane [1986]. More technical and detailed expositions can be found in books on logic.

NONSTANDARD ANALYSIS

14.63. No system of first-order properties of $\mathbb{N}$ or $\mathbb{R}$ can uniquely determine $\mathbb{N}$ or $\mathbb{R}$. Any first-order theory that can be modeled by $\mathbb{N}$ or $\mathbb{R}$ can also be modeled by some system of "numbers" that includes infinitely large members. More precisely, we have this proposition:

Skolem's example (1934). Let $\mathfrak{L}$ be a first-order language that includes infinitely many free variables, the relation symbol " $<$ ", and the constant symbols "1," "2," "3," ... (and possibly other symbols as well). Let Σ be a set of axioms in that language. Suppose that $\mathbb{N}$ (respectively, $\mathbb{R}$) is the domain for some model of Σ, giving the symbols " $<$ " and "1," "2," "3," ... their usual meanings. Then there exist other models for Σ, which are not isomorphic to $\mathbb{N}$ (respectively, $\mathbb{R}$). In fact, there exists a model that contains an "infinitely large number" — i.e., a number that is greater than all the numbers 1,2,3,... .

Proof. Let c be a constant symbol that is not already in use in the language $\mathfrak{L}$; let $\mathfrak{L}' = \mathfrak{L} \cup \{c\}$. By 14.45, the system $(\mathfrak{L}', \Sigma)$ is consistent. Now let Λ be the set of axioms

$$1 < c, \qquad 2 < c, \qquad 3 < c, \qquad \ldots ,$$

and (for each positive integer n) let Λ_n be the first n of these axioms. The first-order theory $(\mathfrak{L}', \Sigma \cup \Lambda_n)$ has a model given by $D = \mathbb{N}$ (respectively, $D = \mathbb{R}$), with c interpreted as $n+1$.

Each finite subset of $\Sigma \cup \Lambda$ is contained in some $\Sigma \cup \Lambda_n$ and therefore has a model. By the Compactness Principle (UF17) in 14.61, $\Sigma \cup \Lambda$ has a model. The interpretation of c in that model is an infinitely large number.

14.64. *Historical remarks and overview.* In the late 17th century, when Newton and Leibniz were first inventing calculus, part of their theory involved *infinitesimals* — i.e., numbers that are infinitely small but nonzero. A basic idea of Leibniz, now known as **Leibniz's Principle**, was that the larger number system (involving real numbers, infinitesimals, infinitely large numbers, etc.) should somehow have the "same properties" as the real number system, but he did not know how to make this principle precise. Among other things, it was not entirely clear just which "properties" would fit Leibniz's principle — and in retrospect it is clear that some properties *cannot* fit Leibniz's Principle. Indeed, the conventional real number system is Dedekind complete and therefore Archimedean, which essentially means that it lacks infinitesimals (see 10.3); the lack of infinitesimals is not one of the "properties" Leibniz had in mind.

In those days mathematics was more computational and did not involve rigor as we know it today. During the next couple of centuries, mathematicians gradually added more rigor to their ways of thinking. During the 19th century, Cauchy developed a theory of limits, and then Weierstrass restated this theory in terms of epsilon-delta arguments, finally putting calculus on a firm foundation. Mathematicians could not find a justification for infinitesimals and no longer needed them for calculus, so infinitesimals gradually fell out of favor and were used less and less. The epsilon-delta arguments gained wide acceptance and are used in nearly all calculus textbooks today.

Early in this century, Skolem observed that there must exist nonstandard models of arithmetic and analysis. In 1960, Abraham Robinson developed this idea in much greater detail and made Leibniz's Principle precise and rigorous: We specify a particular first-order language $\mathcal{L}$ that is suitable for discussing properties of $\mathbb{R}$ or properties of ${}^*\mathbb{R}$; then the same properties (expressible in that language) will be valid in either model; this is the **Transfer Principle**. Troublesome properties such as Dedekind completeness cannot be expressed directly in the first-order language. It is possible to formulate a property in the first-order language that specializes to Dedekind completeness when interpreted in $\mathbb{R}$, but that same property does not yield Dedekind completeness when interpreted in ${}^*\mathbb{R}$. This is discussed further in 14.66.

Robinson gave many applications including a rigorous justification for an infinitesimal calculus much like the one envisioned by Newton and Leibniz. Thus *nonstandard analysis* was born. Important contributions were also made by Zakon, Los, Luxemburg, Kiesler, Loeb, and others. We might call this the "monomorphism" school of thought (in contrast with IST, discussed below); all of the papers involve a monomorphism mapping $S \mapsto {}^*S$. The monomorphism theory can be presented either

(i) "axiomatically" — i.e., we list the properties that a monomorphism must have, or

(ii) "constructively" — i.e., we describe how to form a monomorphism, though perhaps with the use of ultrapowers or other tools that Errett Bishop would not have called "constructive." For instance, it can be shown that the hyperreal line ${}^*\mathbb{R}$ (introduced in 10.20.a) is a model of $\mathbb{R}$ that satisfies Leibniz's Transfer Principle, although we shall not prove that fact in this book.

The "axiomatic" approach is more concise but it can only be justified by the "constructive" approach.

In 1977 Edward Nelson published a much simpler axiomatic approach, which codifies the main ideas of nonstandard analysis without ever mentioning monomorphisms. The approach is called "IST," which stands for "Internal Set Theory" and also stands for "Idealization, Standardization, and Transfer," Nelson's three main axioms. Nelson's approach is perhaps the closest yet to the original conception of Newton and Leibniz. We summarize it in 14.67.

The term "nonstandard analysis" may be misleading; a more descriptive term would be "hyperfinite reasoning." As it is commonly used, nonstandard analysis is a type of reasoning that allows us to treat infinite sets much like finite sets; that should be particularly evident in the Enlargement Principle in 9.54. This "hyperfinite reasoning" was developed and used first in analysis, but it also applies to other branches of mathematics.

It is provable (in logic) that nonstandard mathematics is a **conservative** extension of conventional (standard) mathematics. This means that any statement that can be formulated in conventional mathematics and proved in nonstandard mathematics can also be proved in conventional mathematics. In fact, any proof in nonstandard mathematics can be converted into a proof in conventional mathematics, in a largely mechanical fashion.

Generally, the standard proof requires clever choices of ultrafilters or other abstract tools. One advantage of nonstandard mathematics is that cleverness with ultrafilters is not required: The relevant properties of ultrafilters are already built into the basic principles of nonstandard mathematics — the Transfer Principle, etc. Thus, the working mathematician is freed to concentrate on other difficulties that are more specific to the problem.

The advantage of nonstandard mathematics is that its intuition is sometimes helpful; some proofs in nonstandard mathematics may be easier to find or to understand than the corresponding proofs in conventional mathematics. Indeed, Leibniz and Newton had infinitesimals in mind when they invented calculus; surely this is a testimony to the usefulness of the intuition of nonstandard mathematics.

Leibniz viewed the derivative dy/dx as the quotient of two infinitesimals — hence his notation, which is still in use today. In today's conventional mathematics (i.e., in the Cauchy-Weierstrass epsilon-delta approach), it is customary to define the derivative as a *limit* of quotients, not as an actual quotient. In general, limit arguments can be converted to nonstandard terms — for instance, a function $f : \mathbb{R} \to \mathbb{R}$ satisfies $\lim_{x\to x_0} f(x) = L$ if and only if, whenever ε is a nonzero infinitesimal, then ${}^*f(x_0 + \varepsilon)$ is infinitely close to L. (We proved a result of this sort in 10.37.) However, it must be noted that a similar type of intuition can also be obtained from other tools, e.g., convergent nets; see 15.14(A).

14.65. *Construction of superstructures.* The "constructive" approach to monomorphisms, alluded to in 14.64(ii), does not require so much formal logic, but it uses more set theory. We define the monomorphism $S \mapsto {}^*S$ as a mapping from $\mathcal{S}_\infty(\mathbb{R})$ into $\mathcal{S}_\infty({}^*\mathbb{R})$ or, more generally, from $\mathcal{S}_\infty(X)$ into $\mathcal{S}_\infty({}^*X)$ where X is any infinite set of interest. The *superstructures* $\mathcal{S}_\infty(\mathbb{A})$ are defined as follows:

Let $\mathbb{A}$ be an infinite set. Recursively define

$$\mathcal{S}_0(\mathbb{A}) = \mathbb{A}, \qquad \mathcal{S}_n(\mathbb{A}) = \mathcal{P}\left(\bigcup_{k=0}^{n-1} \mathcal{S}_k(\mathbb{A})\right), \qquad \mathcal{S}_\infty(\mathbb{A}) = \bigcup_{k=0}^{\infty} \mathcal{S}_k(\mathbb{A}),$$

where $\mathcal{P}(X) = \{\text{subsets of } X\}$. (We emphasize that this procedure is only iterated countably many times, unlike the procedures in 5.53 and 5.54.) The set $\mathcal{S}_\infty(\mathbb{A})$ is called the **superstructure** over $\mathbb{A}$. The members of $\mathbb{A}$ are called the **individuals** or **atoms** in this context; the members of $\mathcal{S}_\infty(\mathbb{A}) \setminus \mathbb{A} = \bigcup_{k=1}^{\infty} \mathcal{S}_k(\mathbb{A})$ are called the **entities** of the superstructure.

The collection of entities is closed under most set-theoretical operations. For instance, if a is an entity, then $\bigcup a = \{p : p \in q \text{ for some } q \in a\}$ is also an entity. Any subset of an entity is an entity. If a and b are entities then (a, b) is an entity, since we follow the convention of reading (a, b) as $\{\{a\}, \{a, b\}\}$. If $f : X \to Y$ is a function and X, Y are entities, then f is an entity; here we identify a function with its graph. Even the Axiom of Choice is modeled by the collection of entities: If $g : X \to Y$ is a function, X is an entity, Y is an entity, and $g(x)$ is an entity for each $x \in X$, then there exists some function $f : X \to \bigcup Y$ such that $f(x) \in g(x)$ for each x and f is an entity.

$\mathcal{S}_\infty(\mathbb{R})$ is only a set, not a proper class, and thus it is much smaller than the von Neumann universe V developed in 5.53. Nevertheless, $\mathcal{S}_\infty(\mathbb{R})$ is large enough to include all of the objects used in real analysis. Indeed, it includes all subsets of $\mathbb{R}$, sets of subsets of $\mathbb{R}$, and all sets of subsets of subsets of $\mathbb{R}$, etc. It includes ordered pairs, functions, etc. The superstructure $\mathcal{S}_\infty(\mathbb{R})$ is a transitive set, and it is also closed under intersection, finite union, and finite Cartesian products. However, $\mathcal{S}_\infty(\mathbb{R})$ is *not* closed under countable union or countable Cartesian products: If $S_1, S_2, S_3, \ldots$ are entities, then $S_1 \cup S_2 \cup S_3 \cup \cdots$ and $S_1 \times S_2 \times S_3 \times \cdots$ are *not* necessarily entities.

For each entity S in $\mathcal{S}_\infty(\mathbb{R})$, there is a corresponding entity ${}^*S$ in $\mathcal{S}_\infty({}^*\mathbb{R})$. The correspondence is complicated, so we shall not describe it in full detail; it can be found in books on nonstandard analysis. However, we shall mention a couple of its features: (i) When $S \in \mathcal{S}_1(\mathbb{R})$, then $^*S$ is just the ultrapower described in 9.41. (ii) In 2.7 we subscribed to the convention that the forward image function associated with a given function f is generally denoted by the same letter f; thus $f(S) = \{f(s) : s \in S\}$. However, that convention *cannot* be applied to the monomorphism mapping: in general $^*S$ is *not* equal to $\{^*T \ : \ T \in S\}$.

14.66. The mapping $S \mapsto {}^*S$, from $\mathcal{S}_\infty(\mathbb{R})$ into $\mathcal{S}_\infty(^*\mathbb{R})$, is not surjective. A member of its range — i.e., an object of the form $^*S$ — is called a **standard** object in $\mathcal{S}_\infty(^*\mathbb{R})$; any other member of $\mathcal{S}_\infty(^*\mathbb{R})$ is called a **nonstandard** object. Since the mapping $S \mapsto {}^*S$ is injective, the collection of all standard objects is an isomorphic copy of $\mathcal{S}_\infty(\mathbb{R})$ inside the nonstandard universe $\mathcal{S}_\infty(^*\mathbb{R})$. Some important sets are not standard — for instance, the set of all infinitesimals and the set of all standard sets.

An **internal object** is a *member* of a standard set — i.e., it is an object x that satisfies $x \in {}^*S$ for some standard entity $^*S$. The internal object x may or may not be standard. An **external object** is a member of $\mathcal{S}_\infty(^*\mathbb{R})$ that is not internal.

For any set A in the standard universe, we have $\mathcal{P}(^*A) \supseteq {}^*\mathcal{P}(A)$, and in many cases $\mathcal{P}(^*A) \supsetneqq {}^*\mathcal{P}(A)$. Observe that $\mathcal{P}(^*A)$ is the collection of *all* subsets of $^*A$; it can be shown that $^*\mathcal{P}(A)$ is the collection of all *internal* subsets of $^*A$.

Internal sets arise naturally in uses of the Transfer Principle. For instance, consider the following true statement.

> For each set $S \subseteq \mathbb{R}$ — that is, for each $S \in \mathcal{P}(\mathbb{R})$ — if S has an upper bound in $\mathbb{R}$, then S has a least upper bound in $\mathbb{R}$. Thus: $\mathbb{R}$ is Dedekind complete.

If we use the Transfer Principle naively, without understanding what we are doing, we might end up with this:

> (*Incorrect and false statement.*) For each set $T \subseteq {}^*\mathbb{R}$ — that is, for each $T \in \mathcal{P}(^*\mathbb{R})$ — if T has an upper bound in $^*\mathbb{R}$, then T has a least upper bound in $^*\mathbb{R}$. Thus, $^*\mathbb{R}$ is Dedekind complete.

But in 10.19 we saw that $^*\mathbb{R}$ generally is *not* Dedekind complete. If we apply the Transfer Principle correctly, we obtain the folowing statement, which is less interesting but has the advantage of being correct and true.

> For each set $T \in {}^*\mathcal{P}(\mathbb{R})$ — that is, for each *internal* set $T \subseteq {}^*\mathbb{R}$ — if T has an upper bound in $^*\mathbb{R}$, then T has a least upper bound in $^*\mathbb{R}$.

14.67. *A sketch of IST.* In 1977 Edward Nelson published a new approach to nonstandard analysis, which reaches the applications quickly without first building up so much machinery. In Nelson's theory, we start from conventional mathematics and then add a new word: **standard**. This word is left undefined — i.e., it is not defined in terms of our old vocabulary — but the permitted uses of this word are governed by several new axioms and syntactic instructions.

Actually, we need *two* new words, but only "standard" is an undefined word added to our mathematical ("inner") system. We also add the word "classical" to our metamathematical ("outer") system, and it has a very simple definition: An expression is **classical** if it does not, either explicitly or implicitly, involve the word "standard."

Other words can be defined using the term "standard." For instance, a real or complex number ε is an **infinitesimal** if $|\varepsilon| < y$ for every positive standard number y. Note that any expression that uses the word "infinitesimal" is a nonclassical expression, since it implicitly involves the word "standard."

We shall not list the new axioms and syntactic instructions here; they can be found in books on IST — for instance, Robert [1988].

The term "standard" can *only* be used in accordance with the new axioms and syntactic instructions; it cannot be bandied about freely. For instance, in conventional ZF set theory, the Axiom of Comprehension says that if P is a property formulated in the first-order language and X is a set, then $\{x \in X : P(x) \text{ is true}\}$ is also a set. In IST's version of set theory, we must add the further restriction that the property P is classical — i.e., P can be formulated in the first-order language without mentioning the words "standard," "infinitesimal," etc. If P does not meet that specification, then $\{x \in X : P(x) \text{ is true}\}$ may be a *proper class*, rather than a set.

14.68. *IST's new systems of numbers.* Because of the restrictions made on the language, "classical" mathematics cannot be used to prove the existence or nonexistence of infinitely large or infinitely small numbers. However, IST's new axioms can be used to prove the existence of such numbers. The resulting number system is not called the *hyperreal numbers*; rather, it is called the *real numbers.* In effect, we view the real number system as containing not only the classical real numbers discussed by Weierstrass et al., but also some other numbers — infinitesimals, etc. — that we had not noticed before. Presumably they were there all along. This new use of old terminology may prove disconcerting to some beginners.

Here are a few comparisons between IST and conventional mathematics:

a. The ordinary Principle of Induction says that every nonempty subset of $\mathbb{N}$ has a smallest element. This is true in IST, but we must emphasize that what is needed is a sub*set* of $\mathbb{N}$, not just a sub*class*.

b. The real number system is Dedekind complete — i.e., if S is a subset of $\mathbb{R}$ that is bounded above by a real, then S has a least upper bound. This remains valid in IST, but in that setting we require that S be a sub*set* of $\mathbb{R}$ and not just a sub*class* of $\mathbb{R}$.

c. The real number system is an Archimedean field. That is, if a and b are positive real numbers, then there exists a positive integer n such that $nb > a$. In IST this remains valid, but we point out that in IST the positive integer n might be infinitely large — i.e., its reciprocal might be an infinitesimal.

The "new" real number system of IST should not be confused with the hyperreal number system introduced in 10.18. Both are ordered fields that include infinitesimals, but those infinitesimals are "detected" using different methods of reasoning, in the context of different kinds of logic and set theory. The hyperreal number system, introduced in 10.18 in the setting of conventional set theory, is neither Dedekind complete nor Archimedean.

14.69. *The connection between monomorphisms and IST.* The terminology of IST is rather different from the terminology of monomorphisms. Most papers in nonstandard analysis today use either monomorphisms or IST, but not both. The two schools of thought are nearly equivalent: Most ideas expressed in either system can be translated into the other system, by a method that, though not effortless, is mostly mechanical.

Nelson's IST can be described "constructively" in terms of monomorphisms (i.e., not in Nelson's terms), roughly as follows: Since $^*\mathbb{R}$ contains an isomorphic copy of $\mathbb{R}$, let's just work with that isomorphic copy and forget about the original copy; then we have one less object to worry about. When we want to work with real numbers and related objects, we'll just work with subsets of $^*\mathbb{R}$. We have some axioms that describe our number system, including axioms about infinitesimals and other nonstandard objects. The axioms for $^*\mathbb{R}$ are quite similar to classical axioms for $\mathbb{R}$ — which is not surprising, in view of the Transfer Principle we have already discussed — but $^*\mathbb{R}$ does contain infinitesimals and other nonstandard objects. Now let's change our notation and drop the asterisk; we'll write our number system as "$\mathbb{R}$," and call it "the real numbers." The result is a new "real number system," which contains infinitesimals and other new objects. Similarly, the symbols $\mathbb{N} = \{\text{natural numbers}\}$ and $\mathbb{Z} = \{\text{integers}\}$ both take new meanings; both these sets now contain infinite members. The objects that are called "standard" in IST are those in the image of the monomorphism mapping.

Summary of Some Consistency Results

14.70. *The quest for certainty.* Is it possible to build mathematics on a completely reliable foundation? When consistency of a theory is proved by a fully detailed but finite argument, as in 14.10, we say that **absolute consistency** is established. Truths may vary from one logical system to another, depending on our axiom system, but can we at least prove that our axiom system is absolutely consistent?

In some very simple cases we can. For instance, the axioms of classical propositional logic can be proved absolutely consistent (though this may not be readily apparent from this book's exposition since we have mixed that elementary result with more advanced results).

In 14.7 we described how Gödel used the consistency of ZF to prove the consistency of ZF + AC + GCH. This result — and earlier, more elementary results of a similar nature — show that we can sometimes "bootstrap" our way up, using some weak consistency results to prove other, stronger consistency results.

Mathematicians around 1900, encouraged by a few successes in proving consistency, began to hope that all of mathematics could be put on a firm foundation. The search for absolute proofs of consistency was promoted especially by David Hilbert, and so it is sometimes known as **Hilbert's program**; see the discussion by Kreisel [1976].

However, in 1931 Gödel published his Incompleteness Theorems; the second of these tells us that Hilbert's program *cannot* be carried out for larger portions of mathematics. In fact, it cannot be carried out for any system that is sufficiently sophisticated to yield all of arithmetic.

Gödel's proof is too complicated to present here; even a precise statement of his results is too complicated to present here. However, here is a partial statement of his results:

Let $\mathcal{T}$ be a formal theory that includes arithmetic, and assume that the axioms of $\mathcal{T}$ can be described in a mechanical fashion (i.e., *recursively* — we shall not give a precise definition of this term). Assume the language of $\mathcal{T}$ includes only countably many symbols. Then:

> **Gödel's Second Incompleteness Theorem.** The statement "$\mathcal{T}$ is consistent" can be encoded as a formula expressed in the language of $\mathcal{T}$; but that formula is not a theorem of $\mathcal{T}$ — i.e., that formula *about* $\mathcal{T}$ cannot be proved *inside* $\mathcal{T}$.

The proof uses many of the same ideas indicated in 14.62.

14.71. *Fundamental uncertainty.* One particular consequence of Gödel's Second Incompleteness Theorem is:

> Absolute consistency cannot be proved for any formal system that includes arithmetic. In particular, absolute consistency cannot be proved for ZF.

Thus, mathematics has lost its innocence: A kind of uncertainty is inherent in the foundation of mathematics and cannot be exorcised. The age of supreme confidence, which began when Isaac Newton explained the movements of the heavens with a few simple equations, ended with Kurt Gödel's Second Incompleteness Theorem.

This should not be entirely surprising. On a more fundamental level, we cannot use the basic techniques of reasoning to prove that the *basic techniques of reasoning* are reliable. Such circular reasoning would be worthless. Perhaps we are all really quite mad and merely imagine ourselves to be rational. Even a mathematician must accept certain some things on faith or learn to live with uncertainty.

And there is even disagreement on such matters of faith! For instance, as we noted in 6.4, most mathematicians are comfortable with proof by contradiction and the Law of the Excluded Middle (and we use such proofs freely in this book), but to constructivists such proofs are taboo.

Actually, many researchers in branches of mathematics outside of logic are not aware that the age of certainty has ended. Gödel's Second Incompleteness Theorem has no effect on such simple certainties as $2 + 2 = 4$, and it does not diminish the extraordinary success mathematics has brought to technology. Very clever computations were required to create radio and television, or to land human beings on the moon and then return them safely to earth. Those computations do not actually rely on ZF set theory. For all practical purposes, applied mathematics is a system that "works," and it will continue to work even if an inconsistency someday is found in the most popular formalization of abstract mathematics. It is convenient, and perhaps reassuring, to explain differential equations in terms of set theory, but it is not absolutely necessary. If our explanation of subsets is someday discovered to be inconsistent, researchers in differential equations will not immediately resign from their jobs *en masse.*

14.72. If we cannot establish absolute consistency, what is the next best thing? What shall we use for the basis of our reasoning?

First of all, there is **empirical consistency** — i.e., that which is based on evidence, but not on a complete proof. For instance, mathematicians have now used ZF for nearly a century and have not yet proved any contradictions from it. Thus, ZF seems to be consistent, based on empirical evidence, if we accept conventional rules of reasoning. Similarly,

$$\mathrm{ZF} + \mathrm{AC} + \mathrm{IC} \qquad \text{is empirically consistent (so far),}$$

where AC is the Axiom of Choice and IC is the statement that "there exists an inaccessible cardinal" (introduced in 2.21). However, the accumulated evidence is less for this axiom system than for ZF, as this system has not been studied quite so extensively.

Second, there is **relative consistency**: We may prove that *if* one set of formulas Σ is consistent, *then* another set Φ is consistent — or, more briefly,

$$\mathrm{Con}(\Sigma) \quad \Rightarrow \quad \mathrm{Con}(\Phi).$$

The problem of consistency is not removed altogether — we may not be certain of the consistency of Σ *or* Φ — but the problem may be relocated to a more manageable place: in many cases, Σ is simpler or more intuitively "believable" than Φ. When relative consistency is proved, then Φ inherits all of Σ's plausibility.

Although the absolute consistency of ZF is not attainable, at least ZF is empirically consistent. Hence we may use the consistency of ZF as a plausible *hypothesis* in our relative consistency arguments.

14.73. *The relative consistency of Choice.* The most famous consistency result relative to ZF is Gödel's result: $\mathrm{Con(ZF)} \Rightarrow \mathrm{Con(ZF + AC + GCH)}$, which we have already discussed in 14.7. By Gödel's result, it is "safe" to add the Axiom of Choice and the Continuum Hypothesis to set theory — if this generates any contradictions, then contradictions were already present in ZF anyway. The various pathologies generated by Choice — such as the Banach-Tarski Decomposition, in 6.16 — may cause us to question that axiom and to consider alternatives, but none of those pathologies actually lead to a contradiction (unless ZF is already corrupt), and so none of them can actually force us to give up the Axiom of Choice.

We emphasize that Gödel's result does not say that ZF implies AC. Quite the contrary; in 1963 Cohen showed that

$$\text{if ZF is consistent, then ZF + not-AC + not-CH is also consistent.}$$

Thus ZF does *not* imply AC or CH (the Continuum Hypothesis).

14.74. *Equiconsistent alternatives to ZF.* Since Cohen's breakthrough, several other mathematicians have subsequently proved the relative consistency of more specific negations of the Axiom of Choice. They have proved that certain statements that contradict AC, such as

$$\begin{aligned} \mathrm{BP} &= \text{"every subset of } \mathbb{R} \text{ has the Baire property"} \\ \mathrm{LM} &= \text{"every subset of } \mathbb{R} \text{ is Lebesgue measurable"} \end{aligned}$$

(discussed, respectively, in Chapters 20 and 21) are consistent with certain weak consequences of Choice, such as

DC	= Dependent Choice	UF	= Ultrafilter Principle
CC	= Countable Choice	HB	= Hahn-Banach Theorem

(discussed in Chapters 6 and and 12). Specifically, if ZF is consistent, then so is each of the following sets of axioms:

- ZF + BP + LM (Solovay [1964/1965])
- ZF + UF + not-CC (Halpern and Levy [1971])
- ZF + DC + HB + "all ultrafilters are fixed" (Pincus and Solovay [1977])
- ZF + DC + BP (Shelah [1984])
- ZF + DC + BP + not-LM (Stern [1985])

These consistency results will not be proved in this book; the ambitious reader will find the proofs elsewhere in the literature. Many of these results were proved by forcing. Earlier in this chapter we have suggested some of the ingredients of consistency proofs, but those suggestions are far from actually being a proof. The consistency of conventional set theory (ZF + AC) and the consistency of Shelah's alternative (ZF + DC + BP) will be assumed throughout this book; the importance of Shelah's alternative will be explained in 14.77.

The Axiom of Choice and its negation cannot coexist in one proof, but they can certainly coexist in one mind. It may be convenient to accept AC on some days — e.g., for compactness arguments — and to accept some alternative reality, such as ZF + DC + BP on other days — e.g., for thinking about complete metric spaces.

Each of the axiom systems listed above includes ZF, and thus each system is **equiconsistent** with ZF — i.e., its consistency is *equivalent* to the consistency of ZF. Another way to say this is that the various axiom systems have the same **consistency strength**. Hence each of these alternative set theories is as plausible as conventional set theory. The theologian Kierkegaard might have put it this way: To believe that ZF is consistent requires a certain leap of faith, but to believe the consistency of any of these larger systems of axioms requires no larger leap of faith.

14.75. *Axiom systems that are not equiconsistent with ZF.* Let IC be the statement that "there exists an inaccessible cardinal." As we have remarked earlier, ZF + AC + IC is empirically consistent (so far). That is, no contradictions have yet been established as consequences of ZF + AC + IC, even after considerable scrutiny by set theorists. Hence it is reasonable to take Con(ZF + AC + IC) as a hypothesis in certain arguments (where "Con" stands for "consistency of").

Obviously Con(ZF + AC + IC) $\Rightarrow$ Con(ZF + AC). However, Con(ZF + AC) does *not* imply Con(ZF + AC + IC). That fact can be proved along the following lines: We can use an inaccessible cardinal to form a set $\mathcal{M}$ such that we can prove *inside ZF + AC + IC* that $\mathcal{M}$ is a model of the axioms of ZF + AC (in the sense of 14.6). The Soundness Theorem can also be proved in this setting; thus, inside ZF + AC + IC we can prove the consistency of ZF + AC. Now Gödel's Second Incompleteness Theorem tells us that Con(ZF + AC) $\Rightarrow$ Con(ZF + AC + IC) cannot be proved inside ZF + AC, or even inside ZF + AC + IC. For further details of this argument, see Shoenfield [1967], page 306.

Thus, ZF + AC + IC is *not* equiconsistent with ZF, or ZF + AC, or all the other axiom systems listed in 14.74. We say that ZF + AC + IC has *greater consistency strength* than does ZF or ZF + AC.

Results of Solovay [1970], Shelah [1984, 1985], and Stern [1985] imply that the following sets of axioms are equiconsistent — i.e., consistency of any one of these sets implies consistency of all the others.

- ZF + AC + IC
- ZF + CC + LM
- ZF + DC + LM
- ZF + DC + LM + BP
- ZF + DC + LM + not-BP

These results show that Lebesgue measurability in $\mathbb{R}$ — a set of "ordinary" size — is inextricably connected to questions about inaccessible cardinals — i.e., sets so enormous that they are very hard to imagine.

This equiconsistency result came as something of a surprise to mathematicians. There is an extensive analogy between Lebesgue measurability (in measure theory) and the Baire property (in topology), as shown by Oxtoby [1980] and Morgan [1990]. However, the analogy breaks down when we study equiconsistency: ZF + DC + BP is equiconsistent with ZF, but ZF + DC + LM is not.

Quasiconstructivism and Intangibles

14.76. Dependent Choice, introduced in 6.28, is generally considered to be *constructive*. In fact, it is the strongest form of Choice that is accepted by most schools of constructivism. It is fairly strong, as we can see from the fact that the constructivists have been able to translate so much of classical mathematics into their viewpoint (see 6.8).

Most existence proofs in this book use conventional set theory — that is, Zermelo-Fraenkel set theory plus the Axiom of Choice (ZF + AC). Much of this book is concerned with mere existence proofs, but we also give explicit examples when we can. There is no standard meaning for "explicit example," but many mathematicians would attach a narrower meaning to that phrase than they do to "existence proof."

How can we make the phrase "explicit example" more precise? Some possible interpretations are:

- an object that can be constructed in the sense of Bishop, as in 6.2;
- an object that can be constructed in the sense of Gödel, as in 5.54; or

- an object that is definable in the sense of predicate logic, as in 6.11.

However, all these interpretations involve ideas that are quite far away from mainstream mathematics; they would require us to restructure our entire language and methods of reasoning.

Instead, we shall now propose a compromise between constructivist mathematics and mainstream mathematics, which should be easily understood by most analysts and other "ordinary" mathematicians — i.e., it is not accessible only to logicians. By **quasiconstructive** mathematics we shall mean mathematics that permits the use of conventional rules of reasoning plus ZF + DC, but no stronger forms of Choice. (This kind of reasoning was called "agnostic" mathematics by Garnir [1974], referring to the fact that the user of this mathematics assumes neither the Axiom of Choice nor its negations such as BP or LM.) In this book, an **explicit example** of an object will mean[2] a quasiconstructive proof of existence of that object. Most of this book uses classical, conventional set theory (ZF + AC), but occasionally we shall work with the more restrictive viewpoint of quasiconstructive mathematics to study the presence or absence of explicit examples. We emphasize that constructive mathematics (in the sense of Bishop) is even more restrictive: It prohibits proofs by contradiction and the Axiom of Regularity, both of which are permitted in quasiconstructive mathematics.

14.77. By an **intangible**, we shall mean an object that "exists" but has no "explicit examples" — i.e., an object whose existence can be proved in conventional mathematics (ZF + AC) but not in quasiconstructive mathematics (ZF + DC). Throughout this book, we shall assume the consistency of ZF; we shall use that assumption to *prove* that certain objects are intangibles.

For instance, certain subsets of a topological space have the *Baire property*, discussed in Chapter 20. Let BP be the statement that "every subset of $\mathbb{R}$ has the Baire property." As we stated without proof in 14.74, the consistency of ZF implies the consistency of ZF + DC + BP. Later in this book we shall prove that ZF + AC implies not-BP. Thus,

subsets of $\mathbb{R}$ *that lack the Baire property are intangibles*

— such sets "exist," but there are no "explicit examples" of such sets.

Strictly speaking, it is not the particular set S lacking the Baire property that is intangible; rather, it is *the property of being such a set* that is intangible. If we could quasiconstructively get our hands on a *particular* set S that lacked the Baire property, then S would not be an intangible! We can make our language more precise by saying something like this: "the negation of the Baire property is an intangible property." But it is more convenient and more intuitively appealing to say, somewhat imprecisely, that "the subsets of $\mathbb{R}$ that lack the Baire property are intangibles;" the intended meaning should be clear.

Actually, in some situations we *can* "get our hands on" a particular intangible object — but not quasiconstructively. Consider the following illustration:[3] Let $p_1(x)$ and $p_2(x)$

[2] Admittedly, the phrase "explicit example" has been used in different ways in the literature, and some mathematicians may not agree with the particular meaning attached to that phrase by this book. However, it is this author's feeling that many mathematicians will agree with it. At any rate, the definition given here has the advantage that it leads to some interesting theorems.

[3] Suggested by an anonymous referee.

be, respectively, the statements "$x = 0$" and "the Axiom of Choice is true." (The latter predicate actually does not depend on x.) Then in conventional mathematics (ZF + AC), the set $S = \{x : p_1(x) \text{ and } p_2(x)\}$ is nonempty. In fact, it has exactly one element, and we know what that element is: 0. The issue here is not whether we can *find* a particular member of S, but rather, whether we can *prove* that the object belongs to S. In quasi-constructive mathematics, it is clear that S is either $\varnothing$ or $\{0\}$, but we cannot determine which. Thus, 0 is an intangible — or more precisely, the property of being a member of S is an intangible property. This illustration shows that our simple and precise definitions of "explicit example" and "intangible" match only approximately with the usual imprecise intuitive meanings of those terms.

Later we shall prove that these objects are also intangibles: free ultrafilters, nontrivial universal nets, well orderings of $\mathbb{R}$, inequivalent complete norms on a vector space, finitely additive probabilities that are not countably additive, and members of $(\ell_\infty)^* \setminus \ell_1$. Some classical texts may give one the impression that these peculiar objects are somehow already present and the Axiom of Choice is used merely to "detect" them. However, it is more accurate to say that these objects are *created by* our acceptance of the Axiom of Choice. They disappear if we replace conventional set theory (ZF + AC) with some of its alternatives (such as ZF + DC + BP).

A special role is played by the negation of the Baire property: It is a "weaker" intangible than most other intangibles that we shall consider in this book, in the following sense. If we assume that the other intangible object exists, we can (and shall, elsewhere in this book) use it to prove the existence of a non-BP set, without recourse to formal logic or to any form of Choice stronger than DC. Hence Shelah's result, Con(ZF) $\Rightarrow$ Con(ZF + DC + BP), is the only result we need from formal logic for most of our intangibility proofs.

14.78. *Intangibles in a wider sense.* The "intangibles" defined above might be named more descriptively as "AC/DC intangibles" — i.e., objects whose existence is implied by ZF + AC but not by ZF + DC. That is the only kind of intangible considered in most of this book. However, many more objects are "intangible" if that term is given the broader meaning of "anything that exists but lacks explicit examples," where "exist" and "explicit example" are given any reasonable interpretations. Different interpretations yield different kinds of intangibles. We now mention four kinds of "intangibles" not covered by our AC/DC theory:

a. If we replace conventional set theory ZF + AC with Shelah's ZF + DC + BP, then many of the classic pathological objects of analysis cease to exist. However, they are replaced by a new collection of intangibles, for BP is also a nonconstructive postulate of existence. Indeed, it asserts that every subset of $\mathbb{R}$ can be represented in the form $M \triangle G$ for some meager M and open G, without giving any clue about how to find such a representation.

b. As we remarked in 6.3 and 6.4, the Axiom of Choice is not the only source of nonconstructive reasoning in conventional mathematics; two other sources are the Axiom of Foundation and the Law of the Excluded Middle. These sources yield their own collections of intangibles.

c. For some objects, even the Axiom of Choice is too weak to prove existence. The following example assumes familiarity with some concepts of algebra and functional analysis introduced later in this book. Let K be a compact Hausdorff space, and consider the algebra $C(K, \mathbb{C}) = \{\text{continuous functions from } K \text{ into } \mathbb{C}\}$. An **algebra norm** on $C(K, \mathbb{C})$ is a vector space norm $\| \quad \|$ with the additional property that $\|fg\| \leq \|f\| \, \|g\|$, where the product fg is defined pointwise: $(fg)(x) = f(x)g(x)$ for $x \in K$. The sup norm on $C(K, \mathbb{C})$ is complete, and any other complete algebra norm on $C(K, \mathbb{C})$ is equivalent to the sup norm. Do there exist any incomplete algebra norms on $C(K, \mathbb{C})$? This question cannot be answered in ZF+AC; it can be answered affirmatively or negatively depending on what further axioms we add to ZF+AC. For a full discussion, see Dales and Woodin [1987].

d. Lebesgue unmeasurable sets constitute a different sort of "intangible." Indeed, all the axiom systems mentioned in the last few paragraphs are equiconsistent with ZF, as we noted in 14.74. However, the axiom systems described in 14.75 have greater consistency strength. In particular, that is the case for ZF + DC + LM, where LM is the statement that "every subset of $\mathbb{R}$ is Lebesgue measurable." Thus,

- the existence of Lebesgue nonmeasurable sets can be proved in conventional set theory (ZF + AC) — and in fact we shall give such a proof in 21.22;
- empirically, it seems to be impossible to give an explicitly constructible example of such a nonmeasurable set — i.e., mathematicians have been unable (so far) to prove the existence of Lebesgue nonmeasurable sets using just ZF + DC, so the axiom system ZF + DC + LM seems to be consistent; but
- *proving* the impossibility of such an explicit construction (i.e., proving the consistency of ZF + DC + LM) requires a stronger assumption than our usual gospel of Con(ZF).

Part C

TOPOLOGY AND UNIFORMITY

This Page Intentionally Left Blank

Chapter 15

Topological Spaces

15.1. *Remarks.* We now resume the study of topological spaces, which we began in Chapters 5 and 9. Our study will also use some material from Chapter 7.

Many of the most basic properties of topological spaces are actually valid in the more general setting of pretopological spaces, so we shall begin in that setting. Admittedly, that setting is more general than one usually encounters in textbooks on topology. However, (i) the greater generality gives us some extra insights into topological spaces, as in 15.10; and (ii) it doesn't really require extra work, since the properties of pretopological spaces that we shall study are properties of topological spaces that we would have had to study anyway.

Pretopological Spaces

15.2. Let X be a set. A **neighborhood system for a pretopology** on X is a system of filters $\{\mathcal{N}(x) : x \in X\}$ such that x is a member of each member of $\mathcal{N}(x)$. A member of $\mathcal{N}(x)$ will be called a **neighborhood** of x; thus $\mathcal{N}(x)$ is called the **filter of neighborhoods** or the **neighborhood filter** at x. A set X equipped with such a neighborhood system is called a **pretopological space**. A convergence is defined on X as follows: for nets,

$$x_\alpha \to x \quad \Longleftrightarrow \quad x_\alpha \text{ is eventually in each neighborhood of } x$$

or, equivalently, for proper filters,

$$\mathcal{F} \to x \quad \Longleftrightarrow \quad \mathcal{F} \supseteq \mathcal{N}(x).$$

In particular, the neighborhood filter itself converges to x. Clearly, the convergence is centered and isotone, as defined in 7.34.

15.3. *Some basic properties.* Let X be a pretopological space. Then:

a. The convergence on X is Hausdorff (in the sense of 7.36) if and only if any two distinct points in X have disjoint neighborhoods.

Hints: The convergence has nonunique limits if and only if there exist distinct points $y, z \in X$ and a proper filter $\mathcal{F}$ such that $\mathcal{F} \to y$ and $\mathcal{F} \to z$ — that is, such that $\mathcal{F} \supseteq \mathcal{N}(y)$ and $\mathcal{F} \supseteq \mathcal{N}(z)$. In view of 7.18(E), that can happen if and only if every member of $\mathcal{N}(y)$ meets every member of $\mathcal{N}(z)$.

b. X has the "star property:"

> If $z \in X$ and (x_α) is a net that does not converge to z, then (x_α) has a subnet (y_β) that stays out of some neighborhood of z; hence no subnet of (y_β) converges to z.

Also, X has the "sequential star property:"

> If $z \in X$ and (x_n) is a sequence that does not converge to z, then (x_n) has a subsequence (y_n) that stays out of some neighborhood of z; hence no subsequence of (y_n) converges to z (and in fact, no subnet of (y_n) converges to z).

Hints: Assume the net $(x_\alpha : \alpha \in \mathbb{A})$ does not converge to z. Then there is some neighborhood N of z such that x_α is not eventually in N. Thus, $\mathbb{B} = \{\alpha \in \mathbb{A} : x_\alpha \notin N\}$ is a frequent subset of $\mathbb{A}$, and so the frequent subnet $(x_\alpha : \alpha \in \mathbb{B})$ has the desired properties. For the sequential result, recall from 7.16.d that any frequent subnet of a sequence is actually a subsequence.

Remark. We shall see in 21.43 that some complete lattices have order convergences that lack the sequential star property and therefore are not pretopological convergences.

c. If (x_α) and (y_β) are nets in a set $S \subseteq X$, both converging to some limit $z \in X$, then (x_α) and (y_β) are subnets of a single net in S which also converges to z.

Hint: Let the given nets have eventuality filters $\mathcal{F}$ and $\mathcal{G}$. Then $\mathcal{F} \cap \mathcal{G}$ is a proper filter that contains $\{S\} \cup \mathcal{N}(z)$.

d. Let $p : X \to Y$ be a mapping from one pretopological space into another. Then p is convergence-preserving (defined as in 7.33) if and only if p has this property:

> Whenever N is a neighborhood of $p(x)$ in Y, then $p^{-1}(N)$ is a neighborhood of x in X.

15.4. *Definitions.* Let $(X, \lim)$ be a pretopological convergence space, with neighborhood filters $\mathcal{N}(x)$. We define two maps from $\mathcal{P}(X)$ into itself, the **convergence closure operator** and the **convergence interior operator**, by

$$\begin{aligned} \mathrm{cl}(S) &= \{z \in X : S \text{ is a member of some filter that converges to } z\} \\ &= \{z \in X : \text{some net that converges to } z \text{ is eventually in } S\} \\ &= \{z \in X : S \text{ meets every neighborhood of } z\}, \qquad \text{and} \end{aligned}$$

$$\begin{aligned} \mathrm{int}(S) &= \{z \in X : S \text{ is a member of every filter that converges to } z\} \\ &= \{z \in X : \text{every net that converges to } z \text{ is eventually in } S\} \\ &= \{z \in X : S \text{ is a neighborhood of } z\}. \end{aligned}$$

Then the closure and interior are related by:

$$\mathrm{int}(X \setminus S) \quad = \quad X \setminus \mathrm{cl}(S).$$

Thus, closures and interiors are dual notions, in the sense of 1.7. In practice, however, closures and interiors are commonly used in different ways. Typically, the closure of a set

S is used if S itself does not contain enough points for some purpose — e.g., if S is not closed under some sort of operation of taking limits. The interior of a set S may be used as part of an argument to show that S or some other related set is nonempty, and thus to prove the existence of certain mathematical objects.

15.5. *Further properties of pretopological closures.*

a. $\mathrm{cl}(\varnothing) = \varnothing$, $S \subseteq \mathrm{cl}(S)$, and $S \subseteq T \Rightarrow \mathrm{cl}(S) \subseteq \mathrm{cl}(T)$.

b. $\mathrm{cl}(S \cup T) = \mathrm{cl}(S) \cup \mathrm{cl}(T)$. More generally, $\mathrm{cl}\left(\bigcup_{j=1}^n S_j\right) = \bigcup_{j=1}^n \mathrm{cl}(S_j)$ for any finite n. (For a slight generalization, see 16.23.c.)

c. $\mathrm{cl}(S) \setminus \mathrm{cl}(T) = \mathrm{cl}(S \setminus T) \setminus \mathrm{cl}(T) \subseteq \mathrm{cl}(S \setminus T)$.

Hint: From $S = (S \cap T) \cup (S \setminus T)$, we obtain $\mathrm{cl}(S) = \mathrm{cl}(S \cap T) \cup \mathrm{cl}(S \setminus T)$. Intersect both sides of this equation with the complement of $\mathrm{cl}(T)$, to obtain $\mathrm{cl}(S) \setminus \mathrm{cl}(T) = \mathrm{cl}(S \setminus T) \setminus \mathrm{cl}(T)$. This argument is taken from Kuratowski [1948].

Remarks. The closure of a pretopological space does not necessarily satisfy the *idempotence* condition $\mathrm{cl}(\mathrm{cl}(S)) = \mathrm{cl}(S)$. That is the one condition it still needs to be a Moore closure (see 4.5.a) or a *topological* closure (see 5.19, 15.6, 15.7, and 15.10(E)).

15.6. *Example: a pretopological closure that is not idempotent.* We exhibit a space in which $\mathrm{cl}(\mathrm{cl}(S))$ may differ from $\mathrm{cl}(S)$.

The underlying set will be $\mathbb{R}^2$. For each $(x, y) \in \mathbb{R}^2$ and each number $\varepsilon > 0$, let

$$K_\varepsilon(x,y) \quad = \quad \left\{(x,y') \in \mathbb{R}^2 \ : \ |y-y'| \le \varepsilon\right\} \bigcup \left\{(x',y) \in \mathbb{R}^2 \ : \ |x-x'| \le \varepsilon\right\}.$$

(This is a plus-shaped set centered at (x, y); each of its four arms has length ε.) Now define the neighborhood filter $\mathcal{N}(x, y)$ to be the filter $\{S \subseteq \mathbb{R}^2 \ : \ S \supseteq K_\varepsilon(x, y)$ for some $\varepsilon > 0\}$. The resulting convergence is as follows: A proper filter $\mathcal{F}$ on $\mathbb{R}^2$ converges to a limit (x, y) if and only if $K_\varepsilon(x, y) \in \mathcal{F}$ for every $\varepsilon > 0$. Equivalently, a net (x_α, y_α) in $\mathbb{R}^2$ converges to a limit (x, y) if and only if for each $\varepsilon > 0$, we have eventually $(x_\alpha, y_\alpha) \in K_\varepsilon(x, y)$. Finally, let $S = \{(x, y) \in \mathbb{R}^2 \ : \ x > 0 \text{ and } y > 0\}$. Verify that

$$\begin{array}{rcl} \mathrm{cl}(S) & = & \{(x,y) \in \mathbb{R}^2 \ : \ x \ge 0 \text{ and } y \ge 0 \text{ and } (x,y) \ne (0,0)\}, \\ \text{but} \quad \mathrm{cl}(\mathrm{cl}(S)) & = & \{(x,y) \in \mathbb{R}^2 \ : \ x \ge 0 \text{ and } y \ge 0\} \quad \text{is strictly larger.} \end{array}$$

Topological Spaces and Their Convergences

15.7. *Definitions.* Let $(X, \mathcal{T})$ be a topological space, as defined in 5.12. Recall from 5.16.a that a set $N \subseteq X$ is a *neighborhood* of a point p if $p \in G \subseteq N$ for some open set G. Then

$\mathcal{N}(p) = \{N : N \text{ is a neighborhood of } p\}$ is a filter, called the *neighborhood filter* at p. The $\mathcal{N}(p)$'s form a system of neighborhood filters for a pretopology, as defined in 15.2. Thus,

a topological space is a special type of pretopological space.

Convergence is defined as in 15.2. However, in a topological space $(X, \mathcal{T})$, the definition of convergence can be reformulated in terms of open sets. It is easy to show that for filters,

$\mathcal{F} \to z$ if and only if every open set containing z is a member of $\mathcal{F}$.

The equivalent condition for nets is:

$x_\alpha \to z$ if and only if for each open set G containing z, eventually $x_\alpha \in G$.

The convergence given by either of these rules is the **convergence determined by the topology**. Every topological space is understood to be equipped with this convergence, unless some other arrangement is specified. A convergence rule that can be determined by some topology is called a **topological convergence**.

Of course, any sequence is also a net, and so all our results for nets will apply to sequences. We shall work with sequences rather than with nets or filters whenever possible, since sequences are conceptually simpler.

15.8. *A few basic properties of topological convergences.* Let $(X, \mathcal{T})$ be a topological space. Show that

a. A set is open if and only if it is a neighborhood of each of its points.

b. A set $S \subseteq X$ is open if and only if it has this property for nets:

Whenever $x_\alpha \to z$ and $z \in S$, then eventually $x_\alpha \in S$.

An equivalent property in terms of filters is:

Whenever $\mathcal{F} \to z$ and $z \in S$, then $S \in \mathcal{F}$.

From either of these characterizations, we see that the topology $\mathcal{T}$ can be recovered from its convergence rule; two distinct topologies on X cannot have the same convergence rule. (However, not every convergence rule is determined by a topology. In 15.10 we shall characterize just which convergences are topological.)

c. A set $S \subseteq X$ is closed if and only if it has this property: If $x_\alpha \to z$ and eventually $x_\alpha \in S$, then $z \in S$. An equivalent property is: If $\mathcal{F} \to z$ and $S \in \mathcal{F}$, then $z \in S$.

d. In a topological space, the topological closure and interior (defined in 5.16.b and 5.16.c) are the same as the pretopological convergence closure and interior (defined in 15.4).

Remark. Any topological closure is idempotent (see 5.19), but not every pretopological convergence closure is idempotent (see 15.6). We shall see in 15.10(E) that a pretopological convergence closure is a topological closure *if and only if* it is idempotent.

15.9. *Elementary examples of convergence.* Some of the following examples are based on 5.15. Let X be any set. Then:

a. Recall that the indiscrete topology on X is $\{\varnothing, X\}$. With this topology, every net in X and every filter on X converge to every point of X.

b. Recall that the discrete topology on X is $\mathcal{P}(X)$. With this topology, a net (x_α) converges to a limit z if and only if eventually $x_\alpha = z$, and a filter $\mathcal{F}$ converges to a limit z if and only if $\mathcal{F}$ is the ultrafilter fixed at z.

c. Recall that the cofinite topology on X is

$$\mathcal{T} \quad = \quad \{S \subseteq X \ : \ \text{either } S \text{ is empty or } \complement S \text{ is finite}\}.$$

Show that if X has the cofinite topology and (x_n) is a sequence in X with the property that no point $\xi \in X$ appears in the sequence infinitely many times, then (x_n) converges to every point of X.

d. Let Y be a topological space, and let $X \subseteq Y$. Then a net (x_α) converges to a limit p in the relative topology on X (introduced in 5.15.e and 9.20) if and only if (i) $x_\alpha \to p$ in Y, and (ii) p and all the x_α's lie in the set X.

e. A net (x_α) converges to a limit z with respect to the topology generated by a collection of sets $\mathcal{G}$ if and only if

$$z \in G \in \mathcal{G} \quad \Rightarrow \quad \text{eventually } x_\alpha \in G.$$

In particular, if $z \notin \bigcup_{G \in \mathcal{G}} G$, then *every* net in X converges to z.

f. Let (X, d) be a pseudometric space; the topology of such spaces was described in 5.15.g. Show that a filter $\mathcal{F}$ converges to a limit p in this space if and only if $B_d(p, \varepsilon) \in \mathcal{F}$ for each $\varepsilon > 0$. Equivalently, a net (x_α) converges to p if and only if for each $\varepsilon > 0$ we have eventually $d(x_\alpha, p) < \varepsilon$.

In particular, let $\mathbb{R}$ have its pseudometric topology; then a net of real numbers (r_α) converges to a real number s if and only if for each $\varepsilon > 0$ we have eventually $|r_\alpha - s| < \varepsilon$.

Observe that the convergence in any pseudometric space (X, d) can be characterized in terms of convergence of distances, which are real numbers:

$$x_\alpha \to p \ \text{ in } \ (X, d) \qquad \Longleftrightarrow \qquad d(x_\alpha, p) \to 0 \ \text{ in } \ \mathbb{R}.$$

15.10. Theorem characterizing topological convergences (optional). Let X be a convergence space whose convergence is centered and isotone (as defined in 7.34). Then the following conditions are equivalent.

(A) The convergence on X is topological — i.e., given by a topology.

(B) **(Iterated Net Condition.)** Let $(y_\delta : \delta \in D)$ be a net in X converging to a limit z. For each $\delta \in D$, let $(x_\varepsilon^\delta : \varepsilon \in E_\delta)$ be a net in X converging to y_δ. Let $F = \prod_{\delta \in D} E_\delta$ have the product ordering, and let $D \times F$ have the product ordering. Then the net $(x_{f(\delta)}^\delta : (\delta, f) \in D \times F)$ converges to z.

(C) **(Cook-Fischer Iterated Filter Condition.)** Let $\mathcal{G}$ be a filter on a set I, and let $v : I \to X$ be some function. Assume the filterbase $v(\mathcal{G}) = \{v(G) :$

$G \in \mathcal{G}\}$ converges to some point z in X. For each $i \in I$, suppose $s(i)$ is a filter on X converging to $v(i)$. Then the filter $\mathcal{K} = \bigcup_{G\in\mathcal{G}} \bigcap_{i\in G} s(i)$ converges to z.

(D) **(Kowalsky's Conditions.)** The convergence is pretopological (as defined in 15.2). Furthermore, suppose $\mathcal{G}$ is a filter on X converging to some point z in X. For each $x \in X$, assume that $s(x)$ is a filter on X converging to x. Then the filter $\mathcal{K} = \bigcup_{G\in\mathcal{G}} \bigcap_{i\in G} s(i)$ converges to z.

(E) **(Gherman's Conditions.)** The convergence is pretopological. Moreover, the closure operator defined in 15.4 is idempotent — i.e., it satisfies $\mathrm{cl}(\mathrm{cl}(S)) = \mathrm{cl}(S)$.

Bibliographical remarks. Earlier, more complicated versions of parts of this theorem were given by Kelley [1955/1975] and Cook and Fischer [1967]; those versions assumed a "star property" like that in 15.3.b. The star property assumption was removed independently, in different fashions, by Aarnes and Andenæs [1972] and Gherman [1980]. It should be emphasized that our definition of "isotone" is based on Aarnes-Andenæs subnets — i.e., we assume condition 7.31(*). Kelley studied nets without assuming that condition and without considering filters. With his formulation the star property cannot be omitted; that was shown by Aarnes and Andenæs [1972].

Hints for (A) $\Rightarrow$ (B). Let N be an open neighborhood of z; it suffices to show that eventually $x^{\delta}_{f(\delta)} \in N$.

Outline of (B) $\Rightarrow$ (C). We shall begin by constructing a net that is somewhat like the canonical net of $v(\mathcal{G})$, but is also parametrized by elements of I. Let $\mathcal{G}$ be ordered by reverse inclusion (see 7.4), let X and I have the universal ordering (see 3.9.g), and let products have product ordering. Let

$$D \quad = \quad \{(i,G) \in I \times \mathcal{G} \quad : \quad i \in G\}.$$

Then D is a frequent subset of $I \times \mathcal{G}$, hence directed. For $\delta = (i,G)$ in D, let $y_\delta = v(i)$. The net $(y_\delta : \delta \in D)$ converges to z since its eventuality filter includes $v(\mathcal{G})$.

For each $\delta = (i,G)$ in D, let $\mathcal{S}_\delta = s(i)$; thus $\mathcal{S}_\delta$ is a filter on X that converges to $v(i) = y_\delta$. The canonical net of $\mathcal{S}_\delta$ is

$$\left(x^\delta_\varepsilon \;:\; \varepsilon \in E_\delta\right), \quad \text{where} \quad E_\delta \;=\; \{(w,S) \in X \times \mathcal{S}_\delta \;:\; w \in S\} \quad \text{and} \quad x^\delta_{(w,S)} = w;$$

this net also converges to $v(i) = y_\delta$.

Define $F = \prod_{\delta\in D} E_\delta$ as in the statement of (B). Then the net $(x^\delta_{f(\delta)} : (\delta, f) \in D \times F)$ converges to z by the assumed condition (B). Hence its eventuality filter $\mathcal{E}$ also converges to z. We wish to show that $\mathcal{K} \to z$; it suffices to show that $\mathcal{K} \supseteq \mathcal{E}$.

Let $A \in \mathcal{E}$; we are to show that $A \in \mathcal{K}$. Since A is an eventual set of $(x^\delta_{f(\delta)} : (\delta, f) \in D \times F)$, there is some $\delta_A \in D$ and $f_A \in F$ such that $\delta \succcurlyeq \delta_A,\ f \succcurlyeq f_A \;\Rightarrow\; x^\delta_{f(\delta)} \in A$. Say $\delta_A = (i_A, G_A)$.

Temporarily fix any $i \in G_A$, and let $\delta = (i, G_A)$. Then δ is a member of D that satisfies $\delta \succcurlyeq \delta_A$ and therefore $x^\delta_{f(\delta)} \in A$ for all $f \succcurlyeq f_A$. Thus $x^\delta_\varepsilon \in A$ for all $\varepsilon \in E_\delta$ such that

$\varepsilon \succcurlyeq f_A(\delta)$. Thus the net $\left(x_\varepsilon^\delta : \varepsilon \in E_\delta\right)$ is eventually in A, so A is a member of that net's eventuality filter, which is $\mathcal{S}_\delta = s(i)$.

Thus $A \in \bigcap_{i \in G_A} s(i) \subseteq \mathcal{K}$.

Outline of (C) $\Rightarrow$ (D). Easily, condition (C) implies the iterated filter condition in (D). It suffices to show (C) implies the convergence is pretopological. Fix any $z \in X$; let $\mathcal{N}(z)$ be the intersection of all the filters that converge to z; we wish to show $\mathcal{N}(z) \to z$. Let $I = \{\text{filters on } X \text{ that converge to } z\}$ and $G = \{I\}$. Define $s(i) = i$ and $v(i) = z$ for all $i \in I$. The hypotheses of the Cook-Fischer condition are satisfied, and therefore $\mathcal{K} \to z$. Unwinding the notation, we find that $\mathcal{K} = \mathcal{N}(z)$.

Outline of (D) $\Rightarrow$ (E). Let $I = \mathrm{cl}(S)$; we wish to show $\mathrm{cl}(I) = I$. Clearly, $I \subseteq \mathrm{cl}(I)$. Let $z \in \mathrm{cl}(I)$; it suffices to show $z \in I$.

When $x \in I$, then S meets every element of $\mathcal{N}(x)$; let $\mathcal{N}(x) \vee \{S\}$ denote the filter generated by $\mathcal{N}(x) \cup \{S\}$ (see 5.5.i). Define $s : X \to \{\text{filters on } X\}$ as follows:

$$s(x) \quad = \quad \begin{cases} \mathcal{N}(x) & \text{if } x \notin I \\ \mathcal{N}(x) \vee \{S\} & \text{if } x \in I. \end{cases}$$

In either case, $s(x)$ is a filter that converges to x.

Since $z \in \mathrm{cl}(I)$, some filter $\mathcal{G}$ converges to z and contains I. By the assumed condition (D), Kowalsky's iterated filter $\mathcal{K} = \bigcup_{G \in \mathcal{G}} \bigcap_{i \in G} s(i)$ converges to z. Since $I \in \mathcal{G}$, we have $S \in \bigcap_{i \in I} s(i) \subseteq \mathcal{K}$. Since $S \in \mathcal{K}$ and $\mathcal{K} \to z$, we have $z \in \mathrm{cl}(S) = I$.

Hint for (E) $\Rightarrow$ (A): Let $\xrightarrow{o}$ denote the convergence originally given on X; we are to prove that $\xrightarrow{o}$ is a topological convergence. Condition (E) tells us that the convergence closure operator defined in 15.4 satisfies Kuratowski's axioms 5.19 and thus is the closure operator for some topology $\mathcal{T}$ on X. Since $\mathrm{int}(\complement S) = \complement(\mathrm{cl}(S))$, the convergence interior operator defined in 15.4 is the interior operator for that topology $\mathcal{T}$. Let $\xrightarrow{\mathcal{T}}$ be the convergence of that topology. Now, both $\xrightarrow{o}$ and $\xrightarrow{\mathcal{T}}$ are pretopological, and they have the same interior operator, hence the same neighborhood filters. Being pretopological, they satisfy

$$\mathcal{F} \xrightarrow{o} z \qquad \Longleftrightarrow \qquad \mathcal{F} \supseteq \mathcal{N}(z) \qquad \Longleftrightarrow \qquad \mathcal{F} \xrightarrow{\mathcal{T}} z.$$

That is, the two convergences are the same. Hence $\xrightarrow{o}$ is a topological convergence.

More about Topological Closures

15.11. The closure operator is isotone — i.e., it satisfies $S \subseteq T \Rightarrow \mathrm{cl}(S) \subseteq \mathrm{cl}(T)$ and therefore it satisfies

$$\bigcap_{\lambda \in \Lambda} \mathrm{cl}(S_\lambda) \supseteq \mathrm{cl}\left(\bigcap_{\lambda \in \Lambda} S_\lambda\right) \qquad \text{and} \qquad \bigcup_{\lambda \in \Lambda} \mathrm{cl}(S_\lambda) \subseteq \mathrm{cl}\left(\bigcup_{\lambda \in \Lambda} S_\lambda\right).$$

as we noted in 4.29.c. Neither of these inclusions is necessarily reversible, as we shall now show with simple examples. For both examples, take $X = \mathbb{R}$ with its usual topology; let $\mathbb{Q} = \{\text{rational numbers}\}$. Then

$$\bigcap_{\lambda \in \mathbb{R}} \mathrm{cl}\,(S_\lambda) \quad \supsetneqq \quad \mathrm{cl}\left(\bigcap_{\lambda \in \mathbb{R}} S_\lambda\right) \qquad \text{if } S_\lambda = \mathbb{R} \setminus \{\lambda\};$$

$$\bigcup_{\lambda \in \mathbb{Q}} \mathrm{cl}\,(S_\lambda) \quad \subsetneqq \quad \mathrm{cl}\left(\bigcup_{\lambda \in \mathbb{Q}} S_\lambda\right) \qquad \text{if } S_\lambda = \{\lambda\}.$$

15.12. *Relativization of closures.* Let Y be a topological space, let $X \subseteq Y$, and let X be equipped with the relative topology. Let cl_Y and cl_X denote closures in the topology of Y and the topology of X. Then for any set $S \subseteq X$, we have

$$\mathrm{cl}_X(S) \quad = \quad X \cap \mathrm{cl}_Y(S).$$

15.13. A subset S is **dense** in a topological space X if $\mathrm{cl}(S) = X$. A topological space X is **separable** if it has a countable dense subset. Show that

- **a.** A set $S \subseteq X$ is dense if and only if it meets every nonempty open subset of X.
- **b.** If G is an open subset of a topological space X, and Y is a dense subset of X, then $G \subseteq \mathrm{cl}(G \cap Y)$.
- **c.** The intersection of finitely many open dense sets is open and dense.
- **d.** Any subset of a separable metric space is separable.
- **e.** Any open subset of any separable space is separable.
- **f.** However, separability is not a hereditary property — i.e., not every subspace of a separable space is necessarily separable.

  Example. Let ξ be some particular member of an uncountable set X (for instance, take $0 \in \mathbb{R}$). Let X be given the topology $\mathcal{T} = \{S \subseteq X : \xi \in S \text{ or } S = \varnothing\}$. Show that X is separable, but the relative topology on $X \setminus \{\xi\}$ is not separable.
- **g.** Let (X, d) be a separable metric space. Then there is a sequence (x_n) in X with the property that every point in X is the limit of some subsequence of (x_n). In fact, we can choose the subsequence canonically (i.e., without any arbitrary choices).

  Hints: Repetitions are permitted. If (u_k) is a countable dense set, let (x_n) be the sequence

  $$u_1, \quad u_1, u_2, \quad u_1, u_2, u_3, \quad u_1, u_2, u_3, u_4, \quad u_1, u_2, u_3, u_4, u_5, \quad \ldots$$

  Given any point $p \in X$, we can choose a subsequence (x_{n_i}) of (x_n) canonically as follows: Take $n_1 = 1$. Thereafter, let n_i be the first integer greater than n_{i-1} that satisfies $d(x_{n_i}, p) < \frac{1}{i}$.

Continuity

15.14. *Definition.* Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be any topological spaces, let $x_0 \in X$, and let $f : X \to Y$ be a function. Then the following conditions are equivalent; if any (hence all) are satisfied, we say f is **continuous at the point x_0**.

(A) f is "convergence-preserving" at x_0. That is, whenever (x_α) is a net converging to a limit x_0 in X, then also $f(x_\alpha) \to f(x_0)$ in Y. (Compare this with condition 15.14(E), which is intuitively similar but removes the concept of "time" from our convergence.)

(B) Whenever $\mathcal{B}$ is a filterbase converging to a limit x_0 in X, then the filterbase $\mathcal{B} = \{f(B) \ : \ B \in \mathcal{B}\}$ converges to the limit $f(x_0)$ in Y.

(C) The inverse image of each neighborhood of $f(x_0)$ is a neighborhood of x_0.

If the topologies on X and Y are given by gauges D and E, then an equivalent condition is:

(D) For each pseudometric $e \in E$ and each number $\varepsilon > 0$, there exists some finite set $D' \subseteq D$ and some number $\delta > 0$ such that

$$\max_{d \in D'} d(x_0, x) \ < \ \delta \qquad \Rightarrow \qquad e\left(\varphi(x_0), \varphi(x)\right) \ < \ \varepsilon.$$

We emphasize that the choice of δ and D' may depend on all of ε, e, and x_0, but not on x; this should be contrasted with the definition of *uniform continuity* in 18.8(C). Of course, the preceding condition simplifies slightly if X is a pseudometric space with singleton gauge $D = \{d\}$ — or, more generally, if D is a gauge that is directed (as defined in 4.4.c).

If $X = Y = \mathbb{R}$, and $\mathbb{H} = {}^*\mathbb{R}$ is the hyperreal line constructed as in 10.20.a, then the conditions above are also equivalent to this condition:

(E) Whenever ξ is a hyperreal number that is infinitely close to x_0 (defined in 10.18.c), then ${}^*f(\xi)$ is infinitely close to $f(x_0)$. (Here ${}^*f : \mathbb{H} \to \mathbb{H}$ is defined as in 9.49. Compare this condition with 15.14(A).)

15.15. Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be any topological spaces, and let $f : X \to Y$ be a function. Then the following conditions are equivalent; if any (hence all) are satisfied, we say f is **continuous**.

(A) Inverse images of open sets are open; that is, $T \in \mathcal{T} \ \Rightarrow \ f^{-1}(T) \in \mathcal{S}$. (This definition of continuity was used in 9.8.)

(B) The inverse image of each closed set is closed.

(C) For each set $S \subseteq X$, we have $f(\mathrm{cl}(S)) \subseteq \mathrm{cl}(f(S))$.

(D) For each set $T \subseteq Y$, we have $\mathrm{cl}(f^{-1}(T)) \subseteq f^{-1}(\mathrm{cl}(T))$.

(E) f is continuous at each point x_0 in X, as defined in several equivalent ways in 15.14. In particular, using the formulation in 15.14(A), we obtain the condition that f is convergence-preserving — i.e., whenever (x_α) is a net converging to any limit x_0 in X, then also $f(x_\alpha) \to f(x_0)$ in Y. (This generalizes 7.33 and 15.3.d.)

A **homeomorphism** from one topological space to another is a continuous bijection whose inverse is also continuous. Thus it is an isomorphism in the category of topological spaces.

15.16. *Additional characterization.* A mapping $f : X \to Y$, from one topological space into another, is continuous if and only if f is "locally continuous" in the following sense: each point in X has a neighborhood N such that the restriction $f\big|_N : N \to Y$ is continuous (where N is given the relative topology, defined in 5.15.e).

15.17. *Degenerate examples of continuity.*

a. Any map from a topological space into an indiscrete space is continuous.

b. Any map from a discrete space into a topological space is continuous.

15.18. *Exercise.* If $f : X \to Y$ is a continuous map from one topological space into another and $S \subseteq X$ is *connected* (defined in 5.12), then $f(S) \subseteq Y$ is also connected.

15.19. *Remarks.* To say that f is continuous is to say, roughly, that f carries any set of points near x to a set of points near $f(x)$. Here are two ways in which this notion is important:

(a) In many applied mathematics problems, some data x is based on a measurement, and a decision or consequence $f(x)$ is then computed. Any measurement of x inevitably involves small errors. If f is continuous, then at least the small errors in x will not have catastrophic effects on the decisions and consequences $f(x)$. On the other hand, if f is discontinuous, then even a tiny error in x may make our computed value of $f(x)$ highly erroneous, and thus the computation may be altogether worthless.

(b) In some problems we may *intentionally* introduce an error: x may represent a very difficult problem, while x' may represent a "nearby" problem that is much easier to solve. For instance,

- perhaps x is π, while x' is 3.1416; or
- perhaps x is a complicated function, while x' is a polynomial or other simple function that approximates x; or
- perhaps x stands for a complicated differential equation, while x' denotes a similar equation obtained by dropping one troublesome nonlinear term that, hopefully, only represents a small quantity.

Let $y = f(x)$ be the solution that we are not able to find, and let $y' = f(x')$ be the solution that we are able to find. If f is continuous at x, then y' should be near to y. Of course, in some of the examples cited — functions, differential equations, etc. — the appropriate notion of "near" may be quite complicated, hence the relevant topologies may be quite complicated.

For real-valued functions of a real variable, this informal definition of continuity is sometimes suggested in calculus books: "A function is continuous if its graph is an unbroken curve — i.e., if its graph can be drawn without lifting the pencil from the page." But this presupposes that the function can be drawn at all. Some functions are just too pathological to be drawn with any reasonable degree of accuracy; see for instance 25.19.

15.20. We now caution the reader about some subtle distinctions concerning continuity.

Let $f : X \to Y$ be some mapping (not necessarily continuous) from one topological space into another, and let $S \subseteq X$. The phrase "f is continuous on S" has two possible interpretations:

(i) f is continuous at each point of S, or

(ii) the restriction $f\big|_S$ is continuous.

These are *not* the same! It is easy to prove that (i) $\Rightarrow$ (ii). When the set S is *open*, then we can easily prove that (ii) $\Rightarrow$ (i) also (*exercise*). However, in general (ii) does not imply (i). Indeed, this is obvious in one extreme case: if S is just a singleton, then (ii) is always true, but if f is not continuous then we can choose the singleton so that (i) is false.

In deciding whether a function is continuous, we only need to consider the behavior of that function *on its domain*, not elsewhere. Thus, the sign function (defined as in 2.2.c) is not continuous on $\mathbb{R}$, but its restriction to $\mathbb{R} \setminus \{0\}$ is continuous — i.e., the function

$$f(x) \;=\; \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{$*$}$$

is continuous, since its domain is $\mathbb{R}\setminus\{0\}$.

The definition of "continuity" that we have given in this chapter is the standard one. It (or an equivalent definition) is used by all research mathematicians who work with continuity. Unfortunately, some calculus textbooks do not conform to this usage. These books concern themselves only with real-valued functions defined on subintervals of $\mathbb{R}$, and so they use ad hoc definitions that are easily manageable in that context. They may go astray when dealing with functions defined on more complicated sets. For instance, some calculus textbooks would assert that the function f defined in $(*)$ is not continuous, *because* it is undefined at 0. The student who wishes to proceed to higher mathematics will first need to *unlearn* the not-quite-correct definitions of continuity given in these calculus books.

15.21. *Definitions of one-sided limits and one-sided continuity.* Let X be a subinterval of $[-\infty, +\infty]$, equipped with its relative topology; let Y be any Hausdorff topological space (or more generally, any Hausdorff pretopological space). Let $f : X \to Y$ be some function. Let $x_0 \in X$; assume that X contains some other points higher than x_0. We say that a point $y_0 \in Y$ is the **limit from the right** of f at x_0, or y_0 is the **right-hand limit** of f at x_0, if it satisfies this condition:

> for each neighborhood N of y_0, there is some number $\delta > 0$ such that $x \in (x_0, x_0 + \delta) \;\Rightarrow\; f(x) \in N$

or, equivalently, this condition:

> whenever (x_n) is a decreasing sequence in X that converges to x_0, then the sequence $(f(x_n))$ converges to y_0 in Y.

(Another equivalent condition is obtained if we use nets instead of sequences.) These conditions can be abbreviated

$$y_0 = \lim_{x \downarrow x_0} f(x) \qquad \text{or} \qquad y_0 = \lim_{x \to x_0+} f(x) \qquad \text{or} \qquad y_0 = f(x_0+).$$

We say f is **continuous from the right** at x_0, or **right-continuous** at x_0, if $f(x_0) = f(x_0+)$.

Analogously, we may define the **limit from the left**, or the **left-hand limit**, written $\lim_{x \uparrow x_0} f(x)$ or $\lim_{x \to x_0-} f(x)$ or $f(x_0-)$; we say f is **continuous from the left** at x_0 or **left-continuous** at x_0 if $f(x_0) = f(x_0-)$.

The right- and left-hand limits are called **one-sided limits**.

Exercises.

a. Suppose the domain of f is an interval $[a, b]$. Then:

(i) $y_0 = \lim_{x \to a} f(x)$ means the same thing as $y_0 = \lim_{x \downarrow a} f(x)$, and

(ii) f is continuous at a if and only if f is right-continuous at a.

Also show analogous results at b, with left-hand limits and left-continuity.

b. Suppose the interval X contains some points below x_0 and also some points above x_0. Then:

(i) $y_0 = \lim_{x \to x_0} f(x)$ if and only if both the one-sided limits exist and are equal to y_0, and

(ii) f is continuous at x_0 if and only if f is both right- and left-continuous at x_0.

c. Suppose $J \subseteq \mathbb{R}$ is an open interval and $g : J \to \mathbb{R}$ is an increasing function. Then g has one-sided limits

$$g(t+) = \lim_{s \downarrow t} g(s), \qquad g(t-) = \lim_{s \uparrow t} g(s)$$

at every t, and g is discontinuous at at most countably many points of J.

Hints: See 7.40.c, to prove that $g(t+)$ and $g(t-)$ both exist. For the cardinality result use an argument similar to 10.40.

d. Let $J \subseteq \mathbb{R}$ be an interval with $\sup(J) \notin J$. Assume that the right-hand limit $f(x+) = \lim_{u \downarrow x} f(u)$ exists at every $x \in J$. Define $g : J \to \mathbb{R}$ by setting $g(x) = f(x+)$ for all x. Show that g is right-continuous on J and $g = f$ at every point where f is right-continuous.

15.22. Let X be a topological space, and let $f : X \to [-\infty, +\infty]$ be some function. Then the following conditions are equivalent; if any (hence all) of them are satisfied, we say f is **lower semicontinuous** (abbreviated **l.s.c.**):

(A) $f(x) \leq \liminf f(x_\alpha)$ whenever $x_\alpha \to x$ in X.

(B) For each $b \in [-\infty, +\infty]$, the set $\{x \in X : f(x) > b\}$ is open — i.e., the set $\{x \in X : f(x) \leq b\}$ is closed.

(C) For each $b \in \mathbb{R}$, the set $\{x \in X : f(x) > b\}$ is open — i.e., the set $\{x \in X : f(x) \leq b\}$ is closed.

Proof of equivalence. The proofs of (A) $\Rightarrow$ (B) $\Rightarrow$ (C) are easy. It suffices to show (C) $\Rightarrow$ (A). Suppose that (C) holds but $f(x) > \liminf x_\alpha$ for some net (x_α) converging to some limit x in X. Then (regardless of whether one, both, or neither of the numbers $f(x)$, $\liminf x_\alpha$ is finite) there is some finite number r such that $f(x) > r > \liminf x_\alpha$. By (C) the set $\{u \in X : f(u) \leq r\}$ is closed. It contains all the x_α's after some α_0, but not the point x, contradicting the fact that $x_\alpha \to x$.

Dual notion. Let X be a topological space, and let $f : X \to [-\infty, +\infty]$ be some function. Then the following conditions are equivalent; if any (hence all) of them are satisfied, we say f is **upper semicontinuous** (abbreviated **u.s.c.**):

(A) $f(x) \geq \limsup f(x_\alpha)$ whenever $x_\alpha \to x$ in X.

(B) For each $b \in [-\infty, +\infty]$, the set $\{x \in X : f(x) < b\}$ is open — i.e., the set $\{x \in X : f(x) \geq b\}$ is closed.

(C) For each $b \in \mathbb{R}$, the set $\{x \in X : f(x) < b\}$ is open — i.e., the set $\{x \in X : f(x) \geq b\}$ is closed.

Remarks. Semicontinuity is a sort of "almost-continuity" condition. It can often be used in proofs in place of continuity, especially when limits are replaced with upper or lower limits; see 7.46.

15.23. *Further properties of semicontinuity.* Let X be a topological space.

a. A function $f : X \to [-\infty, +\infty]$ is l.s.c. if and only if $-f$ is u.s.c.

b. A function $f : X \to [-\infty, +\infty]$ is continuous if and only if it is both l.s.c. and u.s.c.

c. Any pointwise infimum of continuous functions on X (or, more generally, any pointwise infimum of u.s.c. functions) is u.s.c.

d. Any pointwise supremum of continuous functions on X (or, more generally, any pointwise supremum of l.s.c. functions) is l.s.c.

A partial converse to the last result is given in 16.16(D). Compare also 12.21.d.

More about Initial and Product Topologies

15.24. *Convergence in initial topologies.* Let $(X, \mathcal{S})$ have the initial topology determined by some mappings $\varphi_\lambda : X \to (Y_\lambda, \mathcal{T}_\lambda)$ — that is, suppose $\mathcal{S}$ is the weakest topology that makes all the φ_λ's continuous (see 9.15 and 9.16). (It is also sometimes known as the *weak topology*.) Show that

a. A set $N \subseteq X$ is a neighborhood of a point $p \in X$ if and only if there exists a finite family of sets $T_j \in \mathcal{T}_{\lambda_j}$ $(j = 1, 2, 3, \ldots, m)$ such that $p \in \bigcap_{j=1}^{m} \varphi_{\lambda_j}^{-1}(T_j) \subseteq N$.

(*Hint*: Use the characterization of neighborhoods in 5.23.b and the characterization of a generating collection of sets in 9.16.)

b. A net (x_α) converges to a limit p in $(X, \mathcal{S})$ if and only if $\varphi_\lambda(x_\alpha) \to \varphi_\lambda(p)$ in each $(Y_\lambda, \mathcal{T}_\lambda)$.

15.25. Following are some important instances of convergence in initial topologies.

a. Convergence in the relative topology was characterized in 15.9.d.

b. If $X = \prod_{\lambda \in \Lambda} Y_\lambda$ is a product of topological spaces, with the product topology, then members of X may be viewed as functions f defined on Λ, satisfying $f(\lambda) \in Y_\lambda$ for each λ. Then $f_\alpha \to f$ in X if and only if $f_\alpha(\lambda) \to f(\lambda)$ in Y_λ for each λ. Convergence in the product topology is sometimes called **pointwise convergence**, or **componentwise convergence**, or **coordinatewise convergence**.

c. Let $\mathcal{S}$ be the supremum of a collection of topologies $\mathcal{T}_\lambda$ on a set X. (This is the initial topology obtained by taking all the φ_λ's equal to the identity map.) Then $x_\alpha \to p$ in $(X, \mathcal{S})$ if and only if $x_\alpha \to p$ in each $(X, \mathcal{T}_\lambda)$.

d. Let (X, D) be a gauge space. The gauge topology $\mathcal{T}_D$ is the supremum of the pseudometric topologies $\{\mathcal{T}_d : d \in D\}$. Hence $x_\alpha \to p$ in (X, D) if and only if $d(x_\alpha, p) \to 0$ for each $d \in D$. (We emphasize that this condition does *not* say $\sup_{d \in D} d(x_\alpha, p) \to 0$.)

15.26. Let $X = \prod_{\lambda \in \Lambda} Y_\lambda$ be a product of topological spaces, with the product topology. Show that

a. A set $S \subseteq X$ is a neighborhood of a point $u \in X$ (in the product topology) if and only if there exist a finite set $L \subseteq \Lambda$ and open sets $T_\lambda \subseteq Y_\lambda$ ($\lambda \in L$) such that

$$u \quad \in \quad \left(\prod_{\lambda \in L} T_\lambda\right) \times \left(\prod_{\lambda \in \Lambda \setminus L} Y_\lambda\right) \quad \subseteq \quad S.$$

(This is a special case of the neighborhood characterization at the beginning of 15.24.) Thus, S must be "fat" in all but finitely many directions.

b. Use the preceding characterization of neighborhoods together with 15.8.a to show that each of the coordinate projection mappings $\pi_\lambda : X \to Y_\lambda$ is an *open mapping* — i.e., show that if $S \subseteq X$ is an open set, then $\pi_\lambda(S)$ is an open subset of Y_λ.

c. The coordinate projections need not be *closed* mappings — i.e., if $S \subseteq X$ is a closed set, it does not necessarily follow that $\pi_\lambda(S)$ is a closed subset of Y_λ. For instance, when $\mathbb{R}^2$ has its usual topology, then $\{(x, y) \in \mathbb{R}^2 : xy \geq 1\}$ is a closed set, but its projection onto the first coordinate is $\{x \in \mathbb{R} : x \neq 0\}$, which is not closed.

d. Let $\varphi : X \to Y$ be any mapping from one set to another (without any topology or other structure necessarily specified). If S is a topological space, then we can define a mapping $S^\varphi : S^Y \to S^X$ by setting

$$S^\varphi(\lambda) = \lambda \circ \varphi \; : \; X \to S \qquad \text{for any} \qquad \lambda : Y \to S \quad \text{in} \quad S^Y.$$

Show that S^φ is continuous, when S^Y and S^X are equipped with their product topologies. *Hint*: 15.25.b.

e. Let Ω be a set. Identify each set $S \subseteq \Omega$ with its characteristic function $1_S : \Omega \to \{0,1\}$; then the 1_S's are members of 2^Ω. Show that the order convergence of sets S_α described in 7.48 is the same as the convergence of the 1_{S_α}'s given by the product topology on 2^Ω (where $2 = \{0,1\}$ has the discrete topology, as usual).

15.27. Theorem. If $\{X_\alpha : \alpha \in A\}$ is a collection of separable topological spaces with $\mathrm{card}(A) \leq \mathrm{card}(\mathbb{R})$, then $\prod_{\alpha\in A} X_\alpha$ (with the product topology) is separable.

Proof. The product topology is not affected if we replace the index set A with another index set of the same cardinality; hence we may assume $A \subseteq \mathbb{R}$. Let $P = \prod_{\alpha\in A} X_\alpha$. For each $\alpha \in A$, let $x_1^\alpha, x_2^\alpha, x_3^\alpha, \ldots$ be a dense sequence in X_α. Let $\mathcal{J}$ be the collection of all closed subintervals of $\mathbb{R}$ that have rational endpoints and positive, finite length. For each positive integer m, each finite sequence $J_1, J_2, \ldots, J_m$ of *disjoint* members of $\mathcal{J}$, each finite sequence $n_1, n_2, \ldots, n_m$ of positive integers, and each $\alpha \in A$, define

$$p_{J_1,\ldots,J_m,n_1,\ldots,n_m}(\alpha) \quad = \quad \begin{cases} x_{n_1}^\alpha & \text{if } \alpha \in J_1 \\ x_{n_2}^\alpha & \text{if } \alpha \in J_2 \\ \vdots & \\ x_{n_m}^\alpha & \text{if } \alpha \in J_m \\ x_1^\alpha & \text{if } \alpha \in A \setminus \bigcup_{i=1}^m J_i. \end{cases}$$

The function $p_{J_1,\ldots,J_m,n_1,\ldots,n_m}$ maps each α to some member of X_α, so $p_{J_1,\ldots,J_m,n_1,\ldots,n_m}$ is actually a member of P. There are only countably many such functions p, since $\mathcal{J}$ and $\mathbb{N}$ are countable. It suffices to show the functions p are dense in P. Let any nonempty open set $G \subseteq P$ be given; it suffices to show G contains one of the functions p. For each $\beta \in A$, let $\pi_\beta : P \to X_\beta$ be the βth coordinate projection; by 15.26.a we know that $G \supseteq \bigcap_{i=1}^m \pi_{\alpha_i}^{-1}(V_i)$ for some distinct numbers $\alpha_1, \ldots, \alpha_m \in A$ and some nonempty open sets $V_i \subseteq X_{\alpha_i}$. Choose disjoint sets $J_1, J_2, \ldots, J_m \in \mathcal{J}$ such that $\alpha_i \in J_i$, and choose some numbers n_i such that $x_{n_i}^{\alpha_i} \in V_i$. Unwinding all the notation, verify that $p_{J_1,\ldots,J_m,n_1,\ldots,n_m}$ is a member of G. This proof follows Willard [1970]; further references are also given by Willard.

15.28. Let $X = \prod_{\lambda\in\Lambda} Y_\lambda$ be a product of topological spaces, equipped with the product topology. A point in X may be seen as an "ordered Λ-tuple" $(x_\alpha, x_\beta, x_\gamma, \ldots)$ (see 1.32). Hence a mapping $h : X \to Z$, from X into some other topological space Z, may be written as $z = h(x_\alpha, x_\beta, x_\gamma, \ldots)$. Ordinary continuity from X (with the product topology) to Z is sometimes called **joint continuity**, to emphasize that the variables $x_\alpha, x_\beta, x_\gamma, \ldots$ are being considered together, not separately. A slightly weaker condition is *separate continuity*; we say that the mapping $z = h(x_\alpha, x_\beta, x_\gamma, \ldots)$ is **separately continuous** if z is continuous as a function of each one of the arguments x_λ whenever all the other arguments are held fixed.

Examples.

a. Let $(X, \mathcal{T})$ be a topological space, and let $d : X \times X \to [0, +\infty)$ be a pseudometric on X (not necessarily associated with the topology $\mathcal{T}$). Let $X \times X$ have the product topology, and let $\mathbb{R}$ have its usual topology. Then d is separately continuous from $X \times X$ into $\mathbb{R}$ if and only if d is jointly continuous. (*Hint*: 2.12.e.) Thus, the phrase "a continuous pseudometric" is not really ambiguous.

b. Define $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$f(x,y) \quad = \quad \begin{cases} \dfrac{xy}{x^2+y^2} & \text{when}(x,y) \neq (0,0) \\ 0 & \text{when } (x,y) = (0,0). \end{cases}$$

Show f is separately continuous but not jointly continuous. *Hints*: 15.14(A) and 15.25.b.

c. In 1.17 we defined the extended real numbers and their arithmetic operations. For many purposes — particularly in the theory of measure and integration — it is convenient to define the product of 0 and $\pm\infty$ to be 0. That causes confusion for some students, because it seems to be contrary to what they would expect from their experience with calculus. We shall now take a closer look at this.

Most of the multiplication rules would make multiplication a jointly continuous operation from $[-\infty, +\infty] \times [-\infty, +\infty]$ into $[-\infty, +\infty]$. That is, if (x_α) and (y_α) are nets converging to some limits x and y, then $x_\alpha y_\alpha \to xy$. For instance, if $x_\alpha \to 3$ and $y_\alpha \to +\infty$, then $x_\alpha y_\alpha \to 3 \cdot (+\infty) = +\infty$. This behavior is very reassuring: it tells us that $\pm\infty$ are very much like ordinary real numbers.

The only exceptions are when we multiply 0 times $\pm\infty$. If $x_\alpha \to 0$ and $y_\alpha \to \pm\infty$, then the product $x_\alpha y_\alpha$ could converge to anything or not converge at all. For instance, take the directed set to be $\mathbb{N}$, so that our nets are sequences. Then

$$\frac{1}{n} \cdot n^2 \to +\infty, \qquad \frac{1}{n} \cdot n \to 1, \qquad \frac{1}{n^2} \cdot n \to 0,$$

and $(\frac{1}{n} \sin n) \cdot n$ does not converge at all. This state of affairs can be summarized as follows:

> Multiplication, considered as a binary operation on $[-\infty, +\infty]$, is jointly continuous everywhere except at the ordered pairs $(0, \pm\infty)$ and $(\pm\infty, 0)$. It cannot be made jointly continuous at those ordered pairs, no matter how we define the products at those pairs. Still, for some purposes in the theory of measure and integration (not involving limits of this sort), it is convenient to define $0 \cdot \infty = 0$ and accept multiplication as an operation that is discontinuous at that ordered pair.

15.29. *A topological equivalent of choice (optional).* We shall show that AC (introduced in 6.12, 6.20, and 6.22) is equivalent to the following assertion, from Schechter [1992]:

> **(AC19) Product of Closures.** For each λ in some index set Λ, let S_λ be a subset of some topological space X_λ. Then $\mathrm{cl}(\prod_{\lambda\in\Lambda} S_\lambda) = \prod_{\lambda\in\Lambda} \mathrm{cl}(S_\lambda)$.

In this equation, the first "cl" denotes closure in the product topology on $\prod_{\lambda\in\Lambda} X_\lambda$, while the second "cl" denotes closure in X_λ.

Actually, the inclusion $\mathrm{cl}(\prod_{\lambda\in\Lambda} S_\lambda) \subseteq \prod_{\lambda\in\Lambda} \mathrm{cl}(S_\lambda)$ is provable in ZF (i.e., without AC). To see this, just note that $\prod_{\lambda\in\Lambda} \mathrm{cl}(S_\lambda) = \bigcap_{\lambda\in\Lambda} \pi_\lambda^{-1}(\mathrm{cl}(S_\lambda))$ is closed (where π_λ is the λth coordinate projection). Thus, it remains to show that the inclusion

(AC20) $$\mathrm{cl}(\textstyle\prod_{\lambda\in\Lambda} S_\lambda) \supseteq \prod_{\lambda\in\Lambda} \mathrm{cl}(S_\lambda)$$

is equivalent to AC. Refer to 6.12.

To prove (AC3) $\Rightarrow$ (AC20), let any $f \in \prod_{\lambda\in\Lambda} \mathrm{cl}(S_\lambda)$ be given; we wish to show that $f \in \mathrm{cl}(\prod_{\lambda\in\Lambda} S_\lambda)$. It suffices to show that $\prod_{\lambda\in\Lambda} S_\lambda$ meets every neighborhood of f. Let G be any neighborhood of f; then $f \in \prod_{\lambda\in\Lambda} G_\lambda \subseteq G$ where G_λ is some open subset of X_λ. Since $f(\lambda) \in \mathrm{cl}(S_\lambda)$, the set S_λ meets every neighborhood of $f(\lambda)$ in X_λ. Thus the set $G_\lambda \cap S_\lambda$ is nonempty. *Choose* any element $z_\lambda \in G_\lambda \cap S_\lambda$. Now the function z, defined by $z(\lambda) = z_\lambda$, is an element of $G \cap \prod_{\lambda\in\Lambda} S_\lambda$.

To prove (AC20) $\Rightarrow$ (AC3), let Λ, S_λ, ξ_λ, X_λ, X, ξ be as in 6.24. Let X_λ be equipped with the indiscrete topology, i.e., in which the only open sets are $\varnothing$ and X_λ. Then $\xi_\lambda \in X_\lambda = \mathrm{cl}(S_\lambda)$. Hence the function ξ is an element of $\prod_{\lambda\in\Lambda} \mathrm{cl}(S_\lambda)$. Now apply (AC20); this tells us $\mathrm{cl}(\prod_{\lambda\in\Lambda} S_\lambda)$ is nonempty, and therefore the set $\prod_{\lambda\in\Lambda} S_\lambda$ is nonempty.

Quotient Topologies

15.30. *Definition.* Let $(X, \mathcal{S})$ be a topological space, let Q be a set, and let $\pi : X \to Q$ be a surjective mapping. The resulting **quotient topology** (or *identification topology*) on Q is defined to be

$$\mathcal{T} \quad = \quad \{T \subseteq Q \ : \ \pi^{-1}(T) \in \mathcal{S}\}.$$

We saw in 5.40.b that this collection $\mathcal{T}$ is a topology on Q. (In fact, 5.40.b shows that $\mathcal{T}$ is a topology regardless of whether π is surjective, but surjectivity of π is part of the definition of a quotient topology.)

When Q is equipped with the quotient topology, then π will be called a **topological quotient map** (or *topological identification map*). The terminology stems from the fact that Q is the quotient set of X, determined by the mapping π (see 3.11). Alternatively, points of Q are obtained by identifying with each other (i.e., merging) those points of X that have the same image under π.

In general, convergence of nets and filters in the quotient topology does not have a simple characterization analogous to that of 15.24.b. A partial result in that direction is given in 22.13.e.

Our treatment of quotients is based partly on Dugundji [1966].

15.31. *Basic properties of the quotient topology.*

- **a.** Let $\pi : X \to Y$ be a topological quotient map. Then a set T is open in Y if and only if $\pi^{-1}(T)$ is open in X. (This is just a restatement of the definition.)
- **b.** Let $\pi : X \to Y$ be a topological quotient map. Then a set T is closed in Y if and only if $\pi^{-1}(T)$ is closed in X.

c. (Composition property.) If $\pi : X \to Q$ is a topological quotient map and $g : Q \to Z$ is some mapping such that the composition $g \circ \pi : X \to Z$ is continuous, then g is continuous.

In fact, a continuous surjective map $\pi : X \to Q$ is a topological quotient map if and only if it has that composition property. For this reason the quotient topology is sometimes called the **final topology** — it has some properties analogous to the initial topology (introduced in 9.15 and 9.16), but with the arrows reversed.

d. Let X be a topological space and let $\pi : X \to Q$ be a surjective mapping. Then the quotient topology on Q makes π continuous. In fact, the quotient topology is the *strongest* (i.e., largest) topology on Q that makes π continuous.

e. Recall that a mapping is *open* if the forward image of each open set is open, or *closed* if the forward image of each closed set is closed.

Show that if $\pi : X \to Y$ is a continuous surjective map that is either open or closed, then π is a topological quotient map.

Several of the most important topological quotient maps are open maps (see 16.5 and 22.13.e), but this is not a property of all topological quotient maps.

f. Let $\pi : X \to Q$ be a topological quotient map. Recall from 4.4.e that the π-saturation of a set $S \subseteq X$ is the set $\pi^{-1}(\pi(S)) \subseteq X$. Show that

> π is an open map if and only if the π-saturation of each open subset of X is open.
>
> π is a closed map if and only if the π-saturation of each closed subset of X is closed.

g. *Example.* If $X = \prod_{\lambda \in \Lambda} Q_\lambda$ is a product of topological spaces with the product topology, then each of the coordinate projections $\pi_\lambda : X \to Q_\lambda$ is a topological quotient map. *Hint*: 15.26.b.

h. *Example.* Let (X, d) and (Q, e) be pseudometric spaces. Let $\pi : X \to Q$ be a surjective mapping that is distance-preserving — i.e., that satisfies $e(\pi(x_1), \pi(x_2)) = d(x_1, x_2)$. Then the mapping π is open, closed, and a topological quotient map.

More generally, let (X, D) and (Q, E) be gauge spaces, with gauges $D = \{d_\lambda : \lambda \in \Lambda\}$ and $E = \{e_\lambda : \lambda \in \Lambda\}$ parametrized by the same index set Λ. Suppose $\pi : X \to Q$ is a surjective mapping that is "distance-preserving" in the following sense:

$$e_\lambda(\pi(x_1), \pi(x_2)) \quad = \quad d_\lambda(x_1, x_2) \qquad \text{for all } x_1, x_2 \in X \text{ and } \lambda \in \Lambda.$$

Then π is open, closed, and a topological quotient map. A slight specialization of this result is given in 16.21.

Neighborhood Bases and Topology Bases

15.32. Let X be a topological space (or, more generally, a pretopological space), and let

$x \in X$. A **base of neighborhoods** at x, or a **neighborhood base** at x, is any filterbase $\mathcal{B}$ that generates the neighborhood filter $\mathcal{N}(x)$. In other words, it is any collection $\mathcal{B} \subseteq \mathcal{N}(x)$ with the property that every member of $\mathcal{N}(x)$ is a superset of some member of $\mathcal{B}$.

15.33. *Examples of neighborhood bases.* Let $(X, \mathcal{T})$ be a topological space, and let $x \in X$. Let $\mathcal{N}(x)$ be the neighborhood filter at x.

a. Trivially, $\mathcal{N}(x)$ itself is a neighborhood base at x.

b. Another neighborhood base at x is given by $\mathcal{N}(x) \cap \mathcal{T}$, the collection of *open* neighborhoods of X.

More generally, an *open neighborhood base* means any neighborhood base, all of whose members are open sets. Thus, it is a neighborhood base $\mathcal{B} \subseteq \mathcal{N}(x) \cap \mathcal{T}$.

c. A *closed neighborhood base* means a neighborhood base, all of whose members are closed sets. A topological space is called **regular** if every point has a closed neighborhood base. Regular spaces will be investigated further in 16.13.

Exercise. Every gauge space (X, D) is regular. *Hint:* Let B_d and K_d denote the open and closed d-balls, as in 5.15.g. If N is a neighborhood of x, then $N \supseteq \bigcap_{d \in C} B_d(x, r) \supseteq \bigcap_{d \in C} K_d(x, r/2)$ for some finite set $C \subseteq D$ and some $r > 0$.

Other examples of neighborhood bases will be given in Chapters 26 through 28.

15.34. A topological space is **first countable** if each neighborhood filter $\mathcal{N}(x)$ is generated by some countable filterbase $\mathcal{B}(x)$ — i.e., if for each x there is some countable collection $\mathcal{B}(x) \subseteq \mathcal{N}(x)$ such that each member of $\mathcal{N}(x)$ is contained in some member of $\mathcal{B}(x)$.

As shown by some of the exercises below, in a first countable space, sequential arguments are sufficient for many purposes; nets are very seldom needed. However, the Principle of Countable Choice is needed for many of these sequential arguments — i.e., the proofs may require a sequence of arbitrary choices, since there is no "canonical sequence" analogous to the canonical nets developed in 7.11.

Sequential arguments are also sufficient in a few special situations in spaces that are not first countable; that is the content of the deep theorems 17.50 and 28.36. For a more elementary example of sequences sufficing in a space that is not first countable, consider the characterizations of closures and continuity when X is an infinite set equipped with the cofinite topology (see 5.15.c and 15.9.c).

Exercises.

a. Any pseudometric space (X, d) is a first countable space. A countable, open neighborhood base at x is given by the open balls $B(x, \frac{1}{n}) = \{u \in X : d(u, x) < \frac{1}{n}\}$ for $n \in \mathbb{N}$.

In particular, $\mathbb{R}$ is first countable.

Remark. Actually, first countable is only a *very slight* generalization of pseudometrizable. Most spaces of interest to analysts are subsets of topological vector spaces; among such spaces, first countable is the same as pseudometrizable — see 26.32.

b. In any first countable space X, $\mathrm{cl}(S)$ is equal to the *sequential closure* of S — that is, the set
$$\{x \in X \;:\; x \text{ is a limit of some sequence in } S\}.$$

c. In a first countable space, if some subnet of a sequence (x_m) converges to a limit z, then some subsequence $(x_{k_n} : n = 1, 2, 3, \ldots)$ also converges to z.

Hints: Let $\{B_1, B_2, B_3, \ldots\}$ be a neighborhood base at z; we may assume $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$. (Why?) Let $k_0 = 0$. Thereafter, show that there exists an integer k_n that satisfies both $k_n > k_{n-1}$ and $x_{k_n} \in B_n$.

d. Let X and Y be topological spaces; assume X is first countable. Then a mapping $p : X \to Y$ is continuous if and only if it preserves sequential convergences — i.e., if and only if it satisfies
$$x_n \to x \text{ in } X \qquad \Rightarrow \qquad p(x_n) \to p(x) \text{ in } Y$$
— regardless of whether Y is first countable.

Hints: Assume p preserves sequential convergences. Let N be a neighborhood of $p(x_0)$ in Y; we wish to show that $p^{-1}(N)$ is a neighborhood of x_0 in X. Let $\{B_1, B_2, B_3, \ldots\}$ be a neighborhood base at x_0 in X; we may assume $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$; we wish to show that $p^{-1}(N)$ contains some B_j. Suppose not. Then there exist points $x_j \in B_j \setminus p^{-1}(N)$. The sequence (x_j) converges to x_0, hence $p(x_j) \to p(x_0)$, hence for j sufficiently large we have $p(x_j) \in N$, a contradiction.

15.35. Another way to describe topologies is in terms of bases for topologies. Let $(X, \mathcal{T})$ be a topological space, and let $\mathcal{B} \subseteq \mathcal{T}$. We say[1] that $\mathcal{B}$ is a **base for the topology** $\mathcal{T}$ if

every member of $\mathcal{T}$ is a union of members of $\mathcal{B}$.

(The union may be of finitely or infinitely many members of $\mathcal{B}$. It may also be of *no* members of $\mathcal{B}$; thus we automatically get $\varnothing$ as a union.) Trivially, the topology $\mathcal{T}$ itself is a base; other bases are sometimes convenient. Some examples are:

a. In a pseudometric space, the collection of all open balls forms a base for the topology.

b. In any poset, the principal lower sets form a base for the lower set topology.

c. Let X be a chain. Then a base for the interval topology (defined in 5.15.f) is given by the sets of the forms
$$S_a = \{x \in X : x > a\}, \quad S^b = \{x \in X : x < b\}, \quad S_a^b = \{x \in X : a < x < b\},$$
for points $a, b \in X$. Note that these sets are full, as defined in 4.4.a. Show that if G is an open subset of X, then the full components of G (defined as in 4.4.a(ii)) are also open.

d. Let $X = \prod_{\lambda \in \Lambda} Y_\lambda$ be a product of topological spaces. By a **basic rectangle** we shall mean a subset of X of the form $\prod_{\lambda \in \Lambda} G_\lambda$ where

(i) each G_λ is an open subset of Y_λ, and

[1] *Caution*: Some texts define "base" a little differently.

(ii) $G_\lambda \neq Y_\lambda$ for at most finitely many λ'.s

(When Λ is a finite set, condition (ii) can be omitted from this definition, since it is satisfied automatically.) Show that the basic rectangles form a base for the product topology (hence their name). These basic rectangles may also be called *basic open rectangles*, to distinguish them from another sort of "basic rectangle" introduced in 21.40.

15.36. *Further properties of bases.*

a. Let $(X, \mathcal{T})$ be a topological space, and let $\mathcal{B}$ be a collection of subsets of X. Then $\mathcal{B}$ is a base for the topology $\mathcal{T}$ if and only if for each $x \in X$, the collection of sets $\mathcal{B}(x) = \{B \in \mathcal{B} \ : \ x \in B\}$ is an open neighborhood base at x.

b. Let X be a set (without any topology specified yet), and let $\mathcal{B}$ be a collection of subsets of X. Then $\mathcal{B}$ is a base for *some* topology $\mathcal{T}$ on X if and only if this condition is satisfied: for each $x \in X$, the collection of sets $\mathcal{B}(x) = \{B \in \mathcal{B} \ : \ x \in B\}$ is a filterbase on X. In that case, the resulting topology is uniquely determined; it is $\mathcal{T} = \{S \subseteq X \ : \ S \text{ is a union of members of } \mathcal{B}\}$.

c. Let $\mathcal{G}$ be any collection of subsets of X. Let $\mathcal{B} = \{B \subseteq X : B \text{ is an intersection of finitely many members of } \mathcal{G}\}$. (Here $X \in \mathcal{B}$, since by convention X is the intersection of *no* members of $\mathcal{G}$.) Show that $\mathcal{B}$ and $\mathcal{G}$ generate (in the sense of 5.23.b) the same topology $\mathcal{T}$, and $\mathcal{B}$ is a base for that topology. Thus, the topology $\mathcal{T}$ generated by $\mathcal{G}$ is equal to the collection of all unions of finite intersections of members of $\mathcal{G}$.

d. Let $(X, \mathcal{T})$ be a topological space with base $\mathcal{B}$. Show that a net (x_α) converges to a limit z in $(X, \mathcal{T})$ if and only if for each $B \in \mathcal{B}$ that contains z, we have eventually $x_\alpha \in B$.

15.37. *Cardinality and metric spaces.*

a. If (X, d) is a separable metric space, then $\text{card}(X) \leq \text{card}(2^{\mathbb{N}}) = \text{card}(\mathbb{R})$.

Proof. Let (x_n) be a sequence such as in 15.13.g. For each $p \in X$, there is a subsequence (x_{n_i}) that converges to p, and in fact we can choose it canonically. Thus we obtain a function $p \mapsto (n_i)$, from X into {strictly increasing sequences of positive integers}, with $x_{n_i} \to p$. Obviously the mapping is injective, so $\text{card}(X) \leq \text{card}(\{\text{strictly increasing sequences of positive integers}\}) \leq \text{card}(\mathcal{P}(\mathbb{N}))$.

b. Every separable metric space has a countable base.

Hints: If (x_n) is a dense sequence, show that the set of open balls $\{B(x_n, 1/k) : n, k \in \mathbb{N}\}$ is a base.

c. Let X be a separable metric space satisfying $\text{card}(X) = \text{card}(\mathbb{R})$. (For instance, $\mathbb{R}$ itself is one such space; many others occur in analysis as well.) Use the preceding results to show that *most subsets of X are neither open nor closed* — i.e., show that

$$\text{card}(\{S \subseteq X \ : \ S \text{ is open or closed}\}) \quad < \quad \text{card}(\{\text{subsets of } X\}).$$

d. Every open subset of $\mathbb{R}$ can be written as a union of countably many disjoint open intervals.

Hints: Let $G \subseteq \mathbb{R}$ be open. By an argument similar to 4.4.a(ii), show that G is a union of disjoint open intervals. There are at most countably many of these intervals, since each contains a rational number and there are only countably many rational numbers.

e. Let $J \subseteq \mathbb{R}$ be an interval (possibly all of $\mathbb{R}$). Let $\mathcal{A}$ be the collection of all unions of finitely many subintervals of J (where singletons are considered to be intervals, and the empty set is also a member of $\mathcal{A}$ by convention). Then $\mathcal{A}$ is an algebra of subsets of J, and the σ-algebra that it generates is the σ-algebra of Borel sets.

Cluster Points

15.38. *Definition.* Let X be a topological space. Let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in X, and let $\mathcal{F}$ be its eventuality filter. Let $\mathcal{B}$ be any filterbase that generates $\mathcal{F}$ (e.g., we may take $\mathcal{B} = \mathcal{F}$). Let $z \in X$. Then the following conditions are equivalent. If any, hence all, are satisfied, we say z is a **cluster point** of (x_α) or $\mathcal{F}$ or $\mathcal{B}$.

(A) z is a limit of some superfilter of $\mathcal{F}$.

(B) z is a limit of some subnet of (x_α) — that is, an AA subnet (as defined in 7.15.a).

(C) z is a limit of some Kelley subnet of (x_α) (defined as in 7.15.b).

(D) z is a limit of some Willard subnet of (x_α) (defined as in 7.15.c.

(E) Some proper filter contains both $\mathcal{F}$ and $\mathcal{N}(z)$.

(F) Every member of $\mathcal{F}$ meets every member of $\mathcal{N}(z)$.

(G) Every member of $\mathcal{B}$ meets every member of $\mathcal{N}(z)$.

(H) $z \in \bigcap_{B \in \mathcal{B}} \mathrm{cl}(B)$.

(I) $z \in \bigcap_{F \in \mathcal{F}} \mathrm{cl}(F)$.

(J) $z \in \bigcap_{\alpha \in \mathbb{A}} \mathrm{cl}\Big(\{x_\beta : \beta \succcurlyeq \alpha\}\Big)$. That is, z is in the closure of each tail set of the net (x_α).

(K) Each neighborhood of z is a frequent set for the net (x_α).

(The interchangeability of the three types of subnets follows from 7.19.) Note that the set of cluster points is always a closed set since it is an intersection of closed sets.

Remarks. For some purposes, a cluster point can be used as an "almost limit" — i.e., it has many of the properties of limits; it can sometimes be used in place of a limit when a limit is not available.

Caution: Some mathematicians have another meaning for the term "cluster point:" Let S be a subset of a topological space X; then z is a *cluster point* of S if $z \in \mathrm{cl}(S \setminus \{z\})$. That meaning will not be used in this book, however.

15.39. *Exercise.* Show that if $\mathcal{E}$ is proper filter on a topological space X, then $\mathcal{C} = \{\mathrm{cl}(E) : E \in \mathcal{E}\}$ is a filterbase on X, and the filter generated by $\mathcal{C}$ has the same cluster points as $\mathcal{E}$ does.

15.40. The definition of "cluster point" also applies to sequences, since they are also nets. By definition, a point z is a cluster point of a sequence (x_n) if some subnet of that sequence converges to z — but that subnet is not necessarily a sequence. The point z may possibly satisfy a stronger condition.

We say z is a **sequential cluster point** of the sequence (x_n) if the following equivalent conditions hold. (Equivalence follows from 7.27.)

(A) z is a limit of a *subsequence* of (x_n).

(B) z is a limit of a sequence (u_k) that is a subnet of (x_n) — that is, an AA subnet (as defined in 7.15.a).

(C) z is a limit of a sequence (u_k) that is a Kelley subnet of (x_n) (defined as in 7.15.b).

(D) z is a limit of a sequence (u_k) that is a Willard subnet of (x_n) (defined as in 7.15.c).

Further observation. In general, the cluster points of a sequence and the sequential cluster points of that sequence need not be the same. But they are the same if the topological space is first countable — see 15.34.c.

More about Intervals

15.41. *Proposition.* Let $(X, \leq)$ be a chain. Then the topological convergence determined by the order interval topology (defined in 5.15.f) is the same as the order convergence (as characterized in 7.38, 7.40.d, or 7.41).

15.42. *Corollary.* Let $(X, \leq)$ be a chain, equipped with the convergence described above. Let (x_α) be a net in X, and let $z \in X$. Then z is a cluster point of (x_α) if and only if these three conditions are satisfied for all σ and τ in X:

(i) if $z > \sigma$, then frequently $x_\alpha > \sigma$;

(ii) if $z < \tau$, then frequently $x_\alpha < \tau$; and

(iii) if $\sigma < z < \tau$, then frequently $\sigma < x_\alpha < \tau$.

15.43. *Important examples.* The metric topologies and order interval topologies are the same, on $\mathbb{R}$ or on $[-\infty, +\infty]$.

15.44. *Optional example.* Let $\mathcal{F}$ be a free ultrafilter on $\mathbb{N}$; thus ${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}}/\mathcal{F}$ is a chain ordered field. Let ${}^*\mathbb{R}$ be equipped with its order interval topology and order convergence.

Show that every convergent sequence in $^*\mathbb{R}$ is eventually constant.

Hints: Suppose not. Say (u_n) is a sequence converging to v in $^*\mathbb{R}$, and (u_n) is not eventually equal to v. Replacing (u_n) with a subsequence, we may assume that none of the u_n's are equal to v. By 10.20.c, the sequence $\big((u_n - v)^{-1} : n \in \mathbb{N}\big)$ is order bounded; obtain a contradiction. This result is from Takeuchi [1984].

15.45. *Proposition.* Let X and Y be chains equipped with their interval topologies. Let $f : X \to Y$ be some mapping. Then

(A) f is continuous and increasing, if and only if

(B) f is sup-preserving and inf-preserving.

Proof. (B) $\Rightarrow$ (A) was proved in a slightly more general setting in 7.40.h. For (A) $\Rightarrow$ (B), let S be a nonempty subset of X, and suppose $\sigma = \sup(S)$. Consider S itself as a directed set. The identity $i_S : S \to S$ is a net that increases to the limit σ. Hence $f \circ i_S : S \to Y$ is a net that increases to the limit $f(\sigma)$. Thus $f(\sigma) = \sup(f(S))$, so f is sup-preserving. Similarly we show that f is inf-preserving.

15.46. *Corollaries: relativization and the interval topology.* Let $(Y, \mathcal{T})$ be a chain equipped with the order interval topology, and let $X \subseteq Y$. Let $\mathcal{I}$ and $\mathcal{R}$ be, respectively, the order interval topology determined on X by its ordering, and the relative topology determined on X by $(Y, \mathcal{T})$. Show that

a. $\mathcal{R} \supseteq \mathcal{I}$.

b. $\mathcal{R} = \mathcal{I}$ if and only if the inclusion $X \xrightarrow{\subseteq} Y$ is continuous when X and Y are equipped with the interval topologies; that occurs if and only if the inclusion $X \xrightarrow{\subseteq} Y$ is sup-preserving and inf-preserving.

c. If X is full (as defined in 4.4.a), then $\mathcal{R} = \mathcal{I}$. (In particular, the interval topology on $\mathbb{R}$ is the same as the relative topology that $\mathbb{R}$ inherits from $[-\infty, +\infty]$.)

d. If X is not full, then the relative topology on X may or may not be the same as the interval topology. For instance, take $Y = \mathbb{R}$; show that

(i) The two topologies agree in the case of $X = \mathbb{Z}$.

(ii) The two topologies disagree, in the case of $X = [0,1] \cup (2,3]$. *Hint*: X is disconnected when given the relative topology, but X is order isomorphic to $[0,2]$.

e. A set $X \subseteq \mathbb{R}$ (equipped with the relative topology) is connected if and only if it is full (as defined in 4.4.a) — i.e., if and only if X is an interval.

f. *Optional.* Let Y be a chain equipped with the order topology. Then Y is connected if and only if Y is Dedekind complete and between any two elements of Y there is another element of Y.

15.47. *Exercises on continuity in* $\mathbb{R}$. Prove that

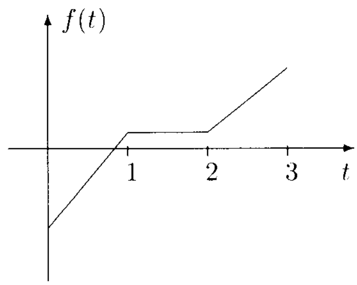
a. If $J \subseteq \mathbb{R}$ is an interval and $f : J \to \mathbb{R}$ is continuous, then $f(J)$ is an interval. *Hint:* 15.46.e and 15.18.

b. **Intermediate Value Theorem.** Let $f : [a,b] \to \mathbb{R}$ be a continuous function. If m is a number between $f(a)$ and $f(b)$, then there exists at least one number $c \in [a,b]$ such that $f(c) = m$.

c. *A partial converse.* Let $f : [a,b] \to \mathbb{R}$ be an increasing function whose range is an interval. Then f is continuous. *Hint:* 15.21.c.

15.48. The Intermediate Value Theorem, as stated above, is not constructive; we may be unable to *find* the number c whose existence is asserted by that theorem. Indeed, for a weak ("Brouwerian") counterexample, consider the mysterious "Goldbach number" Γ described in 10.46. It is a number that is known to be quite close to zero, and in fact it can be approximated as accurately as one may wish, but we do not yet know whether this number is positive, negative, or zero. Use it to define a piecewise-affine function f as in the following diagram. This function is well-defined and continuous, and we can evaluate it with as much accuracy as we may wish. It satisfies $f(0) < 0 < f(3)$. Finding an exact solution $c \in [0,3]$ of $f(c) = 0$ would tell us much about Γ: if $c < 1$, then $\Gamma > 0$; if $1 \le c \le 2$, then $\Gamma = 0$; if $c > 2$, then $\Gamma < 0$. At present, we are unable to find c exactly.

$$f(t) = \begin{cases} -1 + (\Gamma + 1)t & (0 \le t \le 1) \\ \Gamma & (1 \le t \le 2) \\ (1 - \Gamma)(t - 2) + \Gamma & (2 \le t \le 3) \end{cases}$$



However, the following variant of the Intermediate Value Theorem *is* constructively provable:

> **Approximate Intermediate Value Theorem.** If $f : [a,b] \to \mathbb{R}$ is continuous and m is a number between $f(a)$ and $f(b)$, then for each $\varepsilon > 0$ we can find some number $c \in [a,b]$ such that $|f(c) - m| < \varepsilon$.

(We shall not present the proof here. For further discussion see Troelstra and Dalen [1988, pages 292–293].)

The viewpoint of numerical analysis is somewhat similar to that of constructivism, and again gives reason for replacing the classical Intermediate Value Theorem by the Approximate Intermediate Value Theorem. Our values for the function f or the constant m may be based on measurements or on the results of some numerical computation; they are not completely accurate. Our computed value of $f(x)$ or m may differ only very slightly from the "true" value, but that may yield large errors in the computed value of the *exact* solution c of the equation $f(c) = m$. For instance, take $f(x) = 0.001x$ and $m = 0$; let us try to solve $f(c) = m$ for c. The desired solution c changes *greatly* if we change m to 0.001 or if we change the function to $f(x) = 0.001 + 0.001x$. Thus (with $\approx$ denoting "approximately equal"), we may not be able to find an approximation $\widehat{c} \approx c$ to the exact solution c of $f(c) = m$, but we can find a number $\widehat{c}$ such that $f(\widehat{c}) \approx m$.

Chapter 16

Separation and Regularity Axioms

16.1. *Preview.* This chapter considers conditions under which points and/or closed sets can be separated by open sets and/or continuous functions. The chart below shows the relations between the main separation conditions. The chart has implications downward; for instance, every T_3 space is also a T_2 space, and every preregular space is also a symmetric space.

pseudometrizable	metrizable
paracompact (partitions of unity)	paracompact and T_0
normal and symmetric (shrinkings and Urysohn functions, plus symmetric)	T_4 = normal and T_1
completely regular (gauges, uniformities)	$T_{3.5}$ = Tychonov (has Hausdorff compactifications)
regular (closed neighborhood bases; extensions by continuity)	T_3 = regular and separated
preregular (limits are unique, up to topological distinguishability)	T_2 = Hausdorff (has unique limits)
symmetric (the closures of points form a partition of X)	T_1 = Fréchet (points are closed)
(arbitrary topological space)	T_0 = Kolmogorov (points are topologically distinguishable)

The two entries in each row of the chart are closely related: a space satisfies the condition in the right column if and only if the space is Kolmogorov and satisfies the condition in the left column in the same row. For instance, a topological space is Hausdorff if and only if it is both Kolmogorov and preregular. We can move from the left column to the right column by taking the Kolmogorov quotient of a space, as in 16.5.

Normality is an interesting condition by itself — we shall introduce it in 16.26 — but it fits in the chart only in conjunction with the symmetry condition, since normal by itself does not imply completely regular.

The long list of conditions may be daunting to beginners, or it may seem like hair-splitting to some readers. The beginner will find it helpful to concentrate on pseudometrizable spaces and completely regular spaces; those will play the greatest role in this book. Most spaces arising in applications in analysis are at least $T_{3.5}$ spaces, but the abstract theory can be developed more clearly if we classify properties according to the various axioms in the chart. It is possible to decrease the emphasis on some of these properties, but it is not possible to omit them altogether. For instance, some textbooks omit mentioning T_1 spaces at all, but give as an exercise the fact that points in a T_2 space are closed.

The terminology T_0, T_1, T_2, etc., follows the literature, but the reader is cautioned that the literature varies slightly on its definitions of T_3 and $T_{3.5}$; some mathematicians interchange some of the terms in the two columns in our chart. Even mathematicians who agree with our terminology may use it in different ways; for instance, the phrases "Tychonov space," "completely regular T_0 space," and "completely regular Hausdorff space" are used interchangeably in the literature; they all describe the same thing.

Most of the separation and regularity axioms are well known and can be found in any topology book. However, "symmetric spaces" and "preregular spaces" are not so well known. They are the same (respectively) as the "R_0 spaces" and "R_1 spaces" introduced by Davis [1961]. Symmetric spaces were introduced earlier by Shanin [1943]. This book owes a debt to Murdeshwar [1983], who investigated those spaces systematically.

Our choice of emphases is determined by the needs of later chapters. For instance, many topology books concern themselves solely with Hausdorff spaces. This book considers the non-Hausdorff case as well, because one of the best ways to describe a weak topology on a topological vector space (generally Hausdorff, in applications) is as the supremum of a collection of pseudometric topologies (each of which is not Hausdorff).

KOLMOGOROV (T-ZERO) TOPOLOGIES AND QUOTIENTS

16.2. Let x, y be points in a topological space $(X, \mathcal{T})$. Then the following conditions are equivalent. When one, hence all, of these conditions holds, we shall say that x and y are **topologically indistinguishable**. This is clearly an equivalence relation on X.

(A) The topology $\mathcal{T}$ cannot distinguish between x and y. That is, every open set that contains either of x, y also contains the other.

(B) Every closed set that contains either of x, y also contains the other.

(C) $\mathrm{cl}(\{x\}) = \mathrm{cl}(\{y\})$.

(D) $\mathcal{N}(x) = \mathcal{N}(y)$. That is, any neighborhood of either point is also a neighborhood of the other point.

(E) Any filter or net that converges to either of the points x, y must also converge to the other.

(F) Any filter or net that has either of the points x, y as a cluster point must also have the other as a cluster point.

16.3. *Definition.* A topological space $(X, \mathcal{T})$ is called a $\boldsymbol{T_0}$ **space**, also known as a **Kolmogorov space**, if it satisfies either of the following equivalent conditions:

(A) If x and y are topologically indistinguishable (in the sense of 16.2), then $x = y$.

(B) Given any distinct points $x_1, x_2 \in X$, at least one of these two conditions is satisfied:

  (i) There exists an open set G_1 that contains x_1 but not x_2.

  (ii) There exists an open set G_2 that contains x_2 but not x_1.

(Compare the last condition with 16.7(C) and 16.11(D).)

16.4. *Examples.*

a. Let (X, d) be a pseudometric space. Then two points $x, y \in X$ are topologically indistinguishable if and only if $d(x, y) = 0$. Hence X is a Kolmogorov space if and only if the pseudometric d is a metric. This result is generalized in 16.17.

b. The indiscrete topology on a set X is not a Kolmogorov space if $\text{card}(X) \geq 2$.

c. The knob topology on X (see 5.34.c) is not a Kolmogorov topology if $\text{card}(X) \geq 3$.

16.5. Let $(X, \mathcal{S})$ be a topological space. Then "x is topologically indistinguishable from y," defined as in 16.2, is an equivalence relation on X. Let Q be the resulting quotient set — i.e., the set of equivalence classes. Thus, Q is obtained by identifying with each other (i.e., merging into one point) any points of X that are topologically indistinguishable in $(X, \mathcal{S})$.

Let $\pi : X \to Q$ be the quotient mapping (see 3.11). Let Q be equipped with the quotient topology — i.e., the strongest topology on Q that makes π continuous (see 15.30). Then:

a. Every closed subset of X is π-saturated (see 4.4.e).

b. Every open subset of X is π-saturated.

c. The quotient map $\pi : X \to Q$ is both open and closed. *Hint*: 15.31.f

d. The forward image mapping $S \mapsto \pi(S)$ gives a bijection from {open subsets of X} onto {open subsets of Q} and from {closed subsets of X} onto {closed subsets of Q}. This bijection preserves unions and intersections. The lattice of open sets of X (described in 5.21) is lattice isomorphic to the lattice of open sets of Q.

Let $x, x_1, x_2 \in X$ and let S, T be π-saturated subsets of X. Then

e. $x \in S \iff \pi(x) \in \pi(S)$.

f. $S \subseteq T \iff \pi(S) \subseteq \pi(T)$.

g. S and T are disjoint if and only if $\pi(S)$ and $\pi(T)$ are disjoint.

h. S is an X-neighborhood of x if and only if $\pi(x)$ is a Q-neighborhood of $\pi(S)$.

i. $x_1 \in \text{cl}_X(\{x_2\})$ if and only if $\pi(x_1) \in \text{cl}_Q(\{\pi(x_2)\})$.

j. The quotient space Q constructed in this fashion is a Kolmogorov space.

We may call Q the **Kolmogorov quotient space** of X or $\boldsymbol{T_0}$ **quotient space** of X. It preserves many of the properties of X, and so for many purposes we can replace X with Q. For this reason, much of the mathematical literature does not concern itself with spaces that are not Kolmogorov.

Symmetric and Fréchet (T-One) Topologies

16.6. *Definition and proposition.* Let X be a topological space (not necessarily Kolmogorov). We shall say X is a **symmetric space** if it satisfies any of the following equivalent conditions:

(A) The relation $x \in \mathrm{cl}(y)$ is a *symmetric* relation between x and y — that is, $x \in \mathrm{cl}(y) \iff y \in \mathrm{cl}(x)$.

(B) If G is an open neighborhood of x, then $G \supseteq \mathrm{cl}(\{x\})$.

(C) If F is a closed set and $x \in X \setminus F$, then the closed sets F and $\mathrm{cl}(\{x\})$ are disjoint.

(D) The set $\{u \in X : u \text{ is topologically indistinguishable from } x\}$ is equal to $\mathrm{cl}(\{x\})$, for each $x \in X$. (See 16.2 for definition.)

(E) The sets of the form $\mathrm{cl}(\{x\})$, for $x \in X$, form a partition of X; that is, any two such sets $\mathrm{cl}(\{x_1\})$ and $\mathrm{cl}(\{x_2\})$ are either identical or disjoint.

The proof of equivalence is an easy exercise.

16.7. *Definition and proposition.* A topological space $(X, \mathcal{T})$ is called a $\boldsymbol{T_1}$ **space** if any of the following equivalent conditions are satisfied:

(A) For each $x \in X$, the singleton $\{x\}$ is a closed set.

(B) X is a Kolmogorov, symmetric space (defined in 16.3 and 16.6).

(C) Given any distinct points $x_1, x_2 \in X$, both of these conditions are satisfied:

  (i) There exists an open set G_1 that contains x_1 but not x_2.

  (ii) There exists an open set G_2 that contains x_2 but not x_1.

(Compare the last condition with 16.3(B) and 16.11(D).)

A T_1 topology is sometimes called a **Fréchet topology**. However, this usage is uncommon in functional analysis books, since the term "Fréchet space" has another meaning; see 26.14.

16.8. *Examples.*

a. Finite sets are usually equipped with the discrete topology. Show that the only T_1 topology on a finite set is the discrete topology.

b. (*Optional exercise.*) Let $(X, \mathcal{T})$ be a T_1 topological space. If the topological closure is also an algebraic closure (defined in 4.8), show that X has the discrete topology — that is, $\mathcal{T} = \mathcal{P}(X)$.

c. The indiscrete topology on a set X with more than one point is a symmetric space (in fact, a pseudometrizable space), but not Kolmogorov.

d. The set $\mathbb{N}$, equipped with either the lower set topology $\mathcal{U}$ or the upper set topology $\mathcal{V}$ given in 5.15.d, is Kolmogorov but not T_1; hence it is not a symmetric space.

e. (*Optional.*) More generally: Let $(X, \preccurlyeq)$ be a preordered set. Then the lower set topology on X is a Kolmogorov topology if and only if $\preccurlyeq$ is a partial order. The lower set topology is T_1 if and only if $\preccurlyeq$ is the equality relation $(=)$, in which case the resulting topology is the discrete topology.

f. Let X be a topological space, and let Q be its Kolmogorov quotient (as defined in 16.5). Then X is a symmetric space $\iff$ Q is a symmetric space $\iff$ Q is T_1.

PREREGULAR AND HAUSDORFF (T-TWO) TOPOLOGIES

16.9. *Proposition and notation.* Let X be a topological space (not necessarily Kolmogorov). Let $x, y \in X$. Then the following conditions are equivalent; in the next section we shall abbreviate this relationship as $x \Diamond y$.

(A) x is a cluster point of $\mathcal{N}(y)$.

(B) y is a cluster point of $\mathcal{N}(x)$.

(C) Some filter (or net) converges to both x and y.

(D) Some proper filter contains both $\mathcal{N}(x)$ and $\mathcal{N}(y)$.

(E) Every neighborhood of x meets every neighborhood of y.

Clearly, $\Diamond$ is both reflexive and symmetric. However, we do not assert that it is transitive; i.e., we do not assert that $\Diamond$ is an equivalence relation on X.

Note that if $x \in \text{cl}(\{y\})$, then $x \Diamond y$, since the constant net y then converges to both x and y.

16.10. *Definition and proposition.* Let X be a topological space (not necessarily Kolmogorov). Let $\Diamond$ be defined as in 16.9. We shall say X is a **preregular space** if it satisfies any of the following equivalent conditions:

(A) $x \Diamond y \;\Rightarrow\; y \in \text{cl}(\{x\})$.

(B) $x \Diamond y \iff y \in \text{cl}(\{x\})$. In other words, $\text{cl}(\{x\})$ is equal to the set $\{y \in X : x \Diamond y\}$. (Note that this condition implies the symmetry condition 16.6(A).

Also, it implies that $\diamond$ is a transitive relation. That fact, together with the observations in 16.9, tell us that $\diamond$ is an equivalence relation.)

(C) If x is one of the limits of a filter or a net, then the set of all limits of that filter or net is equal to $\mathrm{cl}(\{x\})$.

(D) Each filter or net has at most one limit, up to topological indistinguishability. In other words, if x and y are two limits of a filter or a net, then x and y are topologically indistinguishable (as defined in 16.2).

Proof of equivalence is left as an exercise.

16.11. *Definition and proposition.* A topological space $(X, \mathfrak{T})$ is called a $\boldsymbol{T_2}$ **space**, or a **Hausdorff space**, if any of the following equivalent conditions are satisfied:

(A) When X is equipped with the topological convergence (defined as in 15.2 or 15.7), then X is Hausdorff in the sense of convergence spaces (defined in 7.36) — i.e., any net or filter has at most one limit.

(B) X is a preregular Kolmogorov space.

(C) Let W be any topological space, let W_0 be a dense subset of W, and let $f_0 : W_0 \to X$ be continuous. Then f_0 has at most one extension $f : W \to X$ that is continuous.

(D) Any two distinct points in X have disjoint neighborhoods. In other words, given any distinct points $x_1, x_2 \in X$, there exist *disjoint* open sets G_1 and G_2 such that

  (i) G_1 contains x_1 but not x_2, and

  (ii) G_2 contains x_2 but not x_1.

(Compare the last condition with 16.3(B) and 16.7(C).)

Hausdorff spaces, or T_2 spaces, are also sometimes known as **separated spaces**. That term should not be confused with *separable spaces*, introduced in 15.13.

Proof of equivalence. The equivalence of (A), (B), (D) and the implication (A) $\Rightarrow$ (C) are easy exercises (with 15.3.a as a hint). It remains for us to present a proof of of (C) $\Rightarrow$ (D) — or more precisely, a proof that not-(D) implies not-(C). Suppose x_0, x_1 are distinct points in X that do not have disjoint neighborhoods. Let $W = \{x_0, x_1\}$, equipped with the relative topology. The inclusion $j : W \xrightarrow{\subseteq} X$ is then continuous. The open subsets of W are $\varnothing$, W, and perhaps one of $\{x_0\}, \{x_1\}$, but not both of these singletons. By relabeling if necessary, we may assume that $\{x_1\}$ is not an open subset of W. Then $W_0 = \{x_0\}$ is dense in W, since every neighborhood of x_1 meets W_0. The inclusion $j_0 : W_0 \xrightarrow{\subseteq} X$ is continuous. Two different continuous extensions of it are the map j (already noted) and the constant function x_0 (since any constant function is continuous).

16.12. *Exercises.*

a. If X is an infinite set, then the cofinite topology on X (defined in 5.15.c) is T_1 but not T_2; hence it is a symmetric space but not preregular.

b. Let X be a topological space, and let Q be its Kolmogorov quotient (as defined in 16.5). Then X is preregular $\iff$ Q is preregular $\iff$ Q is T_2.

c. *Technical lemma.* Let (Y, d) be a pseudometric space, and assume the resulting topology is not the indiscrete topology. Let $y_0 \in Y$. Then there exist another point $y_1 \in Y$ and open disjoint sets $S_0, S_1 \subseteq Y$ such that $y_0 \in S_0$ and $y_1 \in S_1$. (This exercise will be used in 18.20.)

Hint: First show that there exists some y_1 such that $d(y_0, y_1) > 0$.

Regular and T-Three Topologies

16.13. *Definition and proposition.* Let X be a topological space (not necessarily Kolmogorov). We shall say X is **regular** if any, hence all, of the following equivalent conditions are satisfied:

(A) Any point and any closed set not containing that point are contained in disjoint open sets.

(B) If $x \in G$ and G is open, then there exists an open set H such that $x \in H \subseteq \mathrm{cl}(H) \subseteq G$.

(C) Each point has a neighborhood base consisting of closed sets.

The proof of equivalence is an easy exercise.

A topological space is sometimes called $\boldsymbol{T_3}$ if it is both Kolmogorov and regular.

16.14. *Exercises.*

a. Any regular space is also a preregular space (hence any T_3 space is also a T_2 space).

b. Let X be a topological space, and let Q be its Kolmogorov quotient (as defined in 16.5). Then X is regular $\iff$ Q is regular $\iff$ Q is T_3.

c. (*Optional.*) Look in Steen and Seebach [1970], and find an example of a topological space that is T_2 but not T_3.

16.15. Theorem on extension by continuity. Let X and Y be topological spaces; assume Y is regular. For each $x \in X$, let $\mathcal{N}(x)$ be the neighborhood filter at x. Let D be a dense subset of X, and let $f : D \to Y$ be continuous. Then

$$D \cap \mathcal{N}(x) \quad = \quad \{D \cap N \ : \ N \in \mathcal{N}(x)\}$$

is a filter on D for each $x \in X$. Moreover, f can be extended to a continuous function $F : X \to Y$ if and only if for each $x \in X$ the filterbase

$$f\Big(D \cap \mathcal{N}(x)\Big) \quad = \quad \Big\{f(D \cap N) \ : \ N \in \mathcal{N}(x)\Big\}$$

is convergent in Y, in which case we can take $F(x)$ (for $x \in X \setminus D$) to be any of the limits of $f(D \cap \mathcal{N}(x))$.

Note that if Y is T_3, then F is uniquely determined.

Proof of theorem (following Dugundji [1966]). That $D \cap \mathcal{N}(x)$ is a filter on D follows easily from the fact that D is dense in X.

Suppose f has a continuous extension F. Then for each $x \in X$ the filter $\mathcal{N}(x)$ converges to x, hence the filterbase $F(\mathcal{N}(x))$ converges to $F(x)$, and it is easy to verify that the filterbase $f(D \cap \mathcal{N}(x))$ also converges to $F(x)$.

Conversely, suppose that $f(D \cap \mathcal{N}(x))$ is convergent for each $x \in X$; we shall prove the existence of a continuous extension. When $x \in D$, then one of the limits of $f(D \cap \mathcal{N}(x))$ is $f(x)$, since f is continuous on D. For each $x \in X$, let $F(x)$ be any one of the limits of $f(D \cap \mathcal{N}(x))$; it suffices to show that F is continuous. Fix any $x_0 \in X$; we shall show F is continuous at x_0. Let Y_0 be any neighborhood of $F(x_0)$ in Y; it suffices to show that $F^{-1}(Y_0)$ is a neighborhood of x_0 in X. Since Y is regular, there is some open set Y_1 with $F(x_0) \subseteq Y_1 \subseteq \mathrm{cl}(Y_1) \subseteq Y_0$. Since $f(D \cap \mathcal{N}(x_0)) \to F(x_0)$, there is some open neighborhood X_0 of x_0 in X such that $f(D \cap X_0) \subseteq Y_1$. It suffices to show that $F(X_0) \subseteq Y_0$.

For each $z \in X_0$, we claim that the collection

$$\mathcal{B}_z \quad = \quad \left\{ Y_1 \cap f(D \cap M) \ : \ M \in \mathcal{N}(z) \right\}$$

is a filterbase on Y. To see this, observe that if $z \in X_0$ and $M \in \mathcal{N}(z)$, then $X_0 \cap M$ is a neighborhood of z, hence it contains a nonempty open set, hence it meets D (which is dense); thus $X_0 \cap D \cap M$ is nonempty. Then $f(X_0 \cap D \cap M) \subseteq Y_1 \cap f(D \cap M)$, so $Y_1 \cap f(D \cap M)$ is nonempty. Our claim follows easily.

Any limit of the filterbase $f(D \cap \mathcal{N}(z))$ is clearly also a limit of the filterbase $\mathcal{B}_z$. Thus, in particular, $\mathcal{B}_z \to F(z)$. Since all the members of $\mathcal{B}_z$ are subsets of Y_1, it follows that $F(z) \in \mathrm{cl}(Y_1) \subseteq Y_0$. This shows $F(X_0) \subseteq Y_0$ and completes the proof.

Completely Regular and Tychonov (T-Three and a Half) Topologies

16.16. Theorem. Let $(X, \mathcal{T})$ be a topological space. Then the following conditions are equivalent:

(A) $\mathcal{T}$ is gaugeable — i.e., given by a gauge, in the sense of 5.15.h.

(B) $\mathcal{T}$ is uniformizable — i.e., given by a uniformity, in the sense of 5.33.

(C) $\mathcal{T}$ is **completely regular** — i.e., it has this property: for each point z and each closed set F not containing that point, there exists a continuous function $\varphi : X \to [0, 1]$ such that $\varphi(z) = 1$ and φ vanishes everywhere on F.

(D) Every lower semicontinuous function $f : X \to [-\infty, +\infty]$ is the pointwise supremum of a collection of continuous functions from X into $[-\infty, +\infty]$.

Hints: The proof of (A) $\Rightarrow$ (B) is an easy exercise that was already posed at the end of 5.33. For (B) $\Rightarrow$ (A), let $\mathcal{U}$ be a given uniformity on X, and let $V_1 \subseteq X \times X$ be any given symmetric entourage. Let $V_0 = X \times X$. By 5.35.c, for each positive integer n we may choose some symmetric entourage $V_{n+1} \subseteq V_n$ such that $V_{n+1}^3 \subseteq V_n$. (This *choosing* is an application of the Principle of Dependent Choice, introduced in 6.28.) Apply Weil's Lemma 4.44; there is some pseudometric d on X that satisfies

$$\{(x,y) \in X^2 \ : \ d(x,y) < 2^{-n}\} \quad \subseteq \quad V_n \quad \subseteq \quad \{(x,y) \in X^2 \ : \ d(x,y) \leq 2^{-n}\}$$

for all positive integers n. Let D be the collection of all pseudometrics d that are determined in this fashion, for all choices of V_1 and the V_n's. Show that the uniformity determined by D is equal to $\mathfrak{U}$.

To prove (C) $\Rightarrow$ (A), for each continuous function $f : X \to [0,1]$ define a corresponding pseudometric $d_f(x,y) = |f(x) - f(y)|$. Show that the gauge consisting of all such d_f's yields the topology $\mathfrak{T}$.

For (A) $\Rightarrow$ (C), note that if $z \notin F$, then $r = \mathrm{dist}_d(z, F) > 0$ for some pseudometric d in the gauge. Then $\varphi(x) = \min\{1, \frac{1}{r}\mathrm{dist}_d(x, F)\}$ has the required properties.

For (D) $\Rightarrow$ (C), let $V = \complement F$; then V is open. Hence its characteristic function 1_V is lower semicontinuous. Also, $1_V(z) = 1$. By (D), then, there is some continuous function $g : X \to [-\infty, +\infty]$ such that $g \leq 1_V$ and $a = g(z) > 0$. Then $\varphi = \min\{1, \frac{1}{a}g^+\}$ has the required properties.

For (C) $\Rightarrow$ (D), let $f : X \to [-\infty, +\infty]$ be a lower semicontinuous function. Replacing f by $\frac{\pi}{2}\arctan f$, we may assume $\mathrm{Range}(f) \subseteq [-1,1]$. (explain). It suffices to show that, for each $x_0 \in X$ and each number $a < f(x_0)$, there is some continuous function $g : X \to [-1,1]$ such that $g \leq f$ and $g(x_0) \geq a$. If $a \leq -1$, then take g to be the constant function -1, and we are done. Thus, we may assume $-1 < a < f(x_0)$. Since f is lower semicontinuous, there is some open neighborhood V of x_0 on which $f \geq a$. Since X is completely regular, there is some continuous function $h : X \to [0,1]$ that vanishes at x_0 and takes the value 1 everywhere on $\complement V$. Then the function $g(x) = a - (a+1)h(x)$ satisfies the requirements. This argument follows Bourbaki [1966].

16.17. *More definitions.* A completely regular Kolmogorov space is also known as a **Tychonov space**, or sometimes as a $\boldsymbol{T_{3.5}}$ **space**. Another characterization of Tychonov spaces will be given in 17.23.

Let (X, D) be a gauge space, let $\mathfrak{U}$ be the resulting uniformity, and let $\mathfrak{T}$ be the resulting topology (see 5.33). Then two points $x, y \in X$ are topologically indistinguishable if and only if $d(x,y) = 0$ for every $d \in D$. Show that the following conditions are equivalent:

(A) The gauge D is separating, as defined in 2.13.

(B) The topology $\mathfrak{T}$ is Kolmogorov — that is, any two distinct points are topologically distinguishable — so $(X, \mathfrak{T})$ is a Tychonov space.

(C) $\bigcap_{U \in \mathfrak{U}} U$ is the diagonal set $I = \{(x,x) : x \in X\}$. That is, if x, y are any two distinct points in X, then there exists some $U \in \mathfrak{U}$ such that $(x,y) \notin U$. A uniformity satisfying this condition is said to be a **separating uniformity**.

16.18. *Remarks: regular versus completely regular.* It is an easy exercise that every completely regular space is regular; thus any Tychonov space is also T_3.

There exist topological spaces that are regular but not completely regular; likewise, spaces that are T_3 but not $T_{3.5}$. All known examples of such spaces are very complicated; we shall not present one here. The ambitious reader can find such examples in Steen and Seebach [1970].

In fact, most topologies used in analysis are completely regular. Two elementary topologies that are *not* completely regular are the lower set topology on $\mathbb{N}$ (see 16.8.d) and the cofinite topology on an infinite set (see 16.12.a), but these are not regular either, and they are somewhat contrived: they are not typical of the topologies used in analysis.

16.19. *Remarks: gauge versus uniformity.* By Weil's Theorem, gauge spaces (X, D) and uniform spaces $(X, \mathcal{U})$ are in some sense "the same thing." Technically, there is a slight difference: A uniform space is equipped with a uniform equivalence class of gauges, whereas a gauge space is equipped with one particular gauge D from that uniform equivalence class. In practice, it is often convenient to represent a uniformity $\mathcal{U}$ by working with some particular gauge D that determines that uniformity. Although the gauge is more specific than the uniformity, most of the properties of interest to us are actually uniform properties — i.e., they are preserved if we replace the gauge with any other uniformly equivalent gauge. Thus, when we discuss a gauge space (X, D), in most cases we are actually concerned with the associated uniform space $(X, \mathcal{U})$. *For most purposes we can and will use gauges and uniformities interchangeably.* Each has its conceptual advantages.

16.20. *Further exercise.* Let $\mathcal{T}$ be a completely regular topology. Then the largest gauge that is compatible with $\mathcal{T}$ (as defined in 5.15.h) is the set of all pseudometrics $d : X \times X \to [0, +\infty)$ that are jointly continuous — i.e., continuous when $X \times X$ is given its product topology and $[0, +\infty)$ is given its usual topology.

Contrast this with 18.12; also see the specialization in 26.31.

16.21. Let X be a topological space, and let Q be its Kolmogorov quotient (as defined in 16.5). Then X is completely regular $\Longleftrightarrow$ Q is completely regular $\Longleftrightarrow$ Q is $T_{3.5}$. In fact, a gauge on either space can be used to produce a corresponding gauge on the other space; the pseudometrics d on X and e on Q correspond to each other by this formula:

$$d(x_1, x_2) \quad = \quad e\Big(\pi(x_1), \pi(x_2)\Big).$$

Two points x_1, x_2 in X are topologically indistinguishable if and only if $d(x_1, x_2) = 0$ for every pseudometric $d \in D$. This is a slight specialization of the observations in 15.31.h.

If X is a pseudometric space — i.e., if the gauge on X consists of just a single pseudometric — then Q is a metric space.

PARTITIONS OF UNITY

16.22. *Definition.* Let X be a topological space. A collection $\mathcal{S} = \{S_\alpha : \alpha \in A\}$ of subsets of X is called

point finite if each point of X belongs to only finitely many S_α's;

locally finite (or *neighborhood finite*) if each point of X has a neighborhood that meets at most finitely many S_α's.

16.23. *Basic properties.*

a. Any finite collection of sets is locally finite.

On the other hand, a locally finite collection of sets need not be finite. For a trivial example, let X be an infinite set with the discrete topology, and consider the singletons of X. For an example in a more familiar setting, let X be the real line with its usual topology; then the intervals $[n, n+1]$ (for integers n) form an infinite collection of sets that is locally finite.

b. Any locally finite collection of sets is point finite.

On the other hand, a point finite collection of sets need not be locally finite. For a trivial example, let X be an infinite set with the indiscrete topology; consider the singletons of X. For an example in a more familiar setting, let X be the real line with its usual topology; then each point of X is in at most one of the open intervals $[\frac{1}{n+1}, \frac{1}{n}]$ (for integers $n > 0$), but any neighborhood of 0 contains infinitely many of those intervals.

c. If $\mathcal{S} = \{S_\alpha : \alpha \in A\}$ is a locally finite collection of sets, then $\{\mathrm{cl}(S_\alpha) : \alpha \in A\}$ is also locally finite, and $\bigcup_{\alpha\in A} \mathrm{cl}(S_\alpha) = \mathrm{cl}\left(\bigcup_{\alpha\in A} S_\alpha\right)$. (This generalizes 15.5.b slightly.)

d. If $\mathcal{G}$ is an open cover of X and X can also be covered by a locally finite open refinement of $\mathcal{G}$, then X can also be covered by a locally finite open precise refinement of $\mathcal{G}$ (with definitions as in 1.26).

Hint: Let $\mathcal{G} = \{G_\beta : \beta \in A\}$ be the given cover, and let $\mathcal{S} = \{S_\alpha : \alpha \in B\}$ be a locally finite open refinement that covers X — that is, $\mathcal{S}$ covers X, and each S_α is contained in some G_β. Use the Axiom of Choice to define a function $\gamma : A \to B$ such that $S_\alpha \subseteq G_{\gamma(\alpha)}$. Now let $T_\beta = \bigcup_{\alpha\in\gamma^{-1}(\beta)} S_\alpha$; then $\{T_\beta : \beta \in B\}$ is a locally finite open cover of X and $T_\beta \subseteq G_\beta$ for each β.

16.24. *Definition.* Let X be a topological space. A **partition of unity** on X is a collection $\{f_\alpha : \alpha \in A\}$ of continuous functions from X into [0,1], satisfying $\sum_{\alpha\in A} f_\alpha(x) = 1$ for each $x \in X$, and such that the sets

$$f_\alpha^{-1}\Big((0,1]\Big) \quad = \quad \{x \in X \ : \ f_\alpha(x) > 0\} \qquad\qquad (\alpha \in A)$$

form a locally finite collection. Note that the sets $f_\alpha^{-1}\Big((0,1]\Big)$ must then form a cover — i.e., their union is equal to X.

The partition of unity $\{f_\alpha : \alpha \in A\}$ is said to be **subordinated** to a given cover $\{T_\beta : \beta \in B\}$ if each set $f_\alpha^{-1}\Big((0,1]\Big)$ is contained in some T_β. The partition of unity $\{f_\alpha\}$ is **precisely subordinated** to the given cover $\{T_\beta\}$ if, moreover, it is parametrized by the same index set (that is, $A = B$), and $f_\alpha^{-1}\Big((0,1]\Big) \subseteq T_\alpha$ for each α.

Some conditions for the existence of partitions of unity will be considered in 16.26(D) and 16.29.

16.25. *Basic properties of partitions of unity.*

a. *Typical use of partitions of unity.* Let $\{f_\alpha : \alpha \in A\}$ be a partition of unity that is precisely subordinated to a covering $\{T_\alpha : \alpha \in A\}$. For each α, let $g_\alpha : X \to \mathbb{R}$ be some given continuous function. Show that a continuous function $g : X \to \mathbb{R}$ can be defined by $g(x) = \sum_{\alpha \in A} f_\alpha(x) g_\alpha(x)$.

We say that g is formed by **patching together** the g_α's. Note that for each x, $g(x)$ is a convex combination of finitely many $g_\alpha(x)$'s. In many cases of interest, g inherits many of the properties of the g_α's. See for instance 18.6.

b. *Availability of precise partitions.* If X has a partition of unity subordinated to a given cover $\{T_\beta : \beta \in B\}$, then X also has a partition of unity that is precisely subordinated to that cover (as defined in 16.24).

Hint: This is similar to 16.23.d. Let $\{f_\alpha : \alpha \in A\}$ be the given partition of unity. Use the Axiom of Choice to define a function $\gamma : A \to B$ such that $f_\alpha^{-1}\big((0,1]\big) \subseteq T_{\gamma(\alpha)}$. Show that the functions $g_\beta = \sum_{\alpha \in \gamma^{-1}(\beta)} f_\alpha$ satisfy the requirements.

c. *Making the sum come out right.* Let $\{f_\alpha : \alpha \in A\}$ be a collection of continuous functions from X into $[0,\infty)$, such that the sets $f_\alpha^{-1}\big((0,\infty)\big)$ form a locally finite cover of X. If $\sum_\alpha f_\alpha = 1$, then $\{f_\alpha\}$ is a partition of unity. If we do not have $\sum_\alpha f_\alpha = 1$, we can modify the f_α's to obtain a partition of unity, as follows:

Using the fact that the sets $f_\alpha^{-1}\big((0,\infty)\big)$ form a locally finite cover, prove that the function $s(x) = \sum_{\alpha \in A} f_\alpha(x)$ is continuous and positive. Then define $g_\alpha(x) = f_\alpha(x)/s(x)$. The g_α's form the desired partition of unity.

Normal Topologies

16.26. *Definition and proposition.* A topological space X is **normal** if it satisfies any of the following equivalent conditions:

(A) Any two disjoint closed sets are contained in disjoint open sets.

(B) **(The Shrinking Lemma.)** Let $\mathcal{T} = \{T_\alpha : \alpha \in A\}$ be an open cover of X that is point finite (see 16.22). Then $\mathcal{T}$ has a **shrinking** — i.e., there exists an open cover $\mathcal{S} = \{S_\alpha : \alpha \in A\}$ such that $\mathrm{cl}(S_\alpha) \subseteq T_\alpha$ for each α.

(C) **(Urysohn's Lemma.)** If A and B are disjoint closed subsets of X, then there exists a continuous function $\sigma : X \to [0,1]$ that takes the value 0 everywhere on A and the value 1 everywhere on B.

(D) For each locally finite open cover of X, there exists a partition of unity precisely subordinated to that cover.

Proof of (B) $\Rightarrow$ (A). Let F_1 and F_2 be disjoint closed sets. Then $\complement F_1$ and $\complement F_2$ form an open cover of X, which is finite and hence point finite. Let G_1, G_2 be a shrinking — i.e., let $\{G_1, G_2\}$ be an open cover with $\mathrm{cl}(G_i) \subseteq \complement F_i$ for each i. Then $\complement\mathrm{cl}(G_1), \complement\mathrm{cl}(G_2)$ are open disjoint sets that contain F_1, F_2, respectively.

Proof of (A) $\Rightarrow$ (B). Let $\mathcal{T} = \{T_\alpha : \alpha \in A\}$ be a given open, point finite cover of X. By the Well Ordering Principle (see 6.20), let $\preccurlyeq$ be a well ordering of A. We shall define the sets S_α by transfinite recursion (see 3.40).

Let any $\beta \in A$ be given. Assume that open sets S_α have been chosen for all $\alpha \prec \beta$, satisfying $\mathrm{cl}(S_\alpha) \subseteq T_\alpha$ and also satisfying

$$\left(\bigcup_{\alpha \prec \beta} S_\alpha\right) \cup \left(\bigcup_{\alpha \succcurlyeq \beta} T_\alpha\right) \quad = \quad X.$$

(In other words, replacing the T_α's with S_α's for all $\alpha \prec \beta$ does not cost us the property of having a cover of X.) We now wish to choose S_β. Let

$$F_\beta \quad = \quad X \setminus \left[\left(\bigcup_{\alpha \prec \beta} S_\alpha\right) \cup \left(\bigcup_{\alpha \succ \beta} T_\alpha\right)\right].$$

Then F_β is a closed subset of T_β. Since X is normal, there is some open set S_β such that $F_\beta \subseteq S_\beta \subseteq \mathrm{cl}(S_\beta) \subseteq T_\beta$. Then $(X \setminus F_\beta) \cup S_\beta = X$, completing our recursive construction.

To show that the sets S_γ form a cover of X, let any $x \in X$ be given. Since the T_γ's are a point finite cover, there are just finitely many indices $\gamma_1 \prec \gamma_2 \prec \cdots \prec \gamma_n$ that satisfy $x \in T_{\gamma_j}$. Then x must be a member of $\bigcup_{j=1}^n S_{\gamma_j}$.

Proof of (C) $\Rightarrow$ (A). If such a function σ exists, then $\sigma^{-1}([0, \frac{1}{3}))$ and $\sigma^{-1}((\frac{2}{3}, 1])$ are disjoint open sets containing A and B, respectively.

Proof of (A) $\Rightarrow$ (C). Let $F_0 = A$. Use (A) to find some closed set F_1 such that $F_0 \subseteq \mathrm{int}\,(F_1) \subseteq F_1 \subseteq X \setminus B$.

Let D be the set of all dyadic rationals in [0,1]; that is,

$$D \quad = \quad \left\{2^{-k}m \in [0,1] \ : \ m, k \in \mathbb{N} \cup \{0\}\right\}.$$

We shall now recursively construct closed sets F_r for $r \in D$, chosen so that $r < s \Rightarrow F_r \subseteq \mathrm{int}\,(F_s)$. The construction will be in stages. After no stages, we already have the sets F_0 and F_1. After k stages, we shall have closed sets F_r for all r in the finite sequence

$$S_k \quad = \quad \left(0\,,\ \frac{1}{2^k}\,,\ \frac{2}{2^k}\,,\ \frac{3}{2^k}\,,\ \ldots\,,\ \frac{2^k-2}{2^k}\,,\ \frac{2^k-1}{2^k}\,,\ 1\right).$$

To construct the next stage, let r and s be any two consecutive numbers in S_k. Since $F_r \subseteq \mathrm{int}\,(F_s)$, we may use (A) to choose some closed set $F_{(r+s)/2}$ that satisfies

$$F_r \quad \subseteq \quad \mathrm{int}\,(F_{(r+s)/2}) \quad \subseteq \quad F_{(r+s)/2} \quad \subseteq \quad \mathrm{int}(F_s).$$

(This *choosing* is an application of the Principle of Dependent Choice 6.28.) In this fashion we choose closed sets F_t for all $t \in S_{k+1} \setminus S_k$; that completes the recursion. Now we may define

$$\varphi(x) \quad = \quad \begin{cases} \inf\{r \in D \ : \ x \in F_r\} & \text{if } x \in F_1, \\ 1 & \text{if } x \notin F_1. \end{cases}$$

Observe that

$$\{x \in X \ : \ \varphi(x) < p\} \quad \subseteq \quad \mathrm{int}\,(F_p) \quad \subseteq \quad F_p \quad \subseteq \quad \{x \in X \ : \ \varphi(x) \leq p\}.$$

It remains to show that φ is continuous; we leave the details as an exercise.

Proof of (D) $\Rightarrow$ (C). If F_1 and F_2 are disjoint closed sets, then $\{\complement F_1, \complement F_2\}$ is an open cover of X. Let φ_1, φ_2 be a precisely subordinated partition of unity; then φ_j vanishes on F_j $(j = 1, 2)$. Hence φ_1 takes the value 1 on F_2 and vanishes on F_1.

Proof that (B) *and* (C) *together imply* (D). Let $\mathcal{T} = \{T_\alpha : \alpha \in A\}$ be a given locally finite open cover. Let $\mathcal{S} = \{S_\alpha : \alpha \in A\}$ be a shrinking of that cover, as in (B). For each α, by (C) there is some continuous function $f_\alpha : X \to [0, 1]$ that takes the value 1 on S_α and the value 0 outside $\mathrm{cl}(T_\alpha)$. Now form a partition of unity $\{g_\alpha\}$ as in 16.25.c.

16.27. *Remarks and examples.* Normality is most useful when it occurs in conjunction with the symmetric condition. Normality does not imply symmetric. For instance, the spaces $X = \{1, 2\}$ with topology $\mathcal{T} = \{\varnothing, \{1\}, \{1, 2\}\}$ and $Y = \{1, 2, 3\}$ with topology $\mathcal{U} = \{\varnothing, \{1\}, \{1, 2, 3\}\}$ are both normal; neither is a symmetric space. (The topology $\mathcal{T}$ is Kolmogorov; the topology $\mathcal{U}$ is not.)

It is easy to see that any normal symmetric space is also completely regular. But normal plus symmetric is strictly stronger than completely regular; in 17.39 we shall give an example of a space that is $T_{3.5}$ but not normal.

16.28. *Examples and corollaries.*

a. Any pseudometric space (X, d) is normal. *Hint*: If F_0, F_1 are disjoint nonempty closed sets, let

$$G_j \quad = \quad \{x \in X \quad : \quad \mathrm{dist}(x, F_j) \ < \ \mathrm{dist}(x, F_{1-j})\} \qquad (j = 1, 2).$$

Other important examples. In 17.7.g we shall show that any compact preregular space is normal.

b. A normal, T_1 space is commonly known as a $\boldsymbol{T_4}$ **space**. Show that any normal, symmetric space is completely regular, hence $T_4 \Rightarrow T_{3.5}$.

Paracompactness

16.29. *Definition.* Recall from 1.26 the definitions of "refinement" and "precise refinement."

Let X be a preregular topological space. Then the following conditions are equivalent. If any, hence all, are satisfied, we say X is **paracompact**.

(A) Every open cover of X has a locally finite refinement.

(B) Every open cover of X has a precise locally finite refinement.

(C) For each open cover of X, there exists a partition of unity subordinated to that cover.

(D) For each open cover of X, there exists a partition of unity precisely subordinated to that cover.

Remarks. It is clear from 16.26(D) that

every paracompact space is normal

— at least, using our definition of "paracompact." Hence any paracompact space is also regular and completely regular. (In 17.38 we give an example of a T_4 space that is not paracompact.)

The reader is cautioned that the definition of "paracompact" varies in the literature. Most mathematicians make either regularity or Hausdorffness (that is, regular or T_2) a part of the definition. Note that either of these implies preregular, so any space that satisfies either of those definitions of "paracompact" is also paracompact by our definition. On the other hand, a few mathematicians omit any such assumptions and simply take condition (A) as the definition of "paracompact;" they then speak about spaces that are "both paracompact and regular."

Proof of theorem. Obviously (B) $\Rightarrow$ (A) and (D) $\Rightarrow$ (C). The implications (A) $\Rightarrow$ (B) and (C) $\Rightarrow$ (D) follow from 16.23.d and 16.25.b. The implications (C) $\Rightarrow$ (A) or (D) $\Rightarrow$ (B) follow from the definition of partition of unity — i.e., if $\{f_\alpha : \alpha \in A\}$ is a partition of unity, then $f_\alpha^{-1}((0,1]) \ : \ \alpha \in A\}$ is a locally finite collection of sets.

In view of 16.26(D), it suffices to show that conditions (B) and preregular imply normal. We shall show, by one argument, that

(i) preregular and (B) imply regular, and

(ii) regular and (B) imply normal.

Let some disjoint sets A and B be given, where B is a closed set; for (i) we assume A is a singleton, and for (ii) we assume A is a closed set. We are to show that

(*) A and B are contained in disjoint open sets.

Temporarily fix any $b \in B$. We shall first show that

(**) the sets A and $\{b\}$ are contained in some disjoint open sets G_b and H_b, respectively.

This is clear in case (ii). In case (i) we have $\mathrm{cl}(\{b\}) \subseteq B$ and hence $a \notin \mathrm{cl}(\{b\})$; our claim now follows from 16.10(D). In either case, we have established (**).

Cover X by the open sets H_b and $X \setminus B$. By our assumption (B), this open cover has a precise locally finite refinement consisting of open sets $J_b \subseteq H_b$ and $N \subseteq X \setminus B$. Let

$J = \bigcup_{b\in B} J_b$. Then J and $\complement\mathrm{cl}(J)$ are disjoint open sets. We shall show that $B \subseteq J$ and $A \subseteq \complement\mathrm{cl}(J)$.

Since $\{J_b : b \in B\} \cup \{N\}$ is a cover of X, and $N \subseteq X \setminus B$, we must have $B \subseteq J$. On the other hand, since the collection of sets $\{J_b\}$ is locally finite, we have

$$\mathrm{cl}(J) = \bigcup_{b\in B} \mathrm{cl}(J_b) \subseteq \bigcup_{b\in B} \mathrm{cl}(H_b) \subseteq \bigcup_{b\in B} \complement G_b = \complement\left(\bigcap_{b\in B} G_b\right);$$

hence $A \subseteq \bigcap_{b\in B} G_b \subseteq \complement\mathrm{cl}(J)$. This completes our proof of (*).

16.30. Dowker's Sandwich Theorem (optional). Let X be a paracompact space. Let $a, b : X \to \mathbb{R}$ be functions such that a is lower semicontinuous, b is upper semicontinuous, and $a(x) > b(x)$ for every $x \in X$. Then there is some continuous $g : X \to \mathbb{R}$ such that $a(x) > g(x) > b(x)$ for every $x \in X$. (*Remark.* It is interesting to compare this result with (HB6) in 12.31.)

Proof of theorem. For real numbers r let $G_r = \{x \in X : a(x) > r > b(x)\}$. The sets G_r ($r \in \mathbb{R}$) form an open cover of X. Let $\{f_r : r \in \mathbb{R}\}$ be a subordinated partition of unity; show that $g(x) = \sum_{r\in\mathbb{R}} r f_r(x)$ has the required properties.

16.31. Theorem (A. H. Stone, 1948). Every pseudometric space is paracompact.

Proof. Our proof follows Rudin [1969]. Let (X, d) be a pseudometric space. Let $B(x, r)$ denote the open ball with center x and radius r — that is, the set $\{u \in X : d(x, u) < r\}$.

We know (X, d) is regular, by 16.16. Let $\{C_\alpha : \alpha \in A\}$ be an open cover of X; we shall construct a locally finite open refinement $\{D_{\alpha,n} : \alpha \in A,\ n \in \mathbb{N}\}$. By the Well Ordering Principle (in 6.20), let $\preccurlyeq$ be a well ordering of the set A. Let $A \times \mathbb{N}$ be given this lexicographical ordering: $(a, n) \prec (a', n')$ if either

(a) $n < n'$, or

(b) $n = n'$ and $a \prec a'$.

Then $A \times \mathbb{N}$ is also well ordered; this is a special case of 3.44.a(ii). We shall define the $D_{\alpha,n}$'s recursively in that order — that is, first we define all the $D_{\alpha,1}$'s (in order of α's); then we define all the $D_{\alpha,2}$'s; then all the $D_{\alpha,3}$'s; etc.

To define $D_{\alpha,n}$, assume that all the preceding D's have already been defined (an assumption that is trivially satisfied when we begin the defining process, since then there are no preceding D's). Let $E_{\alpha,n}$ be the set of all x's that satisfy:

(i) α is the first member of A satisfying $x \in C_\alpha$,

(ii) n is large enough so that $B(x, 2^{2-n}) \subseteq C_\alpha$, and

(iii) x is not a member of any previous D.

Then let

$$D_{\alpha,n} = \bigcup_{x\in E_{\alpha,n}} B(x, 2^{-n}). \qquad (*)$$

Each set $D_{\alpha,n}$ is a union of open balls, hence is open. Also, each of those balls is contained in C_α, hence the collection of open sets $\{D_{\alpha,n} : \alpha \in A, n \in \mathbb{N}\}$ refines the open cover $\{C_\alpha : \alpha \in A\}$. To show that $\{D_{\alpha,n}\}$ covers X, let any $x \in X$ be given. Then there is a first $\alpha \in A$ satisfying $x \in C_\alpha$, and there is some n large enough so that $B(x, 2^{2-n}) \subseteq C_\alpha$. Then $x \in \bigcup_{j \leq n} \bigcup_{\beta \in A} D_{\beta,j}$.

It remains only to show that $\{D_{\alpha,n}\}$ is locally finite. Let any $\xi \in X$ be given; we shall exhibit a neighborhood G of X that meets only finitely many of the $D_{\alpha,n}$'s. Choose some λ and m such that ξ is in the open set $D_{\lambda,m}$. Then choose some positive integer j large enough so that $B(\xi, 2^{-j}) \subseteq D_{\lambda,m}$. We shall use $G = B(\xi, 2^{-m-j})$. The proof consists of showing that (1) G does not meet any $D_{\alpha,n}$ with $n \geq m + j$, and (2) for each positive integer $n \leq m + j - 1$ there is at most one α such that G meets $D_{\alpha,n}$.

Proof of (1). The set $E_{\alpha,n}$ is disjoint from $D_{\lambda,m}$; this follows from (iii) and the fact that $n \geq m + j > m$. For any $y \in E_{\alpha,n}$, we have $d(\xi, y) \geq 2^{-j}$ since $B(\xi, 2^{-j}) \subseteq D_{\lambda,m}$. Since $m + j \geq j + 1$ and $n \geq j + 1$, we have $G \cap B(y, 2^{-n}) \subseteq B(\xi, 2^{-j-1}) \cap B(y, 2^{-j-1}) = \varnothing$. The set $D_{\alpha,n}$ is the union of such balls $B(y, 2^{-n})$ and therefore is disjoint from G.

Proof of (2). Suppose $p \in G \cap D_{\alpha,n}$ and $q \in G \cap D_{\beta,n}$ where $n < m + j$ and $\alpha \neq \beta$; we shall obtain a contradiction. We may assume $\alpha \prec \beta$. By the definition $(*)$, we have $d(p, p') < 2^{-n}$ and $d(q, q') < 2^{-n}$ for some $p' \in E_{\alpha,n}$ and $q' \in E_{\beta,n}$. By (i), since $\beta \succ \alpha$, we have $q' \notin C_\alpha$. By (ii), $B(p', 2^{2-n}) \subseteq C_\alpha$. Thus $d(p', q') \geq \operatorname{dist}(p', \complement C_\alpha) \geq 2^{2-n}$, and therefore

$$d(p,q) \;\geq\; d(p',q') - d(p,p') - d(q,q') \;>\; 2^{-n}(4-1-1) \;=\; 2^{-n+1} \;\geq\; 2^{-m-j+2}.$$

This contradicts the fact that p, q both lie in $G = B(\xi, 2^{-m-j})$ and completes the proof of the theorem.

Hereditary and Productive Properties

16.32. *Definitions.* A property P is **hereditary** if, whenever Y is a topological space with property P and X is a subset of Y equipped with the relative topology, then X also has property P. For instance, Hausdorff is a hereditary property, since any subspace of a Hausdorff space is also Hausdorff when equipped with the relative topology.

A property P is **productive** if, whenever $X = \prod_{\lambda \in \Lambda} Y_\lambda$ is a product of topological spaces and the Y_λ's all have property P, then X (equipped with the product topology) also has property P.

A property P is an **initial property** if, whenever X has the initial topology determined by a collection of mappings $f : X \to Y_\lambda$, and the Y_λ's are topological spaces with property P, then X also has property P. Note that any initial property is also a hereditary property and a productive property, since the relative and product topologies are special cases of initial topologies.

16.33. *Exercises and remarks.*

a. All the following separation axioms are initial properties: symmetric, preregular, regular, completely regular. The verification of these facts are fairly straightforward exercises; we shall omit the details.

b. All the separation axioms T_n, for $n = 0, 1, 2, 3, 3.5$, are hereditary and productive properties. In fact, if X has the initial topology determined by a collection of mappings $f : X \to Y_\lambda$, and that collection of mappings *separates points* of X, and the Y_λ's have one of the properties $T_0, T_1, T_2, T_3, T_{3.5}$, then X also has that property.

c. Normalcy and paracompactness are not hereditary; we shall prove that via an example in 17.40.a.

d. Normalcy is not productive; we shall prove that by an example in 17.40.b.

e. Paracompactness is not productive. Indeed, let X be the real line equipped with the topology generated by all sets of the form $\{x \in \mathbb{R} : a \leq x < b\}$, for $a, b \in \mathbb{R}$. (This is called the **right half-open interval topology**, or the **lower limit topology**.) It can be shown that X is a paracompact Hausdorff space, but $X \times X$ (with the product topology) is not paracompact. (In fact, $X \times X$ is not normal; this gives another proof that normalcy is not productive.) We omit the details of the proof, which can be found in topology books.

Chapter 17

Compactness

17.1. *Preview.* In $\mathbb{R}^n$, a set is compact if and only if it is closed and bounded. That notion is generalized in this and the next few chapters. The following chart shows relations between some of the main relatives of compactness.

Characterizations in Terms of Convergences

17.2. *Definition and exercise.* Let $(X, \mathcal{T})$ be a topological space. We say that X is **compact** if any of the following equivalent conditions are satisfied. (Examples will be given later in the chapter.)

(A) Every open cover of X has a finite subcover. That is, if $\mathcal{G} = \{G_\lambda : \lambda \in \Lambda\}$ is a cover of X consisting of open sets, then some finite subcollection $\{G_{\lambda_1}, G_{\lambda_2}, \ldots, G_{\lambda_n}\}$ is also a cover of X. (This is the most common definition of compactness in the mathematical literature.)

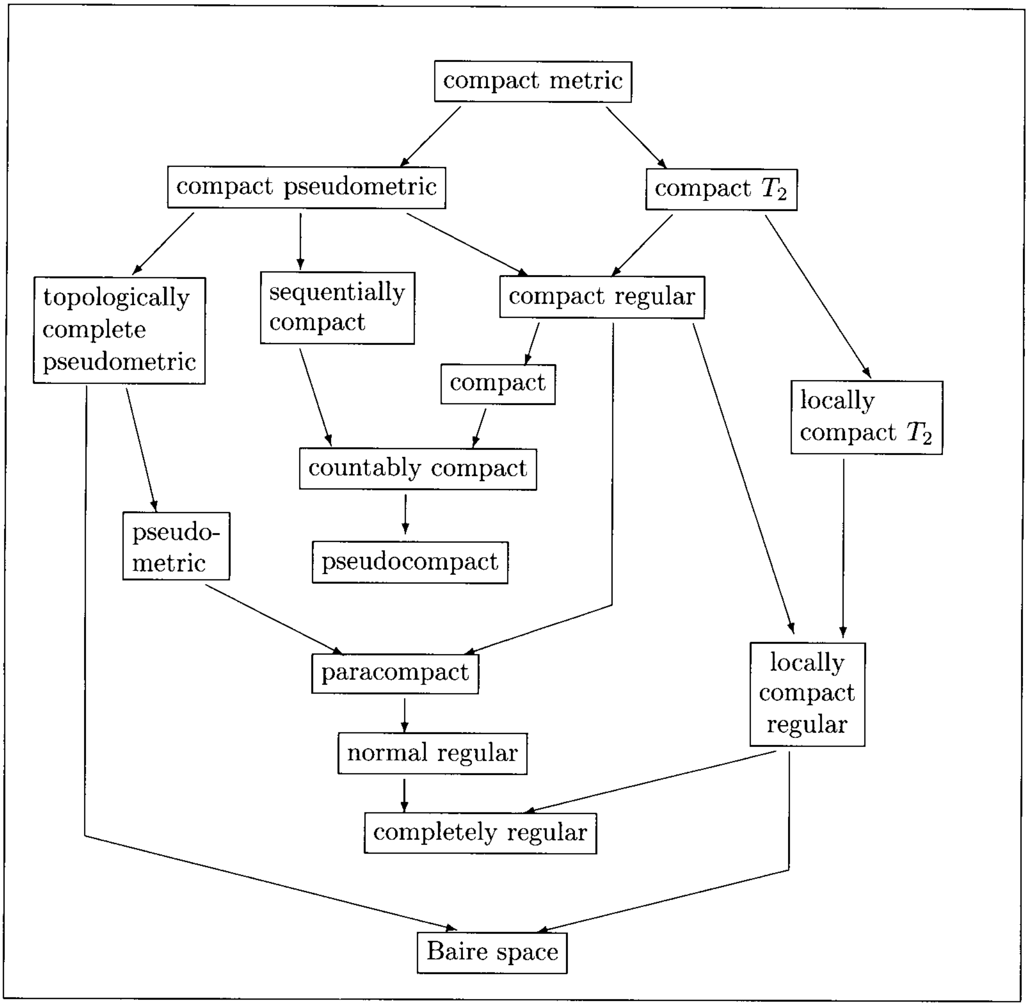
(B) Every filter subbase consisting of closed subsets of X is fixed. That is, whenever $\mathcal{S} = \{S_\lambda : \lambda \in \Lambda\}$ is a collection of closed subsets of X that has the finite intersection property, then $\bigcap_{\lambda \in \Lambda} S_\lambda$ is nonempty.

(C) Every net in X has a cluster point — i.e., every net has a convergent subnet.

(D) If $\mathcal{F}$ is a proper filter on X, then the set $\bigcap_{F \in \mathcal{F}} \mathrm{cl}(F) = \{\text{cluster points of } \mathcal{F}\}$ is nonempty.

Remarks. In view of 7.19 and 15.38, it does not matter which kind of subnet we use in (C). Also, the equation $\bigcap_{F \in \mathcal{F}} \mathrm{cl}(F) = \{\text{cluster points of } \mathcal{F}\}$ was noted in 15.38 (regardless of compactness).

Proof of equivalence. The equivalence of (A) and (B) follows from taking complements — i.e., take $G_\lambda = X \setminus S_\lambda$. The equivalence of (C) and (D) is just the correspondence between nets and proper filters (see 7.9, 7.11, and 7.31). For (D) $\Rightarrow$ (B), let $\mathcal{F}$ be the filter generated by $\mathcal{S}$. For (B) $\Rightarrow$ (D), let $\mathcal{S} = \{\mathrm{cl}(F) : F \in \mathcal{F}\}$.



17.3. *More definitions.* A subset K of a topological space Y is said to be a **compact set** if K is a compact space when equipped with the relative topology induced by Y. This notion is so important that we shall now reformulate all four of the conditions stated in 17.2; the formulations below are occasionally more convenient than those given in 17.2. A set $K \subseteq Y$ is a compact set if one (hence all) of the following conditions is satisfied.

(A) Whenever $\{G_\lambda : \lambda \in \Lambda\}$ is a collection of open subsets of Y with union containing K, then $\bigcup_{\lambda \in L} G_\lambda \supseteq K$ for some finite set $L \subseteq \Lambda$.

(B) Whenever $\{F_\lambda : \lambda \in \Lambda\}$ is a collection of closed subsets of Y such that the collection $\{K \cap F_\lambda : \lambda \in \Lambda\}$ has the finite intersection property, then $\bigcap_{\lambda \in \Lambda} F_\lambda$ meets K.

(C) Every net in K has a cluster point in K — i.e., every net in K has a subnet that converges to a point in K.

(D) If $\mathcal{F}$ is a proper filter on Y and $K \in \mathcal{F}$, then $\mathcal{F}$ has a superfilter that converges to some point in K.

Although these conditions refer to the topology of Y, they do not actually depend on Y, except insofar as it determines the relative topology on K. Thus, if $K \subseteq Y \cap Z$, where Y and Z are two topological spaces that determine the same relative topology on K, then K is compact $\iff$ K is a compact subset of Y $\iff$ K is a compact subset of Z.

17.4. The preceding definitions of compactness and their proof of equivalence did not require the Axiom of Choice or any weakened form of Choice.

Following is another characterization of compactness; this statement is equivalent to the Ultrafilter Principle.

> **(UF18)** Let X be a topological space. Then X is compact if (and only if) every ultrafilter on X converges to some limit — or equivalently, if (and only if) every universal net in X converges to some limit.

In fact, this statement is equivalent to the Ultrafilter Principle with or without the parenthesized "and only if" part. It follows from the definitions of "ultrafilter" and "compact" that if X is compact, then every ultrafilter on X converges; this implication does not require any arbitrary choices and thus is valid in ZF. We assert that our earlier versions of UF are equivalent to the remaining statement that

> (*) X is compact if every ultrafilter on X converges.

Indeed, (*) follows easily from (UF3) in 7.24, together with the definition of "compact." A proof that (*) implies (UF1) will be given in 17.22.

17.5. Let $f : \Lambda \to X$ be a mapping from a set Λ (without any topology necessarily specified) into a compact Hausdorff space X.

(i) Suppose (λ_α) is a universal net in Λ. Then $(f(\lambda_\alpha))$ is a universal net in X, which converges to a unique limit. We say that f **converges along the universal net** (λ_α) to that limit.

(ii) Equivalently, let $\mathcal{U}$ be an ultrafilter on Λ. Then $\{f(U) : U \in \mathcal{U}\}$ is a filterbase on X, and the filter it generates is an ultrafilter. That ultrafilter converges to a unique limit in X. We say that f **converges along the ultrafilter** $\mathcal{U}$ to that limit. Let us denote that limit by $\lim_{\mathcal{U}} f$. We may restate its definition: $\lim_{\mathcal{U}} f$ is the unique point in X with the property that each neighborhood N of $\lim_{\mathcal{U}} f$ contains some set $f(U)$ with $U \in \mathcal{U}$. (This notion is discussed also by Bourbaki [1966].)

We have $\lim_{\mathcal{U}} f = \lim_{\mathcal{U}} g$ whenever f and g are $\mathcal{U}$-equivalent in the sense of 9.41, and so $\lim_{\mathcal{U}}$ is in fact well defined on the quotient space $X^\Lambda/\mathcal{U}$ — i.e., on the ultrapower $^*X$.

17.6. (*Optional.*) *The Ultrafilter Principle implies the Hahn-Banach Theorem.* We shall show that (UF1) $\Rightarrow$ (HB1). Let $(\Delta, \preccurlyeq)$ be a directed set. Let $\mathcal{D}$ be the filter of eventual subsets of Δ — that is, let

$$\mathcal{D} \quad = \quad \Big\{ S \subseteq \Delta \quad : \quad S \supseteq \{\delta \in \Delta : \delta \succcurlyeq \delta_0\} \text{ for some } \delta_0 \in \Delta \Big\}.$$

By (UF1), let $\mathcal{U}$ be an ultrafilter on Δ that extends $\mathcal{D}$. If $f : \Delta \to \mathbb{R}$ is a bounded function, then f may be viewed as a map into the compact Hausdorff space $[a,b]$ for some $a, b \in \mathbb{R}$. Hence we may define $\lim_{\mathcal{U}} f \in \mathbb{R}$ as in 17.5. Obviously the map $f \mapsto \lim_{\mathcal{U}} f$ is positive and linear. It is easy to verify that if $f : \Delta \to \mathbb{R}$ is a bounded function such that the net $\{f(\delta) : \delta \in \Delta\}$ converges to a limit L, then that limit is equal to $\lim_{\mathcal{U}} f$. Thus $\lim_{\mathcal{U}}$ is a Banach limit.

Remark. Actually, Pincus [1972] showed that the Hahn-Banach Theorem is *strictly* weaker than the Ultrafilter Principle, but the proof of that fact is beyond the scope of this book.

Basic Properties of Compactness

17.7. *Elementary examples and properties.*

a. Any finite subset of any topological space is compact. In particular, $\varnothing$ is compact.

b. The union of finitely many compact sets is compact.

c. Let X be a topological space. Then

$$\mathcal{I} \quad = \quad \{S \subseteq X \ : \ S \subseteq K \text{ for some compact set } K\}$$

is an ideal on X. Thus, for some purposes, we may view the members of $\mathcal{I}$ as "small" subsets of X, in the sense of 5.3.

d. Let $\mathcal{S}$ and $\mathcal{T}$ be two topologies on a set X. Then the weaker topology has more compact sets — or at least as many. That is, if $\mathcal{S} \subseteq \mathcal{T}$, then every $\mathcal{T}$-compact set is also $\mathcal{S}$-compact. It is possible for $\mathcal{S}$ and $\mathcal{T}$ to yield the same collections of compact sets even if if $\mathcal{S} \subsetneqq \mathcal{T}$; see the second and third examples below.

(i) The discrete topology on X is the strongest topology, so it should have the fewest compact sets. Show that a subset of X is compact for the discrete topology if and only if that subset is finite.

(ii) The indiscrete topology on X is the weakest topology, so it has the most compact sets. In fact, with the indiscrete topology, every subset of X is compact.

(iii) The cofinite topology is strictly stronger than the indiscrete topology (unless $\operatorname{card}(X) < 2$), but the cofinite topology also makes every subset of X compact.

e. If $(x_1, x_2, x_3, \ldots)$ is a *sequence* converging to a limit x_0 in a topological space, then the set $\{x_0, x_1, x_2, x_3, \ldots\}$ is compact. (This result does not generalize to nets.)

f. In any topological space, the intersection of a closed set and a compact set is compact. In a compact topological space, any closed set is compact. In a Hausdorff topological space, any compact set is closed.

g. Any compact preregular space is paracompact (hence normal and completely regular).

Proof. Given an open cover, any finite subcover *is* a locally finite refinement.

h. The continuous image of a compact set is compact. That is, if $f : X \to Y$ is a continuous map from one topological space into another, and $K \subseteq X$ is compact, then $f(K)$ is compact.

i. Any upper semicontinuous function from a compact set into $[-\infty, +\infty]$ assumes a maximum.

j. Dini's Monotone Convergence Theorem. Let $(g_\alpha : \alpha \in A)$ be a net of continuous functions (or more generally, upper semicontinuous functions) from a compact topological space X into $\mathbb{R}$. Assume that $g_\alpha \downarrow 0$ pointwise — i.e., assume that for each $x \in X$ the net $(g_\alpha(x))$ is decreasing and converges to 0. Then the convergence is uniform — i.e., $\lim_{\alpha \in A} \sup_{x \in X} g_\alpha(x) = 0$.

Hint: Let $\varepsilon > 0$ be given. If none of the closed sets $F_\alpha = \{x \in X : f_\alpha(x) \geq \varepsilon\}$ is empty, show that the collection of F_α's has the finite intersection property.

17.8. Proposition. Let $(X, \leq)$ be a chain ordered set (for instance, a subset of $[-\infty, +\infty]$), and let $\mathcal{T}$ be the interval topology on X (defined in 5.15.f). Then $(X, \mathcal{T})$ is compact if and only if $(X, \leq)$ is order complete.

Furthermore, if $(x_\alpha : \alpha \in \mathbb{A})$ is a net in an order complete chain, then lim inf x_α is the smallest cluster point of the net, and lim sup x_α is the largest cluster point of the net.

Proof. First, suppose that X is order complete. It follows easily from 15.42 that $\liminf x_\alpha$ and $\limsup x_\alpha$ are cluster points of (x_α). It follows from 7.47.a that any other cluster points must lie between those two.

On the other hand, suppose X is not order complete; we shall show X is not compact. Assume D is a nonempty subset of X such that $\sup(D)$ does not exist in $(X, \leq)$. Consider D itself as a directed set; we shall show that the inclusion map $i : D \to X$ is a net with no cluster point. To put our notation in a more familiar form, we shall write the net as $(i_\delta : \delta \in D)$, where in fact $i_\delta = \delta$. Consider any $z \in X$; we shall show z cannot be a cluster point of X. We consider two cases:

(i) z is not an upper bound of D. In this case there is some $\delta_0 \in D$ with $\delta_0 > z$. The set $\{x \in X \ : \ x < \delta_0\}$ contains z but is not a frequent set for the net (i_δ), so z is not a cluster point.

(ii) z is an upper bound of D, but is not the least upper bound. Thus D has some upper bound $b < z$. Then the set $\{x \in X : x > b\}$ contains z but is not a frequent set for the net (i_δ), so z is not a cluster point.

17.9. *Corollaries.*

a. The extended real line $[-\infty, +\infty]$ is compact when equipped with its usual topology. (That topology will be discussed further in 18.24.)

b. A subset of $\mathbb{R}$ is compact if and only if it is closed and bounded. In particular, any interval $[a, b] \subseteq \mathbb{R}$ (where $-\infty < a < b < +\infty$) is compact.

17.10. *Compactness and Hausdorff spaces.*

a. Let S be a subset of a Hausdorff topological space. Then S is compact if and only if S is closed and S is contained in a compact set.

b. Let S be a subset of a compact Hausdorff space. Then S is compact if and only if S is closed.

c. If X is a compact space, Y is a Hausdorff space, and $f : X \to Y$ is continuous, then f is a closed mapping — i.e., the image of a closed subset of X is a closed subset of Y.

If, furthermore, f is a bijection, then f^{-1} is also continuous — that is, f is a homeomorphism.

d. No Hausdorff topology on a set can be strictly weaker than a compact topology on that set. In other words, it is not possible for a set to have two topologies $\mathcal{S} \subsetneqq \mathcal{T}$ where $\mathcal{S}$ is Hausdorff and $\mathcal{T}$ is compact.

17.11. We shall say that a topological space X is **locally compact** if each point has a compact neighborhood. Following are some examples.

a. Any compact space is locally compact.

b. Any set with the discrete topology is a locally compact Hausdorff space.

c. $\mathbb{R}$ is locally compact.

Preview of further results. In 17.17 we shall see that $\mathbb{R}^n$ is a locally compact Hausdorff space, when equipped with the product topology. In 27.17 we shall see that no infinite dimensional Hausdorff topological vector space is locally compact.

In 17.14.d we shall see that any locally compact preregular space is completely regular.

Regularity and Compactness

17.12. If X is a symmetric space and $x \in X$, then $\mathrm{cl}(\{x\})$ is compact.

Proof. Any open cover of $\mathrm{cl}(\{x\})$ has a finite subcover; that is immediate from 16.6(B).

17.13. Let X be a preregular space, and let K be a compact subset of X. Then:

a. If $p \in X$ with $\mathrm{cl}(\{p\})$ disjoint from K, then K and $\mathrm{cl}(\{p\})$ are contained in disjoint open sets.

Proof. For each $x \in K$, we have $x \notin \mathrm{cl}(\{p\})$. By 16.10, p and x are contained in disjoint open sets A_x and B_x, respectively. Then $\{B_x : x \in K\}$ is an open cover of the compact set K, so it has a finite subcover; we have $K \subseteq B = \bigcup_{x \in I} B_x$ for some finite set $I \subseteq K$. Then p is in the open set $A = \bigcap_{x \in I} A_x$.

b. If L is closed and compact, and K and L are disjoint, then K and L are contained in disjoint open sets.

Proof. For each $p \in L$, the sets $\mathrm{cl}(\{p\})$ and K are contained in disjoint open sets A_p and B_p, respectively. The sets A_p form an open cover of L. Choose a finite subcover of the A_p's, and take their union for an open set containing L. The intersection of the corresponding B_p's is an open set containing K.

c. $\mathrm{cl}(K) = \bigcup_{x \in K} \mathrm{cl}(\{x\})$.

Proof. We have $\mathrm{cl}(K) \supseteq \bigcup_{x \in K} \mathrm{cl}(\{x\})$ if K is *any* subset of *any* topological space. For the reverse inclusion, let $p \in \mathrm{cl}(K)$; we wish to show that $p \in \bigcup_{x \in K} \mathrm{cl}(\{x\})$. Since $p \in \mathrm{cl}(K)$, we know that K meets every neighborhood of p. Hence by the preceding exercise, $\mathrm{cl}(\{p\})$ is not disjoint from K. Say $x \in \mathrm{cl}(\{p\}) \cap K$. By 16.10, then, also $p \in \mathrm{cl}(\{x\})$, as required.

d. If K is contained in an open set G, then $\mathrm{cl}(K) \subseteq G$ also.

Proof. Use the preceding exercise and 16.6(B).

e. $\mathrm{cl}(K)$ is compact. More generally, if $K \subseteq T \subseteq \mathrm{cl}(K)$, then T is compact.

Hints: Any open cover of T is also an open cover of K; use the preceding exercise.

f. If $S \subseteq K$, then $\mathrm{cl}(S)$ is compact.

Proof. $\mathrm{cl}(S) = \mathrm{cl}(S) \cap \mathrm{cl}(K)$ is the intersection of a closed set and a compact set; apply 17.7.f.

17.14. Let X be a locally compact preregular space. Then:

a. (*Neighborhoods of points.*) Each point has a neighborhood base consisting of closed compact sets (and hence X is regular).

Proof. Any $x \in X$ has a compact neighborhood, hence (by 17.13.e) has a closed compact neighborhood K. Then K is a compact preregular space, hence (by 17.7.g) K is a regular space. Hence x has a neighborhood basis in K consisting of closed sets. Since K is a neighborhood of x in X, those same sets also form a neighborhood basis for x in X. Those sets are also closed and compact in X.

b. (*Neighborhoods of compact sets.*) Let $K \subseteq G$ with K compact and G open in X. Then there exists some open set H whose closure is compact, such that $K \subseteq H \subseteq \mathrm{cl}(H) \subseteq G$.

Proof. By 17.14.a, each $x \in K$ is contained in some open set A_x whose closure is a compact set contained in G. Then the compact set K has open cover $\{A_x : x \in K\}$, hence some finite subcover $\{A_x : x \in F\}$. Let $H = \bigcup_{x \in F} A_x$. Then $\mathrm{cl}(H) = \bigcup_{x \in F} \mathrm{cl}(A_x)$ by 15.5.b or 16.23.c; hence $\mathrm{cl}(H)$ is a compact subset of G.

c. (*Local partitions of unity.*) Suppose $K \subseteq \bigcup_{j=1}^{n} G_j$, where K is compact and the G_j's are open. Then there exist continuous functions $\varphi_j : X \to [0,1]$ such that $\sum_{j=1}^{n} \varphi_j = 1$ on K, and each φ_j vanishes outside some compact subset of G_j.

Proof. Let $G = \bigcup_{j=1}^{n} G_j$. By two applications of 17.14.b, we may find open sets G', G'' such that

$$K \subseteq G'' \subseteq \mathrm{cl}(G'') \subseteq G' \subseteq \mathrm{cl}(G') \subseteq G$$

and $\mathrm{cl}(G'), \mathrm{cl}(G'')$ are compact sets. The set $\mathrm{cl}(G')$, equipped with the relative topology, is compact and a preregular space, hence paracompact (see 17.7.g). Let $T_j = G'' \cap G_j$ for $j = 1, 2, \ldots, n$, and let $T_0 = \mathrm{cl}(G') \setminus K$. These sets are relatively open in $\mathrm{cl}(G')$, and they form a cover of $\mathrm{cl}(G')$. Let $(S_j : j = 0, 1, 2, \ldots, n)$ be a shrinking of (T_j) — that is, let the S_j's be another cover of $\mathrm{cl}(G')$ consisting of relatively open sets, such that $\mathrm{cl}(S_j) \subseteq T_j$. Form a partition of unity on $\mathrm{cl}(G')$ that is precisely subordinated to (S_j); say $\varphi_0, \varphi_1, \varphi_2, \ldots, \varphi_n$ are continuous functions from $\mathrm{cl}(G')$ into $[0,1]$ such that $\sum_{j=0}^{n} \varphi_j = 1$ and φ_j vanishes outside S_j. Since $K \subseteq \mathrm{cl}(G') \setminus S_0$, we must have $\sum_{j=1}^{n} \varphi_j = 1$ on K. For $1 \leq j \leq n$, extend φ_j to all of X by defining $\varphi_j = 0$ outside of $\mathrm{cl}(G')$. Note that φ_j vanishes outside $\mathrm{cl}(S_j)$, which is a compact subset of G_j. It suffices to show that φ_j is continuous on X. Note that X is the union of the open sets G' and $\complement\mathrm{cl}(G'')$, and φ_j is continuous on each of those sets, since φ_j vanishes on the latter set.

d. *Corollary.* Any locally compact preregular space is completely regular.

Proof. Let any open set G and any point $x \in G$ be given. Then $K = \mathrm{cl}(\{x\})$ is a compact subset of G, by 16.6(B) and 17.12. Apply the preceding exercise with $n = 1$.

17.15. *Definition and proposition.* Let S be a subset of a topological space X. Following are three closely related conditions on S:

(A) $\mathrm{cl}(S)$ is compact. (It is then customary to say that S is **relatively compact**.)

(B) S is a subset of a compact set. (As we noted in 17.7.c, the sets satisfying this condition form an ideal.)

(C) Every net in S (or every proper filter on X that contains S) has a cluster point in X.

We have the following implications:

In any topological space, (A) $\Rightarrow$ (B) $\Rightarrow$ (C). (Obvious.)

In any preregular space, (B) $\Rightarrow$ (A), and so those two conditions are equivalent. (Proved in 17.13.f.)

In any regular space, (C) $\Rightarrow$ (A), and so all three conditions are equivalent. (Proved in the paragraphs below.)

Proof. Assume regular and (C). Let $\mathcal{G}$ be any proper filter on X with $\mathrm{cl}(S) \in \mathcal{G}$; we must show $\mathcal{G}$ has a cluster point. Assume not; we shall obtain a contradiction.

For each $x \in X$, since x is not a cluster point of $\mathcal{G}$, there is some neighborhood N_x of x that is disjoint from some member of $\mathcal{G}$, and hence $X \setminus N_x \in \mathcal{G}$. These conditions are preserved if we replace N_x with a smaller neighborhood of x; since X is regular, we may assume N_x is closed. Then the set $G_x = X \setminus N_x$ is open and a member of $\mathcal{G}$.

Since $\mathcal{G}$ is closed under finite intersection, for any finite set $A \subseteq X$ the set $\mathrm{cl}(S) \cap \bigcap_{a \in A} G_a$ is a member of $\mathcal{G}$ and hence nonempty. By 5.17.e, the set $S \cap \bigcap_{a \in A} G_a$ is nonempty. Thus the collection of sets $\mathcal{S} = \{S\} \cup \{G_x : x \in X\}$ has the finite intersection property and therefore is contained in a proper filter $\mathcal{F}$. By our assumption on S, that filter $\mathcal{F}$ has some cluster point $\xi \in X$. Now N_ξ is a neighborhood of ξ, hence N_ξ meets every member of $\mathcal{F}$, hence N_ξ meets G_ξ, a contradiction.

Tychonov's Theorem

17.16. Recall that the Axiom of Choice, in one form, asserts that a product $\prod_{\lambda\in\Lambda} S_\lambda$ of nonempty sets is nonempty (see (AC3) in 6.12). That result bears some resemblance to:

> **(AC21) Tychonov Product Theorem.** Any product $\prod_{\lambda\in\Lambda} Y_\lambda$ of compact topological spaces is compact.

Here it is understood that the product space is equipped with the product topology. In contrast with (AC3), however, the Tychonov Product Theorem does *not* assert that the product is nonempty. (An empty set is a perfectly acceptable compact topological space!) Thus, it may be surprising that the Tychonov Product Theorem is equivalent to the Axiom of Choice.

Proof of (AC3) $\Rightarrow$ (AC21). We shall make use of (UF18), which we already know to be a consequence of the Axiom of Choice. Let $(x_\alpha : \alpha \in \mathbb{A})$ be a universal net in $\prod_{\lambda\in\Lambda} Y_\lambda$; we must show that $(x_\alpha : \alpha \in \mathbb{A})$ is convergent. Let $\pi_\lambda : X \to Y_\lambda$ be the λth coordinate projection. The net $(\pi_\lambda(x_\alpha) : \alpha \in \mathbb{A})$ is universal in Y_λ. Let $S_\lambda = \{y \in Y_\lambda : \pi_\lambda(x_\alpha) \to y\}$. Then each S_λ is nonempty. Hence $\prod_{\lambda\in\Lambda} S_\lambda$ is nonempty, by (AC3). If $z \in \prod_{\lambda\in\Lambda} S_\lambda$, then $x_\alpha \to z$.

The proof of (AC21) $\Rightarrow$ (AC3) will be given in 17.20.

17.17. *Corollary.* Let n be a positive integer, and let $\mathbb{R}^n$ have its product topology. A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded, where "bounded" has its usual meaning (see 3.15 and 2.12.a). Hence $\mathbb{R}^n$ is locally compact.

17.18. *An auxiliary construction.* This observation will be used occasionally — e.g., in 26.9 and in 26.10.

Let Ω be an open subset of $\mathbb{R}^n$. Then there exists a sequence (G_n) of open sets whose union is Ω, such that each G_n is contained in a compact subset of Ω. We remark:

- **a.** One way to construct such a sequence (G_n) is as follows: The rational numbers are countable (see 2.20.f). Consider all the open balls $G = B(x, r)$ with the property that r and all the coordinates of x are rational numbers, and the closure of $B(x, r)$ is contained in Ω. There are only countably many such balls; let them be the G_n's.
- **b.** In most applications of such a sequence, the particular choice of the G_n's is not important. Any other such sequence H_n will do just as well, because (*exercise*) each G_n is contained in the union of finitely many of the H_n's, and vice versa.

Compactness and Choice (Optional)

17.19. *Remarks.* This subchapter is optional. It is concerned with showing that certain propositions imply either the Axiom of Choice or weakened forms of Choice. Many mathematicians take the viewpoint that the Axiom of Choice is simply "true;" with that viewpoint, this subchapter is of no interest.

17.20. *Compactness equivalents of AC.* We shall prove that the Axiom of Choice is equivalent, not only to Tychonov's Theorem, but also to several other principles that are seemingly weaker:

> **(AC22)** Any product of compact gauge spaces is compact.
>
> **(AC23)** Any product of knob spaces is compact.
>
> **(AC24)** Any product of T_1 compact topological spaces is compact.
>
> **(AC25)** Any product of topological spaces, each equipped with the cofinite topology, is compact.

(Gauge topologies and knob topologies were introduced in 5.15.h and 5.34.c, respectively.)

Intermediate proofs. Any knob space is a compact gauge space, and any space with the cofinite topology is a compact T_1 space. Thus, the proofs of (AC21) $\Rightarrow$ (AC22) $\Rightarrow$ (AC23) and (AC21) $\Rightarrow$ (AC24) $\Rightarrow$ (AC25) are obvious.

Proof of (AC23) $\Rightarrow$ (AC3) and (AC25) $\Rightarrow$ (AC3). This argument is from Kelley [1950]. Define S_λ, ξ_λ, etc., as in 6.24. Equip each Y_λ with either the knob topology or the cofinite topology. In either case, the set S_λ is closed. Hence $\mathcal{F}$ is a filterbase consisting of closed subsets of X. By assumption, X is compact; hence the intersection of the members of $\mathcal{F}$ is nonempty — completing the proof indicated in 6.24.

Further remarks. The Axiom of Choice is equivalent to Tychonov's Theorem, if we use any of the usual definitions of compactness, given in 17.2. An alternate approach is taken by Comfort [1968]. Comfort suggests a different definition of compactness, which is more complicated than the usual definitions but has this interesting property: we can prove that the product of Comfort-compact spaces is Comfort-compact without using the Axiom of Choice. But we haven't really eliminated AC; it turns out that AC is equivalent to the statement that a space is compact (in the usual sense) if and only if it is Comfort-compact.

17.21. We have established that the Axiom of Choice is needed to prove Tychonov's Theorem — i.e., that the product of arbitrarily many arbitrary compact sets is compact. But it is not needed for certain weakened forms of Tychonov's Theorem. For instance, arbitrary choices are not needed for:

> **Tychonov Theorem (finite version).** Let $Y_1, Y_2, \ldots, Y_n$ be compact topological spaces. Then $X = Y_1 \times Y_2 \times \cdots \times Y_n$, with the product topology, is also compact.

Proof. It suffices to prove this for $n = 2$, and then apply induction. Let $\big((u_\alpha, v_\alpha) \,:\, \alpha \in \mathbb{A}\big)$ be any net in $Y_1 \times Y_2$. Then $(u_\alpha \,:\, \alpha \in \mathbb{A})$ is a net in the compact space Y_1; hence it has a convergent subnet. By 7.19, $(u_\alpha : \alpha \in \mathbb{A})$ also has a convergent Kelley subnet $(u_{\alpha(\beta)} : \beta \in \mathbb{B})$. Now, $(v_{\alpha(\beta)} : \beta \in \mathbb{B})$ is a net in the compact space Y_2; hence it has a convergent Kelley subnet $(v_{\alpha(\beta(\gamma))} : \gamma \in \mathbb{C})$. Then $\big((u_{\alpha(\beta(\gamma))}, v_{\alpha(\beta(\gamma))}) : \gamma \in \mathbb{C}\big)$ is a convergent net in $Y_1 \times Y_2$, and it is a subnet of the given net. (*Optional exercise*: Shorten this proof, using the notational convention of 7.21.)

17.22. *Compactness equivalents of UF.* The Ultrafilter Principle was introduced in 6.32. We shall now show that it is equivalent to (UF18) (introduced in 17.4) and the following principles.

> **(UF19)** Any product of compact Hausdorff spaces is compact.
>
> **(UF20) (Stone-Čech Compactification Theorem.)** Let X be a completely regular Hausdorff space. Then there exists a topological space $\beta(X)$ (called the **Stone-Čech compactification** of X) with these properties: (i) $\beta(X)$ is a compact Hausdorff space, (ii) X is a dense subset of $\beta(X)$, and (iii) if K is another compact Hausdorff space and $f : X \to K$ is a continuous map, then f extends uniquely to a continuous map $F : \beta(X) \to K$.
>
> **(UF21)** Let $2 = \{0, 1\}$ be equipped with the discrete topology. Then for any set X, the set 2^X (with the product topology) is compact.

Remarks. (UF19) and (UF21) are just (AC21) and (AC23) specialized to the case of Hausdorff spaces. Most topological spaces of interest in applications are Hausdorff, hence most applications of (AC21) or (AC23) actually follow from (UF19) or (UF21).

A set of the form 2^X, with the product topology, is sometimes known as a **Cantor space**. However, that name is more often used for the "middle thirds" set $C = C_0 \cap C_1 \cap C_2 \cap \cdots$, where $C_0 = [0, 1]$, $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$, etc. Actually, it can be proved that the middle thirds set is homeomorphic to $2^{\mathbb{N}}$, but we shall omit the proof.

From property (UF20)(iii) it follows easily that the Stone-Čech compactification is *unique* up to homeomorphism.

Proof of (UF18) $\Rightarrow$ (UF19). Just modify the proof of (AC3) $\Rightarrow$ (AC21) given in 17.16. If each Y_λ is a compact *Hausdorff* space, then each S_λ is a singleton, and so the Axiom of Choice is not needed to prove $\prod_{\lambda \in \Lambda} S_\lambda$ is nonempty.

Proof of (UF19) $\Rightarrow$ (UF20). Let $I = [0, 1]$ and let $C(X, I) = \{$continuous functions from X into $I\}$. Any $x \in X$ determines an evaluation mapping $T_x : C(X, I) \to I$, defined by $T_x(f) = f(x)$ for each $f \in C(X, I)$. Use the fact that X is a completely regular Hausdorff space, to show that the mapping $x \mapsto T_x$ from X into $I^{C(X,I)}$ is injective and is a homeomorphism onto its range. Identifying X with its image, we may view X as a subset of $I^{C(X,I)}$. By (UF7), $I^{C(X,I)}$ is a compact Hausdorff space. Let $\beta(X)$ be the closure of X in $I^{C(X,I)}$; then $\beta(X)$ is compact and X is a dense subset.

The uniqueness of the extension F follows from the fact that X is dense in $\beta(X)$. To prove the existence of the extension F, let any continuous $f : X \to K$ be given. Whenever $\varphi \in C(K, I)$, then $\varphi \circ f \in C(X, I)$. Hence, if λ is any mapping from $C(X, I)$ into I, then

$\varphi \mapsto \lambda(\varphi \circ f)$ is a mapping from $C(K, I)$ into I, which we shall denote by $F(\lambda)$. The mapping $\lambda \mapsto F(\lambda)$, from $I^{C(X,I)}$ into $I^{C(K,I)}$, is easily seen to be a continuous extension of f. Moreover, $F(\beta(X)) = F(\mathrm{cl}(X)) \subseteq \mathrm{cl}(F(X)) = \mathrm{cl}(f(X)) \subseteq K$.

The equivalence of (UF19) with other forms of UF apparently was first proved by Loś and Ryll-Nardzewski [1954]. However, the proof given above is based on Mycielski [1964].

Proof of (UF20) $\Rightarrow$ (UF21). Let Λ be any set, and let $X = 2^\Lambda$ be equipped with the product topology. Let $\beta(X)$ be its Stone-Čech compactification. Then each coordinate projection $\pi_\lambda : X \to 2$ (for $\lambda \in \Lambda$) extends uniquely to a continuous function $P_\lambda : \beta(X) \to 2$. Define a mapping $P : \beta(X) \to 2^\Lambda$ coordinatewise, so that $\pi_\lambda \circ P = P_\lambda$. Then P is a continuous function from the compact space $\beta(X)$ onto X. By 17.7.h, X is compact.

The equivalence of (UF20) with other forms of UF was apparently first announced by Rubin and Scott [1954]; the proof given here is based on Gillman and Jerison [1960].

Proof of (UF21) $\Rightarrow$ (UF1) (based on Mycielski [1964]). Let X be any given set, and let $\mathcal{E}_0$ be a given proper filter on X; we wish to show that $\mathcal{E}_0$ is included in an ultrafilter. Let $\Sigma = \{\text{subsets of } X\}$. Then $\mathcal{P}(\Sigma) = \{\text{subsets of } \Sigma\}$ may be identified with $2^\Sigma = \{$mappings from Σ into $\{0,1\}\}$, as usual. Any $\mathcal{F} \in \mathcal{P}(\Sigma) = 2^\Sigma$ is a collection of subsets of X, and any $S \in \Sigma$ is a subset of X. Let 2^Σ have the product topology; then the Sth coordinate projection $\pi_S : 2^\Sigma \to 2$, defined by

$$\pi_S(\mathcal{F}) \quad = \quad \begin{cases} 1 & \text{if } S \in \mathcal{F} \\ 0 & \text{if } S \notin \mathcal{F}, \end{cases}$$

is continuous (as with any product topology). Now define the sets

$$\begin{array}{rcllll} D & = & \{\mathcal{F} \in 2^\Sigma & : & \mathcal{F} \text{ is a proper filter on } X\}, & \\ E & = & \{\mathcal{F} \in 2^\Sigma & : & \mathcal{F} \supseteq \mathcal{E}_0\}, & \text{and} \\ \Gamma_S & = & \{\mathcal{F} \in 2^\Sigma & : & S \in \mathcal{F} \text{ or } X \setminus S \in \mathcal{F}\} & \end{array}$$

for each $S \subseteq X$. Show that

$$\begin{array}{rcl} D & = & \displaystyle\bigcap_{A,B \subseteq X} \{\mathcal{F} \in 2^\Sigma \ : \ \pi_A(\mathcal{F})\pi_B(\mathcal{F}) - \pi_{A \cap B}(\mathcal{F}) = \pi_\varnothing(\mathcal{F}) = 0\}, \\ E & = & \displaystyle\bigcap_{A \subseteq X} \{\mathcal{F} \in 2^\Sigma \ : \ [1 - \pi_A(\mathcal{F})]\pi_A(\mathcal{E}_0) = 0\}, \qquad \text{and} \\ \Gamma_S & = & \{\mathcal{F} \in 2^\Sigma \ : \ \pi_S(\mathcal{F}) + \pi_{X \setminus S}(\mathcal{F}) \geq 1\}. \end{array}$$

Since the π_S's are continuous, conclude that D, E, Γ_S are closed. Hence $\Phi_S = D \cap E \cap \Gamma_S$ is also closed.

Note that $\Phi_S = \{\mathcal{F} : \mathcal{F}$ is a proper filter on X that includes $\mathcal{E}_0$ and contains at least one of S or $X \setminus S\}$. Using 5.5.i and the "Finite Axiom of Choice" (in 6.14), show that the collection of closed sets $\{\Phi_S : S \subseteq X\}$ has the finite intersection property. By our assumption of (UF21), 2^Σ is compact; hence there exists some $\mathcal{F} \in \bigcap_{S \subseteq X} \Phi_S$. Then $\mathcal{F}$ is an ultrafilter extending $\mathcal{E}_0$.

17.23. *Corollary* (*optional*). Let X be a topological space. Then X has a Hausdorff compactification (i.e., X is homeomorphic to a dense subset of a compact Hausdorff space) if and only if X is a Tychonov space (defined as in 16.17).

17.24. *Proposition* (*optional*). Let X be a chain, equipped with the interval topology (as defined in 5.15.f). Then X is a T_4 space (i.e., a normal Hausdorff space).

Proof (taken from Gillman and Jerison [1960]). Any order convergence is Hausdorff, as we noted in 7.40.g. We shall show that X is completely regular, and use that fact to help us prove X is normal.

Let Y be the MacNeille completion of X, as described in 4.36.c. Then Y is a chain that is order complete. Let Y have the interval topology. By 17.8, Y is a compact Hausdorff space. Hence Y is a Tychonov space — i.e., a $T_{3.5}$ space. Any subspace of a Tychonov space is another Tychonov space, when equipped with the relative topology. By 15.46.b, the relative topology on X coincides with the interval topology. Thus, X is a Tychonov space.

To show X is normal, let any disjoint closed sets A, B be given. We shall define a continuous function $\varphi : X \to [0,1]$ that takes the value 0 on A and 1 on B. It suffices to show how to define φ on the open set $G = X \setminus (A \cup B)$. Let $\{G_\lambda : \lambda \in \Lambda\}$ be the convex components of G, as defined in 4.4.a(ii). Those components are also open, as noted in 15.35.c. We shall define φ separately on each G_λ.

Each G_λ is a convex open subset of the chain X. Since the G_λ's are disjoint open sets, we have $\mathrm{cl}(G_\lambda) \setminus G_\lambda \subseteq A \cup B$. We claim that, moreover,

$$\mathrm{cl}(G_\lambda) \setminus G_\lambda \text{ contains at most two points.}$$

To see this, suppose x_1, x_2, x_3 are three distinct points of $\mathrm{cl}(G_\lambda) \setminus G_\lambda$; say $x_1 < x_2 < x_3$. Since G_λ is convex and the x_i's do not belong to G_λ, no x_i can lie between two members of G_λ. Thus each x_i must lie above or below all of G_λ. Hence at least two of the x_i's (perhaps all three) lie on the same side of G_λ. Say x_1, x_2 both lie below all members of G_λ (the proof is similar for the other case). Then x_1 lies in the open set $\{u \in X : u < x_2\}$, which is disjoint from G_λ, contradicting the assumption that $x_1 \in \mathrm{cl}(G_\lambda)$. This proves our claim.

To make φ continuous, it suffices to define φ on each open set G_λ so that (i) φ is continuous at each point of G_λ, and (ii) if (x_α) is a net in G_λ which converges to some point $z \in \mathrm{cl}(G_\lambda) \setminus G_\lambda$, then $\varphi(x_\alpha) \to \varphi(z)$.

If $\mathrm{cl}(G_\lambda) \setminus G_\lambda$ is empty or contains only points of A, take $\varphi = 0$ on G_λ. If $\mathrm{cl}(G_\lambda) \setminus G_\lambda$ contains only points of B, take $\varphi = 1$ on G_λ.

Finally, suppose $\mathrm{cl}(G_\lambda) \setminus G_\lambda$ consists of one point $z_0 \in A$ and another point $z_1 \in B$. Since X is completely regular, there exists a continuous function $\varphi_\lambda : X \to [0,1]$ that vanishes on all of A and satisfies $\varphi_\lambda(z_1) = 1$. Take $\varphi = \varphi_\lambda$ on G_λ. This completes our construction of φ, and our proof of the normality of X.

17.25. *Compactness in logic* (*optional*). We now explain why the term "compactness" is used in naming the Compactness Principle of Sentential Logic, (UF16) in 14.61. This explanation follows Johnstone [1987].

Our terminology follows that of 14.24 and 14.51. Let $\mathcal{P}$ be the collection of primitive propositions for a sentential calculus, and let $\mathcal{T}$ be the collection of all sentences (i.e., compound propositions formed from elements of $\mathcal{P}$).

An interpretation of the language is an assigning of "true" or "false" to each member of $\mathcal{P}$. By labeling "true" and "false" respectively as 1 and 0, we may identify the set of all interpretations with the set $2^{\mathcal{P}}$. For each $T \in \mathcal{T}$, define the set

$$U(T) \quad = \quad \{f \in 2^{\mathcal{P}} \; : \; f(T) = 1\} \quad = \quad \{\text{models of } T\}.$$

Show that

a. Each $U(T)$ is nonempty. This can be proved by considerations of finite Boolean algebras, without use of the Axiom of Choice or any of its weaker relatives.

b. $U(T_1) \cap U(T_2) = U(T_1 \wedge T_2)$.

c. $\mathcal{B} = \{U(T) : T \in \mathcal{T}\}$ is a base for a topology on $2^{\mathcal{P}}$. (Recall the relevant definitions and properties in 15.35 and 15.36. Thus each open subset of $2^{\mathcal{P}}$ is a union of $U(T)$'s.)

d. The topology determined by that base is the same as the product topology, where 2 has the discrete topology.

Hints: Refer to the characterizations of convergences in 15.25.b and 15.36.d. We have $f_\alpha \to f$ in the topology determined by the base $\mathcal{B}$ if and only if

$(*)$ for each $T \in \mathcal{T}$ such that $f(T) = 1$, we have eventually $f_\alpha(T) = 1$.

On the other hand, $f_\alpha \to f$ in the product topology if and only if

$(**)$ for each $P \in \mathcal{P}$, we have $f_\alpha(P) \to f(P)$.

To see that $(*) \Rightarrow (**)$, use either $T = P$ or $T = \neg P$, depending on whether $f(P)$ is 1 or 0. To see that $(**) \Rightarrow (*)$, let $P_1, P_2, \ldots, P_n$ be the primitive propositions used in forming T; then for all α sufficiently large we have $f_\alpha(P_j) = f(P_j)$ for all j, and therefore (since f_α and f are Boolean homomorphisms) $f_\alpha(T) = f(T)$.

e. Let $\Sigma \subseteq \mathcal{T}$. Show that Σ is unsatisfiable (that is, Σ implies a contradiction in each interpretation of the language) if and only if $\{U(\neg T) : T \in \Sigma\}$ is a cover of $2^{\mathcal{P}}$. Note that it is then an *open* cover.

f. Show that the Compactness Principle (UF16) is equivalent to the statement that the product topology on $2^{\mathcal{P}}$ is compact — i.e., that every open cover of $2^{\mathcal{P}}$ has a finite subcover.

Compactness, Maxima, and Sequences

17.26. *Definitions: A few more kinds of compactness.*

a. A topological space X is **pseudocompact** if either of the following equivalent conditions holds:

(A) Every continuous function from X into $\mathbb{R}$ is bounded above.

(B) Every continuous function from X into $[-\infty, +\infty]$ assumes a maximum on X.

Clearly, any compact space is pseudocompact.

Proof of equivalence. Obviously (B) $\Rightarrow$ (A). For (A) $\Rightarrow$ (B), let $f : X \to [-\infty, +\infty]$ be continuous. We may replace $f(x)$ with $\max\{0, f(x)\}$; hence we may assume $f \geq 0$. Let $\sigma = \sup f(X)$. If σ is merely a supremum, and not a maximum, define $g : X \to [0, +\infty)$ by $g(x) = \tan(\pi f(x)/2\sigma)$. Show that g is continuous but not bounded above.

b. A topological space X is **countably compact** if any of the following equivalent conditions holds:

(A) Every covering of X by countably many open sets has a finite subcover.

(B) If $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$, where the F_i's are nonempty closed sets, then $\bigcap_{i=1}^{\infty} F_i$ is nonempty.

(C) Every sequence in X has a cluster point, i.e., has a convergent subnet.

Proof of equivalence. For (A) $\Rightarrow$ (B), show that if $\bigcap_{i=1}^{\infty} F_i = \varnothing$, then the sets $G_i = X \setminus F_i$ form a countable open cover with no finite subcover. For (B) $\Rightarrow$ (C), the set of cluster points of any sequence (x_n) is the set $\bigcap_{i=1}^{\infty} \mathrm{cl}(\{x_i, x_{i+1}, x_{i+2}, \ldots\})$, by 15.38. For (C) $\Rightarrow$ (A), if $\{G_1, G_2, G_3, \ldots\}$ is a countable open cover of X with no finite subcover, form a sequence (x_n) with no cluster point by choosing $x_n \in X \setminus (G_1 \cup G_2 \cup \cdots \cup G_n)$. (This uses Countable Choice, not the full strength of AC.)

c. A topological space X is **sequentially compact** if every sequence in X has a convergent subsequence — or equivalently, if every sequence in X has a convergent subnet that happens to be a sequence. (This equivalence follows from 7.27 or 15.40.)

17.27. *Proposition.* The product of countably many sequentially compact spaces (when equipped with the product topology) is sequentially compact.

Remarks and proof. The argument used here is a **diagonal subsequence argument**. A similar argument is used in several other contexts in mathematics.

Let $Y_1, Y_2, Y_3, \ldots$ be sequentially compact spaces; we wish to show $X = \prod_{n=1}^{\infty} Y_n$ is sequentially compact. Let $\pi_n : X \to Y_n$ be the nth coordinate projection. Let $(x_1, x_2, x_3, \ldots)$ be a given sequence in X; we wish to produce a convergent subsequence (v_k). Recursively define $u_{n,j}$'s in X as follows. For $j = 1, 2, 3, \ldots$, let $u_{0,j} = x_j$. Now, after a sequence $(u_{n-1,1},\ u_{n-1,2},\ u_{n-1,3},\ \ldots)$ has been specified in X, let $(u_{n,1},\ u_{n,2},\ u_{n,3},\ \ldots)$ be a subsequence of it with the property that the sequence $(\pi_n(u_{n,1}),\ \pi_n(u_{n,2}),\ \pi_n(u_{n,3}),\ \ldots)$ is convergent in Y_n. This completes the recursion. Now define a **diagonal subsequence** $v_k = u_{k,k}$ $(k = 1, 2, 3, \ldots)$. For each n, verify that $(v_n,\ v_{n+1},\ v_{n+2},\ \ldots)$ is a subsequence of $(u_{n,1},\ u_{n,2},\ u_{n,3},\ \ldots)$, and therefore the sequence $(\pi_n(v_1),\ \pi_n(v_2),\ \pi_n(v_3),\ \ldots)$ is convergent in Y_n.

17.28. *Proposition.* If $\mathrm{card}(U) \geq \mathrm{card}(\mathbb{R})$, then 2^U is not sequentially compact.

Proof. Let $\pi_j : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ be the jth coordinate projection, so $\pi_j(a_1, a_2, a_3, \ldots) = a_j$. Let $A = \{$strictly increasing sequences of positive integers$\}$. Let $E_n = \{a \in A : n$ occurs at an even position in the sequence $a\} = \bigcup_{k=1}^{\infty} \pi_{2k}^{-1}(n)$.

Any member of A may be viewed as a subset of $\mathbb{N}$; thus $\mathrm{card}(A) = \mathrm{card}(\mathcal{P}(\mathbb{N})) = \mathrm{card}(\mathbb{R}) \leq \mathrm{card}(U)$. Let $\varphi : A \to U$ be some injective function. Define a sequence (f_n) in 2^U by taking $f_n : U \to \{0,1\}$ to be the characteristic function of the set $\varphi(E_n)$. We claim that (f_n) has no convergent subsequence.

Indeed, let (f_{n_j}) be any subsequence of (f_n). Then (n_j) is a member of A. Let us denote it by a, and let $u = \varphi(a)$. Observe that

$$\begin{aligned} f_{n_j}(u) = 1 \quad &\iff \quad u \in \varphi(E_{n_j}) \quad \iff \quad a \in E_{n_j} \\ &\iff \quad n_j = n_{2k} \text{ for some } k \quad \iff \quad j \text{ is even,} \end{aligned}$$

since the functions φ and $j \mapsto n_j$ are injective. Thus the sequence $\big(f_{n_j}(u) : j \in \mathbb{N}\big)$ does not converge. This proof follows the presentation of Wilansky [1970].

17.29. *Example of the inadequacy of frequent subnets.* In 7.19 and 15.38 we saw that Willard subnets, Kelley subnets, and AA subnets can be used interchangeably for most purposes in topology. We now show that frequent subnets (defined in 7.16.c) *cannot* be used interchangeably with those other types of subnets.

Let X be any topological space that is compact but not sequentially compact (e.g., the space in 17.28). Let (x_n) be a sequence in X that has no convergent subsequence. Then (x_n) has a convergent subnet (u_β). Any frequent subnet of (x_n) is a subsequence (see 7.16.d). Any net that is equivalent to (u_β) must have the same limit(s) as (u_β). Thus, (u_β) *is a subnet of* (x_n)*, but it is not equivalent (in the sense of 7.17.c) to a frequent subnet of* (x_n).

Other examples of the inadequacy of frequent subnets have been given by Wolk [1982] and other papers cited by Wolk.

17.30. *Relations between different kinds of compactness.*

a. Any countably compact, first countable space is sequentially compact. (Recall that a topological space is *first countable* if the neighborhood filter at each point has a countable filterbase.) *Hint*: 15.34.c.

b. Any countably compact space is pseudocompact.

Hint: Let $f : X \to \mathbb{R}$ be continuous; consider whether the set $\bigcap_{n=1}^{\infty}\{x \in X : f(x) \geq n\}$ is nonempty.

c. (*Optional.*) Any paracompact, pseudocompact space is compact.

Proof. Suppose X is paracompact but not compact. Let $\mathcal{G} = \{G_\alpha : \alpha \in A\}$ be an open cover with no finite subcover. Let $\{f_\alpha : \alpha \in A\}$ be a partition of unity that is precisely subordinated to the given cover. Then the sets $H_\alpha = \{x \in X : f_\alpha(x) \neq 0\} \subseteq G_\alpha$ also form an open cover with no finite subcover. Recursively choose a sequence (x_n) in X and a sequence $(\alpha(n))$ in A such that $x_n \notin H_{\alpha(1)} \cup \cdots \cup H_{\alpha(n-1)}$ and $x_n \in H_{\alpha(n)}$. Define

$$g(x) \quad = \quad \sum_{n=1}^{\infty} n \frac{f_{\alpha(n)}(x)}{f_{\alpha(n)}(x_n)}.$$

Then $g : X \to \mathbb{R}$ is continuous but $g(x_n) \geq n$, so X is not pseudocompact.

17.31. *Remarks.* We have considered three types of compactness that can be described in terms of convergences. A topological space is

> *compact* if every net has a convergent subnet;
>
> *countably compact* if every sequence has a convergent subnet;
>
> *sequentially compact* if every sequence has a convergent subsequence.

It is easy to see that any compact space is countably compact, and any sequentially compact space is countably compact. In general, no other implications hold between these three kinds of compactness — the example in 17.38 shows that sequential compactness does not imply compactness, and the example in 17.28 shows that compactness does not imply sequential compactness. However, under certain additional hypotheses, all three kinds of compactness are equivalent, as we shall see in 17.33 and 17.51.

Optional. It is interesting to consider a fourth type of compactness. Say that a topological space is

> **supersequentially compact** if every net has a convergent subnet that is a sequence.

Clearly, this implies the other three kinds of compactness. Supersequential compactness is not necessarily a useful notion; we introduce it only to illustrate certain ideas about subnets. It turns out that supersequential compactness depends, not on the topology of X, but only on the cardinality of X and on which definition of "subnet" we use. Let X be a nonempty set. Then:

a. If we use Aarnes-Andenæs subnets, then every finite set X is supersequentially compact, no matter what topology we equip it with. Indeed, if $(x_\alpha : \alpha \in \mathbb{A})$ is a net in a finite set X, then there is at least one point $p \in X$ such that the constant sequence $(p, p, p, \ldots)$ is an Aarnes-Andenæs subnet of $(x_\alpha : \alpha \in \mathbb{A})$. The sequence $(p, p, p, \ldots)$ converges to p, no matter how X is topologized.

b. If X is an infinite set, or if we use Kelley subnets, then X is not supersequentially compact. In fact, regardless of convergences, there exists a net (x_α) in X that has no subnets that are sequences; this was established in 7.28.

17.32. Lebesgue's Covering Lemma. Let (X, d) be a compact metric space — or more generally, a countably compact pseudometric space. Let $\{G_\lambda : \lambda \in \Lambda\}$ be an open cover of X. Then there exists a number $\rho > 0$ with the following property: Each open ball $B(x, \rho)$ is contained in one of the G_λ's. (Such a number ρ is called a **Lebesgue number** for the covering.)

Proof. Suppose there is no such ρ. Then there exist open balls $B(x_n, \rho_n)$ with $\rho_n \downarrow 0$, and such that $B(x_n, \rho_n)$ is not contained in any G_λ. The sequence (x_n) has a cluster point z in X. Since $\{G_\lambda\}$ is an open cover, we have $B(z, r) \subseteq G_\mu$ for some $r > 0$ and $\mu \in \Lambda$. Since z is a cluster point of (x_n), there exist n's arbitrarily large with $x_n \in B(z, \frac{1}{2}r)$. For sufficiently large n we have also $\rho_n < \frac{1}{2}r$, and hence $B(x_n, \rho_n) \subseteq B(z, r) \subseteq G_\mu$, a contradiction.

Remarks. The existence of Lebesgue numbers does *not* imply compactness, even in a metric space. For instance, let $\mathbb{Z} = \{\text{the integers}\}$ have its usual metric; then $\mathbb{Z}$ is not compact but every open cover of $\mathbb{Z}$ has a Lebesgue number. The existence or nonexistence of Lebesgue numbers is discussed further by Arala-Chaves [1985].

17.33. Theorem (Gross and Hausdorff, 1914). Let X be a pseudometric space. Then the following conditions are equivalent.

(A) X is compact — i.e., every net has a convergent subnet.

(B) X is sequentially compact — i.e., every sequence has a convergent subsequence.

(C) X is countably compact — i.e., every sequence has a convergent subnet.

(D) Every upper semicontinuous function from X into $[-\infty, +\infty]$ assumes a maximum.

(E) X is pseudocompact — i.e., every continuous, real-valued function assumes a maximum.

Proof. The implications (A) $\Rightarrow$ (C) and (B) $\Rightarrow$ (C) and (D) $\Rightarrow$ (E) are obvious. The implication (C) $\Rightarrow$ (B) is a special case of an exercise in 17.30.a. The implication (A) $\Rightarrow$ (D) is an easy result that was noted in 17.7.i. We proved (C) $\Rightarrow$ (E) in 17.30.b. We shall complete the proof in two different ways, according to the background of the reader. If the reader is familiar with paracompactness and Stone's Theorem (16.31), then (E) $\Rightarrow$ (A) follows from 17.30.c, and we are done. For readers not familiar with Stone's Theorem, we shall give more elementary proofs of (E) $\Rightarrow$ (B) and (C) $\Rightarrow$ (A) below.

Outline of (E) $\Rightarrow$ (B). Assume (E), and suppose (x_n) is a sequence in X with no convergent subsequence; we shall obtain a contradiction. Replacing (x_n) with a subsequence, we may delete repetitions — i.e., we may assume the x_n's are all distinct. For each n, since x_n is not a limit of a subsequence of the sequence, the number $r_n = \frac{1}{2}\inf_{m\neq n} d(x_n, x_m)$ is positive. Hence the balls $B(x_n, r_n)$ are all disjoint. Let

$$f(u) \quad = \quad \begin{cases} n \max\left\{0,\ 1 - \dfrac{d(u, x_n)}{r_n}\right\} & \text{when } u \in B(x_n, r_n) \quad (n \in \mathbb{N}) \\ 0 & \text{when } u \notin \bigcup_{n=1}^{\infty} B(x_n, r_n). \end{cases}$$

Show that f is continuous and not bounded above.

Outline of (C) $\Rightarrow$ (A). Let $\{G_\lambda : \lambda \in \Lambda\}$ be any open cover of X. Let ρ be a Lebesgue number for the cover (see 17.32). Each open ball $B(x, \rho)$ (for $x \in X$) is contained in some G_λ, so it suffices to show that X can be covered by finitely many of the $B(x, \rho)$'s. Suppose not. Recursively choose (using (DC2), in 6.28) a sequence (x_n) in X such that $x_{n+1} \notin B(x_1, \rho) \cup B(x_2, \rho) \cup \cdots \cup B(x_n, \rho)$. The sequence (x_n) has some cluster point $z \in X$. The open ball $B(z, \frac{1}{2}\rho)$ is a frequent set for the sequence (x_n), so it contains x_m, x_n for some distinct numbers m and n. Then $d(x_m, x_n) < \rho$, a contradiction.

17.34. *Proposition.* Any compact pseudometric space is separable.

Proof. The open balls of radius $\frac{1}{n}$ form an open cover; hence the space X can be covered by finitely many of them. Say

$$X = B\left(x_{n,1}, \frac{1}{n}\right) \cup B\left(x_{n,2}, \frac{1}{n}\right) \cup \cdots \cup B\left(x_{n,k_n}, \frac{1}{n}\right).$$

Then $\{x_{n,j} : n \in \mathbb{N}, 1 \le j \le k_n\}$ is a countable dense subset of X.

17.35. *Existence of a closest point.* Let (X, d) be a pseudometric space. Let $K \subseteq X$ be a compact set (or, more generally, a pseudocompact set). Let $x \in X \setminus K$. Then there exists at least one point in K that is closest to x. That is, there exists some $q \in K$ (not necessarily unique) such that $d(x, q) = \mathrm{dist}(x, K)$.

Hint: The function $q \mapsto d(x, q)$ is continuous and real-valued.

Remark. Other conditions for existence and/or uniqueness of a closest point will be given in 22.39(D), 22.45, and 28.41(E).

17.36. Fundamental Theorem of Algebra. Let $P(z)$ be a polynomial of degree $n > 0$ with complex coefficients — i.e., suppose that

$$P(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{n-1} z^{n-1} + a_n z^n$$

where the a_k's are complex numbers, $n > 0$, and $a_n \ne 0$.

Then $P(z)$ has a root in the complex numbers — i.e., we have $P(\zeta) = 0$ for some $\zeta \in \mathbb{C}$.

Proof. Show that P is continuous. Also, show that $\lim_{|z| \to \infty} |P(z)| = \infty$ (that is, show $\lim_{r \to \infty} \inf\{|P(z)| : z \in \mathbb{C}, |z| \ge r\} = \infty$). From these two facts, plus the fact that closed bounded subsets of $\mathbb{C}$ are compact, conclude that $|P(\cdot)|$ assumes a minimum at some point $\zeta \in \mathbb{C}$. Replacing the function $P(z)$ with the function $P(z + \zeta)$, we may assume $\zeta = 0$ — that is, we may assume $|P(0)| \le |P(z)|$ for all complex numbers z. We wish to show that $P(0) = 0$.

We may assume $P(0) = a_0 \ne 0$. Let k be the first positive integer for which $a_k \ne 0$; thus $P(z) = a_0 + a_k z^k + z^{k+1} Q(z)$ for some polynomial Q. As in 10.29, the nonzero number $-a_0/a_k$ has n distinct nth roots in $\mathbb{C}$; let w be any one of those. Since Q is continuous, for $|t|$ sufficiently small we have $|tw^{k+1} Q(tw)| < |a_0|$. Choose such a value of t, satisfying also $0 < t < 1$. Then

$$|P(tw)| = \left|a_0 + a_k (tw)^k + (tw)^{k+1} Q(tw)\right| = \left|(1 - t^k) a_0 + (tw)^{k+1} Q(tw)\right|$$

$$\le (1 - t^k)|a_0| + t^k |tw^{k+1} Q(tw)| < (1 - t^k)|a_0| + t^k |a_0| = |P(0)|,$$

a contradiction.

This proof goes back at least as far as Argand (1806). It has been rediscovered many times; some references are given by Brualdi [1977].

Corollary. Let $P(z)$ be a polynomial of degree n with complex coefficients. Then $P(z) = a_n (z - \zeta_1)(z - \zeta_2) \cdots (z - \zeta_n)$ for some complex numbers $\zeta_1, \zeta_2, \ldots, \zeta_n$.

Pathological Examples: Ordinal Spaces (Optional)

17.37. Let Ω be the first uncountable ordinal. (see 5.44); thus Ω is the set of all countable ordinals. Let Ω be equipped with its usual ordering. Also, let $K = \Omega \cup \{\zeta\}$ for some object ζ that is not a member of Ω; make K into a chain by taking $\omega < \zeta$ for all $\omega \in \Omega$.

Let the sets Ω and K be equipped with their interval topologies (see 5.15.f); thus they are normal Hausdorff spaces (see 17.24). We shall use these and related spaces for some pathological examples — e.g., to show that paracompactness is not hereditary or productive. Such results are not really essential to abstract analysis; they are included here merely to round out our introduction to general topology.

(We remark that the particular choice of ζ does not matter, so long as $\zeta \notin \Omega$. Hence we may take $\zeta = \Omega$, and thus $K = \Omega \cup \{\Omega\}$ is the next ordinal after Ω; this is customary in the study of ordinals. However, we shall *not* specify ζ in that fashion, because it is not necessary to do so for the purposes below, and it may be distracting to have Ω appearing as both an element and a subset of K.)

This presentation is based on Steen and Seebach [1970].

17.38. *Basic properties.* With Ω and K as above:

a. Ω is an initial ordinal (see 5.47), hence a limit ordinal (see 5.46.j), hence has no greatest element.

b. Any sequence in Ω has a supremum in Ω.

Proof. The supremum of any collection of ordinals in Ω is their union. If (x_n) is a sequence in Ω, then $\bigcup_{n=1}^{\infty} x_n$ is a countable union of countable sets, so it is countable (see 6.26). Thus it is a countable ordinal, and hence a member of Ω.

c. A neighborhood base for the point ζ in the space K is given by the sets of the form $N_\omega = \{x \in K : x > \omega\}$, for points $\omega \in \Omega$.

d. Ω is dense in K (hence our notation).

Proof. For each $\omega \in \Omega$, there is some strictly larger element $\omega' \in \Omega$; then $\omega' \in N_\omega \cap \Omega$. Thus the set Ω meets every neighborhood of the point ζ.

e. K is compact, but Ω is not compact.

Proof. K is order complete, but Ω is not order complete. Now apply 17.8.

f. Ω is first countable, but K is not.

Proof. Let any $x \in \Omega$ be given. Then a neighborhood base at x is given by the sets $\{\omega \in \Omega : \omega > v\}$ for points $v < x$ (of which there are only countably many) and the set $\{\omega \in \Omega : \omega < x^+\}$, where x^+ is the first element after x.

The space K is not first countable, because ζ does not have a countable neighborhood base. Indeed, if $(B_j : j \in \mathbb{N})$ were such a basis, then each B_j would have to contain some set of the form $N_{\omega(j)}$. Since K is a Hausdorff space, the sole member of $\bigcap_{j=1}^{\infty} N_{\omega(j)}$ is the point ζ. Let σ be the supremum of the $\omega(j)$'s, and let τ be some strictly larger member of Ω. Then $\tau \in \bigcap_{j=1}^{\infty} N_{\omega(j)}$, a contradiction.

g. Ω is countably compact.

Proof. Let (x_n) be a sequence in Ω, and let σ be its supremum. Then (x_n) is also a sequence in the smaller ordinal $\{\omega \in \Omega : \omega \leq \sigma\}$, which is compact by 17.8, hence countably compact. Thus (x_n) has a subnet that converges to a limit in $\{\omega \in \Omega : \omega \leq \sigma\}$, hence in Ω.

h. Ω is pseudocompact, sequentially compact, and normal, but not paracompact.

Proof. 17.30.b, 17.30.a, 17.24, and 17.30.c.

17.39. *Proposition.* Let Ω and K be as above. Then $\Omega \times K$, with the product topology, is $T_{3.5}$ but not normal.

Proof. The notation (x, y) will refer to ordered pairs, not intervals.

Both Ω and K are $T_{3.5}$ spaces, hence are $K \times K$ and $\Omega \times K$. The set $\{(x, x) : x \in K\}$ is closed in $K \times K$, hence the set $A = \{(\alpha, \alpha) : \alpha \in \Omega\} = \{(x, x) : x \in K\} \cap (\Omega \times K)$ is closed in $\Omega \times K$. Also the set $B = \Omega \times \{\zeta\}$ is closed in $\Omega \times K$. These two closed sets are disjoint. Suppose that there exist disjoint open sets $U \supseteq A$ and $V \supseteq B$, in $\Omega \times K$; we shall obtain a contradiction.

For each $x \in \Omega$, we have $(x, \zeta) \in B \subseteq V$, hence $(x, \zeta) \notin \mathrm{cl}(U)$. Therefore the set $\{(x, \lambda) : x < \lambda < \zeta\}$ is not contained in U. Let $\lambda(x)$ be the first member of Ω that satisfies $x < \lambda$ and $(x, \lambda) \notin U$.

Let x_0 be the first member of Ω, and thereafter let $x_{n+1} = \lambda(x_n)$. The sequence (x_n) has a supremum σ in Ω, as we noted in 17.38. Since $\lambda(x) > x$ for each x, the sequence (x_n) is increasing, and therefore converges to σ. Hence $(x_n, x_{n+1}) \to (\sigma, \sigma) \in A \subseteq U$. However, (x_n, x_{n+1}) is in $\complement U$, which is a closed set. This is a contradiction.

17.40. *Corollaries.*

a. $\Omega \times K$ is not normal, but it is a subset of the compact Hausdorff space $K \times K$. Thus neither normal nor paracompact is a hereditary property.

b. $\Omega \times K$ is a product of two normal spaces, but it is not normal. Thus the property of being normal is not productive.

BOOLEAN SPACES

17.41. *Definitions.* Recall that a set is *clopen* if it is both closed and open. A topological space is **zero-dimensional** if it has a base consisting of clopen sets — i.e., if every open set can be expressed as a union of clopen sets.

A **Boolean space** (or **Stone space**) is a zero-dimensional, compact Hausdorff space.

Remarks. Such a prevalence of clopen sets may seem highly pathological to analysts, since the topological spaces X most commonly used by analysts have *no* clopen subsets other than $\varnothing$ and X itself. However, Boolean spaces are useful in logic and related topics. Also, zero-dimensional spaces (not necessarily compact) will be studied further in Chapter 20 where they will be useful in the study of certain spaces that are *not* zero-dimensional; see especially 20.27 through 20.30.

17.42. *Examples.*

a. Any set with the discrete topology is a zero-dimensional space. Any finite set with the discrete topology is a Boolean space. In particular, $2 = \{0, 1\}$ is a Boolean space.

b. Any subset of a zero-dimensional space is zero-dimensional, when equipped with the relative topology. Any closed subset of a Boolean space is a Boolean space.

c. The set $\mathbb{Q} = \{\text{rational numbers}\}$, topologized as a subset of $\mathbb{R}$, is a zero-dimensional space. A clopen base for it is given by the intervals $(a, b) \cap \mathbb{Q}$ for irrational numbers a, b.

d. The set $\mathbb{R}\backslash\mathbb{Q} = \{\text{irrational numbers}\}$, topologized as a subset of $\mathbb{R}$, is a zero-dimensional space. A clopen base for it is given by the intervals $(a, b) \cap (\mathbb{R} \backslash \mathbb{Q})$ for rational numbers a, b.

e. Any product $X = \prod_\lambda Y_\lambda$ of zero-dimensional spaces is zero-dimensional.

Hints: For each λ, let $\mathcal{B}_\lambda$ be a base for the topology of Y_λ consisting of clopen sets. Show that a clopen base for the product topology on X is given by the sets of the form $\left(\prod_{\lambda\in F} B_\lambda\right) \times \left(\prod_{\lambda\in\Lambda\backslash F} Y_\lambda\right)$ where F is a finite subset of Λ and each B_λ is a member of $\mathcal{B}_\lambda$.

f. Further examples are given by these two principles, which are both equivalent to the Ultrafilter Principle:

(UF22) Any product of Boolean spaces is a Boolean space.

(UF23) 2^A is a Boolean space, for any set A.

Proof of equivalence. Refer to 17.22. It is easy to see that (UF19) $\Rightarrow$ (UF22) $\Rightarrow$ (UF23) $\Rightarrow$ (UF21).

g. (*Optional.*) Let α be some ordinal, and let $X = \{x \leq \alpha : x \text{ is an ordinal}\}$. Let X have the order interval topology (described in 5.15.f). Then it can be shown that X is a Boolean space. We omit the proof.

17.43. The remainder of this subchapter is optional; it will not be used later in this book. In the sections below we shall show that the categories of

Boolean spaces (with continuous maps)

and

Boolean algebras (with Boolean homomorphisms)

are dual to each other, in the sense of 9.55. All of the conclusions stated below are consequences of the Ultrafilter Principle and its various equivalents.

17.44. *Definitions of the dual objects.* For the set Δ described in 9.55, we shall use $2 = \{0, 1\}$, which may be viewed both as a Boolean algebra (with the obvious ordering

$0 < 1$) and as a Boolean space (with the discrete topology). Thus, for any object X in either category, the **dual** set is

$$X^* \quad = \quad \{\text{morphisms from } X \text{ into } 2\}.$$

The **dual of a Boolean algebra** A is the set of two-valued homomorphisms on A, defined as in 13.19. It is a subset of $2^A = \{\text{maps from } A \text{ into } \{0,1\}\}$. We equip 2^A with the product topology and equip A^* with the resulting relative topology. We know that 2^A is a Boolean space (see (UF23), in 17.42.f), and it is easy to show that A^* is a closed subset of 2^A (*exercise*). Thus, by 17.42.b, A^* is a Boolean space. Recall from (UF5) that A^* separates points of A.

The simplest way to define the dual of a Boolean space S is to use clop(S), the algebra of clopen subsets; this is an algebra of sets, and thus a Boolean algebra with fundamental operations $\cup, \cap, \complement, \varnothing, S$. However, for our study of duality, it will be convenient to replace those clopen sets with their characteristic functions. Thus, the **dual of a Boolean space** S is defined to be the set

$$\begin{aligned} S^* &= \{\text{characteristic functions of clopen subsets of } S\} \\ &= \{\text{continuous functions from } S \text{ into } \{0,1\}\}. \end{aligned}$$

(*Exercise.* Prove those two sets are equal; here $\{0,1\}$ has the discrete topology, as usual.) The Boolean algebra S^* has smallest and greatest elements equal to the constant functions 0 and 1; its other fundamental operations are

$$f \vee g = \max\{f,g\}, \qquad f \wedge g = \min\{f,g\}, \qquad \complement f = 1 - f.$$

Since S is a Boolean space, it is easy to see that S^* separates points of S.

17.45. *Dual morphisms.* Let X and Y be objects in either category (Boolean spaces or Boolean algebras), and let $f : X \to Y$ be a morphism in that category (i.e., a continuous map or a Boolean homomorphism). Then we may define a corresponding function $f^* : Y^* \to X^*$ as in 9.55, called the *dual* of f, by defining $f^*(\lambda) = \lambda \circ f : X \to 2$ for each $\lambda : Y \to 2$ in Y^*. We shall show that *it is a morphism in the other category.* Thus the mappings $X \mapsto X^*$ and $f \mapsto f^*$ define a contravariant functor.

a. If $f : X \to Y$ is a Boolean homomorphism between Boolean algebras, then $f^* : Y^* \to X^*$ is a continuous map; that is just a special case of 15.26.d.

b. If $f : X \to Y$ is a continuous map between Boolean spaces, then $f^* : Y^* \to X^*$ is the restriction to Y^* of the inverse image map $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ that was defined in 2.8 and studied further in 9.32. As we saw in 2.8, the inverse image map preserves all the basic set operations, and so it is a homomorphism of algebras of sets. Thus it is a Boolean homomorphism. Then its restriction f^* is, too.

17.46. *Reflexivity.* Let X be an object in either category. Then X^* is a collection of mappings from X into $2 = \{0,1\}$. Conversely, each $x \in X$ may be viewed as a mapping T_x from X^* into 2, defined by $T_x(f) = f(x)$. Since X^* separates points of X, the members of X may be viewed as *distinct* mappings from X^* into 2. Thus T may be viewed as an inclusion $X \xrightarrow{\subseteq} 2^{X^*}$. We shall show that each mapping $T_x : X^* \to 2$ is actually a morphism, and thus a member of X^{**}, so that T maps X into X^{**}. Moreover, T is a morphism in the category in which X and X^{**} are objects. In fact, T is an isomorphism from X onto X^{**}.

a. First, suppose X is a Boolean algebra; thus its dual X^* is a Boolean space and X^{**} is another Boolean algebra. If we unwind the notation, we find that

> $T_x : X^* \to 2$ is the characteristic function of the set $S(x) \subseteq X^*$ defined by the Stone map $S : X \to \mathfrak{S}$ as in 13.21.

S is an isomorphism of Boolean algebras, from X onto $\mathfrak{S} = \text{Range}(S)$, by (UF6). Thus, it remains only to show that a subset of X^* is clopen if and only if it belongs to $\mathfrak{S}$.

First we show that every member of $\mathfrak{S}$ is clopen. It is easy to show that each mapping T_x is continuous from X^* into 2 — e.g., by considering the product topology on 2^X characterized in terms of convergence of nets in 2^X. Hence each set $S(x) = T_x^{-1}(1)$ is the inverse image, under a continuous map, of a clopen set; hence each $S(x)$ is clopen.

Finally, we shall show that any clopen subset of X^* belongs to $\mathfrak{S}$. For each $x \in X$, let $\pi_x : 2^X \to 2$ be the xth coordinate projection. Each π_x is continuous, hence each set of the form $\pi_x^{-1}(0)$ or $\pi_x^{-1}(1)$ is clopen. A *basic rectangle* in 2^X is an intersection of finitely many of these sets; hence it is also clopen. The basic rectangles form a base for the topology of 2^X; hence $\mathcal{B} = \{B \cap X^* : B \text{ is a basic rectangle in } 2^X\}$ is a base for the topology of X^*. Observe that $S(x) = \pi_x^{-1}(1) \cap X^*$ and $S(\complement x) = \complement S(x) = \pi_x^{-1}(0) \cap X^*$ and

$$S(x_1) \cap S(x_2) \cap \cdots \cap S(x_n) \quad = \quad S\left(x_1 \wedge x_2 \wedge \cdots \wedge x_n\right);$$

hence each member of $\mathcal{B}$ can be written in the form $S(x)$ for some $x \in X$. Since $\mathcal{B}$ is a base for the topology of X^*, each open subset of X^* is a union of $S(x)$'s. Since X^* is compact, any clopen subset of X^* is compact, and therefore is a union of finitely many of the $S(x)$'s. But also

$$S(x_1) \cup S(x_2) \cup \cdots \cup S(x_n) \quad = \quad S\left(x_1 \vee x_2 \vee \cdots \vee x_n\right),$$

so each clopen subset of X^* is in the range of S.

b. Conversely, suppose X is a Boolean space. Then its dual X^* is a Boolean algebra, and X^{**} is another Boolean space. It is tedious but straightforward to verify that each mapping $T_x : X^* \to 2$ is a Boolean homomorphism and thus a member of X^{**}; we omit the details. Also, the mapping $T : X \to X^{**}$ given by $x \mapsto T_x$ is continuous. (*Hint*: Use 15.15(E) and 15.25.b.)

It remains to show that $T : X \to X^{**}$ is surjective. This can be proved directly (using the Ultrafilter Theorem or one of its equivalents); such a proof is given by Monk [1989], for instance. However, a slightly shorter and more transparent proof in Halmos [1963] uses 17.46.a: Since X is compact, X^{**} is Hausdorff, and T is continuous, the range of T is compact and therefore closed. Thus it suffices to show that the range of T is dense in X^{**}. Since X^{**} is a Boolean space, its clopen sets form a base. Let G be any nonempty clopen subset of X^{**}; it suffices to show that Range(T) meets

G. The function $1_G : X^{**} \to 2$ is a member of X^{***}. We have already proved in 17.46.a that any Boolean algebra is reflexive, so $X^* = X^{***}$ — or, more precisely, X^* and X^{***} act the same on X^{**}. Thus there is some clopen set $C \subseteq X$ such that the mapping $1_C : X \to 2$ — a member of X^* — acts the same as 1_G on the set X^{**}. Since G is nonempty, it follows easily that 1_G is not the zero function, hence 1_C is not the zero function, hence C is not empty. Choose any $x \in C$. Then $1 = 1_C(x) = T_x(1_C) = 1_G(T_x)$, so $T_x \in G$ — that is, Range(T) meets G.

17.47. *Further duality results.* Many properties of X correspond to properties of X^*. A few of these results are listed below, without proof. (These results are *not* recommended as exercises; some of them are too difficult without further hints. Proofs can be found in Halmos [1963], Monk [1989], and other books and papers cited by those two books.

Let A be a Boolean algebra and let S be a Boolean space, with $A^* = S$ and $S^* = A$. Then:

a. A is finite $\iff$ S is finite $\iff$ the topology on S is discrete $\iff$ the Stone map (introduced in 13.21) is an isomorphism from A to $\mathcal{P}(\mathrm{Ult}(A))$, where $\mathrm{Ult}(A) = \{$Boolean ultrafilters in $A\}$.

b. A is countable $\iff$ S is metrizable.

c. By an **atom** in A we mean an element $x \succ 0$ such that $\{a \in A : 0 \prec a \prec x\}$ is empty. There is a natural correspondence between the atoms of A and the isolated points of S, as follows: If $s \in S$ is an isolated point, then the singleton $\{s\}$ is a clopen subset of S, hence its characteristic function is a member of A; this mapping $s \mapsto 1_{\{s\}}$ gives a bijection from $\{$isolated points of $S\}$ onto $\{$atoms of $A\}$.

d. For any open set $G \subseteq S$ and any ideal $I \subseteq A$, define

$$G^* = \{1_C : C \text{ is a clopen subset of } G\}, \qquad I^* = \bigcup_{\substack{C \text{ clopen}, \\ C \subseteq S,\ 1_C \in I}} C.$$

Then the maps $G \mapsto G^*$ and $I \mapsto I^*$ are inverses of each other, and they give a bijection between the open subsets of S and the ideals in A. The maps $G \mapsto G^*$ and $I \mapsto I^*$ are both order-preserving (where order is given by $\subseteq$); thus they give an isomorphism between the lattice of open subsets of S and the lattice of ideals in A. Furthermore, the set $S \setminus G = \{s \in S : s \notin G\}$ (with the relative topology) is a Boolean space, and the corresponding Boolean algebra is isomorphic to the quotient algebra $A/(G^*)$.

e. A morphism $f : X \to Y$ (in either category — i.e., a continuous map or a Boolean homomorphism) is injective if and only if f^* is surjective.

EBERLEIN-SMULIAN THEOREM

17.48. The presentation in this subchapter is modified from Kelley and Namioka [1976] and Wilansky [1970]. The results of this subchapter may be postponed; they will not be needed until Chapter 28.

17.49. *Proposition.* Let S be a compact topological space, let (M, d) be a metric space, and let M^S be equipped with the product topology. Let Φ be a collection of continuous functions from S into M, and let $\psi : S \to M$ also be continuous. Suppose that $\psi \in \mathrm{cl}(\Phi)$. Then ψ is a cluster point of some sequence in Φ — i.e., ψ is a limit of a subnet of a sequence in Φ.

Proof. First note that one neighborhood base at ψ in M^S is given by the sets of the form

$$G_\sigma \;=\; \left\{ g \in M^S \;:\; \max_{1 \le i \le n} d\left[\psi(s_i), g(s_i)\right] < \frac{1}{n} \right\}$$

$$\text{for } n \in \mathbb{N} \text{ and } \sigma = (s_1, s_2, \ldots, s_n) \in S^n.$$

Hence we may choose some $\varphi_\sigma \in \Phi \cap G_\sigma$. For any $\tau \in S^n$, since ψ and φ_τ are continuous, the set

$$H_\tau \;=\; \{\sigma \in S^n \;:\; \varphi_\tau \in G_\sigma\} \;=\; \{(s_1, s_2, \ldots, s_n) \in S^n \;:\; \max_{1 \le i \le n} d[\psi(s_i), \varphi_\tau(s_i)] < \frac{1}{n}\}$$

is an open neighborhood of τ in S^n. Since S^n is compact, it is contained in the union of the sets H_τ $(\tau \in A_n)$, for some finite set $A_n \subseteq S^n$.

Let $\Phi_n = \{\varphi_\tau : \tau \in A_n\}$. Note that for each positive integer n and each $\sigma \in S^n$, the set G_σ meets Φ_n. In fact, G_σ meets Φ_m for every $m \ge n$, since $G_{\sigma'} \subseteq G_\sigma$ whenever σ' is an extension of σ.

Each Φ_n is a finite set, which we now arrange in any order. Form a sequence (g_k) by taking the element of Φ_1, then the elements of Φ_2, then the elements of Φ_3, etc. Then each G_σ is a frequent set for the sequence (g_k). Since the G_σ's form a neighborhood base $\mathcal{B}$ for ψ, the sequence (g_k) has ψ as a cluster point.

17.50. Eberlein-Smulian Theorem (nonlinear version). Let S be a compact topological space and let (M, d) be a compact metric space. The product M^S will be topologized with the product topology; subsets of M^S will be topologized with the relative topology thereby determined. In particular, this applies to

$$C(S, M) \quad = \quad \{\text{continuous functions from } S \text{ into } M\} \quad \subseteq \quad M^S.$$

Let $\Phi \subseteq C(S, M)$. Then the following conditions are equivalent:

(A) (*Iterated limit condition.*) For every net $(\varphi_\alpha : \alpha \in \mathbb{A})$ in Φ and every net $(s_\beta : \beta \in \mathbb{B})$ in S, we have

$$\lim_{\alpha \in \mathbb{A}} \lim_{\beta \in \mathbb{B}} \varphi_\alpha(s_\beta) \quad = \quad \lim_{\beta \in \mathbb{B}} \lim_{\alpha \in \mathbb{A}} \varphi_\alpha(s_\beta)$$

whenever both sides of the equation exist — i.e., whenever all the indicated limits exist.

(B) (*Sequential iterated limit condition.*) For all sequences (φ_m) in Φ and (s_n) in S, we have

$$\lim_{m\to\infty}\lim_{n\to\infty}\varphi_m(s_n) \quad = \quad \lim_{n\to\infty}\lim_{m\to\infty}\varphi_m(s_n)$$

whenever both sides of the equation exist.

(C) Every net in Φ has a cluster point in $C(S,M)$. That is, Φ is relatively compact in $C(S,M)$.

(D) Every sequence in Φ has a cluster point in $C(S,M)$.

(E) Every sequence in Φ has a sequential cluster point in $C(S,M)$.

(F) $\mathrm{cl}(\Phi) \subseteq C(S,M)$.

(G) For each countable set $V \subseteq \Phi$, the set $\mathrm{cl}(V)$ is metrizable, and $\mathrm{cl}(V) \subseteq C(S,M)$.

Remarks. Note that M^S is a compact Hausdorff space (regardless of the topology of S); hence a subset of M^S is compact if and only if it is closed. In particular, $\mathrm{cl}(\Phi)$ is compact. Those considerations do not depend on the topology of S, but the definition of $C(S,M)$ does depend on that topology, and so do the conditions listed in the theorem. In general, $C(S,M)$ is not closed in M^S; see 18.32.g. Note that M^S and its subsets are completely regular; hence 17.15 is applicable.

Outline of Proof. We shall show (F) $\Rightarrow$ (C) $\Rightarrow$ (D) $\Rightarrow$ (B) $\Rightarrow$ (A) $\Rightarrow$ (F), and also [(A) and (F)] $\Rightarrow$ (G) $\Rightarrow$ (E) $\Rightarrow$ (D).

Proof of (F) $\Rightarrow$ (C). M^S is compact.

Proof of (C) $\Rightarrow$ (D). Obvious.

Proof of (D) $\Rightarrow$ (B). Show that (φ_m) has a cluster point φ and (s_n) has a cluster point s. Then show that if either iterated limit exists, it is equal to $\varphi(s)$.

Proof of (B) $\Rightarrow$ (A). By assumption, the limits

$$p_\alpha = \lim_\beta \varphi_\alpha(s_\beta), \qquad q_\beta = \lim_\alpha \varphi_\alpha(s_\beta), \qquad p_\infty = \lim_\alpha p_\alpha, \qquad q_\infty = \lim_\beta q_\beta;$$

all exist in M; we wish to show that $p_\infty = q_\infty$. Recursively choose sequences $(\alpha(m) : m = 1,2,3,\ldots)$ in $\mathbb{A}$ and $(\beta(n) : n = 1,2,3,\ldots)$ in $\mathbb{B}$, as follows:

Let j be a positive integer. Assume that $\alpha(m)$ and $\beta(n)$ have already been selected for all positive integers $m,n < j$. (This assumption is free when $j = 1$ and no selections have been made yet.) Now choose some $\alpha(j)$ in $\mathbb{A}$ large enough so that

(1) $d[p_{\alpha(j)}, p_\infty] < 1/j$, and

(2) $d[\varphi_{\alpha(j)}(s_{\beta(n)}), q_{\beta(n)}] < 1/j$ for all positive integers $n < j$.

(Again, condition (2) is free when $j = 1$.) Then choose some $\beta(j)$ in $\mathbb{B}$ large enough so that

(3) $d[q_{\beta(j)}, q_\infty] < 1/j$, and

(4) $d[\varphi_{\alpha(m)}(s_{\beta(j)}), p_{\alpha(m)}] < 1/j$ for all positive integers $m \leq j$.

This completes the recursive definition. (We do *not* assert that the sequences $(\varphi_{\alpha(m)})$ and $(s_{\beta(n)})$ are subnets of the given nets (φ_α) and (s_β).) Now apply (B) to the sequences $(\varphi_{\alpha(m)} : m \in \mathbb{N})$ and $(s_{\beta(n)} : n \in \mathbb{N})$; this proves $p_\infty = q_\infty$.

Proof of (A) $\Rightarrow$ (F). Suppose $\psi \in \mathrm{cl}(A)$, and ψ is discontinuous at some point $s_0 \in S$. Then some net (φ_α) in Φ converges to ψ, and some net (s_β) in S converges to s_0 and satisfies $\psi(s_\beta) \not\to \psi(s_0)$. Since M is compact, by replacing (s_β) with a subnet we may assume that $(\psi(s_\beta))$ converges to some limit $z \neq \psi(s_0)$. This contradicts (A). Thus $\mathrm{cl}(\Phi) \subseteq C(S, M)$.

Proof that (F) *and* (A) *together imply* (G). For any $s \in S$ and $\varphi \in V$, let $[\varepsilon(s)](\varphi) = \varphi(s)$; thus we define $\varepsilon(s) \in M^V$. Observe that the "evaluation map" $\varepsilon : S \to M^V$ defined in this fashion is continuous. Since M^V is a compact metric space, it is separable, and hence there is some countable set $T \subseteq S$ such that $\varepsilon(T)$ is dense in $\varepsilon(S)$.

We now claim that

$$\text{if } g, h \in \mathrm{cl}(V) \text{ and } g = h \text{ on } T, \text{ then } g = h \text{ on } S.$$

To see this, fix any $s_0 \in S$. There is some net (t_β) in T such that $\varepsilon(t_\beta) \to \varepsilon(s_0)$. Replacing (t_β) with a subnet, we may also assume $(g(t_\beta))$ and $(h(t_\beta))$ are convergent. By (A), show that $g(s_0) = \lim_\beta g(t_\beta) = \lim_\beta h(t_\beta) = h(s_0)$. This proves our claim.

Now, the restriction map $\mathfrak{R} : g \mapsto g\big|_T$ is continuous from M^S onto M^T. By 17.10.c, that map gives a homeomorphism of $\mathrm{cl}(V)$ onto its image, $\mathfrak{R}(\mathrm{cl}(V))$, which is a subset of the metrizable space M^T.

Proof of (G) $\Rightarrow$ (E). M^S is compact, and any compact metric space is sequentially compact.

17.51. Corollary. Assume the conditions of the preceding theorem. Then Φ is compact $\iff$ Φ is sequentially compact $\iff$ Φ is countably compact.

Outline of proof. Countable compactness is always implied by either of the other two kinds of compactness, so we may assume Φ is countably compact. Then condition (D) of the Eberlein-Smulian Theorem is satisfied, hence all the conditions of that theorem are satisfied. It now suffices to show Φ is closed, for then compactness and sequential compactness follow from parts (C) and (E) of the Eberlein-Smulian Theorem.

Let $\psi \in \mathrm{cl}(\Phi)$; we wish to show $\psi \in \Phi$. By condition (F) of the theorem, we have $\psi \in C(S, M)$. By 17.49, ψ is a cluster point of some sequence (φ_n) in Φ. We know $\mathrm{cl}(\{\varphi_n\})$ is metrizable, by condition (G) of the Eberlein-Smulian Theorem; hence by 15.34.c, ψ is the limit of some subsequence of (φ_n). Since Φ is Hausdorff and countably compact, any sequence in Φ that converges must have its unique limit in Φ; hence $\psi \in \Phi$.

Chapter 18

Uniform Spaces

18.1. *Preview.* We now resume the study of uniform spaces, which we began in Chapters 5 and 9. Our study will also make use of material from Chapters 7 through 17; see especially 16.16.

As shown in the following chart, uniform structure fits between topological structure and the structure provided by distances, in its degree of detail of information about objects. Movement from right to left in this table is given by forgetful functors (discussed in 9.34).

	less $\longleftarrow$ details about the object $\longrightarrow$ more		
structure:	topological	uniform	distances
typical questions:	Is f continuous? Is S compact? Is S topologically complete?	Is f uniformly continuous? Is S complete?	In metric spaces: Is f nonexpansive? Is S bounded?
broader $\uparrow$	subset of 2^X	subset of $2^{X\times X}$	subset of $[0, +\infty)^{X\times X}$
	topology	quasi-uniformity	quasigauge
class of objects $\updownarrow$	completely regular topology	uniformity	gauge (a set of pseudometrics)
	pseudometrizable topology	pseudometrizable uniformity	pseudometric
narrower $\downarrow$	metrizable topology	metrizable uniformity	metric

These functors are not inclusions of subcategories in categories, because the maps are not injective. For instance, each gauge uniquely determines a uniformity. We may "forget" which gauge determined the uniformity; different gauges on a set may determine the same or different uniformities. Similarly, each uniformity uniquely determines a completely regular topology. We may "forget" which uniformity determined the topology; different uniformities on a set may determine the same or different completely regular topologies on that set.

Each category in the table is a full subcategory of the category above it. (Full subcategories were introduced in 9.5.) For instance, completely regular topological spaces are a full subcategory of topological spaces; in either of these two categories the morphisms are the continuous maps.

Lipschitz Mappings

18.2. *Definitions.* Let (X,d) and (Y,e) be metric spaces. A mapping $p : X \to Y$ is said to be **Lipschitz**, or *Lipschitzian*, if $e(p(x_1),p(x_2)) \leq \kappa d(x_1,x_2)$ for some finite constant κ and all $x_1, x_2 \in X$. The smallest such κ is then called the **Lipschitz constant** of p; it is equal to

$$\langle p \rangle_{\mathrm{Lip}} \quad = \quad \sup\left\{\frac{e(p(x_1),p(x_2))}{d(x_1,x_2)} \quad : \quad x_1,x_2 \in X, \quad x_1 \neq x_2\right\}.$$

The set of all Lipschitz mappings from (X,d) into (Y,e) will be denoted $\mathrm{Lip}(X,Y)$.

We say p is **nonexpansive** if $\langle p \rangle_{\mathrm{Lip}} \leq 1$. The mapping is a **strict contraction** if $\langle p \rangle_{\mathrm{Lip}} < 1$. *Caution*: This book will not use the term "*contraction*" by itself. Some mathematicians use that term for nonexpansive mappings; others use it for strict contractions.

18.3. *Examples.*

a. Let S be a nonempty subset of a metric space (X,d). Then the map $x \mapsto \mathrm{dist}(x,S)$, defined in 4.40, is nonexpansive from X into $\mathbb{R}$. (See 4.41.b.)

b. (This example assumes more knowledge of calculus.) Let $p : \mathbb{R} \to \mathbb{R}$ be continuously differentiable. Then p is Lipschitz if and only if p' is bounded, in which case $\langle p \rangle_{\mathrm{Lip}} = \sup_{x\in\mathbb{R}} |p'(x)|$. (We shall generalize this result in 25.24.)

c. $x \mapsto |x|$ is Lipschitz on $\mathbb{R}$, but not continuously differentiable.

d. (*Preview.*) Let $p : X \to Y$ be a linear map from one normed vector space to another. Then p is continuous if and only if p is Lipschitzian. (See 23.1.)

e. Suppose (X,ρ) is a metric space, and $f : X \to X$ is a mapping with the property that for each $x \in X$, the set $\{f^n(x) : n = 0,1,2,3,\ldots\}$ is metrically bounded. Then

$$\beta(x,y) \quad \stackrel{\centerdot}{=} \quad \sup\left\{\rho\Big(f^n(x), f^n(y)\Big) \ : \ n = 0,1,2,3,\ldots\right\}$$

is a metric on X that is larger than or equal to ρ and makes the mapping f nonexpansive from (X,β) into itself. In fact, β is the smallest metric on X that has those two properties. (In 19.47.c we shall consider some conditions under which β is topologically equivalent to ρ.)

18.4. *Definitions.* Let (X,d) and (Y,e) be metric spaces. For $\alpha > 0$ and mappings $p : X \to Y$, let

$$\langle p \rangle_{\alpha} \quad = \quad \sup\left\{\frac{e(p(x_1),p(x_2))}{d(x_1,x_2)^{\alpha}} \quad : \quad x_1,x_2 \in X, \quad x_1 \neq x_2\right\}.$$

We say p is **Hölder continuous with exponent α** if $\langle p\rangle_\alpha < \infty$. We shall denote the class of such functions by $\text{Höl}^\alpha(X,Y)$. Note that $\text{Höl}^1(X,Y) = \text{Lip}(X,Y)$, with $\langle p\rangle_1 = \langle p\rangle_{\text{Lip}}$.

18.5. *Examples and exercises.*

a. A function $p : X \to Y$ is constant if and only if $\langle p\rangle_\alpha = 0$.

b. Let X and Y both be equal to the set $[0, +\infty)$ equipped with the usual metric $d(x,y) = |x-y|$. Let $\alpha, \beta \in (0,1]$. Let $p(x) = x^\beta$. Then

$$\langle p\rangle_\alpha \quad = \quad \begin{cases} 1 & \text{if } \beta = \alpha \\ \infty & \text{if } \beta \neq \alpha. \end{cases}$$

Hint: First show that $((u+h)^\beta - u^\beta)/h^\alpha$ is a nonincreasing function of u, for $u, h > 0$.

c. Let $p : X \to Y$ and $q : Y \to Z$ be Hölder continuous with exponents α and β respectively. Then the composition $q \circ p : X \to Z$ is Hölder continuous with exponent $\alpha\beta$, and in fact $\langle q \circ p\rangle_{\alpha\beta} \leq \langle p\rangle_\alpha^\beta \, \langle q\rangle_\beta$.

In particular, the composition of Lipschitzian functions is a Lipschitzian function; we have $\langle qp\rangle_{\text{Lip}} \leq \langle p\rangle_{\text{Lip}} \, \langle q\rangle_{\text{Lip}}$.

d. Let X and Y be metric spaces, with X metrically bounded. For $\alpha > \beta$ and $p : X \to Y$, show that $\langle p\rangle_\beta \leq \langle p\rangle_\alpha (\text{diam}(X))^{\alpha-\beta}$ and hence $\text{Höl}^\alpha(X,Y) \subseteq \text{Höl}^\beta(X,Y)$.

e. For $\alpha > 1$, the spaces $\text{Höl}^\alpha(\mathbb{R}, Y)$ are not very interesting, for they contain only constant functions. A hint will be given when we generalize this result slightly in 22.18.e.

18.6. Let X and Y be metric spaces. A function $f : X \to Y$ is called **locally Lipschitz** if any (hence all) of the following equivalent conditions are satisfied:

(A) f is Lipschitz on a neighborhood of each point,

(B) f is Lipschitz on each compact set.

(C) f is Lipschitz on a neighborhood of each compact set. More precisely, if K is a compact subset of X, then there is some number $r > 0$ such that the restriction of f to the open set $\{x \in X \ : \ \text{dist}_d(x, K) < r\}$ is Lipschitz.

Exercises.

a. Prove the equivalence.

Hints: It suffices to show (A) $\Rightarrow$ (C). Suppose not. Show that there exist sequences $(x_n), (x_n')$ in X such that $x_n \neq x_n'$, $\text{dist}_d(x_n, K) \to 0$, $\text{dist}_d(x_n', K) \to 0$, and $e\Big(f(x_n), f(x_n')\Big)/d(x_n, x_n') \to \infty$. Passing to subsequences, we may assume (x_n) and (x_n') converge to limits x_∞ and x_∞' in K. Show that $x_\infty = x_\infty'$. Then what?

b. For any open cover of a metric space X, there exists a partition of unity subordinated to that cover, consisting of locally Lipschitzian functions.

Hints: Let $\mathcal{T} = \{T_\alpha : \alpha \in A\}$ be a locally finite open cover that is subordinated to the given cover (see 16.31 and 16.29). For each α, define $f_\alpha : X \to [0, +\infty)$ by $f_\alpha(x) = \text{dist}(x, X \setminus T_\alpha)$. Then define a partition of unity $\{g_\alpha\}$ as in 16.25.c.

c. Let X be a metric space, and let $f : X \to \mathbb{R}$ be continuous. Then f can be approximated uniformly by locally Lipschitz functions — i.e., for any $\varepsilon > 0$ there exists a locally Lipschitz function $g : X \to \mathbb{R}$ satisfying $\sup_{x\in X} |f(x) - g(x)| < \varepsilon$.

Proof. Each $x \in X$ has a neighborhood N_x such that $|f(x)-f(y)| < \varepsilon$ for all $y \in N_x$. Choosing smaller N_x's if necessary, we may assume that the N_x's are all open; then they form an open cover of X. Let $\{p_\alpha : \alpha \in A\}$ be a locally Lipschitzian partition of unity that is subordinated to the cover $\{N_x\}$. For each $\alpha \in A$, let $v(\alpha)$ be some member of X such that $\{u \in X \ : \ p_\alpha(u) \neq 0\} \subseteq N_{v(\alpha)}$. Show that $g(x) = \sum_{\alpha\in A} f(v(\alpha))\, p_\alpha(x)$ has the required properties.

Remarks. This argument also works for functions f, g from any metric space X into any Banach space B. It will be used for differential equations in 30.11.

Uniform Continuity

18.7. *Notation.* Let $(X, \mathcal{U})$ be a uniform space, with uniformity $\mathcal{U}$ determined by gauge D. Let $((x_\alpha, x'_\alpha) \ : \ \alpha \in \mathbb{A})$ be a net in $X \times X$, and let $\mathcal{E}$ be its eventuality filter on $X \times X$. Show that the following conditions are equivalent.

(A) $\mathcal{U} \subseteq \mathcal{E}$.

(B) For each $U \in \mathcal{U}$, we have eventually $(x_\alpha, x'_\alpha) \in U$.

(C) $d(x_\alpha, x'_\alpha) \to 0$ in $\mathbb{R}$ for each $d \in D$. We shall abbreviate this as $D(x_\alpha, x'_\alpha) \to 0$.

We emphasize that the last condition does *not* say $\sup_{d\in D} d(x_\alpha, x'_\alpha) \to 0$.

18.8. *Definition.* Let $(X, \mathcal{U})$ and $(Y, \mathcal{V})$ be uniform spaces, and let D and E be any gauges that determine the uniformities $\mathcal{U}$ and $\mathcal{V}$, respectively. Let $\varphi : X \to Y$ be some function. Then the following conditions are equivalent. If any (hence all) of these conditions hold, we say φ is **uniformly continuous**.

(A) Whenever $V \in \mathcal{V}$, then the set

$$(\varphi \times \varphi)^{-1}(V) \quad = \quad \Big\{(x, x') \in X \times X \quad : \quad \big(\varphi(x), \varphi(x')\big) \in V\Big\}$$

is a member of $\mathcal{U}$. That is, the inverse image of each entourage is an entourage. (This is the definition of uniform continuity used in 9.8.)

(B) Whenever $D(x_\alpha, x'_\alpha) \to 0$ in X, then $E\left(\varphi(x_\alpha), \varphi(x'_\alpha)\right) \to 0$ in Y. (Notation is as in 18.7(C).)

(C) For each number $\varepsilon > 0$ and each pseudometric $e \in E$, there exist some number $\delta > 0$ and some finite set $D' \subseteq D$ such that

$$\max_{d\in D'} d(x_1, x_2) \ < \ \delta \qquad \Rightarrow \qquad e\left(\varphi(x_1), \varphi(x_2)\right) \ < \ \varepsilon.$$

(We emphasize that the choice of δ and D' depends on ε and e but not on x_1 or x_2. Compare this with 15.14(D).)

(D) For each $e \in E$, there exists a finite set $D_e \subseteq D$ and a function $\gamma_e : [0, +\infty) \to [0, +\infty)$ that is continuous and increasing, and satisfies $\gamma_e(0) = 0$ and

$$e\Big(\varphi(x), \varphi(x')\Big) \quad \leq \quad \gamma_e\Big(\max_{d \in D_e} d(x, x')\Big).$$

Such a system of sets D_e and functions γ_e will be called a **modulus of uniform continuity** for φ.

Note that if the gauge D is directed (as defined in 4.4.c), then conditions 18.8(C) and 18.8(D) can be simplified slightly: the sets D' and D_e may be taken to be singletons $\{d\}$.

18.9. *Examples and related properties.*

a. If the uniformity on X is given by a pseudometric d, then sequences suffice in 18.8(B) (regardless of whether Y is pseudometrizable). That is, a mapping $\varphi : X \to Y$ is uniformly continuous if and only if

$$\text{whenever } d(x_n, x'_n) \to 0 \text{ in } X, \text{ then } E\Big(\varphi(x_n), \varphi(x'_n)\Big) \to 0 \text{ in } Y,$$

with notation as in 18.7(C).

b. Any Hölder-continuous function from one metric space into another is uniformly continuous. The converse is false. For instance, define $f : [0, e^{-1}] \to \mathbb{R}$ by

$$f(t) \quad = \quad \begin{cases} 0 & \text{when } t = 0 \\ -1/\ln t & \text{when } 0 < t \leq e^{-1}\,. \end{cases}$$

Show that f is not Hölder continuous with any exponent. It is easy to see that f is continuous; then the uniform continuity of f will follow by a compactness argument in 18.21.

c. Any uniformly continuous function is continuous (where each uniform space is equipped with its uniform topology). This can be proved using uniformities or using gauges; the student is urged to give both proofs.

d. Show that the function $f(t) = 1/t$ is continuous, but not uniformly continuous, on the open interval $(0, 1)$. Use this fact to give two different metrics on $(0, 1)$ that yield different uniformities but that both yield the usual topology.

e. (*Preview.*) Let $p : X \to Y$ be a linear map from one topological vector space to another — or more generally, an additive map from one topological Abelian group to another. Let X and Y be equipped with their usual uniform structures (see 26.37). If p is continuous, then p is uniformly continuous; see 26.36.c.

f. Let X be a set, let $\{(Y_\lambda, E_\lambda) : \lambda \in \Lambda\}$ be a collection of gauge spaces, and let $\varphi_\lambda : X \to Y_\lambda$ be some mappings. Show that the initial uniformity on X determined by the φ_λ's and E_λ's (as in 9.16) is equal to the uniformity on X determined by the gauge $D = \bigcup_{\lambda \in \Lambda} \{e\varphi_\lambda : e \in E_\lambda\}$, where $(e\varphi_\lambda)(x, x') = e\left(\varphi_\lambda(x), \varphi_\lambda(x')\right)$. We may call this the **initial gauge** determined by the φ_λ's and E_λ's (although any other gauge uniformly equivalent to this one will generally do just as well).

An important special case: When $X = \prod_{\lambda \in \Lambda} Y_\lambda$ and the φ_λ's are the coordinate projections, we obtain the **product gauge**.

g. The forgetful functor from uniform spaces to topological spaces preserves the formation of initial objects.

That is, the uniform topology $\tau(\mathcal{U})$ determined by an initial uniformity $\mathcal{U}$ determined by π_λ's and uniformities $\mathcal{V}_\lambda$ is equal to the initial topology determined by the π_λ's and the uniform topologies $\tau(\mathcal{V}_\lambda)$ determined by those uniformities.

18.10. Theorem on uniform continuity of extensions. Let X and Y be uniform spaces, let $X_0 \subseteq X$ be dense, let $\varphi : X \to Y$ be continuous, and suppose that the restriction of φ to X_0 is uniformly continuous. Then φ is uniformly continuous on X. In fact, if some gauges are specified for X and Y, then any modulus of uniform continuity for the restriction of φ to X_0 is also a modulus of uniform continuity for φ on X.

In particular, if φ is continuous and the restriction of φ to X_0 is Hölder continuous or Lipschitzian, then φ is Hölder continuous or Lipschitzian with the same constant.

Hints: Use notation as in 18.8(D). Let any $x, x' \in X$ be given. Choose a net $\Big((x_\alpha, x'_\alpha)\Big)$ in $X_0 \times X_0$ that converges to (x, x'). For each α, we have

$$e\Big(\varphi(x_\alpha), \varphi(x'_\alpha)\Big) \quad \leq \quad \gamma_e\Big(\max_{d \in D_e} d(x_\alpha, x'_\alpha)\Big).$$

Holding e fixed, take limits to obtain a corresponding inequality for (x, x').

18.11. Characterization of uniformly equivalent gauges. Let D and E be gauges on a set X. Then the following conditions are equivalent:

(A) D and E are uniformly equivalent — i.e., they generate the same uniformity.

(B) The identity map $i_X : X \to X$ is uniformly continuous in both directions between the gauge spaces (X, D) and (X, E).

(C) For each net $((x_\alpha, x'_\alpha) \ : \ \alpha \in \mathbb{A})$ in $X \times X$, we have

$$D(x_\alpha, x'_\alpha) \to 0 \qquad \Longleftrightarrow \qquad E(x_\alpha, x'_\alpha) \to 0$$

with notation as in 18.7(C).

Hint: A uniformity, being a proper filter, is the eventuality filter for some net.

18.12. *Further exercise.* Let $\mathcal{U}$ be a uniformity on a set X. Then the largest gauge that is compatible with $\mathcal{U}$ (as defined in 5.32) is the set of all pseudometrics $d : X \times X \to [0, +\infty)$ that are jointly uniformly continuous — i.e., uniformly continuous when $X \times X$ is given its product uniformity and $[0, +\infty)$ is given its usual uniformity. (Compare this with 16.20.)

18.13. If D is any gauge, then D is uniformly equivalent to its max closure and its sum closure, defined as in 4.4.c.

(Hence it is often possible to replace a gauge with a directed gauge; thus in many contexts we may assume a gauge is directed.)

18.14. *Definition.* We shall say β is a **bounded remetrization function** if:

(i) β is a mapping from $[0, +\infty)$ onto a bounded subset of $[0, +\infty)$;

(ii) β is continuous;

(iii) β is increasing; that is, $s \leq t \Rightarrow \beta(s) \leq \beta(t)$;

(iv) $\beta(t) = 0 \iff t = 0$; and

(v) β is subadditive; that is, $\beta(s+t) \leq \beta(s) + \beta(t)$.

Show that

a. $\arctan(t)$, $\tanh(t)$, $\min\{1, t\}$, and $t/(1+t)$ are bounded remetrization functions of t. (*Hint*: See 12.25.e.) Note that $\min\{1, t\}$ is not injective.

b. If β is a bounded remetrization function and d is a (pseudo)metric on a set X, then $e(x, y) = \beta(d(x, y))$ defines a (pseudo)metric $e = \beta \circ d$ on X that is uniformly equivalent to d and is bounded.

c. If β is a bounded remetrization function and D is a gauge on a set X, then $\{\beta \circ d : d \in D\}$ is a gauge on X that is uniformly equivalent to D and is *uniformly* bounded — i.e., we have $\sup\{\beta(d(x,y)) \ : \ x, y \in X, \ d \in D\} < \infty$.

18.15. *Example.* The usual metric on $\mathbb{R}$ is $d(x, y) = |x - y|$. Another metric, bounded and uniformly equivalent to the usual one, is $e(x, y) = \arctan(|x - y|)$. On the other hand, $\rho(x, y) = |\arctan(x) - \arctan(y)|$ is a bounded metric on $\mathbb{R}$ that is equivalent, but not uniformly equivalent, to the usual metric. (All three metrics yield the same topology.)

Pseudometrizable Gauges

18.16. *Finite gauges.* Any finite gauge $\{d_1, d_2, \ldots, d_n\}$ on a set X is uniformly equivalent to a single pseudometric d. There are many ways to define $d(u, v)$. Two commonly used pseudometrics are

$$\sum_{j=1}^{n} d_j(u, v), \qquad\qquad \max_{1 \leq j \leq n} d_j(u, v).$$

Hints: The proof of equivalence may be accomplished most easily using 18.11(C).

For other, more complicated pseudometrics equivalent to these, see 22.11.

18.17. Any countably infinite gauge $\{d_1, d_2, d_3, \ldots\}$ on a set X is uniformly equivalent to a single pseudometric. One such pseudometric is

$$d(u, v) \quad = \quad \sum_{j=1}^{\infty} 2^{-j} \arctan\Big(d_j(u, v)\Big).$$

More generally, we could use any pseudometric of the form

$$d(u,v) \quad = \quad \sum_{j=1}^{\infty} a_j \, \beta\Big(d_j(u,v)\Big)$$

where β is a bounded remetrization function (see 18.14) and the a_j's are positive numbers with finite sum. Any such pseudometric d is called a **Fréchet combination** of the d_j's.

These formulas are admittedly rather complicated, but in general they cannot be replaced by a simpler formula. In many applications, the d_j's themselves are quite simple, and so we may reason in terms of the d_j's instead of d. However, in such applications, we may sometimes use the fact that the structure of X can be given by *some* single pseudometric d, without referring to any particular choice of d.

For instance, an argument involving a net convergence $x_\alpha \to z$ can often be replaced by an argument involving a sequential convergence $x_n \to z$, since the topology is pseudometrizable. Then that convergence $x_n \to z$ can be represented conveniently by the condition that $d_j(x_n, z) \to 0$ for each j.

Fréchet combinations will be used to give pseudometrics for certain product topologies and uniformities, and for uniform convergence on compact sets in certain classes of functions — continuous, smooth, holomorphic, etc.; see 18.18, 26.6, 26.7, 26.8, and 26.10.

Remark. Unlike the finite and countable cases, an uncountable gauge generally is *not* uniformly equivalent to a single pseudometric. We shall prove this by an example in 18.20.

18.18. *Product pseudometrics.* If $(Y_1, e_1), (Y_2, e_2), \ldots, (Y_n, e_n)$ are pseudometric spaces, then the product gauge on $X = \prod_{j=1}^{n} Y_j$ is uniformly equivalent to a single pseudometric. Two pseudometrics commonly used for this purpose are

$$\sum_{j=1}^{n} e_j(u_j, v_j), \qquad\qquad \max_{1 \leq j \leq n} e_j(u_j, v_j)$$

where $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$. This is a special case of 18.16 obtained by taking $d_j(u,v) = e_j(\pi_j(u), \pi_j(v))$.

If $(Y_1, e_1), (Y_2, e_2), (Y_3, e_3), \ldots$ is a sequence of pseudometric spaces, then the product gauge on $X = \prod_{j=1}^{\infty} Y_j$ is uniformly equivalent to a single pseudometric. We can use any pseudometric of the form

$$e(u,v) \quad = \quad \sum_{j=1}^{\infty} a_j \, \beta\Big(e_j(u_j, v_j)\Big)$$

where β is a bounded remetrization function and the a_j's are positive numbers with finite sum; here $u = (u_1, u_2, u_3, \ldots)$ and $v = (v_1, v_2, v_3, \ldots)$. This is a special case of 18.17, obtained by taking $d_j(u,v) = e_j(\pi_j(u), \pi_j(v))$.

The product gauge on an uncountable product generally is *not* uniformly equivalent to a single pseudometric; this will be proved in 18.20.

18.19. *Further examples and consequences.*

a. The product gauge on $\mathbb{R}^{\mathbb{N}} = \{\text{sequences of reals}\}$ is uniformly equivalent to the metric $d(u,v) = \sum_{j=1}^{\infty} 2^{-j} \min\{1, |u_j - v_j|\}$.

b. If $Y_1, Y_2, Y_3, \ldots$ are each equipped with the Kronecker metric, then the product gauge on $\prod_{j=1}^{\infty} Y_j$ is also uniformly equivalent to this simple metric:

$$d\Big((y_n), (y'_n)\Big) \quad = \quad \frac{1}{\min\{n \ : \ y_n \neq y'_n\}}.$$

Note that a product of infinitely many discrete spaces generally is *not* discrete. In particular, $2^{\mathbb{N}}$ is not discrete. In fact, *no* point in $2^{\mathbb{N}}$ is isolated.

18.20. Theorem. Let $X = \prod_{\lambda \in \Lambda} Y_\lambda$ be a product of topological spaces with the product topology, and assume X is nonempty and pseudometrizable. Then (i) each of the Y_λ's has a pseudometrizable topology, and (ii) for all but countably many of the λ's, Y_λ has the indiscrete topology.

(Thus, except for degenerate cases, a product of uncountably many topological spaces is not pseudometrizable. For example, the product topology and uniformity on $\mathbb{R}^{\mathbb{R}}$ or on $\{0,1\}^{\mathbb{R}}$ cannot be given by a single pseudometric.)

Proof. As usual, let $\pi_\lambda : X \to Y_\lambda$ be the λth coordinate projection. Assume the product topology is given by a pseudometric d on X.

To prove (i), fix any μ. Since X is nonempty, $\prod_{\lambda \neq \mu} Y_\lambda$ is also nonempty; fix any $\zeta \in \prod_{\lambda \neq \mu} Y_\lambda$. Define a pseudometric d_μ on Y_μ by $d_\mu(y_\mu, y'_\mu) = d\left((y_\mu, \zeta), (y'_\mu, \zeta)\right)$. Show that this pseudometric yields the given topology on Y_μ.

To prove (ii), assume that the set

$$M \quad = \quad \{\lambda \in \Lambda \ : \ \text{the topology of } Y_\lambda \text{ is not the indiscrete topology}\}$$

is uncountable; we shall obtain a contradiction. Fix any point $\xi \in X$. For $n = 1, 2, 3, \ldots$, let $B_n = \{x \in X \ : \ d(\xi, x) < 1/n\}$. Since B_n is an open set containing ξ, there exist some finite set $J_n \subseteq \Lambda$ and open sets $G_j \subseteq Y_j$ $(j \in J_n)$ such that

$$\xi \quad \in \quad \bigcap_{j \in J_n} \pi_j^{-1}(G_j) \quad \subseteq \quad B_n$$

by 15.26.a. Let $J = \bigcup_{n=1}^{\infty} J_n$; then J is a countable subset of Λ. Since M is uncountable, there is at least one $\mu \in M \setminus J$. Fix any such μ. By 16.12.c there exist some $p \in Y_\mu$ and some disjoint open sets $S, T \subseteq Y_\mu$ such that $\pi_\mu(\xi) \in S$ and $p \in T$. Let η be the point in X defined by

$$\pi_\lambda(\eta) \quad = \quad \begin{cases} \pi_\lambda(\xi) & \text{when } \lambda \neq \mu \\ p & \text{when } \lambda = \mu. \end{cases}$$

Then $S \times \prod_{\lambda \neq \mu} Y_\lambda$ and $T \times \prod_{\lambda \neq \mu} Y_\lambda$ are disjoint open subsets of X that contain ξ and η, respectively. The existence of such sets implies that $d(\xi, \eta) > 0$. However, $\mu \notin J$, so $\pi_j(\xi) = \pi_j(\eta)$ for every $j \in J$. From this it follows that $\eta \in B_n$ for each positive integer n — that is, $d(\xi, \eta) < 1/n$ for each n, a contradiction.

Compactness and Uniformity

18.21. *Proposition.* Let X and Y be uniform spaces, assume X is compact, and suppose $f : X \to Y$ is continuous. Then:

(i) f is uniformly continuous.

(ii) If f is injective and Y is Hausdorff, then the inverse map f^{-1} : Range$(f) \to X$ is also uniformly continuous; thus f is a uniform isomorphism onto its range.

Proof of (i). Assume $(x_\alpha, x'_\alpha) \to I$ in X; we must show that $(f(x_\alpha), f(x'_\alpha)) \to I$ in Y. Suppose not; then there is some pseudometric e in a gauge for Y such that $e(f(x_\alpha), f(x'_\alpha)) \not\to 0$. Replacing $((x_\alpha, x'_\alpha))$ with a subnet, we may assume that $e(f(x_\alpha), f(x'_\alpha)) > \rho$ for some constant $\rho > 0$. Again replacing $((x_\alpha, x'_\alpha))$ with a subnet and using the compactness of X, we may assume (x_α) converges to some limit $v \in X$. Then $x'_\alpha \to v$ also. By continuity, we have $f(x_\alpha) \to f(v)$ and $f(x'_\alpha) \to f(v)$, a contradiction.

Proof of (ii). $f(X)$ is compact, by 17.7.h. Also, f^{-1} : Range$(f) \to X$ is continuous, by 17.10.c. Hence it is uniformly continuous, by the argument of the preceding paragraph.

Remark. There exist metric spaces X that are not compact, but nevertheless have the property that any continuous function from X into another metric space is uniformly continuous. Indeed, one such space is $\mathbb{Z}$, with its usual metric. Such spaces are discussed further by Arala-Chaves [1985].

18.22. *Corollary.* A compact topological space has at most one uniform structure. In other words, if $\mathcal{U}$ and $\mathcal{V}$ are uniformities on a set X that yield the same compact topology, then $\mathcal{U} = \mathcal{V}$.

18.23. *Proposition (optional).* Let $(X, \mathcal{U})$ be a uniform space, let $\mathcal{T}$ be the uniform topology, let $X \times X$ be equipped with the product topology, and let $I = \{(x, x) : x \in X\}$. By a "neighborhood of the diagonal" we shall mean a set $U \subseteq X \times X$ such that $U \supseteq G \supseteq I$ for some G that is open in $X \times X$. Show that

(i) Any entourage is a neighborhood of the diagonal.

(ii) If $(X, \mathcal{T})$ is compact, then every neighborhood of the diagonal is an entourage — i.e., the uniformity $\mathcal{U}$ is equal to the set of neighborhoods of the diagonal.

Proof of (i). Given $U \in \mathcal{U}$, let V be a symmetric entourage satisfying $V \circ V \subseteq U$. For each $x \in X$, we know that $V[x]$ is a neighborhood of x, hence $x \in E_x \subseteq V[x]$ for some open set E_x in X. Show that $G = \bigcup_{x \in X}(E_x \times E_x)$ has the required properties.

Proof of (ii). Let G be an open subset of $X \times X$ with $G \supseteq I$; we must show $G \in \mathcal{U}$. For each $x \in X$, we know that G is a neighborhood of (x, x) in $X \times X$. Hence $G \supseteq B_x \times B_x$, where B_x is some neighborhood of x in X. Then $B_x \supseteq U[x]$ for some entourage $U \in \mathcal{U}$. This U may depend on x; let us write it as U_x to reflect this — that is, $B_x \supseteq U_x[x]$. Let V_x be a symmetric entourage satisfying $V_x \circ V_x \subseteq U_x$. Then $V_x[x]$ is a neighborhood of x, so the compact set X can be covered by finitely many of the sets $\{V_x[x] : x \in X\}$. Say we have $X = \bigcup_{i=1}^n V_{x_i}[x_i]$. Then $W = \bigcap_{i=1}^n V_{x_i}$ is an entourage. It suffices to show $W \subseteq G$. Fix any $(p,q) \in W$. Since $\{V_{x_i}[x_i] : i = 1, 2, \ldots, n\}$ is a cover of X, we have $p \in V_{x_j}[x_j]$ for some j. It is now easy to show that $(p,q) \in U_{x_j}[x_j] \times U_{x_j}[x_j] \subseteq G$. This presentation follows the one of James [1987].

18.24. Let $f : \mathbb{R} \to \mathbb{R}$ be any continuous, bounded, strictly increasing function; three examples are

$$f(x) = \frac{x}{1+|x|}, \qquad f(x) = \arctan x, \qquad f(x) = \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

With any of these functions, $\lim_{x\to-\infty} f(x)$ and $\lim_{x\to+\infty} f(x)$ both exist in $\mathbb{R}$, and so it is natural to take those finite numbers as definitions of $f(-\infty)$ and $f(+\infty)$, respectively. Then $d(x,y) = |f(x) - f(y)|$ defines a metric on the extended real line $[-\infty, +\infty]$. Show that

a. The topology determined on $[-\infty, +\infty]$ in this fashion is the same as the order interval topology introduced in 5.15.f, since both topologies have the same convergences. *Hint*: 15.41.

b. f is a distance-preserving and order-preserving map from $([-\infty, +\infty], d)$ onto the metric space

$$\Big(\ [f_j(-\infty)\ ,\ f_j(+\infty)]\ , \quad \text{usual metric of } \mathbb{R} \ \Big).$$

c. The topological space $[-\infty, +\infty]$, with the order interval topology, is compact.

d. Any two such functions f yield metrics on $[-\infty, +\infty]$ that are uniformly equivalent. Thus, different choices of f yield different metrics, but the particular choice of f does not matter greatly.

Uniform Convergence

18.25. Let S be a set, let (Y, e) be a pseudometric space, let $(\varphi_\alpha : \alpha \in \mathbb{A})$ be a net in $Y^S = \{\text{functions from } S \text{ into } Y\}$, and let $\varphi \in Y^S$. We say that φ_α **converges uniformly** to φ on S if

$$\lim_{\alpha \in \mathbb{A}} \ \sup_{s \in S} \ e\Big[\varphi_\alpha(s), \varphi(s)\Big] \quad = \quad 0.$$

We may refer to this as e-**uniform convergence** when we need to be more specific.

We shall show that uniform convergence is given by a pseudometric on Y^S. Indeed, observe that the convergence is unchanged if we replace e with any uniformly equivalent pseudometric e'. By 18.14 we may assume that e' is bounded. Show that the uniform convergence on S is then given by this pseudometric on Y^S:

$$\rho_{S,e'}(\varphi, \psi) \quad = \quad \sup_{s \in S} \ e'\Big[\varphi(s), \psi(s)\Big].$$

We shall generalize this notion below.

18.26. *Generalization.* Let X be a set, and let $\mathcal{S}$ be a collection of subsets of X. Let $(Y, \mathcal{U})$ be a uniform space. A net (φ_α) in $Y^X = \{\text{functions from } X \text{ into } Y\}$ will be said to **converge $\mathcal{U}$-uniformly on elements of** $\mathcal{S}$ to a limit $\varphi \in Y^X$ if

$$\text{for each } S \in \mathcal{S} \text{ and } U \in \mathcal{U}, \text{ eventually } \{(\varphi_\alpha(s), \varphi(s)) \ : \ s \in S\} \subseteq U.$$

This can be expressed in terms of gauges as well. Let E be any gauge on Y that yields the uniformity $\mathcal{U}$; then $\varphi_\alpha \to \varphi$ in the sense above if and only if

$$\sup_{s \in S} e\Big[\varphi_\alpha(s), \varphi(s)\Big] \quad \to \quad 0 \qquad\qquad \text{for each } e \in E \text{ and } S \in \mathcal{S}.$$

(We emphasize that this condition does *not* require that either of the conditions

$$\sup_{e \in E} e[\varphi_\alpha(s), \varphi(s)] \to 0 \qquad\qquad \text{or} \qquad\qquad \sup_{s \in X} e[\varphi_\alpha(s), \varphi(s)] \to 0$$

must hold.)

We shall show that $\mathcal{U}$-uniform convergence on elements of $\mathcal{S}$ is the topological convergence given by a gauge on Y^X. Indeed, observe that we obtain the same convergence if we replace E with any uniformly equivalent gauge E'. By 18.14 we may assume that E' is bounded. Show that $\mathcal{U}$-uniform convergence on members of $\mathcal{S}$ is then given by the gauge $\{\rho_{S,e'} \ : \ S \in \mathcal{S},\ e' \in E'\}$, with $\rho_{S,e'}$ defined as in 18.25. The resulting uniformity on Y^X could be described as the **uniformity of uniform convergence on members of** $\mathcal{S}$. In some contexts the topology resulting from that uniformity is called the **topology of uniform convergence on members of** $\mathcal{S}$. That terminology is especially prevalent in contexts where the topology and uniformity determine each other uniquely, as described in 26.37.

Of course, different choices of $\mathcal{S}$ may yield the same uniform convergences or different uniform convergences. Here are some important choices of $\mathcal{S}$.

- If $\mathcal{S}$ contains just the singletons $\{x\}$ (for $x \in X$), then uniform convergence on elements of $\mathcal{S}$ is the same thing as pointwise convergence. Thus, the product topology on Y^X is a special case of uniform convergence topologies.

- When $\mathcal{S} = \{X\}$, then the uniform convergence topology is called simply the **topology of uniform convergence on** X.

- When X is a topological space, another important choice is $\mathcal{S} = \{\text{compact subsets of } X\}$, resulting in the **topology of uniform convergence on compact sets**.

Several other important choices of $\mathcal{S}$ will be introduced in 28.9.

When $\mathcal{S}$ is countable — or, more generally, when some countably subcollection of $\mathcal{S}$ covers the union of the members of $\mathcal{S}$ — then the gauge $\{\rho_{S,e'} \ : \ S \in \mathcal{S},\ e' \in E'\}$ can be replaced by a single pseudometric; some examples of this are given in 26.8 and 26.10.

18.27. *Proposition (optional).* A uniformity not only determines, but also is determined by, its uniform convergences. More precisely, let $\mathcal{U}$ and $\mathcal{V}$ be two uniformities on a set Y. Then the following conditions are equivalent:

(A) The uniformity $\mathcal{U}$ is stronger than the uniformity $\mathcal{V}$ — that is, $\mathcal{U} \supseteq \mathcal{V}$.

(B) For every set S, the topology of $\mathcal{U}$-uniform convergence on S is stronger than the topology of $\mathcal{V}$-uniform convergence on S. That is, if (φ_α) is a net in Y^S that converges $\mathcal{U}$-uniformly on S to a limit φ, then $\varphi_\alpha \to \varphi$ $\mathcal{V}$-uniformly also.

Proof. Clearly (A) $\Rightarrow$ (B). To show (B) $\Rightarrow$ (A), we shall take $S = Y \times Y$ — that is, we shall consider functions $\varphi : Y \times Y \to Y$. Let $\big((x_\alpha, x'_\alpha) : \alpha \in \mathbb{A}\big)$ be a net in $Y \times Y$ with eventuality filter equal to $\mathcal{U}$. (For instance, we could use the filter's canonical net; see 7.11.) Let D and E be gauges that determine the uniformities $\mathcal{U}$ and $\mathcal{V}$. Then $D(x_\alpha, x'_\alpha) \to 0$ in Y, and (in view of observations in 18.7) it suffices to show that $E(x_\alpha, x'_\alpha) \to 0$ in Y.

Denote $S = Y \times Y$. Define $\varphi(x, x') = x'$ for all $(x, x') \in S$. Define functions $\varphi_\alpha : S \to Y$ by

$$\varphi_\alpha(x, x') \quad = \quad \begin{cases} x_\alpha & \text{if } x = x_\alpha \text{ and } x' = x'_\alpha \\ x & \text{otherwise.} \end{cases}$$

For each $d \in D$ and $s \in S$ we have $d\big(\varphi_\alpha(s), \varphi(s)\big) \leq d(x_\alpha, x'_\alpha)$, which tends to 0 uniformly for all choices of $s \in S$; thus $\varphi_\alpha \to \varphi$ D-uniformly on S. By assumption (B), then, $\varphi_\alpha \to \varphi$ E-uniformly on S as well. Fix any $e \in E$. Then

$$e(x_\alpha, x'_\alpha) \quad = \quad e\Big(\varphi_\alpha(x_\alpha, x'_\alpha),\ \varphi(x_\alpha, x'_\alpha)\Big) \quad \leq \quad \sup_{s \in S} e\Big(\varphi_\alpha(s),\ \varphi(s)\Big),$$

which tends to 0 as α increases. This proves $E(x_\alpha, x'_\alpha) \to 0$.

EQUICONTINUITY

18.28. Let X be a topological space, let (Y, e) be a pseudometric space, and let $\varphi : X \to Y$ be some mapping. Then the **oscillation** of φ at a point $x_0 \in X$ with respect to the pseudometric e is defined to be the number

$$\operatorname{osc}_e(\varphi, x_0) \quad = \quad \inf_{N \in \mathcal{N}(x_0)} \operatorname{diam}_e\Big(\varphi(N)\Big) \quad = \quad \inf_{N \in \mathcal{N}(x_0)} \ \sup_{x, x' \in N} e\Big(\varphi(x), \varphi(x)\Big),$$

where $\mathcal{N}(x_0)$ is the neighborhood filter at x_0. (We may omit the subscript e if the choice of e is understood.) Thus, the oscillation is a number in $[0, +\infty]$. More generally, if Φ is a collection of functions from X into Y, the oscillation of Φ at x_0 is defined to be the number

$$\operatorname{osc}_e(\Phi, x_0) \;=\; \inf_{N \in \mathcal{N}(x_0)} \ \sup_{\varphi \in \Phi} \operatorname{diam}_e\Big(\varphi(N)\Big) \;=\; \inf_{N \in \mathcal{N}(x_0)} \ \sup_{\varphi \in \Phi} \ \sup_{x, x' \in N} e\Big(\varphi(x), \varphi(x)\Big).$$

We observe that

a. $\{x \in X \ : \ \mathrm{osc}_e(\Phi, x) < \varepsilon\}$ is an open set, and so $\mathrm{osc}_e(\Phi, \cdot) : X \to [0, +\infty]$ is an upper semicontinuous function.

b. For a single function φ, the number $\mathrm{osc}_e(\varphi, x_0)$ is 0 if and only if φ is continuous at x_0. Thus, $\mathrm{osc}_e(\varphi, x_0)$ may be taken as a numerical measurement of the size of the discontinuity of φ at x_0. We shall generalize this result to collections Φ of functions in 18.29(E), below.

18.29. *Definition.* Let X be a topological space, and let $x_0 \in X$. Let $(Y, \mathcal{V})$ be a uniform space, with uniformity determined by a gauge E. Let Φ be a collection of functions from X into Y. Then the following conditions are equivalent. If any (hence all) of them holds, we shall say that Φ is **equicontinuous** at the point x_0.

(A) Whenever $\Big((\varphi_\alpha, x_\alpha) \ : \ \alpha \in \mathbb{A}\Big)$ is a net in $\Phi \times X$ with $x_\alpha \to x_0$, then $E\Big(\varphi_\alpha(x_\alpha), \varphi_\alpha(x_0)\Big) \to 0$ (in the sense of 18.7(C)).

(B) Whenever $x_\alpha \to x_0$ in X, then $x_\alpha(\cdot) \to x_0(\cdot)$ uniformly on Φ, where we define $x_\alpha(\varphi) = \varphi(x_\alpha)$ and $x_0(\varphi) = \varphi(x_0)$.

(C) For each $V \in \mathcal{V}$, there is some neighborhood G of x_0 in X such that

$$x \in G, \quad \varphi \in \Phi \qquad \Rightarrow \qquad \Big(\varphi(x), \varphi(x_0)\Big) \ \in \ V.$$

(D) For each $e \in E$ and $\varepsilon > 0$, there is some neighborhood H of x_0 in X such that

$$x \in H, \quad \varphi \in \Phi \qquad \Rightarrow \qquad e\Big(\varphi(x), \varphi(x_0)\Big) \ \leq \ \varepsilon.$$

(E) The oscillation $\mathrm{osc}_e(\Phi, x_0)$ is equal to 0 for each $e \in E$. (Thus, the oscillation of a collection of functions may be taken as a numerical measurement of the extent to which the collection fails to be equicontinuous.)

A collection of mappings $\Phi : X \to Y$ is said to be **equicontinuous** if it is equicontinuous at every point of X.

18.30. *Some immediate observations about equicontinuity.*

a. If X is metrizable — or, more generally, if X is a first countable topological space — then nets can be replaced by sequences in conditions (A) and (B) above.

b. Let Y have uniformity given by gauge E. Then a collection Φ of functions is equicontinuous from the topological space X to the gauge space (Y, E) if and only if Φ is equicontinuous from X to the pseudometric space (Y, e) for each $e \in E$.

c. Any element of an equicontinuous family is continuous.

d. Any finite collection of continuous functions is equicontinuous.

e. The equicontinuous families form an ideal — i.e., any subset of an equicontinuous family is equicontinuous and the union of finitely many equicontinuous families is equicontinuous.

f. (*Preview.*) Let Φ be a collection of continuous linear maps from one normed space into another. Let $\|\varphi\|$ be the operator norm of φ. Then Φ is equicontinuous if and only if $\sup\{\|\varphi\| : \varphi \in \Phi\} < \infty$. See 23.1 and 23.12.

18.31. *Definition.* Let $(X, \mathcal{U})$ and $(Y, \mathcal{V})$ be uniform spaces, with uniformities determined by gauges D and E, respectively. Let Φ be a collection of functions from X into Y. Then the following conditions are equivalent. If any (hence all) of them holds, we shall say that Φ is **uniformly equicontinuous**:

(A) Whenever $\big((\varphi_\alpha, x_\alpha, x'_\alpha) : \alpha \in \mathbb{A}\big)$ is a net in $\Phi \times X \times X$ that satisfies $D(x_\alpha, x'_\alpha) \to 0$ (in the sense of 18.7(C)), then $E\big(\varphi_\alpha(x_\alpha), \varphi_\alpha(x'_\alpha)\big) \to 0$.

(B) For each $V \in \mathcal{V}$, there is some $U \in \mathcal{U}$ such that

$$(x, x') \in U, \quad \varphi \in \Phi \qquad \Rightarrow \qquad \big(\varphi(x), \varphi(x')\big) \in V.$$

(C) For each number $\varepsilon > 0$ and each pseudometric $e \in E$, there exists some number $\delta > 0$ and some finite set $D' \subseteq D$ such that

$$\max_{d \in D'} d(x, x') < \delta \qquad \Rightarrow \qquad \sup_{\varphi \in \Phi} e\left(\varphi(x), \varphi(x')\right) < \varepsilon.$$

(We emphasize that the choice of δ and D' depends on ε and e but not on x, x', φ.)

(D) For each $e \in E$, there exists a finite set $D_e \subseteq D$ and a function $\gamma_e : [0, +\infty) \to [0, +\infty)$ that is continuous and increasing, and satisfies $\gamma_e(0) = 0$ and

$$\sup_{\varphi \in \Phi} e\Big(\varphi(x), \varphi(x')\Big) \leq \gamma_e\Big(\max_{d \in D_e} d(x, x')\Big).$$

In other words, the φ's have a common modulus of uniform continuity.

Clearly, if Φ is uniformly equicontinuous, then Φ is equicontinuous and each member of Φ is uniformly continuous.

Further exercise. If X and Y are uniform spaces, X is compact, and $\Phi : X \to Y$ is equicontinuous, then Φ is uniformly equicontinuous.

18.32. *Convergence of continuous functions.* Suppose X is a topological space and Y is a uniform space, with uniformity $\mathcal{V}$ determined by gauge E. Let $(\varphi_\alpha : \alpha \in \mathbb{A})$ be a net of functions from X into Y, and let $\varphi \in Y^X$ also. Show that

a. If $\varphi_\alpha \to \varphi$ pointwise and the set $\{\varphi_\alpha : \alpha \in \mathbb{A}\}$ is equicontinuous at x_0, then φ is continuous at x_0. (*Hint*: Take limits in the inequality $\operatorname{osc}_e(\{\varphi_\alpha\}, x_0) \leq \varepsilon$ to obtain $\operatorname{osc}(\varphi, x_0) \leq \varepsilon$.)

b. If $\varphi_\alpha \to \varphi$ pointwise and the set $\{\varphi_\alpha\}$ is equicontinuous, then φ is continuous and $\varphi_\alpha \to \varphi$ uniformly on compact subsets of X.

c. If (φ_α) is equicontinuous, φ is continuous, and $\varphi_\alpha \to \varphi$ pointwise on a dense subset of X, then $\varphi_\alpha \to \varphi$ pointwise everywhere on X.

d. If $\varphi_\alpha \to \varphi$ uniformly on X and each φ_α is continuous, then the net (φ_α) is "asymptotically equicontinuous," in the sense that

$$\lim_{\alpha \in \mathbb{A}} \operatorname{osc}_e\Big(\{\varphi_\beta : \beta \succcurlyeq \alpha\},\ x_0\Big) \quad = \quad 0 \qquad \text{for each } x_0 \in X \text{ and } e \in E.$$

Using that fact (or any other means), show that φ is continuous.

e. If $(\varphi_n : n \in \mathbb{N})$ is a *sequence* of continuous functions converging uniformly to a limit φ, then φ is continuous and the set $\{\varphi_n : n \in \mathbb{N}\}$ is equicontinuous. *Hint*: Use the preceding result on asymptotic equicontinuity, plus 18.30.d.

f. If X is a uniform space, $\varphi_\alpha \to \varphi$ uniformly, and each φ_α is uniformly continuous, then φ is uniformly continuous.

g. *Classical example.* Let $X = Y = [0,1]$ with the usual metric, and let $\varphi_n(x) = x^n$. Then the functions φ_n are continuous on $[0,1]$; the sequence (φ_n) is equicontinuous on $[0,1)$ but not at $x = 1$. The sequence converges pointwise, but not uniformly, to the function

$$\varphi(x) \quad = \quad \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1. \end{cases}$$

This function is continuous on $[0,1)$ but discontinuous at $x = 1$.

18.33. *Equicontinuity and the product topology.* Let Λ be a topological space, let Y be a uniform space, and let Y^Λ be equipped with the product topology. Suppose that $\Phi \subseteq Y^\Lambda$ is equicontinuous. Then:

a. The closure of Φ in Y^Λ (in the product topology) is also equicontinuous. *Hint*: 18.29(D) or 18.29(E).

b. If Λ is separable and Y is metrizable, then the relative topology on Φ is metrizable.
Hints: Let C be a countable dense subset of Λ. Use 18.32.c to show that the restriction mapping $\rho : Y^\Lambda \to Y^C$ takes Φ homeomorphically onto $\rho(\Phi)$.

c. If Λ and Y are both separable and Y is metrizable, then Φ is also separable. *Hint*: Y^C is separable, by 15.27.

18.34. *Equicontinuity and uniform convergence topologies.* Let X be a topological space, let $\mathcal{S}$ be a collection of subsets of X, let Y be a uniform space, and let $Y^X = \{$functions from X into $Y\}$ be equipped with the topology of uniform convergence on elements of $\mathcal{S}$. Then the set $C(X,Y) = \{$continuous functions from X into $Y\}$ is a closed subset of Y^X, provided that either

(i) each $x \in X$ has a neighborhood that is a member of $\mathcal{S}$, or

(ii) X is metrizable (or more generally, first countable), and each convergent sequence in X is contained in some member of $\mathcal{S}$.

Proof. If (f_α) is a net in $C(X,Y)$ that converges uniformly on members of $\mathcal{S}$ to some $f \in Y^X$, then at least each restriction $f\big|_S$ will be continuous for $S \in \mathcal{S}$, by 18.32.d. Now the continuity of f on X follows under hypothesis (i) by 15.16, or under hypothesis (ii) by 15.34.d.

18.35. Arzela-Ascoli Theorem. Let X be a topological space, let Y be a uniform space, and let $C(X,Y) = \{\text{continuous functions from } X \text{ into } Y\}$ be given the topology of uniform convergence on compact subsets of X. Let $\Phi \subseteq C(X,Y)$. If

(1a) Φ is equicontinuous, and

(1b) the set $\Phi(x) = \{\varphi(x) : \varphi \in \Phi\}$ is relatively compact in Y for each $x \in X$

then

(2) Φ is relatively compact in $C(X,Y)$.

We also have this partial converse: Assume X is locally compact or first countable; then (2) $\Rightarrow$ (1).

Remarks. Since the topologies on Y and $C(X,Y)$ are completely regular, the characterizations of "relatively compact" given in 17.15 are applicable.

Proof of (1) $\Rightarrow$ (2). Let $(\varphi_\alpha : \alpha \in A)$ be any universal net in Φ; it suffices to show (φ_α) converges to a limit in $C(X,Y)$. For each fixed x, observe that $(\varphi_\alpha(x) : \alpha \in A)$ is a universal net in Y. Hence (φ_α) converges pointwise to a limit φ. Use the equicontinuity of Φ to show that the limit function is continuous and that the convergence is uniform on compact sets.

Proof of (2) $\Rightarrow$ (1b). Fix any $x \in X$. Then the singleton $\{x\}$ is compact. Any net in $\Phi(x)$ can be represented in the form $(\varphi_\alpha(x) : \alpha \in \mathbb{A})$ for some net $(\varphi_\alpha : \alpha \in \mathbb{A})$ in Φ. That net has a subnet converging uniformly on compact sets, hence converging on x.

Proof of (2) $\Rightarrow$ (1a). Let E be a gauge that determines the uniform structure of Y. Consider any net $((\varphi_\alpha, x_\alpha) : \alpha \in \mathbb{A})$ in $\Phi \times X$ with $x_\alpha \to x_0$; in the case where Y is first countable, we may assume this net is a sequence. We are to show that $E\left(\varphi_\alpha(x_\alpha), \varphi_\alpha(x_0)\right) \to 0$ (in the sense of 18.7).

Suppose, on the contrary, that $e\left(\varphi_\alpha(x_\alpha), \varphi_\alpha(x_0)\right) \not\to 0$ for some $e \in E$. Replacing the given net with a subnet (or replacing the given sequence with a subsequence), we may assume that

$$e\left(\varphi_\alpha(x_\alpha), \varphi_\alpha(x_0)\right) \geq \rho \qquad \text{for some constant } \rho > 0. \tag{$*$}$$

We may assume that all the x_α's and x_0 are contained in some compact set — either because x_0 has a compact neighborhood K that contains all the x_α's for α sufficiently large, or because (x_α) is a sequence converging to x_0 (see 17.7.e). Now use the assumption (2). Again passing to a subnet (at this point we no longer need a subsequence), we may assume that (φ_α) converges to some limit φ in $C(X,Y)$ uniformly on compact sets and thus uniformly on K. Hence

$$e\left(\varphi_\alpha(x_\alpha), \varphi(x_\alpha)\right) \to 0 \qquad \text{and} \qquad e\left(\varphi_\alpha(x_0), \varphi(x_0)\right) \to 0.$$

Also $e\left(\varphi(x_\alpha), \varphi(x_0)\right) \to 0$ since φ itself is continuous. These results contradict $(*)$.

18.36. Corollary (Arzela-Ascoli Theorem, classical version). Let

$$C[0,1] \quad = \quad \{\text{continuous functions from } [0,1] \text{ into } \mathbb{R}\}$$

be given the topology of uniform convergence on $[0,1]$ — i.e., the topology determined by the metric $e(f,g) = \max\{|f(t) - g(t)| : t \in [0,1]\}$. Let $\Phi \subseteq C[0,1]$. Then Φ is relatively compact in $C[0,1]$ if and only if Φ is uniformly bounded (i.e., $\sup_{t\in[0,1]} \sup_{f\in\Phi} |f(t)| < \infty$) and equicontinuous.

Chapter 19

Metric and Uniform Completeness

19.1. *Introductory remarks.* Most applications of Cauchy completeness are in metric spaces, but a more general setting is occasionally useful; we shall develop the concept in the setting of uniform spaces. (A still more general setting is possible; see the remarks in 19.29.)

Some important variants of completeness will be introduced later: Baire spaces, and barrelled and ultrabarrelled spaces.

Many of the ideas in Chapter 18 were based on the category in which the objects were uniform spaces and the morphisms were the uniformly continuous maps. Some of the ideas in the present chapter arise more naturally in the category whose objects are uniform spaces and whose morphisms are *Cauchy continuous maps*, introduced in 19.24.

Cauchy Filters, Nets, and Sequences

19.2. *Definitions.* Let $(X, \mathfrak{U})$ be a uniform space, and let D be any gauge that determines the uniformity $\mathfrak{U}$. Let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in X, and let $\mathcal{F}$ be its eventuality filter. Then the following conditions are equivalent; if any (hence all) of them are satisfied, we say that (x_α) and $\mathcal{F}$ are **Cauchy**.

(A) For each $d \in D$ and each number $\varepsilon > 0$, there is some $F \in \mathcal{F}$ such that $\operatorname{diam}_d(F) \leq \varepsilon$.

(B) For each $d \in D$ and each number $\varepsilon > 0$, there is some $\gamma = \gamma_{d,\varepsilon} \in \mathbb{A}$ such that $\alpha, \beta \succcurlyeq \gamma \;\Rightarrow\; d(x_\alpha, x_\beta) \leq \varepsilon$.

(C) For each $U \in \mathfrak{U}$, there is some $F \in \mathcal{F}$ such that $F \times F \subseteq U$.

(D) (*A two-sided Cauchy condition.*) For each entourage $U \in \mathfrak{U}$, we have eventually $(x_\alpha, x_\beta) \in U$. That is, for each $U \in \mathfrak{U}$, there exists some $\gamma_U \in \mathbb{A}$ such that $\alpha, \beta \succcurlyeq \gamma_U \;\Rightarrow\; (x_\alpha, x_\beta) \in U$.

(E) (*A one-sided Cauchy condition.*) For each entourage $V \in \mathfrak{U}$, there exists some γ in $\mathbb{A}$ with the property that $\alpha \succcurlyeq \gamma \;\Rightarrow\; x_\alpha \in V[x_\gamma]$. (Recall from 5.33 that $V[x] = \{y : (x, y) \in V\}$.)

Proofs. The proofs of equivalence of all the conditions but the last one should be fairly straightforward. To prove 19.2(D) $\Rightarrow$ 19.2(E), take $U = V^{-1}$ and then take $\beta = \xi$. To prove 19.2(E) $\Rightarrow$ 19.2(D), choose V so that $V \circ V^{-1} \subseteq U$.

Further observations. The conditions above are unaffected if we replace the gauge D with any uniformly equivalent gauge. They may also be unaffected in some cases if we replace the gauge with one that is not uniformly equivalent; see 19.25(ii).

Observe that (x_α) and $\mathcal{F}$ are Cauchy for the gauge D if and only if they are Cauchy for each pseudometric $d \in D$.

19.3. Sequences are just a special case of nets, but they are a case important enough to deserve separate mention. Let $(X, \mathcal{U})$ be a uniform space; let D be a gauge that yields the uniformity $\mathcal{U}$. Then a sequence (x_n) in X is Cauchy if

> for each $U \in \mathcal{U}$, there exists some positive integer N_U such that $m, n \geq N_U \Rightarrow (x_m, x_n) \in U$

or, equivalently,

> for each $d \in D$ and each number $\varepsilon > 0$, there is some positive integer $N = N_{d,\varepsilon}$ such that $m, n \geq N \Rightarrow d(x_m, x_n) \leq \varepsilon$.

19.4. *Elementary properties of Cauchy nets.* Let X be a uniform space, with uniformity $\mathcal{U}$ given by gauge D. Show that

- **a.** Any convergent net is Cauchy; any convergent filter is Cauchy.
- **b.** Any subnet of a Cauchy net is Cauchy; any superfilter of a Cauchy filter is Cauchy.
- **c.** Suppose (x_α) is a Cauchy net. Then any subnet of (x_α) has the same set of limits as (x_α). Hence, if some subnet of (x_α) converges, then so does (x_α).
- **d.** If (x_α) is a universal net and some subnet of (x_α) is Cauchy, then (x_α) itself is Cauchy.
- **e.** Let X be a set equipped with the initial uniformity determined by some collection of mappings into uniform spaces, $\pi_\lambda : X \to X_\lambda$ $(\lambda \in \Lambda)$, as defined in 9.15 and 9.16. Then a net $(x_\alpha : \alpha \in \mathbb{A})$ is Cauchy in X if and only if the net $(\pi_\lambda(x_\alpha) : \alpha \in \mathbb{A})$ is Cauchy in X_λ for each $\lambda \in \Lambda$.

  In particular, a net $(f_\alpha : \alpha \in \mathbb{A})$ in a product $\prod_{\lambda \in \Lambda} Y_\lambda$ of uniform spaces is Cauchy if and only if it is Cauchy coordinatewise — i.e., if and only if the net $(f_\alpha(\lambda) : \alpha \in \mathbb{A})$ is Cauchy for each $\lambda \in \Lambda$.

19.5. *Proposition.* If (x_α) and (y_β) are Cauchy nets that have a common subnet, then they also have a common supernet that is Cauchy — i.e., then (x_α) and (y_β) are subnets of a Cauchy net.

This can be reformulated in terms of filters (see 7.18(E)): Let $\mathcal{F}$ and $\mathcal{G}$ be Cauchy filters on a uniform space X. Suppose that every member of $\mathcal{F}$ meets every member of $\mathcal{G}$. Then $\mathcal{F} \cap \mathcal{G}$ is Cauchy.

Proof. The proof is easier in terms of filters. Let $\mathcal{U}$ be the uniformity on X. Let any $U \in \mathcal{U}$ be given; we wish to find some $H \in \mathcal{F} \cap \mathcal{G}$ satisfying $H \times H \subseteq U$. By 5.35.c, choose some symmetric entourage V such that $V \circ V \subseteq U$. Since $\mathcal{F}$ and $\mathcal{G}$ are Cauchy, there exist $F \in \mathcal{F}$ and $G \in \mathcal{G}$ such that $F \times F \subseteq V$ and $G \times G \subseteq V$. Now let $H = F \cup G$; show that $H \times H \subseteq U$.

19.6. *Observation.* A sequence $(x_1, x_2, x_3, \ldots)$ converges to a point x if and only if the sequence $(x_1, x, x_2, x, x_3, x, \ldots)$ is Cauchy.

We shall now prove an analogous result for nets.

Proposition. Let (x_α) be a net in X, and let $z \in X$. Then $x_\alpha \to z$ if and only if (x_α) and the constant net (z) are subnets of some Cauchy net.

This can be reformulated more canonically, using filters:

Let $\mathcal{F}$ be a filter on X, and let $\mathcal{C}$ be the ultrafilter fixed at some point z. Then $\mathcal{F} \to z$ if and only if the filter $\mathcal{F} \cap \mathcal{C}$ is Cauchy.

Proof. The proof is easier in terms of filters. Let $\mathcal{N}(z)$ be the neighborhood filter at z.

First, suppose $\mathcal{F} \to z$. Then $\mathcal{F} \supseteq \mathcal{N}(z)$. Hence $\mathcal{F} \cap \mathcal{C} \supseteq \mathcal{N}(z)$. The filter $\mathcal{N}(z)$ is convergent to z, hence the filter $\mathcal{F} \cap \mathcal{C}$ is convergent to z, hence it is Cauchy.

Conversely, suppose $\mathcal{F} \cap \mathcal{C}$ is Cauchy, and let any $N \in \mathcal{N}(z)$ be given; we shall show that $N \in \mathcal{F}$. By the construction of the uniform topology, we know that $N \supseteq U[z]$ for some entourage U in the uniformity $\mathcal{U}$. Since $\mathcal{F} \cap \mathcal{C}(z)$ is Cauchy, there is some set $K \in \mathcal{F} \cap \mathcal{C}$ satisfying $K \times K \subseteq U$. Since $K \in \mathcal{C}$, we have $z \in K$. Therefore

$$w \in K \quad \Rightarrow \quad (z, w) \in K \times K \subseteq U \quad \Rightarrow \quad w \in U[z] \subseteq N.$$

Thus $K \subseteq N$. Since $K \in \mathcal{F}$, it follows that $N \in \mathcal{F}$.

19.7. *Proposition.* Let $(x_\alpha : \alpha \in \mathbb{A})$ be a net in a uniform space $(X, \mathcal{U})$. Suppose that

(i) for each increasing sequence $\beta(1) \preccurlyeq \beta(2) \preccurlyeq \beta(3) \preccurlyeq \cdots$ in $\mathbb{A}$, the sequence $(x_{\beta(k)} : k \in \mathbb{N})$ is Cauchy.

Then

(ii) the net $(x_\alpha : \alpha \in \mathbb{A})$ is Cauchy.

(We shall use this result in 29.25.) The converse is false — i.e., in general, condition (ii) does not imply condition (i).

Proof of (i) $\Rightarrow$ (ii). We shall show that condition 19.2(E) is satisfied. Indeed, suppose not. Then there is some entourage $V \in \mathcal{U}$ for which there is no corresponding $\xi \in \mathbb{A}$. Hence for each $\xi \in \mathbb{A}$ there exists some $\alpha \succcurlyeq \xi$ such that $x_\alpha \notin V[x_\xi]$. In this fashion we recursively construct an increasing sequence $\alpha(1), \alpha(2), \alpha(3), \ldots$ satisfying $(x_{\alpha(k)}, x_{\alpha(k+1)}) \notin V$. Such a sequence is not Cauchy.

Counterexample to (ii) $\Rightarrow$ (i). Let $X = \mathbb{R}$, with its usual metric. Let $\mathbb{A} = \{(r, n) \in \mathbb{R} \times \mathbb{N} : |r| \le \frac{1}{n}\}$. Let $\mathbb{A}$ be ordered as follows: $(r_1, n_1) \preccurlyeq (r_2, n_2)$ if $n_1 \le n_2$ (regardless of the values of r_1 and r_2). It is easy to see that $\mathbb{A}$ is a directed set. Let $x_{(r,n)} = r$; then it is clear that the net $(x_\alpha : \alpha \in \mathbb{A})$ converges to 0 and thus is Cauchy. However, consider the increasing sequence $(\alpha(k) : k \in \mathbb{N})$ defined by $\alpha(k) = \big((-1)^k, 1\big)$; the sequence $(x_{\alpha(k)}) = (-1, 1, -1, 1, \ldots)$ is not Cauchy.

19.8. *More about Cauchyness in metric spaces.* Let (X, d) be a pseudometric space.

a. Any Cauchy sequence in X is metrically bounded. Any Cauchy net in X is eventually metrically bounded.

b. Any Cauchy sequence (x_n) satisfies $d(x_n, x_{n+1}) \to 0$.

The converse is false. For instance, in $\mathbb{R}$ with its usual metric, the sequence $x_n = \sqrt{n}$ is not Cauchy but satisfies $d(x_n, x_{n+1}) \to 0$.

(*Optional.*) If (X, d) is an ultrametric space and $d(x_n, x_{n+1}) \to 0$, then (x_n) is Cauchy.

c. (Summation property.) If a sequence (y_k) satisfies $\sum_{k=1}^{\infty} d(y_k, y_{k+1}) < \infty$, then (y_k) is Cauchy.

Not every Cauchy sequence has that summation property, but every Cauchy sequence (x_n) has a subsequence (y_k) with that property. In fact, we may choose (y_k) so that $d(y_k, y_{k+1}) < 2^{-k}$ for all k.

Complete Metrics and Uniformities

19.9. *Definitions.* A uniform space $(X, \mathfrak{U})$ is **complete** if each Cauchy net or filter in X has at least one limit. (This condition, which only holds in some uniform spaces, is a converse to 19.4.a, which holds in *any* uniform space.)

It will be helpful to extend this terminology to subsets of X as well. We say that a set $S \subseteq X$ is **complete** (or D is complete on S) if every Cauchy net or filter in S has at least one limit in S.

19.10. *Basic properties of completeness.* Let $(X, \mathfrak{U})$ be a uniform space.

a. Let $S \subseteq X$. Then S itself is a uniform space, when equipped with the trace of $\mathfrak{U}$ (see 9.20). Show that S is complete, when viewed as a subset of X, if and only if S is complete when viewed as a subset of itself.

b. Any closed subset of a complete uniform space is complete.

c. If X is a complete Hausdorff uniform space and $S \subseteq X$, then S is complete if and only if S is closed. (Contrast this with 20.12.)

d. We say that a set $S \subseteq X$ is **sequentially complete** if every Cauchy *sequence* in S converges to at least one limit in S. Observe that any complete uniform space is sequentially complete.

Caution: Some mathematicians who are concerned only with sequences omit the term "sequentially" here. However, completeness and sequential completeness are not equivalent in general.

19.11. *Examples and further properties in pseudometric spaces.* Let (X, d) be a pseudometric space. Then:

a. If X is sequentially complete, then it is complete.

b. **(Summability property.)** (X, d) is complete if and only if it has this property: Whenever (x_n) is a sequence satisfying $\sum_{n=1}^{\infty} d(x_n, x_{n+1}) < \infty$, then (x_n) converges.

c. **(Cauchy's Intersection property.)** (X, d) is complete if and only if it has the following property: If $S_1 \supseteq S_2 \supseteq S_3 \supseteq \cdots$ and the S_n's are nonempty closed subsets of X with $\lim_{n\to\infty} \operatorname{diam}(S_n) = 0$, then $\bigcap_{n=1}^{\infty} S_n$ is nonempty. (The proof of this fact uses Countable Choice (CC).)

d. On any set, the "Kronecker metric" $d(x, y) = 1 - \delta_{xy}$ (defined in 2.12.b) is complete. However, see the contrasting result below:

e. The sets $\mathbb{R}$ and $\mathbb{Z}$ are complete when equipped with the usual metric $d(x, y) = |x - y|$, but not with $d(x, y) = |\arctan(x) - \arctan(y)|$. (*Hint*: 17.9.b and 19.8.a.)

Note that these metrics are topologically equivalent but not uniformly equivalent — i.e., they yield the same topology on $\mathbb{R}$ or on $\mathbb{Z}$, but not the same uniform structure. The topology they yield on $\mathbb{Z}$ is the discrete topology. Thus, not every discrete metric is complete (although "the" discrete metric is complete, as noted in 19.11.d).

f. Use the completeness of $\mathbb{R}$ to show that the metric space $B(\Lambda)$ introduced in 4.41.f is complete. The embedding in 4.41.f shows that *every metric space may be viewed as a subset of a complete metric space.* See also 22.14.

g. As we noted in 18.24, all metrics that yield the usual topology on $[-\infty, +\infty]$ are uniformly equivalent to one another — i.e., they yield the same uniformity. That uniformity is complete.

h. Any knob space (defined as in 5.34.c) is complete.

i. *Technical exercise.* Let ρ and d be metrics on a set X. Suppose that

(i) ρ is complete,

(ii) ρ is topologically stronger than d, and

(iii) every d-Cauchy sequence has a subsequence that is ρ-Cauchy.

Then ρ and d are topologically equivalent and d is complete. (This result will be used in 19.47.)

19.12. Completeness of uniform convergence. If Y is a complete metric space and X is any set, then Y^X is complete when equipped with the uniformity of uniform convergence on X.

More generally, if Y is a complete uniform space and X is any set, then Y^X is complete when equipped with the uniformity of uniform convergence on members of $\mathcal{S}$ (described in 18.26) for any collection $\mathcal{S} \subseteq \mathcal{P}(X)$.

Proof. Let E be a bounded gauge that determines the uniformity of Y. Let Y^X be equipped with the gauge $\{\rho_{S,e} \ : \ S \in \mathcal{S},\ e \in E\}$, with notation as in 18.25 and 18.26. Let $(f_\alpha : \alpha \in \mathbb{A})$ be a Cauchy net in Y^X; we must show that (f_α) converges to a limit in Y^X.

Let $T = \bigcup_{S\in\mathcal{S}} S$. We first show that (f_α) converges pointwise on T. Fix any $t \in T$. Then $t \in S$ for some $S \in \mathcal{S}$. Fix any $e \in E$. Since (f_α) is $\rho_{S,e}$-Cauchy, the net $(f_\alpha(t))$ is e-Cauchy in Y. This applies for each $e \in E$, so the net $(f_\alpha(t))$ is E-Cauchy in Y. Since Y

is assumed complete, there is some limit to which $(f_\alpha(t))$ converges. (There may be more than one, if the gauge space (Y, E) is not Hausdorff.) Let $f(t)$ be any limit of $(f_\alpha(t))$.

Now fix any $S \in \mathcal{S}$, $e \in E$, and $\varepsilon > 0$. Since (f_α) is Cauchy, there is some $\gamma_0 \in \mathbb{A}$ such that $\alpha, \beta \succcurlyeq \gamma_0 \Rightarrow \rho_{S,e}(f_\alpha, f_\beta) \leq \varepsilon$ — i.e., such that

$$\alpha, \beta \succcurlyeq \gamma_0, \quad t \in S \qquad \Rightarrow \qquad e\left(f_\alpha(t), f_\beta(t)\right) \leq \varepsilon.$$

Hold α fixed and let β increase, and take limits. This shows that

$$\alpha \succcurlyeq \gamma_0, \quad t \in S \qquad \Rightarrow \qquad e\left(f_\alpha(t), f(t)\right) \leq \varepsilon.$$

It follows that $f_\alpha \to f$ in the topology of uniform convergence on members of $\mathcal{S}$.

Remarks. In most cases of interest, Y is Hausdorff and $\bigcup_{S \in \mathcal{S}} S = X$. Then the complete uniform structure on Y^X is also Hausdorff, and a set $\Phi \subseteq Y^X$ is complete if and only if it is closed. Thus it becomes important to know which subsets of Y^X are closed; see 18.34.

19.13. Completeness of pointwise convergence. Any product of complete spaces is complete, when equipped with the product uniformity.

In some cases this is easy to verify and does not require the Axiom of Choice. Indeed, the product of finitely many or countably many metric spaces is metrizable, with metrics given as in 18.16 and 18.17. In particular, $\mathbb{R}^n$ is complete, and so is any closed subset of $\mathbb{R}^n$. Since $\mathbb{C}$ is isomorphic to $\mathbb{R}^2$ as a uniform space, $\mathbb{C}$ is also complete.

For arbitrary products, however, the Axiom of Choice is needed. We shall now show that AC (introduced in 6.12, 6.20, 6.22) is equivalent to the two principles below. Recall the definition of "knob space," in 5.34.c.

> **(AC26) Product of Complete Spaces.** For each λ in some set Λ, let Y_λ be a complete uniform space. Then the product $X = \prod_{\lambda \in \Lambda} Y_\lambda$, equipped with the product uniform structure, is also complete.

> **(AC27) Product of Knob Spaces.** Any product of knob spaces, when equipped with the product uniform structure, is complete.

Proof of (AC3) $\Rightarrow$ (AC26). Let $\mathcal{F}$ be a Cauchy filter on X; we wish to show that $\mathcal{F}$ has at least one limit. For each λ, the filterbase $\pi_\lambda(\mathcal{F})$ is Cauchy in Y_λ. Since Y_λ is complete, the set S_λ of limits of $\pi_\lambda(\mathcal{F})$ is nonempty. By the Axiom of Choice (AC3), $R = \prod_{\lambda \in \Lambda} S_\lambda$ is nonempty; then any element of R is a limit of $\mathcal{F}$.

Remark. It should be noted that *if all the Y_λ's are Hausdorff, then the Axiom of Choice is not needed,* since each S_λ is a singleton. In this special case, the argument above establishes the statement (AC26) using just ZF — i.e., set theory without the Axiom of Choice. In particular, AC is not needed to prove that 2^Λ is complete, where $2 = \{0, 1\}$ has the discrete uniform structure.

Proof of (AC26) $\Rightarrow$ (AC27). As we noted in 19.11.h, every knob space is complete.

Proof of (AC27) $\Rightarrow$ (AC3). Let Λ, S_λ, ξ_λ, Y_λ, X, ξ be as in 6.24, and equip X with topology and uniform structure as the product of knob spaces. For each $\lambda \in \Lambda$, the filterbase $\pi_\lambda(\mathcal{F})$

is Cauchy on Y_λ, since it includes the set $\pi_\lambda(T_{\{\lambda\}}) = S_\lambda$. Since S_λ is closed in Y_λ, any limit of $\pi_\lambda(\mathcal{F})$ must lie in S_λ. Since each $\pi_\lambda(\mathcal{F})$ is Cauchy, the filterbase $\mathcal{F}$ is Cauchy on X. By hypothesis, X is complete, so $\mathcal{F}$ has at least one limit ζ in X. We have $\pi_\lambda(\zeta) \in S_\lambda$ for each λ; thus $\zeta \in \prod_{\lambda \in \Lambda} S_\lambda$.

TOTAL BOUNDEDNESS AND PRECOMPACTNESS

19.14. *Definition.* Let $(X, \mathcal{U})$ be a uniform space, and let D be any gauge that determines the uniformity $\mathcal{U}$. A set $S \subseteq X$ is **totally bounded** if

> for each $U \in \mathcal{U}$, there is some finite set $F \subseteq X$ such that $S \subseteq \bigcup_{x \in F} U[x]$

or, equivalently, if

> for each number $\varepsilon > 0$ and each $d \in D$, there exists some finite set $F \subseteq X$ such that $S \subseteq \bigcup_{x \in F} B_d(x, \varepsilon)$.

The proof of equivalence of these two definitions is left as an exercise. A substantially different characterization of totally bounded sets will be given in 19.17.

19.15. *Basic properties of total boundedness.*

a. We obtain an equivalent definition if we replace "finite set $F \subseteq X$" with "finite set $F \subseteq S$" in either of the conditions above.

b. In a pseudometric space, the definition above simplifies slightly. A set is totally bounded if and only if, for each $\varepsilon > 0$, the set can be covered by finitely many balls of radius ε.

c. If X is a uniform space and $S \subseteq X$, then S is also a uniform space (see 9.20). Show that S is totally bounded, as a subset of X, if and only if S is totally bounded as a subset of itself.

d. Let D be a gauge that determines the uniformity $\mathcal{U}$. Then a set $S \subseteq X$ is totally bounded in the uniform space $(X, \mathcal{U}) = (X, D)$ if and only if S is totally bounded in each of the pseudometric spaces (X, d) for $d \in D$. (Hence many questions about total boundedness of uniform spaces can be reduced to questions about total boundedness of pseudometric spaces.)

e. Any finite subset of a uniform space is totally bounded.

f. The totally bounded subsets of a uniform space form an ideal. That is: any subset of a totally bounded set is totally bounded, and the union of finitely many totally bounded sets is totally bounded.

g. If $\{Y_\lambda : \lambda \in \Lambda\}$ is a collection of totally bounded uniform spaces, and X is equipped with an initial uniformity determined by the Y_λ's, then X is also totally bounded.

In particular, any product of totally bounded uniform spaces is totally bounded, when equipped with the product uniform structure.

More particularly, 2^Λ is totally bounded for any set Λ. This argument does not require the use of the Axiom of Choice or any weakened form of Choice; we shall use that observation in a proof in 19.17.

h. Let $S \subseteq X$. If S is totally bounded, then $\mathrm{cl}(S)$ is totally bounded. *Hint*:

$$\mathrm{cl}\left(\bigcup_{x\in F} B_d(x,\varepsilon)\right) = \bigcup_{x\in F} \mathrm{cl}\,(B_d(x,\varepsilon)) \subseteq \bigcup_{x\in F} K_d(x,\varepsilon) \subseteq \bigcup_{x\in F} B_d(x,2\varepsilon),$$

where K_d is the closed ball defined as in 5.15.g.

i. If X is totally bounded, then any universal net (or any ultrafilter) in X is Cauchy. *Hint*: 5.8(E) and/or 7.25.d.

j. Any totally bounded pseudometric space is separable. *Hint*: Use a sequence of ε's that decreases to 0.

k. Let X be a uniform space, and let S be a subset with the property that every sequence in S has a cluster point in X. Then S is totally bounded.

Hints: Suppose not. Then there is some number $\varepsilon > 0$ and some pseudometric d in a determining gauge, such that S cannot be covered by finitely many d-balls of radius ε. Hence we can recursively choose a sequence (x_n) in S such that $x_n \notin \bigcup_{j=1}^{n-1} B_d(x_j,\varepsilon)$. Let z be a cluster point of that sequence. Show that $B_d(z,\varepsilon/2)$ contains infinitely many of the x_n's and hence it contains at least two of them, a contradiction.

19.16. Let X be a uniform space. In this book we shall say that X is **precompact** if every proper filter on X has a Cauchy superfilter — or, equivalently, if every net in X has a Cauchy subnet.

It is easy to see that

a uniform space is compact if and only if it is complete and precompact.

This result does *not* require the Axiom of Choice or any weak form of the Axiom of Choice, unlike the results below.

19.17. The Ultrafilter Principle, introduced in 6.32, is equivalent to the following principles:

(UF24) Let X be equipped with the uniform structure given by a gauge. Then X is precompact if and only if X is totally bounded.

(UF25) Let X be equipped with the uniform structure given by a gauge. Then X is compact if and only if X is complete and totally bounded.

Remark. Most mathematicians use the terms "precompact" and "totally bounded" interchangeably. That is not surprising, since most mathematicians view the Axiom of Choice as "true" and therefore view (UF24) as "true." In this book we have distinguished between "precompact" and "totally bounded" precisely to see the role of the Ultrafilter Principle as a weak form of the Axiom of Choice.

The equivalence of (UF25) with other forms of UF was first announced by Rubin and Scott [1954].

Proof of (UF1) $\Rightarrow$ (UF24). If X is totally bounded, then X is precompact; this follows immediately from (UF1) and 19.15.i.

Conversely, we shall show that precompact implies totally bounded. This part of the proof does not require UF; it can be proved in ZF. Assume X is precompact but not totally bounded; we shall obtain a contradiction. Since X is not totally bounded, there is some number $\varepsilon > 0$ and some pseudometric $d \in D$ such that X cannot be covered by finitely many balls $B_d(x, \varepsilon)$, with centers $x \in X$. Let

$$\mathcal{I} = \{A \subseteq X : A \text{ can be covered by finitely many open balls of radius } \varepsilon\}.$$

Then $\mathcal{I}$ is a proper ideal on X. It is dual to the proper filter $\mathcal{F} = \{X \setminus A : A \in \mathcal{I}\}$.

By assumption, $\mathcal{F}$ has a Cauchy superfilter $\mathcal{G}$. Since $\mathcal{G}$ is Cauchy, it has some member G with diameter less than ε. Then G can be covered by an open ball of radius ε, and so $G \in \mathcal{I}$. This leads to a contradiction.

Proof of (UF24) $\Rightarrow$ (UF25). Immediate from 19.16.

Proof of (UF25) $\Rightarrow$ (UF21). By 19.15.g and the remark in 19.13, we know that 2^X is totally bounded and complete. Hence 2^X is compact. This proves (UF21), which was presented in 17.22.

19.18. An important special case of (UF25) is: *A pseudometric space is compact if and only if it is complete and totally bounded.* This can be proved without using UF; we omit the proof.

19.19. *Definition and exercises.* Let (X, d) be a metric space, and let S be a metrically bounded subset of X. Then we define **Kuratowski's measure of noncompactness**

$$\alpha(S) = \inf\Big\{r : S \text{ can be covered by finitely many sets with diameter } \leq r\Big\}$$

and **Hausdorff's measure of noncompactness**

$$\beta(S) = \inf\Big\{r : S \text{ can be covered by finitely many balls with radius } \leq r\Big\}.$$

Show that

- **a.** The two measures are "equivalent," in this sense: $\beta(S) \leq \alpha(S) \leq 2\beta(S)$. Thus one measure is small if and only if the other is small.
- **b.** $\alpha(S)$ and $\beta(S)$ are zero if and only if S is totally bounded. Thus, in a complete metric space, $\alpha(S)$ and $\beta(S)$ are zero if and only if S is relatively compact.
- **c.** $\beta(S)$ is the distance from $\mathrm{cl}(S)$ to the nearest compact set, in the Hausdorff metric on the space of closed, metrically bounded sets (see 5.18.d).

19.20. Niemytzki-Tychonov Theorem. Let $(X, \mathcal{T})$ be a topological space; assume the topology $\mathcal{T}$ is pseudometrizable. Then $(X, \mathcal{T})$ is compact if and only if every pseudometric yielding the topology $\mathcal{T}$ is complete.

Proof. This proof is taken from Engelking [1977]. The "only if" part follows from 19.18. For the "if" part, assume that $(X, \mathfrak{T})$ is not compact; we shall construct an incomplete pseudometric for $\mathfrak{T}$.

By 17.33, X is not countably compact. By 17.26.b, there exist nonempty closed sets $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$ such that the intersection $\bigcap_{i=1}^{\infty} F_i$ is empty. By 18.14, there is some pseudometric σ on X that yields the given topology and is bounded by 1. Now for $i = 1, 2, 3, \ldots$, define

$$\rho_i(x,y) = |\delta_i(x) - \delta_i(y)| + \sigma(x,y)\min\{\delta_i(x), \delta_i(y)\} \qquad \text{where } \delta_i(x) = \operatorname{dist}_\sigma(x, F_i).$$

Verify that ρ_i is a pseudometric on X. Let $\rho(x,y) = \sum_{i=1}^{\infty} 2^{-i}\rho_i(x,y)$; verify that ρ is a pseudometric on X. Show that ρ and σ have the same convergent sequences; thus ρ yields the topology $\mathfrak{T}$. Also show that $\operatorname{diam}_\rho(F_i) \leq 2^{-i}$; by 19.11.c we conclude that ρ is not complete.

Bounded Variation

19.21. Let (X, d) be a metric space, and let $\varphi : [a,b] \to X$ be some function. The **variation** of φ is the number

$$\operatorname{Var}(\varphi, [a,b]) = \sup\left\{\sum_{j=1}^{n} d\Big(\varphi(t_{j-1}),\, \varphi(t_j)\Big) : a = t_0 < t_1 < \cdots < t_n = b\right\}.$$

Here the supremum is taken over all partitions of $[a,b]$ into finitely many subintervals. (For clarification, we may refer to this as the **variation in the sense of intervals**, or **variation in the classical sense**; another meaning of "variation" is given in 29.5. In 29.34 we discuss the relation between the two notions.) The function φ has **bounded variation** on $[a,b]$ if $\operatorname{Var}(\varphi, [a,b]) < \infty$. Some elementary but important properties are noted below.

- **a.** $\operatorname{Var}(\varphi, [a,q]) + \operatorname{Var}(\varphi, [q,b]) = \operatorname{Var}(\varphi, [a,b])$ for $a < q < b$.
- **b.** $\operatorname{Var}(\varphi, [a,b]) = 0$ if and only if φ is constant on $[a,b]$.
- **c.** If $\sigma : [\widehat{a}, \widehat{b}] \to [a,b]$ is an increasing function or a decreasing function, then $\operatorname{Var}(\varphi \circ \sigma, [\widehat{a}, \widehat{b}]) = \operatorname{Var}(\varphi, [a,b])$.
- **d.** If $\varphi : [a,b] \to X$ has bounded variation and $\gamma : X \to Y$ is Lipschitzian (where X and Y are metric spaces), then $\gamma \circ \varphi : [a,b] \to Y$ has bounded variation. In fact, $\operatorname{Var}(\gamma \circ \varphi, [a,b]) \leq \langle\gamma\rangle_{\mathrm{Lip}} \operatorname{Var}(\varphi, [a,b])$.

  Remark. Although this result is easy to prove, a harder proof will not yield a stronger result; that is evident from the converse given in 19.23.
- **e.** *Corollary.* If $\psi : [a,b] \to Y$ is Lipschitzian, then ψ has bounded variation, with $\operatorname{Var}(\psi, [a,b]) \leq (b-a)\langle f\rangle_{\mathrm{Lip}}$.
- **f.** *Example.* Show that the function $f(t) = t\cos\left(\frac{\pi}{t^2}\right)$ is continuous on $[0,1]$ but does not have bounded variation.

g. A function with bounded variation need not be continuous. For instance, show that any increasing function from φ into $\mathbb{R}$ has bounded variation.

h. A function $\varphi : [a, b] \to \mathbb{R}$ has bounded variation on $[a, b]$ if and only if φ is the difference of two increasing functions. (We emphasize that φ does not need to be continuous.)

Hint for the "only if" part: Assume φ has bounded variation. Show that $p(t) = \mathrm{Var}(\varphi, [a, t])$ and $n(t) = p(t) - \varphi(t)$ are increasing functions.

i. A function $\varphi : [a, b] \to \mathbb{C}$ has bounded variation on $[a, b]$ if and only if $\mathrm{Re}\,\varphi : [a, b] \to \mathbb{R}$ and $\mathrm{Im}\,\varphi : [a, b] \to \mathbb{R}$ both have bounded variation.

More advanced ideas about bounded variation will be covered in 22.19 and thereafter.

19.22. *Proposition.* Suppose $\varphi : [a, b] \to X$ has bounded variation, where X is a *complete* metric space. Then $\varphi(t+) = \lim_{u \downarrow t} \varphi(t)$ exists at every $t \in [a, b)$, and $\varphi(t-) = \lim_{u \uparrow t} \varphi(t)$ exists at every $t \in (a, b]$, and φ is continuous except at countably many points of $[a, b]$. In fact, φ is right or left continuous at each point where the increasing function $t \mapsto \mathrm{Var}(\varphi, [a, t])$ is right or left continuous, respectively.

Hints: If φ is real-valued and increasing, then these results follow from 15.21.c. If φ is real-valued, then these results follow from 19.21.h. Now consider the case where φ takes values in a complete metric space X.

Let $\psi(u) = \mathrm{Var}(\varphi, [a, u])$. Then ψ is an increasing real-valued function on $[a, b]$, so for each $t \in [a, b)$ we have the existence of $\psi(t+) = \lim_{u \downarrow t} \psi(u)$, and moreover $\psi(t+) = \psi(t)$ except at countably many values of t.

When u and v decrease to t, then $\psi(u)$ and $\psi(v)$ both converge to the same limit $\psi(t+)$; since

$$d\left(\varphi(u), \varphi(v)\right) \quad \leq \quad \mathrm{Var}(\varphi, [u, v]) \quad = \quad \psi(v) - \psi(u),$$

it follows that the values of $\varphi(u)$ and $\varphi(v)$ form a Cauchy net. Thus $\varphi(t+) = \lim_{u \downarrow t} \varphi(u)$ exists at every $t \in [a, b)$. If ψ is right continuous at t, then we may apply that argument above with $u = t$ to show that φ is right continuous at t.

19.23. Josephy's Theorem (optional). Let (X, d) and (Y, e) be compact metric spaces. Suppose $\gamma : X \to Y$ has the property that whenever $\varphi : [0, 1] \to X$ has bounded variation, then $\gamma \circ \varphi : [0, 1] \to Y$ has bounded variation. Then γ is Lipschitzian. (This is a converse to 19.21.d.)

Proof. Suppose γ is not Lipschitzian. Then there exists a sequence (x_n, x_n') in $X \times X$ such that $x_n \neq x_n'$ and $e\left(\gamma(x_n), \gamma(x_n')\right) / d\left(x_n, x_n'\right) \to \infty$. Since Y has finite diameter, we have $d(x_n, x_n') \to 0$. Since X is compact, by passing to a subsequence we may assume the sequences (x_n) and (x_n') both converge to some limit x_∞ in X. For simplicity of notation later in the proof, let us also denote $x_0 = x_1'$.

Passing to a further subsequence, we may assume that

$$\frac{e\left(\gamma(x_n), \gamma(x_n')\right)}{d(x_n, x_n')} > n(n+1), \qquad d(x_n, x_\infty) < \frac{1}{n^2}, \qquad d(x_n', x_\infty) < \frac{1}{n^2}.$$

Now let $\delta_n = d(x_n, x'_n)$; then $\delta_n < 2/n^2$. Define $\varphi : [0,1] \to X$ by

$$\varphi(t) \quad = \quad \begin{cases} x_\infty & \text{when } t = 0 \\ x_n & \text{when } t \in \left[\frac{1}{n+1}, \frac{1}{n}\right) \text{ and } t - \frac{1}{n+1} \text{ is a multiple of } \delta_n \\ x'_n & \text{when } t \in \left[\frac{1}{n+1}, \frac{1}{n}\right) \text{ and } t - \frac{1}{n+1} \text{ is not a multiple of } \delta_n \\ x_0 & \text{when } t = 1. \end{cases}$$

We shall show that φ has bounded variation but $\gamma \circ \varphi$ does not. It is clear that φ has left and right limits at each point. Since we wish to show that $\gamma \circ \varphi$ does not have bounded variation, we may (arguing by contradiction) assume that it does; thus we may assume that $\gamma \circ \varphi$ also has left and right limits.

Let ι be either the identity map or γ; by analyzing the function $\iota \circ \varphi$ we shall simultaneously analyze the two functions φ and $\gamma \circ \varphi$. Let ρ be either the metric d or the metric e. The function $\iota \circ \varphi$ has variation given by

$$\begin{aligned} &\operatorname{Var}(\iota \circ \varphi, [0,1]) \\ &= \lim_{N \to \infty} \left\{ \rho\left(\iota(\varphi(0)), \iota(\varphi(\frac{1}{N+1})) \right) + \sum_{n=1}^{N} \operatorname{Var}\left(\iota \circ \varphi, \left[\frac{1}{n+1}, \frac{1}{n} \right] \right) \right\} \\ &= \rho\left(\iota(x_\infty), \lim_{t \downarrow 0} \iota(\varphi(t)) \right) + \sum_{n=1}^{\infty} \operatorname{Var}\left(\iota \circ \varphi, \left[\frac{1}{n+1}, \frac{1}{n} \right] \right). \end{aligned}$$

The term involving $\iota(x_\infty)$ is finite; the question is whether the infinite series converges.

Temporarily fix n, and let us analyze $\operatorname{Var}\left(\iota \circ \varphi, \left[\frac{1}{n+1}, \frac{1}{n}\right]\right)$. Let

$$\tau_j = \frac{1}{n+1} + j\delta_n \qquad \text{for } j = 0, 1, 2, \ldots, J,$$

with nonnegative integer J chosen so that

$$\frac{1}{n+1} = \tau_0 < \tau_1 < \cdots < \tau_J < \frac{1}{n} \le \tau_{J+1}.$$

The function φ is a step function on the interval $\left[\frac{1}{n+1}, \frac{1}{n}\right]$ — it takes the value x_n at each of the points $\tau_0, \tau_1, \tau_2, \ldots, \tau_J$, the constant value x'_n on each of the open intervals $(\tau_0, \tau_1), (\tau_1, \tau_2), \ldots, (\tau_{J-1}, \tau_J), (\tau_J, \frac{1}{n})$, and the value x_{n-1} at $\frac{1}{n}$. Hence

$$\operatorname{Var}\left(\iota \circ \varphi, \left[\frac{1}{n+1}, \frac{1}{n} \right] \right) \quad = \quad (2J+1)\rho\Big(\iota(x_n), \iota(x'_n)\Big) + \rho\Big(\iota(x'_n), \iota(x_{n-1})\Big).$$

To estimate this quantity, we shall use the inequality $\tau_J < \frac{1}{n} \le \tau_{J+1}$, which can be rewritten as $J < \frac{1}{n(n+1)\delta_n} \le J + 1$.

We now analyze the two cases separately: When ι is the identity function, we obtain

$$\mathrm{Var}\left(\varphi, \left[\frac{1}{n+1}, \frac{1}{n}\right]\right) = (2J+1)\delta_n + d(x_n', x_{n-1}) < \left(\frac{2}{n(n+1)} + \delta_n\right) + \left(\frac{1}{n^2} + \frac{1}{(n-1)^2}\right),$$

which is summable over n. On the other hand, when ι is the function γ, we obtain

$$\begin{aligned}\mathrm{Var}\left(\gamma\circ\varphi, \left[\frac{1}{n+1}, \frac{1}{n}\right]\right) &= (2J+1)e\left(\gamma(x_n), \gamma(x_n')\right) + e\left(\gamma(x_n'), \gamma(x_{n-1})\right) \\ &\geq (J+1)\cdot e\left(\gamma(x_n), \gamma(x_n')\right) \geq \left(\frac{1}{n(n+1)\delta_n}\right)\cdot\left(n(n+1)d(x_n, x_n')\right) = 1,\end{aligned}$$

which is not summable over n. This completes the proof.

For generalizations and further references, see Pierce [1994].

Cauchy Continuity

19.24. *Definition.* Let X and Y be uniform spaces. A mapping $f : X \to Y$ is **Cauchy continuous** if it has this property: Whenever (x_α) is a Cauchy net in X, then $(f(x_\alpha))$ is a Cauchy net in Y. An equivalent formulation in terms of filters is: Whenever $\mathcal{F}$ is a Cauchy filter on X, then the filter generated by the filterbase $f(\mathcal{F}) = \{f(F) : F \in \mathcal{F}\}$ is Cauchy on Y.

Cauchy continuity will only be studied briefly here; a deeper study can be found in Lowen-Colebunders [1989].

19.25. *Proposition.* Let $f : X \to Y$ be a map from one uniform space into another. Then

$$f \text{ is uniformly continuous} \quad\Rightarrow\quad f \text{ is Cauchy continuous} \quad\Rightarrow\quad f \text{ is continuous.}$$

(*Hint*: For the second implication, use 19.6.)

Moreover, neither of these implications is reversible. For instance, let $\mathbb{R}$ and the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ have the usual metric $d(x,y) = |x-y|$. Show that

(i) the function $f(x) = \tan(x)$, from $(-\frac{\pi}{2}, \frac{\pi}{2})$ to $\mathbb{R}$, is continuous but not Cauchy continuous, and

(ii) the function $f(x) = x^3$, from $\mathbb{R}$ to $\mathbb{R}$, is Cauchy continuous but not uniformly continuous.

19.26. *Exercise.* Let X and Y be uniform spaces; assume X is complete. Let $f : X \to Y$ be some mapping. Then f is continuous if and only if f is Cauchy continuous.

19.27. **Theorem.** Let X and Y be complete uniform spaces. Let $S \subseteq X$ be dense, and let $p : S \to Y$ be some given function. Then the following conditions are equivalent:

(A) p extends to a continuous function $\widehat{p} : X \to Y$;

(B) p extends to a Cauchy continuous function $\widehat{p} : X \to Y$;

(C) p is Cauchy continuous from S into Y.

Furthermore, if p is uniformly continuous, then so is any continuous extension $\widehat{p}$; in fact, any modulus of uniform continuity for p will also be a modulus of uniform continuity for $\widehat{p}$.

Proof. The conclusion about uniform continuity follows from 18.10. The implication (B) $\Rightarrow$ (C) is trivial. The implication (A) $\Rightarrow$ (B) follows easily from 19.26. Now assume (C); it suffices to prove (A). Fix any $x \in X$; let $\mathcal{N}(x)$ be its neighborhood filter in X. Then $S \cap \mathcal{N}(x) = \{S \cap N : N \in \mathcal{N}(x)\}$ is the neighborhood filter of x in S. That filter converges to x in S, and therefore that filter is Cauchy. Since p is Cauchy continuous, the filterbase $p(S \cap \mathcal{N}(x)) = \{p(S \cap N) : N \in \mathcal{N}(x)\}$ is Cauchy in Y — that is, the filter it generates is Cauchy. Since Y is complete, that filter is convergent. Now we may apply 16.15; this completes the proof.

19.28. *Definition.* Let X be a complete uniform space, and let $f : [a, b] \to X$ be some function. We say f is **piecewise continuous** if it satisfies any of these equivalent conditions. (The proof of equivalence uses 19.27.)

(A) f is continuous except at finitely many points and has left- and right-hand limits at those points.

(B) We can form a partition $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ such that f is uniformly continuous on each open interval (t_{j-1}, t_j).

(C) We can form a partition $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ such that f agrees on each *open* interval (t_{j-1}, t_j) with some X-valued function f_j that is continuous on the *closed* subinterval $[t_{j-1}, t_j]$.

Cauchy Spaces (Optional)

19.29. *Remarks.* Some of the ideas covered in this chapter can be extended to a setting slightly more general than uniform spaces. A **Cauchy space** is a set X equipped with a collection $\mathcal{C}$ of proper filters on X, called the **Cauchy filters**, which satisfy these axioms:

(i) For each $x \in X$, the ultrafilter fixed at x is Cauchy.

(ii) If $\mathcal{F}, \mathcal{G}$ are proper filters, $\mathcal{F}$ is Cauchy, and $\mathcal{F} \subseteq \mathcal{G}$, then $\mathcal{G}$ is Cauchy.

(iii) If $\mathcal{F}, \mathcal{G}$ are Cauchy and each member of $\mathcal{F}$ meets each member of $\mathcal{G}$, then $\mathcal{F} \cap \mathcal{G} = \{A \subseteq X : A \in \mathcal{F} \text{ and } A \in \mathcal{G}\}$ is Cauchy.

A function $p : X \to Y$ from one Cauchy space to another is **Cauchy continuous** if $\mathcal{F}$ is Cauchy in X implies $p(\mathcal{F})$ is Cauchy in Y; this generalizes 19.24.

Any Cauchy space can be made into a convergence space in a natural way: We say that a proper filter $\mathcal{F}$ **converges** to a point x if $\mathcal{F} \cap \mathcal{U}(x)$ is Cauchy; here $\mathcal{U}(x)$ is the ultrafilter fixed at x. Note that, with this definition, any convergent filter is Cauchy.

In the theory of Cauchy spaces, one of the main topics of investigation has been: In what ways may we form completions of Cauchy spaces — i.e., larger Cauchy spaces in which every Cauchy filter converges? That topic is surveyed in Kent and Richardson [1984].

19.30. *Example.* It is easy to see that any *uniform space* is a Cauchy space — i.e., the uniform space's Cauchy filters (defined as in 19.2) satisfy the three axioms of 19.29. Indeed, those axioms follow from 19.4.a, 19.4.b, and 19.5, respectively. The convergence that is then defined from the Cauchy structure, as in 19.29, coincides with the topological convergence determined by the uniformity, as in 5.33 and 15.7; that fact is just 19.6.

19.31. *Example.* Any *lattice group* $(X, \preccurlyeq)$ can be made into a Cauchy space in a natural way: Say that a net $(x_\alpha : \alpha \in \mathbb{A})$ or its filter $\mathcal{F}$ on X is Cauchy if there exists a set $S \subseteq X$ with these three properties:

(i) S is directed downward — i.e., for each $s_1, s_2 \in S$ there exists $s \in S$ with $s \preccurlyeq s_1 \wedge s_2$.

(ii) $0 = \inf(S)$.

(iii) For each $s \in S$, we have eventually $/x_\alpha - x_\beta/ \preccurlyeq s$ — that is, there is some $\gamma_0 \in \mathbb{A}$ such that $\alpha, \beta \succcurlyeq \gamma_0 \Rightarrow /x_\alpha - x_\beta/ \preccurlyeq s$. Equivalently, for each $s \in S$ there is some $F \in \mathcal{F}$ such that $x, x' \in F \Rightarrow /x - x'/ \preccurlyeq s$.

For purposes of the discussion below, we shall then say that S is a "witness" of the Cauchyness of $\mathcal{F}$.

If we define "Cauchy" in this fashion, then it is very easy to see that the "Cauchy" filters satisfy the first two axioms in 19.29. The "Cauchy" filters also satisfy the third axiom in 19.29, as we shall now demonstrate. Let $\mathcal{F}$ and $\mathcal{G}$ be two Cauchy filters such that every member of $\mathcal{F}$ meets every member of $\mathcal{G}$. Say $\mathcal{F}$ and $\mathcal{G}$ have witnesses S and T, respectively (in the sense of the preceding paragraph). Using 8.33.a, verify that $S+T = \{s+t : s \in S, t \in T\}$ is a witness for the Cauchyness of $\mathcal{F} \cap \mathcal{G}$.

In the context of lattice groups, the convergence determined by the Cauchy structure (as in 19.29) turns out to be precisely the order convergence (as defined in 7.38 and 7.40.d and further characterized in 8.44.a); this is easy to verify.

However, order convergence in a lattice group sometimes is not a topological convergence; we see an example of this in 21.43. Thus, the Cauchy structure of a lattice group is not necessarily given by a uniformity.

Completions

19.32. *Definitions.* Let X be a uniform space. By a **completion** (or more specifically, a *uniform completion*) of X we mean a complete uniform space Y with a dense subset that is isomorphic to X. Here "isomorphism" usually means a bijection that is uniformly continuous in both directions; we shall give this term a slightly stronger meaning in 19.36.

For pseudometric spaces and metric spaces, the term "completion" has a more specialized meaning. Let (Λ, d) be a (pseudo)metric space. By a **completion** of Λ we mean a complete (pseudo)metric space Y with a dense subset that is isomorphic to Λ; but here "isomorphism" means a distance-preserving bijection.

19.33. Theorem: Existence of completions of metric spaces. Every metric space has a completion. We shall sketch two proofs of this fact.

a. The first proof is extremely short: Let Λ be any metric space. Then, as we noted in 4.41.f and 19.11.f, Λ can be embedded isometrically in the complete metric space $B(\Lambda)$. Hence the closure of Λ in $B(\Lambda)$ is a completion. This proof has the conceptual drawback that it relies on already knowing $\mathbb{R}$ is complete.

b. The second proof is a bit longer, but it contains enough insight to be worth mentioning. Let (X, d) be any metric space. Show that if (x_n) and (y_n) are two Cauchy sequences in X, then $(d(x_n, y_n))$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, the number $D((x_n), (y_n)) = \lim_{n\to\infty} d(x_n, y_n)$ exists. Show that D, defined in this fashion, is a pseudometric on the set of all Cauchy sequences. Call two Cauchy sequences equivalent if the distance between them is 0; then D becomes a metric on the set of all equivalence classes of Cauchy sequences. Show that that metric is complete and X is a dense subset of the resulting metric space, with embedding given by $x \mapsto (x, x, x, \ldots)$.

c. *Cantor's Construction of* $\mathbb{R}$, *1883* (*optional*). A slight modification of the argument in the preceding paragraph gives us a construction of $\mathbb{R}$. Let X be the set $\mathbb{Q}$ of rational numbers, with its usual definition and properties. Define $d : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ by $d(x, y) = |x - y|$. A sequence (x_n) in $\mathbb{Q}$ will be called "Cauchy" if for each rational number $\varepsilon > 0$ there exists some M such that $m, n \geq N \Rightarrow |x_m - x_n| < \varepsilon$. Now verify lots of things; the quotient space constructed in the preceding paragraph is a Dedekind complete, chain ordered field, and thus it is $\mathbb{R}$.

19.34. *Preliminaries on Kolmogorov quotients.* Before continuing to the next two sections, the reader may find it helpful to briefly review sections 16.5 and 16.21, on Kolmogorov quotients. The quotient is formed from a space by "collapsing together" (i.e., identifying) those points that are indistinguishable from one another. It is easy to see that

a gauge space is complete if and only if its Kolmogorov quotient is complete,

provided that the Kolmogorov quotient is equipped with the gauge determined as in 16.21.

19.35. Lemma. Every pseudometric space has a (distance-preserving) completion.

Proof. Let (S, d) be a pseudometric space. Let Q be its Kolmogorov quotient; then Q is a metric space when metrized as in 16.21. The quotient map $\pi : S \to Q$ is distance-preserving and surjective (but not injective unless S is Hausdorff). Let C be a distance-preserving Hausdorff completion of Q, formed as in 19.33.a or 19.33.b, and let $i : Q \xrightarrow{\subseteq} C$ be the inclusion map. The composition $S \xrightarrow{\pi} Q \xrightarrow{i} C$ is a distance-preserving map into a complete metric space, but in general this map is not injective.

To overcome that drawback, we shall form a new space X that has C as its Kolmogorov quotient — i.e., we shall reverse the process of forming a Kolmogorov quotient. By relabeling if necessary, we may assume C is disjoint from S. We define the set X to be $\big(C \setminus i(Q)\big) \cup S$. To define the pseudometric of X, view X as a modification of C, formed by "uncollapsing" the points that were collapsed together by π. For each $q \in Q$, replace the single point $i(q) \in C$ with a relabeled copy of the set $\pi^{-1}(q) \subseteq S$, all the members of which were separated by distance 0 in S and will be separated by distance 0 in the new space X. Points in $C \setminus i(Q)$ are left unaltered in forming the new space X. The inclusion $S \xrightarrow{\subseteq} X$ is distance-preserving and injective, with X complete.

19.36. Theorem: Existence of completions of uniform spaces. Every uniform space has a completion. Furthermore, the completion can be given by a *distance-preserving* inclusion, in the following sense:

Let S be a uniform space whose uniform structure is given by a gauge D. Then there exists a complete uniform space X with gauge E, such that S is a dense subset of X and the members of D are just the restrictions of the members of E.

If S is Hausdorff, then we may choose X Hausdorff as well.

Proof. The proof may seem long because it involves a great deal of notation; but it is conceptually simple and actually involves very little computation.

For each pseudometric $d \in D$, let (T_d, d) be a completion of the pseudometric space (S, d). Here we use the same letter d for the given pseudometric on S and its extension to the larger space. The letter D will be used to represent not only the original gauge, but also the collection of these extensions.

Let $Y = \prod_{d \in D} T_d$ be equipped with the product uniform structure; then Y is complete, by (AC26) in 19.13. There may be many gauges on Y that give that product uniform structure; one particularly convenient gauge is formed as follows:

For each pseudometric d, we define a corresponding pseudometric on Y, which we shall denote by $\widehat{d}$, as follows: $\widehat{d}(y, y') = d\left(\pi_d(y), \pi_d(y')\right)$, where $\pi_d : Y \to T_d$ is the dth coordinate projection. It follows trivially from 18.11(C) that the product uniform structure on Y is given by the gauge $\widehat{D}$ consisting of all such pseudometrics $\widehat{d}$.

Define an inclusion $i : S \xrightarrow{\subseteq} Y$ by taking $i(s) = (s, s, s, \ldots)$ — that is, each member of S is mapped to the corresponding constant function. Clearly this map is distance-preserving: $d(s, s') = \widehat{d}(i(s), i(s'))$. The closure of $i(S)$ in Y is a distance-preserving completion of S.

If the original uniform space (S, D) is Hausdorff, then the construction above may be modified to yield a *Hausdorff* distance-preserving completion, as follows: Let Q be the Kolmogorov quotient of Y. Then the gauge space $(Q, \widehat{D})$ is complete and Hausdorff. The quotient map $\pi : Y \to Q$ is not necessarily injective, but its restriction to $i(S)$ is injective. Thus, the closure of $\pi(i(S))$ in Q is a distance-preserving Hausdorff completion of S.

19.37. Theorem: Uniqueness of Hausdorff completions. Both of the results below follow easily from 19.27, by an argument similar to the uniqueness proof in 4.38; for the metric space result, use a suitable modulus of uniform continuity.

a. Let X be a Hausdorff uniform space. Then the Hausdorff completion of X is unique up to isomorphism. In other words, if $i_1 : X \xrightarrow{\subseteq} Y_1$ and $i_2 : X \xrightarrow{\subseteq} Y_2$ are two such completions, then the bijection $i_2 \circ i_1^{-1}$: Range$(i_1) \to$ Range(i_2) extends uniquely to a bijection $\widehat{\imath} : Y_1 \to Y_2$ that is uniformly continuous in both directions.

b. Let X be a metric space. Then the metric completion of X is unique up to isomorphism. In other words, if $i_1 : X \xrightarrow{\subseteq} Y_1$ and $i_2 : X \xrightarrow{\subseteq} Y_2$ are two such completions, then the bijection $i_2 \circ i_1^{-1}$: Range$(i_1) \to$ Range(i_2) extends uniquely to a distance-preserving bijection $\widehat{\imath} : Y_1 \to Y_2$.

19.38. *Example and remarks.* The Lebesgue space $L^1[0,1]$, defined in 22.28, is a complete metric space in which $C[0,1] = \{$continuous scalar-valued functions on $[0,1]\}$ is dense; those properties will be proved in 22.30.d and 22.31(i). Thus $L^1[0,1]$ is the completion of $C[0,1]$, where the metric used is $d(f,g) = \int_0^1 |f(t) - g(t)|\,dt$. Although we shall prove that fact as a theorem, it could instead be used as a *definition* of $L^1[0,1]$. It is perhaps the most elementary definition of $L^1[0,1]$; it does not require any measure theory.

However, that definition has several drawbacks. It depends heavily on the topological structure of the interval $[0,1]$, and thus it does not generalize readily to the Lebesgue spaces $L^1(\mu)$. Also, it does not give us easy access to the important theorems that sometimes make $L^1[0,1]$ more useful than $C[0,1]$ — e.g., theorems such as the Monotone and Dominated Convergence Theorems 21.38(ii) and 22.29. Moreover, viewing $L^1[0,1]$ as the completion of $C[0,1]$ does not offer us much insight into the structure of $L^1[0,1]$: It describes members of that space as equivalence classes of Cauchy sequences of members of $C[0,1]$, where the definition of "equivalence" is somewhat complicated; or it identifies $L^1[0,1]$ as a subset of the collection of bounded maps from $C[0,1]$ into $\mathbb{R}$. We would prefer to view the members of $L^1[0,1]$ as maps from $[0,1]$ into $\mathbb{R}$.

We shall follow the usual development of integration theory: We begin with measures and measurable functions (in 9.8, 11.37, and Chapter 21). We use linearity to define the integrals of *simple functions*; then we take limits to obtain the integrals of other measurable functions. The measure μ can be defined on any measurable space Ω; the particular topological properties of $[0,1]$ are not especially relevant in this construction. Two measurable functions from Ω into the scalars are equivalent if they differ only on a set of measure 0. The members of $L^p(\mu)$ are equivalence classes of measurable functions whose integrals are not too big — see 22.28. This approach requires an explanation of "measurable function" and "measure 0," but it does not involve Cauchy sequences and ultimately it is more insightful. For most purposes, we can work with any member of an equivalence class, and so we obtain members of $L^1[0,1]$ as maps from $[0,1]$ into $\mathbb{R}$.

Banach's Fixed Point Theorem

19.39. Theorem (Banach, Caccioppoli). If X is a nonempty complete metric space and $f : X \to X$ is a strict contraction, then f has a unique fixed point ξ.

Moreover, $\xi = \lim_{k\to\infty} f^k(x)$ for every $x \in X$. In fact, we have this estimate of the rate of convergence:

$$d\Big(f^n(x),\xi\Big) \quad\le\quad \frac{\langle f\rangle_{\mathrm{Lip}}^n d\Big(x,f(x)\Big)}{1-\langle f\rangle_{\mathrm{Lip}}}.$$

Hints: Show $d\left(f^j(x),f^{j+1}(x)\right)\le\langle f\rangle_{\mathrm{Lip}}^j d(x,f(x))$ by induction on j. Also,

$$d\Big(f^n(x),f^m(x)\Big) \quad\le\quad \sum_{j=n}^{m-1} d\left(f^j(x),f^{j+1}(x)\right) \qquad \text{for } m>n,$$

by repeated use of the triangle inequality.

Remarks. The Contraction Mapping Theorem is remarkable: It has a short and simple proof, and yet it has many applications; see for instance 19.40.c and 30.9. In some respects it cannot be improved upon; this is made clear by the two converses given in 19.47 and 19.50.

19.40. *Exercises.*

a. Let (X,d) be a metric space. Let $f:X\to X$ be a strict contraction — or, more generally, let f be a self-mapping of X satisfying

$$d\Big(f(x),f(y)\Big) \quad<\quad d(x,y) \qquad \text{whenever} \quad x\ne y. \tag{$*$}$$

Then f has at most one fixed point.

b. It is possible for a map $f:X\to X$ satisfying condition $(*)$ of the previous exercise to have no fixed points — even if (X,d) is a nonempty complete metric space. Show this with $X=[1,+\infty)$ with the usual metric and $f(x)=x+\frac{1}{x}$.

c. Show that the equation $\cos(x)=x$ has a unique solution in $\mathbb{R}$. Then use a calculator to find that solution, correct to five decimal places.

d. *Continuous dependence of fixed points.* Let X be a nonempty complete metric space. Let (f_n) be a sequence of strict contraction self-mappings of X; say ξ_n is the fixed point of f_n. Assume $\sup_{n\in\mathbb{N}}\langle f_n\rangle_{\mathrm{Lip}}<1$, and $f_n\to f_\infty$ pointwise. Then f_∞ is a strict contraction, with fixed point $\xi_\infty=\lim_{n\to\infty}\xi_n$.

e. In the preceding exercise, the assumption $\sup_{n\in\mathbb{N}}\langle f_n\rangle_{\mathrm{Lip}}<1$ cannot be replaced with the weaker assumption that the f_n's and f_∞ are all strict contractions. The following example requires some familiarity with ℓ_2 (defined in 22.25). Let $X=\ell_2$, and let $f_n(x_1,x_2,x_3,\ldots)$ be the sequence whose nth component is $\frac{1}{n}+\frac{n-1}{n}x_n$ and whose other components are 0. Show that $\lim_{n\to\infty}\xi_n$ does not exist in ℓ_2.

f. If we know a strict contraction within some small error, then we also know its fixed point within some small error. More precisely, let f_1 and f_2 be strict contraction self-mappings of a nonempty complete metric space (X,d), and suppose $d\Big(f_1(x),f_2(x)\Big)\le\varepsilon$ for all $x\in X$. Then the distance between the fixed points of f_1 and f_2 is not greater than $\varepsilon/(1-\langle f_1\rangle_{\mathrm{Lip}})$.

19.41. *Proposition: continuous dependence of fixed points.* Let (X, d) be a complete metric space, and let $f : X \to X$ be a strict contraction. Let (f_n) be a sequence of arbitrary self-mappings of X (i.e., not necessarily contractions, or even necessarily continuous); assume that $f_n \to f$ uniformly on X. Let x be the unique fixed point of f, and for each n suppose that x_n is some (not necessarily unique) fixed point of f_n. Then $x = \lim_{n\to\infty} x_n$.

Proof (modified from Vidossich [1974]). Since $f_n \to f$ uniformly,

$$d\Big(x_n, f(x_n)\Big) \quad = \quad d\Big(f_n(x_n), f(x_n)\Big) \quad \to \quad 0 \qquad\qquad \text{as} \quad n \to \infty.$$

Similarly, $d\Big(x_m, f(x_m)\Big) \to 0$ as $m \to \infty$. Then

$$\Big|d(x_m, x_n) - d\Big(f(x_m), f(x_n)\Big)\Big| \;\le\; d\Big(x_m, f(x_m) + d\Big(x_n, f(x_n)\Big) \;\to\; 0$$

as $m, n \to \infty$. Since $d\Big(f(x_m), f(x_n)\Big) \le \kappa\, d(x_m, x_n)$ for some constant $\kappa = \langle f\rangle_{\mathrm{Lip}} < 1$, it follows easily (*exercise*) that $d(x_m, x_n) \to 0$ — that is, the sequence (x_n) is Cauchy.

Say $x_n \to z$. Since f is continuous we have $f(x_n) \to f(z)$. On the other hand, $d\Big(f(x_n), x_n\Big) \to 0$. Hence $x_n \to f(z)$. Therefore $z = f(z)$. Since x is the unique fixed point of f, we have $x = z$.

19.42. Slow Contraction Theorem (optional). Let $\gamma : [0, +\infty) \to [0, +\infty)$ be a nondecreasing function. Assume that

$$\gamma \text{ is upper semicontinuous and } \gamma(t) < t \text{ for each } t > 0 \qquad \text{(Boyd and Wong [1969])}$$

or, more generally, assume that

$$\lim_{n\to\infty} \gamma^n(t) \;=\; 0 \text{ for each } t \in [0, +\infty) \qquad \text{(Dugundji and Granas [1982])}.$$

Let (X, d) be a nonempty, complete metric space. Let f be a self-mapping of X that satisfies $d[f(x), f(y)] \le \gamma[d(x, y)]$ for all $x, y \in X$. Then f is continuous, f has a unique fixed point ξ, and $\lim_{n\to\infty} f^n(x) = \xi$ for each $x \in X$.

Proof (following Dugundji and Granas [1982]). Since $\gamma^n(t) \to 0$ for each $t > 0$, either $\gamma(t) = 0$ for some $t > 0$ or $\{\gamma(t) : t > 0\}$ contains arbitrarily small positive numbers. In either case, since γ is nondecreasing, it follows from $d(f(x), f(y)) \le \gamma(d(x, y))$ that f is continuous.

Next we shall show that any orbit $x, f(x), f^2(x), f^3(x), \ldots$ converges. Fix any $x = x_0 \in X$; let $x_n = f^n(x)$ and $c_n = d(x_n, x_{n+1})$. Then $c_n \le \gamma^n\,[d(x_0, x_1)]$, so $c_n \to 0$.

Suppose (x_n) is not Cauchy. Then there exist $\varepsilon > 0$ and integers $m(k)$ and $n(k)$ such that

$$k \le m(k) < n(k) \qquad \text{and} \qquad b_k = d(x_{m(k)}, x_{n(k)}) \ge \varepsilon$$

for $k = 1, 2, 3, \ldots$. For each k, we may assume that $n(k)$ is chosen as small as possible; hence $d(x_{m(k)}, x_{n(k)-1}) < \varepsilon$. Since

$$0 \quad < \quad b_k \;-\; d(x_{m(k)}, x_{n(k)-1}) \quad \le \quad c_{n(k)-1} \;\to\; 0,$$

it follows that $\lim_{k\to\infty} b_k = \varepsilon$.

Choose some p large enough so that $\gamma^p(2\varepsilon) < \varepsilon/3$. Then choose some k large enough so that $b_k < 2\varepsilon$ and $\sup_{j\geq k} c_j < \varepsilon/3p$. Now

$$\begin{array}{rcl} \varepsilon & \leq & d\left(x_{m(k)}, x_{n(k)}\right) \\ & \leq & \displaystyle\sum_{j=m(k)}^{m(k)+p-1} d\left(x_j, x_{j+1}\right) \;+\; d\left(x_{m(k)+p, n(k)+p}\right) \;+\; \sum_{j=n(k)}^{n(k)+p-1} d\left(x_j, x_{j+1}\right) \\ & < & p\cdot\dfrac{\varepsilon}{3p} \;+\; \gamma^p(2\varepsilon) \;+\; p\cdot\dfrac{\varepsilon}{3p} \qquad < \qquad \varepsilon, \end{array}$$

a contradiction.

Thus any sequence of iterates $(f^n(x))$ converges to a limit. If $f^n(x) \to \xi$ then $f^{n+1}(x) \to \xi$ also, but $f^{n+1}(x) \to f(\xi)$ by continuity of f; thus ξ is a fixed point. For uniqueness, suppose that ξ and η are two fixed points; then $d(\xi,\eta) = d(f^n(\xi), f^n(\eta)) \leq \gamma^n(d(\xi,\eta)) \to 0$.

19.43. *Remarks and further exercise.* The so-called "Slow Contraction Theorem" generalizes Banach's Contraction Mapping Theorem, since we can take $\gamma(t) = t\langle f\rangle_{\mathrm{Lip}}$. However, in the Contraction Mapping Theorem, the sequence $f^n(x)$ converges to the unique fixed point at a geometric rate. In contrast, the convergence in the Slow Contraction Theorem may be very slow.

In fact, it may be *arbitrarily slow.* Assume given *any* sequence $(a_1, a_2, a_3, \ldots)$ of positive numbers decreasing strictly to 0. We shall devise X, f, γ, ξ, x as in the Slow Contraction Theorem, satisfying $d(f^n(x), \xi) = a_n$.

Hints: Let $a_\infty = 0$, and let $X = \mathbb{N}\cup\{\infty\}$. Define

$$d(m,n) \;=\; \begin{cases} \max\{a_m, a_n\} & \text{if } m \neq n \\ 0 & \text{if } m = n \end{cases} \qquad\qquad f(n) \;=\; \begin{cases} n+1 & \text{if } n \in \mathbb{N} \\ n & \text{if } n = \infty. \end{cases}$$

Define $\gamma(r) = a_{n+1}$ when $r \in (a_{n+1}, a_n]$.

19.44. *Further remarks.* In a sense, the Slow Contraction Theorem 19.42 really is *not* more general than the Contraction Mapping Theorem. It is easy to show (*exercise*) that if f satisfies the hypotheses of the Slow Contraction Theorem, then f also satisfies the hypotheses of Meyer's converse to the Contraction Mapping Theorem, given in 19.47, and therefore the given metric can be replaced by a metric that makes f a strict contraction. Thus, the Slow Contraction Theorem may be helpful in the initial discovery stages of some research, but generally it can be replaced by Banach's Theorem at some later stage of that research, and so the Slow Contraction Theorem may go unmentioned in the final presentation of that research.

19.45. Caristi's Theorem (Browder, Caristi, and Kirk) (optional). Let (X,d) be a complete metric space, let $v : X \to [0,+\infty)$ be some lower semicontinuous function, and let $f : X \to X$ be some function such that $d(t, f(t)) \leq v(t) - v(f(t))$ for all $t \in X$. Then f has at least one fixed point.

Proof. Define a partial ordering on X by:

$$t \preccurlyeq u \qquad \text{if} \qquad d(t,u) \leq v(t) - v(u).$$

(This is sometimes called the **Brönsted ordering.**) By assumption, $t \preccurlyeq f(t)$ for all $t \in X$. Let $C \subseteq X$ be a nonempty $\preccurlyeq$-chain; we now note some properties of C:

a. The inclusion map $i : C \xrightarrow{\subseteq} X$ (considered as a map from $(C, \preccurlyeq)$ to (X,d)) is a Cauchy net.

Proof. Let $\tau = \inf_{c \in C} v(c)$; then $\tau \geq 0$. For any $\varepsilon > 0$, there is some $c_\varepsilon \in C$ with $v(c_\varepsilon) < \tau + \varepsilon$. If $c \in C$ and $c \succcurlyeq c_\varepsilon$, then $v(c) \geq \tau$ and $d(c_\varepsilon, c) \leq v(c_\varepsilon) - v(c) < \varepsilon$.

b. If λ is the limit of that net, then λ is a $\preccurlyeq$-upper bound for C.

Proof. For any fixed $c \in C$ and for all $c' \succcurlyeq c$ in C, we have $d(c,c') \leq v(c) - v(c')$. Take limits as c' increases; use the fact that v is lower semicontinuous.

Now let $\mathcal{C}$ be the collection of all nonempty $\preccurlyeq$-chains. Then $\mathcal{C}$ is nonempty since each singleton is a member of $\mathcal{C}$. Use $\subseteq$ to partially order $\mathcal{C}$. By a chain in $\mathcal{C}$ we mean a collection $\mathcal{S} \subseteq \mathcal{C}$ such that any two sets $S_1, S_2 \in \mathcal{S}$ satisfy $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$. It is easy to see that the union of all the members of $\mathcal{S}$ is then a chain in X, and thus a member of $\mathcal{C}$; hence it is the supremum of $\mathcal{S}$ in the poset $(\mathcal{C}, \subseteq)$.

Define $\varphi : \mathcal{C} \to \mathcal{C}$ by $\varphi(C) = C \cup f(\lim C)$; then $\varphi(C) \supseteq C$. By Zermelo's Fixed Point Theorem 5.52, φ has at least one fixed point C_0. Thus $f(\lim C_0) \in C_0$. Since $\lim C_0 \succcurlyeq c$ for all c in C_0, we have in particular $\lim C_0 \succcurlyeq f(\lim C_0)$. On the other hand, since $t \preccurlyeq f(t)$ for all t in X, we have $\lim C_0 \preccurlyeq f(\lim C_0)$. Thus $\lim C_0$ is a fixed point of f.

Remarks. Caristi's Theorem generalizes the Contraction Mapping Theorem, for if f is a strict contraction then we can take $v(t) = d(t, f(t))/(1 - \langle f \rangle_{\text{Lip}})$.

Our proof of Caristi's Theorem follows that of Mańka [1992]. It does not use the Axiom of Choice or any weakened form of Choice. Some analysts may prefer to prove Caristi's Theorem by the method indicated in the remarks in 19.51, although that proof uses Dependent Choice.

Meyers's Converse (Optional)

19.46. *Motivating exercise.* Assume the notations and hypotheses of Banach's Contraction Mapping Theorem 19.39. Then $f^n \to \xi$ uniformly on some neighborhood of ξ — that is, $\sup_{x \in V} d(\xi, f^n(x)) \to 0$ as $n \to \infty$, for some neighborhood V of ξ.

19.47. Meyers's Converse to the Contraction Mapping Theorem. Let f be a continuous self-mapping of a nonempty, complete metric space (X, ρ). Suppose that ξ is a fixed point of f, and $f^n(x) \to \xi$ as $n \to \infty$ for each $x \in X$. Also assume that $f^n \to \xi$ uniformly on some neighborhood of ξ — i.e., assume ξ has some neighborhood V such that

$$\lim_{n \to \infty} \sup_{v \in V} \rho\Big(f^n(v), \xi\Big) \quad = \quad 0.$$

Then there exists a topologically equivalent, complete metric d on X that makes f a strict contraction.

Remarks. This proof is taken from Meyers [1967]. A similar result was discovered independently in Leader [1977]. Both proofs were inspired by the treatment of the compact case given in Janos [1967].

Proof of theorem. The proof is in several steps.

a. By replacing V with a smaller neighborhood of ξ, we may assume also that V is open and that $f(V) \subseteq V$.

Hints: Certainly the theorem's hypotheses on V remain satisfied if we replace V with any smaller neighborhood of ξ. Replacing V with such a neighborhood, we may assume V is open. Now choose k large enough so that $f^k(V) \subseteq V$; then let $W = \bigcap_{j=0}^{k-1} f^{-j}(V)$. The set W has the required properties; we shall relabel it as V.

b. Some easy observations: $\bigcup_{n=0}^{\infty} f^{-n}(V) = X$ and

$$\cdots \supseteq f^{-2}(V) \supseteq f^{-1}(V) \supseteq V \supseteq f(V) \supseteq f^2(V) \supseteq \cdots .$$

For integers n (not necessarily positive), let $K_n = \mathrm{cl}\Big(f^n(V)\Big)$. Show that $f(K_n) \subseteq K_{n+1}$, and

$$\cdots \quad \supseteq \quad K_{-2} \quad \supseteq \quad K_{-1} \quad \supseteq \quad K_0 \quad \supseteq \quad K_1 \quad \supseteq \quad K_2 \quad \supseteq \quad \cdots .$$

Also, $K_n \to \xi$ as $n \to \infty$ — that is, any neighborhood of ξ contains K_n for all n sufficiently large. Hence $\mathrm{diam}(K_n) \to 0$ as $n \to \infty$ — if we use the metric ρ or any other metric that is equivalent to ρ — and $\bigcap_{n\in\mathbb{Z}} K_n = \{\xi\}$.

c. By replacing ρ with an equivalent metric that is also topologically complete, we may assume that f is a nonexpansive mapping — i.e., that $\rho(f(x), f(y)) \leq \rho(x, y)$.

Hints: For any $x, y \in X$, the sequence $\Big(\rho(f^n(x), f^n(y)) \ : \ n = 1, 2, 3, \ldots\Big)$ consists of nonnegative numbers converging to 0; hence a maximum exists:

$$\beta(x, y) \quad = \quad \max\left\{\rho\left(f^n(x),\, f^n(y)\right) \ : \ n = 0, 1, 2, \ldots\right\}.$$

As we noted in 18.3.e, β is a metric, uniformly stronger than ρ, and β makes f nonexpansive. In view of 19.11.i, it suffices to show that ρ is topologically stronger than β. Let any $x \in X$ and $\varepsilon > 0$ be given; we must find a $\delta > 0$ such that $\rho(x, y) < \delta \ \Rightarrow \ \beta(x, y) < \varepsilon$. Choose N large enough so that $\mathrm{diam}_\rho(f^N(V)) < \varepsilon$ and $f^N(x) \in V$. Using the continuity of f in (X, ρ), show that there is some $\delta > 0$ satisfying

$$\rho(x, y) < \delta \qquad \Rightarrow \qquad f^N(y) \in V \quad \text{and} \quad \max_{j<2N} \rho\left(f^j(x), f^j(y)\right) < \varepsilon.$$

Show that this δ has the right properties.

We shall now replace ρ with β (by relabeling), for simplicity of notation.

d. Define

$$\mu(x) \quad = \quad \sup\{n \in \mathbb{Z} \ : \ x \in K_n\} \quad = \quad \begin{cases} \infty & \text{if } x = \xi \\ \text{a finite integer} & \text{if } x \neq \xi. \end{cases}$$

Show that $\mu(f(x)) \geq \mu(x) + 1$ for all $x \in X$. Moreover, for any sequence (x_m) in X,

$$\mu(x_m) \to \infty \qquad \Rightarrow \qquad x_m \to \xi \text{ (in the given topology).}$$

e. Define $r(x,y) = 2^{-\min\{\mu(x),\mu(y)\}}\rho(x,y)$. (We could replace 2 with any constant real number greater than 1, but we shall use 2 for simplicity.) Verify that r is a mapping from $X \times X$ into $[0, +\infty)$ that satisfies

$$r(x,y) = 0 \iff x = y, \qquad r(x,y) = r(y,x), \qquad r\Big(f(x), f(y)\Big) \leq \frac{1}{2} r(x,y).$$

(The last inequality follows from $\mu(f(x)) \geq \mu(x) + 1$ and the fact that f is nonexpansive.)

f. Use r to define a pseudometric $d(x,y) = \inf \sum_i r(a_{i-1}, a_i)$ as in 4.42. Show that $d(x,y) \leq r(x,y)$ and $d\Big(f(x), f(y)\Big) \leq \frac{1}{2} d(x,y)$.

g. To show d is a metric, let any $x, y \in X$ with $x \neq y$ be given; we must show $d(x,y) > 0$ (and in doing so we shall also obtain an estimate that will be useful later). Since we cannot have both x and y equal to ξ, we may assume $x \neq \xi$, and thus $\mu(x) < \infty$. Fix any integer $k \geq \mu(x)$; then $x \notin K_{k+1}$. (For the proof of $d(x,y) > 0$ we may simply take $k = \mu(x)$, but other choices of k will be useful later.)

Consider any sequence $a_0, a_1, a_2, \ldots, a_m$ with $a_0 = x$ and $a_m = y$; we must obtain a positive lower bound for $\sum_{i=1}^m r(a_{i-1}, a_i)$ independent of the choice of the sequence (a_i). We consider two cases. In the first case, $\mu(a_i) \leq k$ for all i. Then

$$\sum_{i=1}^m r(a_{i-1}, a_i) \quad \geq \quad \sum_{i=1}^m 2^{-k}\rho(a_{i-1}, a_i) \quad \geq \quad 2^{-k}\rho(x,y),$$

which is positive. In the second case, let j be the first integer that satisfies $\mu(a_j) \geq k+1$. Then $\min\{\mu(a_{i-1}), \mu(a_i)\} \leq k$ for $i = 1, 2, \ldots, j$, and so

$$\begin{aligned} \sum_{i=1}^m r(a_{i-1}, a_i) \quad &\geq \quad \sum_{i=1}^j r(a_{i-1}, a_i) \quad \geq \quad \sum_{i=1}^j 2^{-k}\rho(a_{i-1}, a_i) \\ &\geq \quad 2^{-k}\rho(x, a_j) \quad \geq \quad 2^{-k}\operatorname{dist}_\rho(x, K_{k+1}), \end{aligned}$$

which is positive since K_{k+1} is a closed set that does not include x. In any case, we obtain

$$d(x,y) \quad \geq \quad 2^{-k} \min\{\rho(x,y), \ \operatorname{dist}_\rho(x, K_{k+1})\} \quad \text{when } \mu(x) \leq k. \qquad (**)$$

h. The metric ρ is stronger than d.

Hints: Suppose that $x_m \xrightarrow{\rho} x$; we must show that $x_m \xrightarrow{d} x$. Fix any positive integer k large enough so that $x \in f^{-k}(V)$. Since $f^{-k}(V)$ is an open set, for m sufficiently large we have $x_m \in f^{-k}(V)$. Since $f^{-k}(V) \subseteq K_{-k}$, we have $\min\{\mu(x), \mu(x_m)\} \geq -k$, and therefore $d(x, x_m) \leq r(x, x_m) \leq 2^k \rho(x, x_m)$.

i. Let (x_m) be a sequence in X. If (x_m) is d-Cauchy, then some subsequence of (x_m) is ρ-Cauchy.

Hints: Suppose not. Then no subsequence of (x_m) is ρ-convergent. In particular, no subsequence of (x_m) is ρ-convergent to ξ. Hence, for some M the number $R = \mathrm{dist}_\rho\Big(\xi,\ \{x_M, x_{M+1}, x_{M+2}, x_{M+3}, \ldots\}\Big)$ is positive.

If the $\mu(x_m)$'s are unbounded, then there is some subsequence (x_{m_j}) such that $\mu(x_{m_j}) \to \infty$. However, then $x_{m_j} \xrightarrow{\rho} \xi$ by 19.47.d.

Thus, $\sup_m \ \mu(x_m) < \infty$. Fix any integer $k > \sup_m \ \mu(x_m)$ large enough so that also $\mathrm{diam}_\rho(K_{k+1}) < \frac{1}{2}R$. Then for any $m \geq M$ we have $\mathrm{dist}_\rho(x_m, K_{k+1}) > \frac{1}{2}R$. Since (x_m) is d-Cauchy, for all m, m' sufficiently large, we have

$$d(x_m, x_{m'}) \quad < \quad 2^{-k-1}R \quad < \quad 2^{-k}\mathrm{dist}_\rho\,(x_m, K_{k+1})\,.$$

It then follows from $(**)$ that $d(x_m, x_{m'}) \geq 2^{-k}\rho(x_m, x_{m'})$. Thus the sequence (x_m) is ρ-Cauchy.

j. The two metrics ρ and d are topologically equivalent, and d is complete. (Immediate from 19.11.i.)

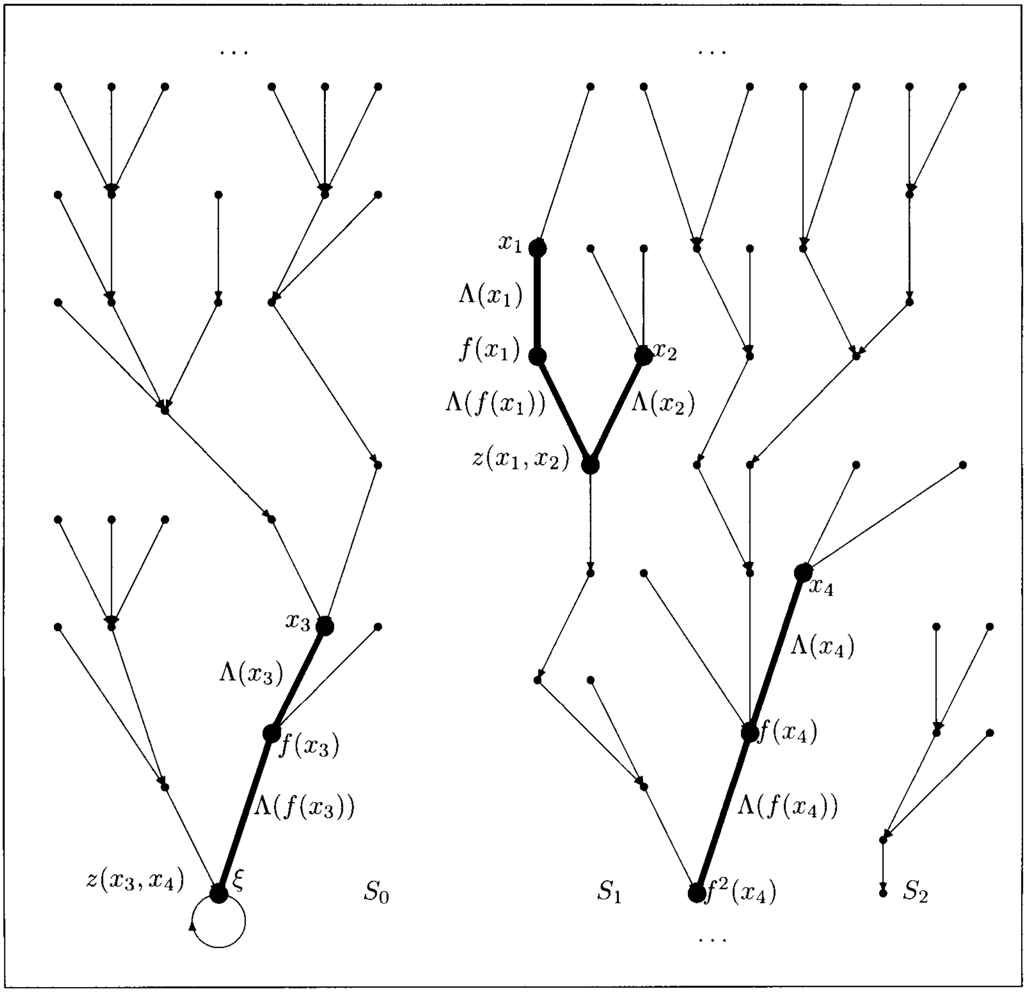
Bessaga's Converse and Brönsted's Principle (Optional)

19.48. *Technical lemma: Bessaga-Brunner Metric.* We introduce a slightly complicated type of metric that will be used in two long proofs below.

Let X be a set, and let $f : X \to X$ and $\Lambda : X \to (0, +\infty)$ be some mappings. (We emphasize that the values of Λ are nonzero.) Assume that $f(\xi) = \xi$, and assume that no iterate f^n $(n \geq 1)$ has any fixed point other than ξ. We shall use f and Λ to define a metric d on X.

For brevity, our arguments will rely on the following diagram. We denote $f^0(x) = x$. We write $x \approx y$ if there exist nonnegative integers p, q such that $f^p(x) = f^q(y)$. It is easy to see that this is an equivalence relation on X. Let S_0 be the equivalence class containing ξ.

We now sketch a graph that shows the action of f on X. We view each element of X as a vertex of the graph, and draw an arrow (i.e., directed line segment) from x to $f(x)$ for each $x \in X$. We shall arrange the graph so that this line segment goes downward — i.e., so that $f(x)$ is below x. Because ξ is the only fixed point of any of the iterates of f, we see that the graph has no closed cycles (i.e., loops) other than the one at ξ. The graph consists of



several components: one for each equivalence class. Each component is a simply connected tree, with root downward and branches upward. Each tree may go infinitely high or only finitely high; the information provided to us is not enough to determine whether some point x_0 has infinitely many predecessors $f^{-1}(x_0)$, $f^{-2}(x_0)$, $f^{-3}(x_0)$, ... or only finitely many. The tree representing S_0 coalesces at its bottom to a single root at ξ. If there is any other equivalence class, then it is represented by a tree that does not coalesce to a single root, but instead continues downward through an infinite succession of branches. We may describe such a tree as "infinitely deep."

Now, for each $x \in X \backslash \{\xi\}$, we label the line segment from x to $f(x)$ with the number $\Lambda(x)$. (The number $\Lambda(\xi)$ will not be used.) We shall take that number $\Lambda(x)$ to be the distance between x and $f(x)$. In the notation of 4.43, we have $r(x, f(x)) = r(f(x), x) = \Lambda(x)$. The function r is only defined on the pairs of points that are adjacent in the diagram.

If x and y are two points in the same equivalence class, then we trace forward through the graph to the first point z where x and y coalesce — i.e., we have $z = f^p(x) = f^q(y)$ where the nonnegative integers p, q are as low as possible, as shown in the diagram. We may write $p = p(x,y)$, $q = q(x,y)$, $z = z(x,y)$, to emphasize the dependence on x and y. There is no shorter route between x and y; there is no other route at all except by retracing some steps. The distance between x and y is thus

$$d(x,y) \quad = \quad d(x,z) + d(y,z) \quad = \quad \sum_{j=0}^{p-1} \Lambda\left(f^j(x)\right) + \sum_{j=0}^{q-1} \Lambda\left(f^j(y)\right) \qquad (\natural)$$

where it is understood that an empty sum is 0 (obtained when $x = z$ and $p = 0$, or when $y = z$ and $q = 0$). For example, in the illustration, we have

$$d(x_1, x_2) \quad = \quad [\Lambda(x_1) + \Lambda(f(x_1))] + \Lambda(x_2)$$

since $p(x_1, x_2) = 2$ and $q(x_1, x_2) = 1$ for this example. The function d defined in this fashion is a metric on the equivalence class (and not just a pseudometric), since we have assumed $\Lambda(x)$ is strictly positive for each x.

If there is more than one equivalence class — i.e., if the diagram contains more than one tree — then further considerations are necessary. We shall define $f^\infty(x) = \xi$ for all $x \in X$. (Intuitively, it is helpful to imagine that each "infinitely deep" tree continues downward and has ξ at its infinitely deep bottom.) The formula $(\natural)$ now becomes meaningful for all points $x, y \in X$ (not necessarily in the same equivalence class), with the understanding that p and q are not necessarily finite. We still choose p and q to be the lowest values that satisfy $f^p(x) = f^q(y)$. We find that

$$p \quad = \quad \begin{cases} \text{some finite number} & \text{if } x \in S_0 \text{ or } x \approx y \\ \infty & \text{if } x \notin S_0 \text{ and } x \not\approx y. \end{cases}$$

Analogous conditions apply to q. We define $d : X \times X \to [0, +\infty]$ as in $(\natural)$. For instance, in the illustration,

$$d(x_3, x_4) \quad = \quad [\Lambda(x_3) + \Lambda(f(x_3))] + \left[\Lambda(x_4) + \Lambda(f(x_4)) + \Lambda(f^2(x_4)) + \cdots\right]$$

since in this example we have $p(x_3, x_4) = 2$ and $q(x_3, x_4) = \infty$. It is easy to show that the function d defined in this fashion is a metric if its values are always finite — i.e., if the sums in $(\natural)$ always converge. That will be true for certain choices of Λ and f considered in 19.50.

19.49. *Motivation for Bessaga's converse.* Let f be a self-mapping of a set X. Suppose that for some k, the map f^k has a unique fixed point ξ. Then ξ is also the unique fixed point of f.

19.50. Bessaga's Converse to the Contraction Mapping Theorem. Let X be a set, let $f : X \to X$ be some mapping, let $f(\xi) = \xi$, and suppose that no iterate f^n $(n \geq 1)$ has any fixed point other than ξ. Then there exists a complete metric Λ on X that makes f a strict contraction.

Proof. We shall define a metric as in 19.48. We shall define $\Lambda : X \to (0, +\infty)$ by taking $\Lambda(x) = 2^{\lambda(x)}$ for a certain integer-valued function λ specified below. (Actually, we have chosen 2 just for simplicity; we could replace it with any constant real number greater than 1.)

We require that $\lambda : X \to \mathbb{Z}$ have the property that

$$\lambda(f(x)) = \lambda(x) - 1 \qquad \text{whenever } f(x) \neq x \text{ (i.e., whenever } x \neq \xi).$$

To show that there exists such a function λ, define equivalence classes and sketch trees as in 19.48. Choose some representative element z_S from each equivalence class S. (This requires some form of the Axiom of Choice, if there are infinitely many equivalence classes.) Define $\lambda(z_S) = 0$ for each equivalence class S. After that, λ is uniquely determined: add 1 when moving up in the tree, and subtract 1 when moving down in the tree.

As in 19.48, we define $f^\infty(x) = \xi$ for all $x \in X$ and define p, q, z, d as in 19.48. In the present application, that yields

$$d(x,y) \quad = \quad \sum_{j=0}^{p-1} 2^{\lambda(x)-j} + \sum_{j=0}^{q-1} 2^{\lambda(y)-j}.$$

These sums converge even if p or q is infinite, so d is a metric. From $\lambda(f(x)) = \lambda(x) - 1$ it follows that $\langle f \rangle_{\mathrm{Lip}} \leq \frac{1}{2}$. For arguments below, we note that $q(x,\xi) = 0$, and hence $d(x,\xi) \leq \sum_{j=0}^{\infty} 2^{\lambda(x)-j} = 2^{\lambda(x)+1}$.

To show that the metric is complete, let (x_n) be a Cauchy sequence; we wish to show that (x_n) converges. If the numbers $\lambda(x_n)$ are not bounded below, then some subsequence (x_{n_k}) satisfies $\lambda(x_{n_k}) \to -\infty$ and hence $d(x_{n_k}, \xi) \to 0$; hence $x_n \to \xi$ since (x_n) is Cauchy. Thus we may assume that $\lambda(x_n)$ is bounded below by some finite constant C. Whenever x_m and x_n are *distinct* members of X, then at least one of the numbers $p(x_m, x_n)$, $q(x_m, x_n)$ is positive, and so

$$d(x_m, x_n) \quad \geq \quad \min\{2^{\lambda(x_m)}, 2^{\lambda(x_n)}\} \quad \geq \quad 2^C.$$

However, (x_n) is Cauchy, so $d(x_m, x_n) < 2^C$ for all m, n sufficiently large. Thus, for all m, n sufficiently large, the points x_m and x_n are not distinct — i.e., the sequence (x_n) is eventually constant and therefore convergent.

19.51. We shall show that the Principle of Dependent Choice, introduced in 6.28, is equivalent to the following principles about complete metric spaces:

> **(DC3) Dancs-Hegedus-Medvegyev Principle.** Let (X, d) be a nonempty, complete metric space. Let $\preccurlyeq$ be a partial ordering on X, which is semicontinuous in the following sense: For each $x \in X$, the set $F(x) = \{y \in X : y \succcurlyeq x\}$ is closed in the metric space (X, d). Assume also that d and $\preccurlyeq$ satisfy the **Picard condition**:
>
> > whenever (x_n) is a sequence in X with $x_1 \preccurlyeq x_2 \preccurlyeq x_3 \preccurlyeq \cdots$, then $d(x_n, x_{n+1}) \to 0$.

Then $(X, \preccurlyeq)$ has a maximal element.

(DC4) Brönsted's Maximal Principle. Let (X, d) be a nonempty, complete metric space, and suppose $r : X \to [0, +\infty)$ is lower semicontinuous. Define a partial ordering $\preccurlyeq$ on X by:

$$x \preccurlyeq y \quad \text{if} \quad d(x,y) \leq r(x) - r(y).$$

Then $(X, \preccurlyeq)$ has a maximal element.

Remarks. Caristi's Theorem (19.45) follows from Brönsted's Theorem (DC4) by a one-line proof: The maximal point is a fixed point. Also, Brönsted's Theorem follows from Caristi's Theorem by a one-line proof, if we are permitted to use the Axiom of Choice: Just take f to be a suitable choice function. Thus, the two theorems are "equivalent" in a sense used by some mathematicians: Each follows *easily* from the other, if we are permitted to use conventional set theory (including the Axiom of Choice). However, Brunner [1987 *Zeitschr.*] has pointed out that the two theorems are *not* equivalent in the sense of set theory, for Brönsted's Theorem is equivalent to DC (as we shall show), whereas Caristi's Theorem actually follows from just ZF, without DC or any other weakened version of Choice. Further discussion of this and related ideas are given by Mánka [1988].

Proof of (DC2) $\Rightarrow$ (DC3). Note that $u \in F(v) \Rightarrow F(u) \subseteq F(v)$. Also note that $x \in F(x)$ for each x, hence $F(x)$ is nonempty. We may replace d with any uniformly equivalent metric; by 18.14 we may assume d is bounded. Then $\mathrm{diam}(F(x))$ is finite.

For any nonempty set $S \subseteq X$ and any point $x \in X$ and any number $\varepsilon > 0$, the set $\{y \in S \,:\, d(x,y) \geq \frac{1}{2}\mathrm{diam}(S) - \varepsilon\}$ is nonempty; this follows easily from the definition of diameter. Hence, given any point $x_{n-1} \in X$, there is some $x_n \in F(x_{n-1})$ satisfying

$$d(x_{n-1}, x_n) \quad \geq \quad \frac{1}{2}\mathrm{diam}\,(F(x_{n-1})) \;-\; 2^{-n}.$$

Using (DC2), we construct a sequence $(x_n : n \in \mathbb{N})$ satisfying this inequality. By the Picard condition, then, $d(x_n, x_{n-1}) \to 0$; hence $\mathrm{diam}(F(x_n)) \to 0$. From $x_n \in F(x_{n-1})$ we obtain $F(x_n) \subseteq F(x_{n-1})$. By 19.11.c, $\bigcap_{n=1}^{\infty} F(x_n)$ contains exactly one point, z. For each $n \in \mathbb{N}$, we have $z \in F(x_n)$ and hence $F(z) \subseteq F(x_n)$; thus $F(z) \subseteq \bigcap_{n=1}^{\infty} F(x_n) = \{z\}$. Therefore z is $\preccurlyeq$-maximal in X.

Proof of (DC3) $\Rightarrow$ (DC4). Easy exercise.

Proof of (DC4) $\Rightarrow$ (DC1). This proof is a slight simplification of a proof given by Brunner [1987 *Zeitschr.*]

Let Φ be a function that contradicts (DC1); we shall use it to construct a contradiction of (DC4). Thus, we assume that we are given a set A and a function

$$\Phi \quad : \quad A \quad \to \quad \{\text{nonempty subsets of } A\}$$

for which there does not exist an infinite choice sequence — i.e., an infinite sequence (a_n) satisfying $a_{n+1} \in \Phi(a_n)$ for all n.

By a *choice sequence of length n* we shall mean a finite sequence

$$x = \Big(x(1), x(2), \ldots, x(n)\Big)$$

that satisfies $x(k+1) \in \Phi(x(k))$ for $k = 1, 2, \ldots, n-1$. Such a sequence may be viewed as a function from the set $\{1, 2, \ldots, n\}$ into A. We shall also consider the empty sequence to be a choice sequence (of length 0); we shall denote it by ξ. By the "Axiom" of Finite Choice (in 6.14), any choice sequence can be extended to a longer choice sequence; thus there is no maximal choice sequence.

Let X be the set of all choice sequences. We observe that X does not contain an infinite $\subseteq$-chain. Indeed, if $\mathcal{C}$ were such a chain, then $\bigcup_{c \in \mathcal{C}} \mathrm{Graph}(c)$ would be the graph of an infinite choice sequence.

For each $x \in X$, let $\lambda(x)$ be the length of x. Also, define the "immediate truncation function" $f : X \to X$ by

$$f(x) = \begin{cases} \Big(x(1), x(2), \ldots, x(n-1)\Big) & \text{if } x = \Big(x(1), x(2), \ldots, x(n)\Big) \\ \xi & \text{if } x = \xi. \end{cases}$$

Then $\lambda(f(x)) = \lambda(x) - 1$ when x is not the empty sequence.

Let $\Lambda(x) = 2^{-\lambda(x)}$. Define equivalence and the functions p, q, z, d as in 19.48. Note that every point in X is equivalent to the empty sequence, and thus $S_0 = X$ is the only equivalence class. Therefore p and q are always finite, the sums in 19.48($\natural$) always converge, and d is a metric on X. We restate its formula here:

$$\begin{aligned} d(x,y) \quad = \quad d(x,z) + d(y,z) \quad &= \quad \sum_{j=0}^{p-1} 2^{-\lambda(x)+j} + \sum_{j=0}^{q-1} 2^{-\lambda(y)+j} \\ &= \Big[2^{-\lambda(z)} - 2^{-\lambda(x)}\Big] + \Big[2^{-\lambda(z)} - 2^{-\lambda(y)}\Big], \qquad (\natural) \end{aligned}$$

where p, q are the smallest nonnegative integers satisfying $f^p(x) = f^q(y)$, and $z = z(x,y)$ is the common value of $f^p(x)$ and $f^q(y)$. The choice sequence $z = z(x,y)$ is the longest common restriction of x and y; it is the empty sequence if x and y begin with different choices.

Define the Brönsted ordering $\preccurlyeq$ as in (DC4), using the function $\Lambda(x) = 2^{-\lambda(x)}$. That is, define

$$x \preccurlyeq y \qquad \text{to mean} \qquad d(x,y) \leq 2^{-\lambda(x)} - 2^{-\lambda(y)}.$$

We claim that the following statements are equivalent:

(A) the sequence y is an extension of x — that is, $z(x,y) = x$;

(B) $d(x,y) = 2^{-\lambda(x)} - 2^{-\lambda(y)}$;

(C) $x \preccurlyeq y$.

Indeed, (A) $\Rightarrow$ (B) follows from ($\natural$), and (B) $\Rightarrow$ (C) is trivial. For (C) $\Rightarrow$ (A), suppose that $x \preccurlyeq y$. Then

$$\Big[2^{-\lambda(z)} - 2^{-\lambda(x)}\Big] + \Big[2^{-\lambda(z)} - 2^{-\lambda(y)}\Big] \quad = \quad d(x,y) \quad \leq \quad 2^{-\lambda(x)} - 2^{-\lambda(y)},$$

hence $2 \cdot 2^{-\lambda(z)} \leq 2 \cdot 2^{-\lambda(x)}$, hence $\lambda(z) \geq \lambda(x)$. But z is a restriction of x; hence in fact $z = x$.

In particular, we note that

$$d\Big(f(y), y\Big) \quad = \quad 2^{-\lambda(y)+1} - 2^{-\lambda(y)} \quad = \quad 2^{-\lambda(y)}$$

if y is not the empty sequence. From the illustration in 19.48 it is clear that the nearest points to any choice sequence x are the extensions y obtained by adding one more term at the end of sequence x, and those sequences satisfy $x = f(y)$; thus their distance from x is $2^{-\lambda(y)} = 2^{-\lambda(x)-1}$. Thus

$$w, x \in X, \quad d(w, x) < 2^{-\lambda(x)-1} \qquad \Rightarrow \qquad w = x. \tag{1}$$

This shows that the topology determined by the metric d is discrete. It follows that any real-valued function defined on X is continuous and hence lower semicontinuous.

Obviously, $(X, \preccurlyeq)$ has no maximal element. It suffices to show that d is complete, for this will contradict (DC4). Let (x_n) be a Cauchy sequence; we shall show that (x_n) converges. In view of 19.4.c, it suffices to show that some subsequence of (x_n) converges; thus we may replace (x_n) with any subsequence. Via such a replacement, we may assume that

$$d(x_j, x_k) \; < \; 2^{-j} \qquad \text{whenever} \quad k > j. \tag{2}$$

This property will be preserved if we replace (x_n) by a further subsequence.

First consider the case in which (x_n) has some subsequence whose lengths are bounded. Such a subsequence is eventually constant, by (1) and (2); hence it is convergent.

We now consider the remaining case, in which $(\lambda(x_n))$ has no bounded subsequence — i.e., the case in which $\lim_{n \to \infty} \lambda(x_n) = \infty$. Replacing (x_n) with a subsequence, we may assume that

$$\lambda(x_1) \; < \; \lambda(x_2) \; < \; \lambda(x_3) \; < \; \cdots \qquad \text{and} \qquad \lambda(x_n) \geq n+1 \tag{3}$$

for all n.

Let $v_j = z(x_j, x_{j+1})$; that is, v_j is the largest common restriction of x_j and x_{j+1}. Then, by (2) and (3) and ($\natural$),

$$\begin{aligned} 2^{-j} \quad > \quad d(x_j, x_{j+1}) \quad &= \quad \left[2^{-\lambda(v_j)} - 2^{-\lambda(x_j)}\right] + \left[2^{-\lambda(v_j)} - 2^{-\lambda(x_{j+1})}\right] \\ &> \quad 2\left[2^{-\lambda(v_j)} - 2^{-\lambda(x_j)}\right] \quad \geq \quad 2\left[2^{-\lambda(v_j)} - 2^{-j-1}\right]. \end{aligned}$$

The inequality $2^{-j} > 2\left[2^{-\lambda(v_j)} - 2^{-j-1}\right]$ simplifies to $\lambda(v_j) > j$. Thus v_j, the common restriction of x_j and x_{j+1}, has length greater than j. That is, the functions v_j, x_j, x_{j+1} all have domains that include the set $\{1, 2, \ldots, j\}$, and those functions all agree on that set. Let w_j be the function on $\{1, 2, \ldots, j\}$ obtained by restricting any of v_j, x_j, x_{j+1} to that set. The function w_j is a choice sequence, since it is a restriction of a choice sequence. Then w_{j+1}, defined analogously, is an extension of w_j, since both these functions are restrictions of x_{j+1}. The sequence $w_1, w_2, w_3, \ldots$ forms an infinite $\subseteq$-chain in X, a contradiction.

Chapter 20

Baire Theory

20.1. *Preview.* The name "Baire" is, unfortunately, associated with four distinct notions, which can easily be confused:

- sets of the first or second category of Baire;
- Baire spaces;
- sets with the Baire property; and
- Baire sets.

All are introduced in this chapter. The first three of these notions are closely related and will be studied extensively in the following pages. The fourth notion is less important for the purposes of this book and will be introduced briefly in 20.34 mainly to prevent the beginner from confusing Baire sets with the other "Baire" notions.

Much of the material in this chapter is taken from Kuratowski [1948], Bourbaki [1966], Engelking [1977], Oxtoby [1980], and Vaughan [1988].

G-Delta Sets

20.2. *Terminology.* In some older topology books, the letters "F" and "G" are reserved for closed sets and open sets, respectively. That convention is no longer widely used. This text does not follow that convention in general, but gives those letters preference whenever convenient.

The following related convention is still widely used: The union of countably many closed sets is called an "$\boldsymbol{F_\sigma}$;" the intersection of countably many open sets is called a "$\boldsymbol{G_\delta}$." Similarly, the union of countably many G_δ's is a $G_{\delta\sigma}$; the intersection of countably many F_σ's is a $F_{\sigma\delta}$.

The letters F and σ come from *fermé* and *sum*, French for "closed" and "sum." The letters G and δ come from *Gebiet* and *Durchschnitt*, German for "open set" and "intersection;" see Hocking and Young [1961].

Exercises.

a. The complement of an F_σ is a G_δ, and conversely. (Thus any results about F_σ's can be restated in terms of G_δ's, or conversely.)

b. The terms "F_δ" and "G_σ" are not useful, since the intersection of countably many closed sets is a closed set, etc. Likewise, the terms "$F_{\sigma\sigma}$" and "$G_{\delta\delta}$" are not useful: the union of countably many F_σ's is another F_σ, etc.

c. Any F_σ is in fact the union of an *increasing* sequence of closed sets, and any G_δ is the intersection of a *decreasing* sequence of open sets.

Hint: If $S = \bigcup_{n=1}^\infty K_n$ where the K_n's are closed, then also

$$S \quad = \quad K_1 \;\cup\; (K_1 \cup K_2) \;\cup\; (K_1 \cup K_2 \cup K_3) \;\cup\; (K_1 \cup K_2 \cup K_3 \cup K_4) \;\cup\; \cdots.$$

d. The intersection of *finitely* many F_σ's is another F_σ; the union of *finitely* many G_δ's is another G_δ.

Hint: If $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ and $B_1 \subseteq B_2 \subseteq B_3 \subseteq \cdots$, show that $(\bigcup_{n=1}^\infty A_n) \cap (\bigcup_{n=1}^\infty B_n) = \bigcup_{n=1}^\infty (A_n \cap B_n)$.

e. In a pseudometric space, every closed set is a G_δ, and every open set is an F_σ.

Hint: If (X, d) is a pseudometric space and H is a closed subset, then H is the intersection of the open sets $\{x \in X : \operatorname{dist}(x, H) < 1/n\}$ (for $n = 1, 2, 3, \ldots$).

Meager Sets

20.3. Let X be a topological space; for sets $S \subseteq X$ let $\complement S = X \setminus S$. The **boundary** (or *frontier*) of a set S is the set $\operatorname{cl}(S) \cap \operatorname{cl}(\complement S)$; it is often denoted by ∂S or $\operatorname{bdry}(S)$ or $\operatorname{fr}(S)$. Note that $\partial S = \partial(\complement S)$. Show that X can be partitioned into the three disjoint sets

$$\operatorname{int}(S), \qquad \operatorname{int}(\complement S), \qquad \partial S,$$

which are open, open, and closed, respectively. Also, S is closed if and only if $\partial S \subseteq S$, and S is open if and only if ∂S is disjoint from S.

If S is a sufficiently "nice" set, then ∂S may be quite small, as in 20.4(C) and 20.4(D). However, slightly "nasty" sets may have boundaries that are quite large. For instance, in the real line, the set of rationals has boundary equal to the entire real line.

20.4. Let $(X, \mathcal{T})$ be a topological space, and let $\complement$ denote complementation in X. Let $S \subseteq X$. Then the following conditions are equivalent. If any (hence all) of them are satisfied, we say S is **nowhere-dense** (or *rare* or *nondense*).

(A) The closure of S has no interior; that is, $\operatorname{int}(\operatorname{cl}(S)) = \varnothing$.

(B) The complement of S contains an open dense set; that is, $\operatorname{cl}(\operatorname{int}(\complement S)) = X$.

(C) S is contained in the boundary of some open set.

(D) S is contained in the boundary of some closed set.

(E) Every nonempty open subset of X contains a nonempty open set that is disjoint from S. (In other words, there aren't any nonempty open sets in which the trace of S is dense. This explains the name "nowhere-dense.")

(F) $S \subseteq \mathrm{cl}(X \setminus \mathrm{cl}(S))$.

20.5. *Further properties and examples.*

a. Any subset of a nowhere-dense set is nowhere-dense.

b. If A and B are nowhere-dense, then $A \cup B$ is nowhere-dense. (Hence the nowhere-dense sets form an ideal; they may be viewed as the "small" sets for some purposes.)

Hints: We may replace A and B with their closures; hence we may assume A and B are closed. Now apply 15.13.c to $\complement A$ and $\complement B$.

c. Let X be a topological space, and let $X^{\mathbb{N}}$ have the product topology. Let $A \subseteq X^{\mathbb{N}}$ be nowhere-dense. By considering our characterization of the topology in terms of basic rectangles, we see that any finite sequence $s = (x_1, x_2, \ldots, x_m)$ in X can be extended to a longer sequence $s' = (x_1, x_2, \ldots, x_m, \ldots, x_n)$ having the property that no infinite sequence extending s' is a member of A.

20.6. Let X be a topological space. A set $S \subseteq X$ is **meager**, or of the *first category of Baire*, if it is the union of countably many nowhere-dense sets. A set that is not meager is called **nonmeager**, or of the *second category of Baire*. Thus, every set is of either the first or second category.

A set T is **comeager** (or *residual* or *generic*) if $X \backslash T$ is meager — equivalently, if T contains the intersection of countably many open dense sets.

Remarks. The collection of meager sets forms an ideal — in fact, it is a σ-ideal; i.e., it is closed under countable union, by 6.26. In the cases of greatest interest, X is a Baire space (see 20.15 and sections thereafter), hence X is not a meager subset of itself, and therefore the meager sets form a *proper* ideal. Thus

we may think of the meager sets as "small" and the comeager sets as "large,"

in the sense of in 5.3. Although "large" is a stronger property than "nonempty," in some situations the most convenient way to prove that some set S is nonempty is by showing the set is "large." That is one of the main ways in which the Baire Category Theorem (20.16) gets used.

20.7. *Some examples in $\mathbb{R}$ and other topological spaces.*

a. If x_0 is an isolated point in a topological space, then any set containing x_0 is nonmeager.

b. In the real line (or more generally, in any T_1 topological space that has no isolated points), every singleton $\{x\}$ is nowhere-dense, so every countable set is meager.

c. A countable subset of $\mathbb{R}$ must have empty interior. This will follow from 20.16, but it can also be proved directly by noting that any nondegenerate interval is uncountable.

d. A countable subset of $\mathbb{R}$ may or may not be nowhere-dense. For instance, $\mathbb{Z}$ is nowhere-dense, but $\mathbb{Q}$ is dense.

Generic Continuity Theorems

20.8. Baire-Osgood Equicontinuity Theorem. Let Ω be a topological space, let (X, d) be a pseudometric space, and let $f_1, f_2, f_3, \ldots$ be continuous functions from Ω to X. Assume that $\lim_{n\to\infty} f_n(\omega)$ exists in X for each $\omega \in \Omega$. Then the set

$$E \quad = \quad \Big\{\omega \in \Omega \ : \ (f_n) \text{ is equicontinuous at } \omega\Big\}$$

is comeager in Ω.

Remarks. By 18.32.a, the limit function $f(\omega) = \lim_{n\to\infty} f_n(\omega)$ is continuous at each point of E. The theorem is used mainly when Ω is a Baire space (discussed in 20.6 and defined in 20.15). In that setting, $\{f_n\}$ is equicontinuous and hence f is continuous, at "most" points of Ω. This theorem is a nonlinear version of the Banach-Steinhaus Uniform Boundedness Principle; some linear (or additive) versions of that principle are given in 23.13 and 27.26(U5).

Proof of theorem. For positive integers j and k, let

$$C_{j,k} \quad = \quad \bigcap_{m,n\geq k} \left\{\omega \in \Omega \ : \ d\Big(f_m(\omega), f_n(\omega)\Big) \leq \frac{1}{j}\right\}.$$

These sets are closed. Since the sequence $(f_m(\omega) : m \in \mathbb{N})$ is Cauchy for each fixed ω, for each fixed j we have $\Omega = \bigcup_{k=1}^{\infty} C_{j,k}$.

Define oscillation as in 18.28. We wish to show that the set $A = \{\omega \in \Omega \ : \ \mathrm{osc}_\Phi(\omega) > 0\}$ is meager. For positive integers j, let

$$B_j \quad = \quad \left\{\omega \in \Omega \ : \ \mathrm{osc}_\Phi(\omega) \geq \frac{1}{j}\right\};$$

then $A = \bigcup_{j=1}^{\infty} B_j = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} (B_j \cap C_{5j,k})$. It suffices to show that each of the closed sets $B_j \cap C_{5j,k}$ has empty interior.

Fix any j and k, and suppose G is a nonempty open subset of $B_j \cap C_{5j,k}$. Fix any $\omega_0 \in G$. The set $\{f_1, f_2, f_3, \ldots, f_k\}$ is a finite set of continuous functions, hence it is equicontinuous at ω_0. Thus ω_0 has some open neighborhood $N \subseteq G$ such that

$$\sup_{\omega\in N} d\Big(f_p(\omega), f_p(\omega_0)\Big) \quad \leq \quad \frac{1}{5j} \qquad\qquad \text{for } p = 1, 2, \ldots, k$$

and therefore

$$\sup_{\omega,\omega'\in N} d\Big(f_p(\omega), f_p(\omega')\Big) \quad \leq \quad \frac{2}{5j} \qquad\qquad \text{for } p = 1, 2, \ldots, k. \qquad (*)$$

Now, we know that $\sup_{m,n\geq k} d\Big(f_m(\omega), f_n(\omega)\Big) \leq \frac{1}{5j}$ for any $\omega \in C_{5j,k}$ by the definition of that set. Hence also $\sup_{m\geq k} d\Big(f_m(\omega), f_k(\omega)\Big) \leq \frac{1}{5j}$ for every $\omega \in C_{5j,k}$, hence for every

$\omega \in N$. Similarly, $d(f_m(\omega'), f_k(\omega')) \leq \frac{1}{5j}$ for all $m \geq k$. Combine these results with $(*)$ (applied at $p = k$) to obtain

$$d\Big(f_m(\omega), f_m(\omega')\Big) \quad \leq \quad \frac{4}{5j} \qquad \text{for all } \omega, \omega' \in N \text{ and } m \geq k.$$

Combine this result with $(*)$ and the fact that $\omega_0 \in B_j$ to obtain

$$\frac{1}{j} \quad \leq \quad \operatorname{osc}_\Phi(\omega_0) \quad \leq \quad \sup_{m \in \mathbb{N}} \operatorname{diam}(f_m(N)) \quad \leq \quad \frac{4}{5j} \quad < \quad \frac{1}{j},$$

a contradiction.

20.9. A Continuous Extension Theorem. Let X be a pseudometric space, and let $A \subseteq X$ be equipped with the relative topology. Let (Y, d) be a complete pseudometric space, and let $f : A \to Y$ be a continuous map. Then f can be extended to a continuous map $\widehat{f} : C \to Y$, for some G_δ-set C with $A \subseteq C \subseteq \operatorname{cl}(A)$. In fact, one such set is

$$C \quad = \quad \Big\{x \in \operatorname{cl}(A) \ : \ \text{the filterbase } f\Big(A \cap \mathcal{N}(x)\Big) \text{ converges}\Big\}.$$

Here $\mathcal{N}(x)$ denotes the neighborhood filter of x in X, and $f(A \cap \mathcal{N}(x)) = \{f(A \cap N) : N \in \mathcal{N}(x)\}$.

Proof of theorem (following Dugundji [1966]). Define C as above. Then A is dense in C, since $C \subseteq \operatorname{cl}(A)$. By 16.15, f can be extended to a continuous function from C to Y. It suffices to show that C is a G_δ-set in X. For each $n \in \mathbb{N}$, let

$$A_n \quad = \quad \left\{x \in \operatorname{cl}(A) \ : \ \operatorname{diam}\Big(f(A \cap U)\Big) < \frac{1}{n} \text{ for some } U \in \mathcal{N}(x)\right\}.$$

Since Y is complete, a filterbase converges in Y if and only if it is Cauchy; from this it follows that $C = \bigcap_{n=1}^{\infty} A_n$.

Note that if U is an open set with $\operatorname{diam}(f(A \cap U)) < \frac{1}{n}$ then $U \cap \operatorname{cl}(A) \subseteq A_n$. Any neighborhood of x contains an open neighborhood of x; from this it follows that A_n is open in $\operatorname{cl}(A)$. Thus $A_n = \operatorname{cl}(A) \cap G_n$ for some set G_n that is open in X. Hence $C = \operatorname{cl}(A) \cap \bigcap_{n=1}^{\infty} G_n$. The set $\operatorname{cl}(A)$ is a G_δ by 20.2.e; thus C is a G_δ.

20.10. Vidossich's Generic Fixed Point Theorem. Let (X, d) be a complete metric space, and let Ψ be a collection of continuous maps from X into X. Let Ψ be equipped with any metrizable topology stronger than the topology of uniform convergence on compact subsets of X.

Let $\Psi_u = \{f \in \Psi \ : \ f \text{ has a unique fixed point}\}$; define a mapping $\tau : \Psi_u \to X$ by letting $\tau(f)$ be the unique fixed point of f. Let Ψ_0 be the collection of those f's in Ψ_u with this further property:

> If (f_n) is a sequence converging in Ψ to f and x_n is a fixed point (not necessarily the only one) of f_n for $n = 1, 2, 3, \ldots$, then $\tau(f) = \lim_{n \to \infty} x_n$.

Then $\Psi_0 \subseteq \Psi^* \subseteq \Psi_u \subseteq \Psi$ for some set Ψ^* that is a G_δ-set in Ψ, and on which τ is continuous.

Remarks. In some cases of interest, Ψ_0 is dense in Ψ. In these cases, Ψ^* is comeager in Ψ, and so we reach this conclusion: "Most" of the continuous functions in Ψ have unique fixed points and have their fixed points depending continuously on the functions.

Proof of theorem. The map τ is continuous from Ψ_0 to X. By 20.9, τ has an continuous extension $\widehat{\tau} : \Psi_1 \to X$, where $\Psi_0 \subseteq \Psi_1 \subseteq \mathrm{cl}(\Psi_0)$ and Ψ_1 is a G_δ-set in Ψ.

We claim that

$$\widehat{\tau}(f) \text{ is a fixed point of } f$$

for each $f \in \Psi_1$. Indeed, since $\Psi_1 \subseteq \mathrm{cl}(\Psi_0)$, we can find a sequence (f_n) in Ψ_0 converging to f in Ψ. Let $x_n = \tau(f_n) = \widehat{\tau}(f_n)$ and $x = \widehat{\tau}(f)$; by the continuity of $\widehat{\tau}$ we have $x_n \to x$. Hence the set $K = \{x, x_1, x_2, x_3, \ldots\}$ is compact, and $f_n \to f$ uniformly on K. By the continuity of f we have $f(x_n) \to f(x)$. Then

$$\begin{aligned} d\Big(x, f(x)\Big) &\leq d(x, x_n) + d\Big(x_n, f_n(x_n)\Big) + d\Big(f_n(x_n), f(x_n)) + d\Big(f(x_n), f(x)\Big) \\ &\leq d(x, x_n) + 0 + \sup_{v \in K} d\Big(f_n(v), f(v)\Big) + d\Big(f(x_n), f(x)\Big) \quad \to \quad 0. \end{aligned}$$

This proves our claim.

It suffices to exhibit a G_δ set Ψ^* contained in $\Psi_1 \cap \Psi_u$, for on such a set we have $\tau = \widehat{\tau}$.

For $f \in \Psi_1$, the set $F(f) = \{\text{fixed points of } f\}$ is nonempty; let $\delta(f) = \mathrm{diam}_d\Big(F(f)\Big)$. Observe that $\delta(\cdot) = 0$ on Ψ_0.

Now consider Ψ_1 as a topological space, equipped with the relative topology. Next we claim that

$$\delta(\cdot) \text{ is continuous at each point of } \Psi_0.$$

In other words,

> for each $\varepsilon > 0$, each $f \in \Psi_0$ has an open neighborhood $V_{f,\varepsilon}$ in Ψ_1 on which $\delta(\cdot) \leq \varepsilon$.

Indeed, suppose not. Since the given topology on Ψ and on Ψ_1 is metrizable, there exists a sequence (f_n) in Ψ_1 converging to f with $\delta(f_n) > \varepsilon$. Then there exist $x_n, y_n \in F(f_n)$ with $d(x_n, y_n) > \varepsilon$. By our definition of Ψ_0, both the sequences (x_n) and (y_n) must converge to $\tau(f)$. But then $d(x_n, y_n) \to 0$, a contradiction. This proves the claim.

Now observe that $W_\varepsilon = \bigcup_{f \in \Psi_0} V_{f,\varepsilon}$ is an open set in Ψ_1 that contains Ψ_0, and on which $\delta(\cdot) \leq \varepsilon$. Hence $\Psi^* = \bigcap_{n=1}^\infty W_{1/n}$ is a G_δ-set in Ψ_1 that contains Ψ_0, and on which $\delta(\cdot) = 0$. Since Ψ^* is a G_δ-set in Ψ_1 and Ψ_1 is a G_δ-set in Ψ, it follows (easy exercise) that Ψ^* is a G_δ-set in Ψ.

Now consider any $f \in \Psi^*$. Then $F(f)$ is a nonempty set with diameter 0 (since $\Psi^* \subseteq \delta^{-1}(0) \cap \Psi_1$). Since X is a metric space, $F(f)$ is a singleton; hence $f \in \Psi_u$. Thus $\Psi^* \subseteq \Psi_1 \cap \Psi_u$. This completes the proof.

20.11. Corollary on Nonexpansive Mappings. Let (X, d) be a complete metric space, with the property that

(!) the identity map $i : X \to X$ can be approximated uniformly on X by a sequence of strict contractions $c_n : X \to X$.

Let Ψ be the set of all nonexpansive self-mappings of X, equipped with the topology of uniform convergence on X. Then there exists a set Ψ^* that is comeager in Ψ, such that each $f \in \Psi^*$ has a unique fixed point $\tau(f) \in X$ and the mapping $\tau : \Psi^* \to X$ is continuous. (Thus, *most* nonexpansive self-mappings of X have unique fixed points, which depend continuously on the mappings.)

Remark. Condition (!) is satisfied by bounded metric spaces that are not too irregularly shaped. For instance, it is satisfied if X is a closed bounded subset of a Banach space such that $X - x_0$ is a star set (in the sense of 12.3) for some $x_0 \in X$. Indeed, in that case we can take $c_n(x) = (1 - \frac{1}{n})x + \frac{1}{n}x_0$.

Proof of corollary. We shall apply 20.10. It suffices to show that the set Ψ_0, defined as in that theorem, is dense in Ψ under the hypotheses of the present corollary. If $f : X \to X$ is any nonexpansive mapping, then f is uniformly approximated by the mappings $c_n \circ f$, which are strict contractions. By 19.41, every strict contraction is a member of Ψ_0.

TOPOLOGICAL COMPLETENESS

20.12. A topological space $(X, \mathcal{T})$ is **topologically complete** (or *completely metrizable*) if its topology is pseudometrizable, and at least one of the pseudometrics that yields the topology $\mathcal{T}$ is complete. In describing a topologically complete space, we do not necessarily specify a particular pseudometric.

Caution: Some mathematicians apply the term "topologically complete" only to spaces that are metrizable — i.e., Hausdorff.

20.13. Alexandroff-Mazurkiewicz Theorem on Topological Completeness. Let (X, d) be a topologically complete Hausdorff space, and let $S \subseteq X$ have the relative topology. Then S is topologically complete if and only if S is a G_δ set in X — i.e., the intersection of countably many open subsets of X.

Proof. Let d be a complete metric on X. We first show that any open set $G \subseteq X$ is topologically complete. Verify that

$$e(s,t) \quad = \quad d(s,t) + \left| \frac{1}{\operatorname{dist}(s, X \setminus G)} - \frac{1}{\operatorname{dist}(t, X \setminus G)} \right| \qquad (s, t \in G)$$

is a complete metric on G that is topologically equivalent to the restriction of d.

Now suppose $S = \bigcap_{n=1}^\infty G_n$ is the intersection of countably many open sets. Then the product $P = \prod_{n=1}^\infty G_n$ has a topology that can be given by a complete metric, by 19.13. Let D be the diagonal set $\{(x_n) \in P : x_1 = x_2 = x_3 = \cdots\}$. Then D is a closed subset of P (why?), hence also complete. Finally, the mapping $s \mapsto (s, s, s, \ldots)$ is a homeomorphism from S onto D.

For the converse, suppose that $S \subseteq X$ is topologically complete. Let e be a complete metric on S that is topologically equivalent to the restriction of d. The identity map $i : (S, d) \to (S, e)$ is continuous, so by 20.9 it extends to a continuous map $\widehat{\imath} : C \to S$, where C is a G_δ-subset of X that contains S and is defined by

$$C \quad = \quad \Big\{ x \in \mathrm{cl}_X(S) \ : \ \text{the filterbase } i\Big(S \cap \mathcal{N}(x)\Big) \text{ converges in } S \Big\}.$$

It suffices to show that $C \subseteq S$. Let $x_0 \in C$; we wish to show that $x_0 \in S$. Each neighborhood of x_0 contains a member of $S \cap \mathcal{N}(x_0)$, so $S \cap \mathcal{N}(x_0)$ converges to x_0. By assumption, X is Hausdorff, so $S \cap \mathcal{N}(x_0)$ converges to no other limit. Since $x_0 \in C$, the filterbase $i(S \cap \mathcal{N}(x_0))$ converges in S. But i is just the identity map, so we have established that the filterbase $S \cap \mathcal{N}(x_0)$ converges in S. Thus its limit, x_0, lies in S.

20.14. *Example.* Show that the set $\mathbb{R} \setminus \mathbb{Q} = \{\text{irrational numbers}\}$, topologized as a subset of $\mathbb{R}$, is topologically complete.

Baire Spaces and the Baire Category Theorem

20.15. Let X be a nonempty topological space. Show that the following conditions on X are equivalent. If X possesses any one (hence all) of these properties, we say X is a **Baire space**.

(A) If $G_1, G_2, G_3, \ldots$ is a sequence of open dense subsets of X, then the set $\bigcap_{n=1}^{\infty} G_n$ is dense in X.

(B) If $F_1, F_2, F_3, \ldots$ is a sequence of closed subsets of X and $\bigcup_{n=1}^{\infty} F_n$ contains a nonempty open set, then at least one of the F_n's contains a nonempty open set.

(C) Any comeager subset of X is dense in X.

(D) Any meager subset of X has empty interior.

(E) Any nonempty open subset of X is nonmeager.

The last condition implies, in particular, that X itself is nonmeager, and hence the meager sets form a *proper* σ-ideal on X.

20.16. For our purposes, the most important result about Baire spaces is

> **(DC5) Baire Category Theorem.** Any complete pseudometric space is a Baire space.

For motivation the reader may wish to glance ahead to applications of this theorem, in 20.29, 23.13, 23.14, 23.15.b, 26.2, 27.18, and 27.25. We shall prove that the Baire Category

Theorem is an equivalent of the Principle of Dependent Choices, which was introduced in section 6.28.

Proof of (DC2) $\Rightarrow$ (DC5). Let (X, d) be a complete pseudometric space, and let any open dense sets $V_1, V_2, V_3, \ldots \subseteq X$ be given. We wish to show $\bigcap_{j=1}^{\infty} V_j$ is dense. Let G_0 be any nonempty open subset of X; we are to show that $\bigcap_{j=1}^{\infty} V_j$ meets G_0. We choose nonempty open sets $G_1, G_2, G_3, \ldots$ as follows: Assume G_{n-1} has already been chosen (this is clear for $n = 1$). Since V_n is open and dense, $G_{n-1} \cap V_n$ is a nonempty open set. Now (using the Principle of Dependent Choice) we may *choose* a nonempty open set G_n satisfying $\mathrm{cl}(G_n) \subseteq G_{n-1} \cap V_n$ and also satisfying

$$\mathrm{diam}(G_n) \quad < \quad \frac{1}{n}. \tag{$**$}$$

Let $K_n = \mathrm{cl}(G_n)$. Then $K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots$, and by Cauchy's Intersection Property (in 19.11.c) we have $\bigcap_{n=1}^{\infty} K_n$ nonempty. Since also $K_n \subseteq V_n$, this completes the proof.

Proof of (DC5) $\Rightarrow$ (DC1) (*optional*). This result is from Blair [1977]; it can also be found in Oxtoby [1980].

Let S be a set, and let $\Phi : S \to \{$nonempty subsets of $S\}$ be some given function. We wish to construct a choice sequence for Φ — i.e., a sequence (s_n) in S such that $s_{n+1} \in \Phi(s_n)$ for each $n \in \mathbb{N}$. Let S have the discrete metric, and let $X = S^{\mathbb{N}}$ have the product topology. Then S and X are both complete (see 19.13 for the latter). Verify that, for each $k \in \mathbb{N}$, the set

$$A_k \quad = \quad \bigcup_{l>k} \bigcup_{s \in S} \bigcup_{t \in \Phi(s)} \{x \in X \ : \ x_k = s, \ x_l = t\}$$

is open and dense in X. By the Baire Category Theorem, $\bigcap_{k=1}^{\infty} A_k$ is nonempty. Choose any $y \in \bigcap_{k=1}^{\infty} A_k$. Let $k_1 = 1$. Thereafter, let k_{i+1} be the first integer satisfying $k_{i+1} > k_i$ and $y_{k_{i+1}} \in \Phi(y_{k_i})$; such an integer exists since $y \in A_{k_i}$. The sequence $s_i = y_{k_i}$ is the desired choice sequence.

20.17. *Remark.* If X is also assumed *separable*, then the theorem above can be proved using just ZF; the Principle of Dependent Choice is not needed. (The details are left as an exercise.)

In particular, $2^{\mathbb{N}}$ is a Baire space, and that can be proved in ZF. This fact enters into some of our arguments about weak forms of Choice.

20.18. *Proposition.* Any locally compact regular space is a Baire space.

Proof. The proof is similar to that in 20.16, except that in place of $(**)$ we impose the condition that $\mathrm{cl}(G_n)$ be compact, and instead of the Cauchy Intersection Property 19.11.c we use 17.14.a and 17.3(B).

20.19. *Example.* Let $\mathbb{R} = \{$real numbers$\}$ and $\mathbb{Q} = \{$rational numbers$\}$ have their usual topologies; thus $\mathbb{Q}$ is a subspace of $\mathbb{R}$. Show that

a. $\mathbb{Q}$ is an F_σ subset of $\mathbb{R}$.

b. If $q \in \mathbb{Q}$, and N is a neighborhood of q in $\mathbb{Q}$, then N contains infinitely many members of $\mathbb{Q}$. Hence the singleton $\{q\}$, considered as a subset of $\mathbb{Q}$, is a closed set with empty interior.

c. $\mathbb{Q}$ is a meager subset of itself.

d. $\mathbb{Q}$ is not a Baire space.

e. $\mathbb{Q}$ is not topologically complete.

f. $\mathbb{Q}$ is not a G_δ subset of $\mathbb{R}$.

Almost Open Sets

20.20. Let X be a topological space. Observe that an equivalence relation $\approx$ can be defined on $\mathcal{P}(X)$ by:

$$A \approx B \qquad \Longleftrightarrow \qquad A \triangle B \text{ is meager.}$$

Here $\triangle$ denotes symmetric difference, as in 1.27.

Let $S \subseteq X$. Then the following conditions are equivalent:

(A) S is equivalent (in the sense defined above) to an open set; i.e., $S = G \triangle M$ for some open set G and some meager set M.

(B) S is equivalent (in the sense defined above) to a closed set; i.e., $S = F \triangle M$ for some closed set F and some meager set M.

(C) There exists a meager set $M \subseteq X$ such that $S \backslash M$ is a clopen subset of the topological space $X \backslash M$ (when that space is equipped with the relative topology).

(*Hint*: The boundary of an open set is meager — see 20.4.)

If any (hence all) of those conditions is satisfied, we say that S has the **Baire property**, or that S satisfies the **condition of Baire**, or that S is **almost open**. (Perhaps a more descriptive term would be "almost clopen.") Almost open sets play important roles in our theory of intangibles (in 14.77) and in our study of closed graph theorems; see 27.25, 27.45, and 29.38.

20.21. *Corollary.* The almost open subsets of a topological space X form a σ-algebra on X. Indeed, it is the smallest σ-algebra that contains both the σ-algebra of Borel sets and the ideal of meager sets; see 5.28.

20.22. Theorem (optional). A set has the Baire property if and only if it is equal to the union of a G_δ set and a meager set. (See also the related remark in 24.35.)

Proof of theorem. Since the sets with the Baire property form a σ-algebra containing all open sets and all meager sets, it follows easily that any union of a G_δ and a meager set is almost open. Conversely, suppose that $S = G \triangle M$ where G is open and M is meager.

Then M is contained in some meager set K that is an F_σ. Now $G \backslash K$ is a G_δ and $S \cap K$ is meager, and the union of these two sets is S.

20.23. *Definition and proposition.* Let X be a topological space, and let $f : X \to \mathbb{R}$ be some function. Then the following conditions are equivalent. If either, hence both, are satisfied, we say that f has the **property of Baire**.

(A) For each open set $G \subseteq \mathbb{R}$, the set $f^{-1}(G)$ is an almost open subset of X. (In other words, f is measurable when X is equipped with its σ-algebra of almost open sets and $\mathbb{R}$ is equipped with its σ-algebra of Borel sets.)

(B) There exists a meager set $M \subseteq X$ such that the restriction of f to $X \setminus M$ is continuous (when $X \setminus M$ is equipped with the relative topology).

Outline of proof of equivalence. The proof of (B) $\Rightarrow$ (A) is an easy exercise; we omit the details. For (A) $\Rightarrow$ (B), let (U_n) be a countable base for the topology of $\mathbb{R}$ — for instance, the open intervals with rational endpoints. Then $f^{-1}(U_n) = G_n \triangle M_n$ where G_n is open and M_n is meager. Then $M = \bigcup_{n=1}^{\infty} M_n$ is meager. Let r be the restriction of f to $X \setminus M$. For each n, the set $r^{-1}(U_n) = G_n \setminus M$ is open in $X \setminus M$; hence r is continuous. This proof is taken from Kuratowski [1948].

20.24. *An application of the Baire theory to Boolean algebras (optional).* Let $(X, \preccurlyeq)$ be any nondegenerate Boolean algebra. Then $(X, \preccurlyeq)$ can be embedded in a complete Boolean algebra $(X, \preccurlyeq)$ in a natural way (so that the inclusion is inf- and sup-preserving), as follows:

Let X^* be the dual of X, defined as in 13.19 and topologized as in 17.44. It is a Boolean space; hence it is a Baire space. Let $\mathcal{A} = \{$almost open subsets of $X^*\}$ and $\mathcal{M} = \{$meager subsets of $X^*\}$. Then $\mathcal{A}$ is an algebra of subsets of X^*, and $\mathcal{M}$ is an ideal in $\mathcal{A}$; hence we can form the quotient Boolean algebra $Y = \mathcal{A}/\mathcal{M}$ and the quotient mapping $\pi : \mathcal{A} \to Y$.

The dual of the Boolean space X^* is the Boolean algebra X^{**}, defined as in 17.44 — i.e., the algebra of all clopen subsets of X^*. The Stone mapping $S : X \to X^{**}$, defined in 13.21 and investigated further in 13.22 and 17.46.a, is an isomorphism of Boolean algebras. Every clopen set is open, and therefore is almost open; thus we have inclusions $X^{**} \xrightarrow{\subseteq} \mathcal{A} \xrightarrow{\subseteq} \mathcal{P}(X^*)$. The composition

$$\kappa : \quad X \quad \xrightarrow{S} \quad X^{**} \quad \xrightarrow{\subseteq} \quad \mathcal{A} \quad \xrightarrow{\pi} \quad Y$$

is a homomorphism of Boolean algebras.

It can be shown that Y is complete and that κ is injective, sup-preserving, and inf-preserving. We omit the details of the proof (which are too long for us to recommend them as an exercise); they can be found in Rasiowa and Sikorski [1963, page 89].

RELATIVIZATION

20.25. Assume that X is a topological space, and $Y \subseteq X$ is equipped with the relative topology. Use cl_X and cl_Y to denote the closures in X and in Y.

Let $S \subseteq Y \subseteq X$. Prove the following list of results. (The list is admittedly long and tedious, but that seems to be unavoidable, and these results are needed for later results such as 20.30 and 27.45.)

a. The following are equivalent:

(A) S is nowhere-dense in Y.

(B) $S \subseteq \mathrm{cl}_Y(Y \setminus \mathrm{cl}_Y(S))$. *Hint*: 20.4(F).

(C) $S \subseteq \mathrm{cl}_X(Y \setminus \mathrm{cl}_X(S))$. *Hint*: 15.12.

The last condition has the advantage that all the closures are with respect to the topology on X; this makes some later results easier to prove.

b. If S is nowhere-dense in Y, then S is nowhere-dense in X.

c. If S is meager in Y, then S is meager in X.

d. Suppose Y is dense in X, and S is nowhere-dense in X. Then S is nowhere-dense in Y.

Proof. Let "cl" denote closure in X. We have $X = \mathrm{cl}(Y)$ and $\mathrm{cl}(S) = \mathrm{cl}(\mathrm{cl}(S))$, hence by 15.5.c

$$X \setminus \mathrm{cl}(S) \quad = \quad \mathrm{cl}(Y) \setminus \mathrm{cl}(\mathrm{cl}(S)) \quad \subseteq \quad \mathrm{cl}(Y \setminus \mathrm{cl}(S)).$$

The right side is closed, so we may replace the left side by its closure — i.e., $\mathrm{cl}(X \setminus \mathrm{cl}(S)) \subseteq \mathrm{cl}(Y \setminus \mathrm{cl}(S))$. By 20.4(F) we have $S \subseteq \mathrm{cl}(X \setminus \mathrm{cl}(S))$. Thus we deduce $S \subseteq \mathrm{cl}(Y \setminus \mathrm{cl}(S))$. Now apply 20.25(C).

e. Suppose Y is dense in X, and S is meager in X. Then S is meager in Y.

f. Suppose Y is dense in X, and S is almost open in X. Then S is almost open in Y.

g. If S is almost open in Y and Y is almost open in X, then S is almost open in X.

Hint: By assumption, $S = M \triangle G$, where M is meager in Y — hence in X — and G is open in Y. Then $G = Y \cap H$, where H is open in X. Thus $S = M \triangle (Y \cap H)$, where all of M, Y, H are almost open in X. Since the almost open subsets of X form a σ-algebra, S is almost open in X.

h. Suppose X is a Baire space and $X \setminus Y$ is meager in X. Then S is almost open in Y if and only if S is almost open in X.

Almost Homeomorphisms

20.26. Recall from 17.41 that a zero-dimensional space is a topological space with a base of clopen sets. A few basic properties and examples were given in 17.42.

20.27. Alexandroff-Urysohn Theorem on the Irrationals (1928). Let X be a nonempty, separable, zero-dimensional, metrizable, topologically complete space, in which no nonempty clopen set is compact. Then X is homeomorphic to $\mathbb{N}^{\mathbb{N}}$.

In particular, $\mathbb{R} \setminus \mathbb{Q} = \{$the irrational numbers$\}$, topologized as a subset of $\mathbb{R}$, is homeomorphic to $\mathbb{N}^{\mathbb{N}}$.

Proof. We may equip X with a complete metric d. By 18.14, we may assume that $\operatorname{diam}(X) < 1$. We first shall show this preliminary result:

($\natural$) Let Y be a nonempty clopen subset of X, and let any $\varepsilon > 0$ be given. Then we may write $Y = Y_1 \cup Y_2 \cup Y_3 \cup \cdots$, where the Y_j's are nonempty, disjoint, clopen sets and $\operatorname{diam}(Y_j) < \varepsilon$.

By assumption, Y is not compact. Since Y is closed, we know by 19.18 that Y is not totally bounded. Thus (replacing ε by some smaller number if necessary) we may assume that Y cannot be covered by finitely many sets that have diameter less than ε. Since X is separable and zero-dimensional, X has a countable clopen base $A_1, A_2, A_3, \ldots$. Then the A_j's that have diameter less than ε also form a countable clopen base. Hence the sets $A_j \cap Y$ that satisfy $\operatorname{diam}(A_j) < \varepsilon$ form a clopen cover of Y. Let those sets be $B_1, B_2, B_3, \ldots$. Let $C_1 = B_1$ and

$$C_n \quad = \quad B_n \setminus (B_1 \cup B_2 \cup \cdots \cup B_{n-1}) \qquad\qquad (n = 2, 3, 4, \ldots).$$

Then the sets C_n are clopen, disjoint, have diameter less than ε, and have union equal to Y. Finally, let the sequence (Y_n) consist of those C_n's that are nonempty. There are infinitely many Y_n's, by our choice of ε. This completes the proof of ($\natural$).

We now define a mapping φ : $\{$finite sequences in $\mathbb{N}\} \to \{$nonempty clopen subsets of $X\}$, by recursion on the length of the finite sequence, as follows. First, let φ map the empty sequence to the whole set X itself.

Now, assume that φ has been defined on all sequences of length k, for some $k \geq 0$. Thus, for each $(n_1, n_2, \ldots, n_k) \in \mathbb{N}^k$, we already have $\varphi(n_1, n_2, \ldots, n_k)$ equal to some nonempty clopen subset of X. Applying ($\natural$), we can partition that nonempty clopen subset into a countably infinite collection of nonempty clopen subsets, each of which has diameter less than 2^{-k-1}. Take those sets to be the values of $\varphi(n_1, n_2, \ldots, n_k, p)$, for $p = 1, 2, 3, \ldots$. This completes the recursive definition of φ.

Next we define a function $\Phi : \mathbb{N}^{\mathbb{N}} \to X$, as follows. For any sequence $\sigma = (n_1, n_2, n_3, \ldots)$ in $\mathbb{N}^{\mathbb{N}}$, consider the sets $S_0(\sigma) = X$ and

$$S_1(\sigma) = \varphi(n_1), \qquad S_2(\sigma) = \varphi(n_1, n_2), \qquad S_3(\sigma) = \varphi(n_1, n_2, n_3), \qquad \ldots$$

These are clopen subsets of X, satisfying $S_0(\sigma) \supseteq S_1(\sigma) \supseteq S_2(\sigma) \supseteq S_3(\sigma) \supseteq \cdots$ and $\operatorname{diam}(S_n(\sigma)) < 2^{-n}$. Since the metric space X is complete, $\bigcap_{n=0}^{\infty} S_n(\sigma)$ consists of a singleton, whose element we now define to be the value of $\Phi(\sigma)$. Using 15.25.b and the fact that the $S_n(\sigma)$'s are clopen, verify that the mapping Φ is actually a homeomorphism from $\mathbb{N}^{\mathbb{N}}$ onto X.

20.28. Lemma. Let X be a nonempty, separable, zero-dimensional, complete metric space. Let $Y \subseteq X$ be a G_δ set that is dense in X, such that $X \setminus Y$ is also dense in X. Then Y is homeomorphic to the irrationals.

Proof. This result is from Mazurkiewicz [1917-1918]. The set Y is a separable metric space. It is zero-dimensional, for if $\{B_\alpha : \alpha \in A\}$ is a clopen base for X, then $\{B_\alpha \cap Y : \alpha \in A\}$

is a clopen base for Y. That Y is complete follows from Alexandroff's Theorem 20.13. We shall apply the Alexandroff-Urysohn Theorem 20.27; it suffices to show that no nonempty clopen subset of Y is compact (when we use the relative topology of Y).

Indeed, suppose K is a nonempty clopen compact set in Y, where we use the relative topology of Y; we shall obtain a contradiction. Since K is compact in Y, it is also compact in X; thus $\mathrm{cl}_X(K) = K$. Since K is open in Y, we have $K = G \cap Y$ for some nonempty set G which is open in X. Then $K = \mathrm{cl}_X(G \cap Y) \supseteq G$ by 15.13.b. Since $X \setminus Y$ is dense in X, the nonempty set G must meet $X \setminus Y$ — contradicting $G \subseteq K \subseteq Y$.

20.29. Theorem. Let X be a nonempty, complete, separable metric space, having no isolated points. Then there exists a meager set $M \subseteq X$ and a homeomorphism f from $X \setminus M$ onto the irrational numbers (where the irrationals are topologized as a subset of $\mathbb{R}$).

Proof. This is from Schechter, Ciesielski, Norden [1993]. For later reference we note that the proof of this theorem does not require the Axiom of Choice; at most, it requires DC.

Since X is a separable metric space, it has a countable base $B_1, B_2, B_3, \ldots$. Let D be the union of the boundaries of the B_j's; then D is meager. We easily verify that $X \setminus D$ is a nonempty, separable, zero-dimensional metric space. Moreover, it is a G_δ subset of a complete metric space; hence it is topologically complete by 20.13.

Let C be any countable dense subset of $X \setminus D$. Then any superset of C is also dense in $X \setminus D$. The set $M = C \cup D$ is meager in $X \setminus D$; hence $X \setminus M$ is dense in $X \setminus D$, by the Baire Category Theorem. Also, M is the union of countably many closed sets, so $X \setminus M$ is a G_δ set in $X \setminus D$. By the preceding lemma, $X \setminus M$ is homeomorphic to the irrationals.

20.30. Corollary. Let X_1 and X_2 be nonempty, complete, separable metric spaces, that have no isolated points. Then there exist meager sets $M_j \subseteq X_j$ $(j = 1, 2)$ such that $X_1 \setminus M_1$ is homeomorphic to $X_2 \setminus M_2$.

TAIL SETS

20.31. *Definitions.* We consider two different notions of "tail sets." We shall relate them in exercise 20.32.c, below. (However, these two notions are unrelated to a third meaning of the term, given in 7.7.)

a. We may sometimes write the set $2^{\mathbb{N}}$ as $\{0, 1\}^{\mathbb{N}}$, particularly if we want to emphasize that we are viewing it as a collection of sequences of 0s and 1s. A set $S \subseteq \{0, 1\}^{\mathbb{N}}$ is a **tail set** in $\{0, 1\}^{\mathbb{N}}$ if it has this property:

> Whenever $x = (x_1, x_2, x_3, \ldots)$ is a member of S, and $y = (y_1, y_2, y_3, \ldots)$ is another sequence of 0s and 1s that differs from x in only finitely many components, then y is also a member of S.

(The idea is that x and y are eventually the same; they have the same "tails.")

b. A **dyadic rational** will mean a number of the form $m/2^n$, for integers m and n. A set $S \subseteq [0,1)$ is a **tail set** in the interval $[0,1)$ if it has this property:

> Whenever x is a member of S and y is another point in $[0,1)$ that differs from x by a dyadic rational, then y is also a member of S.

20.32. *Exercises.* Show that

a. The two kinds of tail sets can also be described as follows: Say that two sequences of 0s and 1s are equivalent if they differ in only finitely many components; or say that two numbers in $[0,1)$ are equivalent if they differ by a dyadic rational. These are equivalence relations on $\{0,1\}^{\mathbb{N}}$ and on $[0,1)$, respectively. In either setting, a set is a tail set if and only if it is a union of equivalence classes.

b. The tail sets in $\{0,1\}^{\mathbb{N}}$ form an algebra of subsets of $\{0,1\}^{\mathbb{N}}$; the tail sets in $[0,1)$ form an algebra of subsets of $[0,1)$.

c. The countable sets

$$A \quad = \quad \left\{x \in \{0,1\}^{\mathbb{N}} \ : \ x_j = 0 \text{ for only finitely many } j\text{'s}\right\},$$

$$B \quad = \quad \left\{x \in \{0,1\}^{\mathbb{N}} \ : \ x_j = 1 \text{ for only finitely many } j\text{'s}\right\}$$

are tail sets in $\{0,1\}^{\mathbb{N}}$. The countable set

$$D \quad = \quad \{y \in [0,1) \ : \ y \text{ is a dyadic rational}\}$$

is a tail set in [0,1). Show that the mapping

$$(x_1, x_2, x_3, \ldots) \qquad \mapsto \qquad \sum_{n=1}^{\infty} \frac{x_n}{2^n} \quad = \quad \frac{x_1}{2} + \frac{x_2}{4} + \frac{x_3}{8} + \cdots$$

is a homeomorphism from $\{0,1\}^{\mathbb{N}} \setminus (A \cup B)$ onto $[0,1) \setminus D$, where $\{0,1\}$ has the discrete topology, $\{0,1\}^{\mathbb{N}}$ has the product topology, $[0,1)$ has its usual topology, and subsets have their relative topologies.

Show that a subset of $\{0,1\}^{\mathbb{N}} \setminus (A \cup B)$ is a tail set in $\{0,1\}^{\mathbb{N}}$ if and only if the corresponding subset of $[0,1) \setminus D$ is a tail set in $[0,1)$.

20.33. Oxtoby's Zero-One Law. In either $\{0,1\}^{\mathbb{N}}$ or $[0,1)$, if S is a tail set that has the Baire property, then S is either meager or comeager.

Proof. This proof is taken from Miller and Živaljević [1984]. We first prove this in $[0,1)$. By assumption, $S = G \triangle M$ where G is open and M is meager. We may assume G is nonempty (else $S = M$ and we are done). Take $P = G \cap M$; then P is a meager subset of G and $S \supseteq G \setminus P$.

Since G is a nonempty open set, it contains a set of the form $[2^{-m}(k-1), 2^{-m}k))$ for some positive integers m, k, which will be held fixed throughout the rest of this proof. Let

$$I_j \quad = \quad \left[\frac{j-1}{2^m}, \frac{j}{2^m}\right) \qquad \text{for} \qquad j = 1, 2, \ldots, 2^m.$$

With this notation, $G \supseteq I_k$. Let $P_k = I_k \cap P$; then P_k is a meager subset of I_k and $S \supseteq I_k \setminus P_k$.

Let P_j be the translate of P_k that is a subset of I_j — that is, let $P_j = P_k + 2^{-m}(j-k)$. Then P_j is also a meager set, since the operation of translation preserves all the relevant topological properties. Then $\bigcup_{j=1}^{2^m} P_j$ is a union of finitely many meager sets, and thus is meager. Since S is a tail set, we have

$$S \quad \supseteq \quad \bigcup_{j=1}^{2^m} (I_j \setminus P_j) \quad = \quad [0,1) \setminus \left(\cup_{j=1}^{2^m} P_j\right),$$

which is comeager. This completes the proof in $[0,1)$.

We can now use 20.32.c to transfer our conclusions to $\{0,1\}^{\mathbb{N}}$ as well. Admittedly, the mapping considered in 20.32.c is not a homeomorphism between $\{0,1\}^{\mathbb{N}}$ and $[0,1)$; it is only a homeomorphism between $\{0,1\}^{\mathbb{N}} \setminus (A \cup B)$ and $[0,1) \setminus D$. However, the exceptional sets A, B, D are meager tail sets and thus have no effect on our conclusion.

Baire Sets (Optional)

20.34. *Definition.* Let Ω be a locally compact Hausdorff space. Then the **Baire σ-algebra** on Ω is

> the σ-algebra $\mathcal{B}_1$ generated by the compact G_δ's in Ω; or, equivalently,
>
> the σ-algebra $\mathcal{B}_2$ generated by the continuous functions from Ω into $\mathbb{R}$ that have compact support — i.e., the smallest σ-algebra on Ω that makes all such functions measurable from Ω to $\mathbb{R}$ (where $\mathbb{R}$ is equipped with its Borel σ-algebra).

The members of this σ-algebra are called the **Baire sets**.

Proof of equivalence. To show that $\mathcal{B}_2 \subseteq \mathcal{B}_1$, let any continuous $f : \Omega \to \mathbb{R}$ with compact support be given. For each real number $r > 0$, the set $\{\omega \in \Omega : f(\omega) \geq r\}$ is compact; also it is the intersection of the sets $\{\omega \in \Omega : f(\omega) > r - \frac{1}{n}\}$ $(n = 1, 2, 3, \ldots)$, which are open. Thus the set $\{\omega \in \Omega : f(\omega) \geq r\}$ is a member of $\mathcal{B}_1$. Similarly, the set $\{\omega \in \Omega : f(\omega) \leq -r\}$ belongs to $\mathcal{B}_1$. It follows easily that f is measurable from $(\Omega, \mathcal{B}_1)$ to $\mathbb{R}$.

To show that $\mathcal{B}_1 \subseteq \mathcal{B}_2$, let K be a compact G_δ. Say $K = G_1 \cap G_2 \cap G_3 \cap \cdots$, where the G_j's are open. Since K is compact, finitely many of the G_j's suffice to cover K. By 17.14.c, there exists a continuous function $f_j : \Omega \to [0,1]$ with compact support, such that $f_j = 1$ on K and f_j vanishes outside G_j. Then each f_j is measurable from $\mathcal{B}_2$ to the reals. The characteristic function of K is the pointwise infimum of the sequence (f_j), so it too is measurable from $\mathcal{B}_2$ to the reals. Thus $K \in \mathcal{B}_2$; it follows that $\mathcal{B}_1 \subseteq \mathcal{B}_2$.

20.35. *Observations and remarks.* Every Baire set is a Borel set. In many commonly used topological spaces, the Baire and Borel sets are the same. For instance, this is true in any compact metric space, by 20.2.e and 17.7.f.

We shall not need the Baire sets later in this book. We have mentioned them only to prevent confusion: Do not confuse the Baire sets with the sets that have the Baire property; these are two different σ-algebras. In general we have

$$\left\{\begin{array}{c}\text{Baire}\\\text{sets}\end{array}\right\} \subsetneqq \left\{\begin{array}{c}\text{Borel}\\\text{sets}\end{array}\right\} \subsetneqq \left\{\begin{array}{c}\text{sets with}\\\text{Baire}\\\text{property}\end{array}\right\}.$$

The definition of "Baire set" varies somewhat in the literature. For instance, some mathematicians prefer to use the σ-ring generated by the compact G_δ's, rather than the σ-algebra (see 5.27). This has certain advantages in the study of regular measures on topological spaces; that is the main setting where Baire sets are important.

Chapter 21

Positive Measure and Integration

Measurable Functions

21.1. *Definitions, review, and remarks.* By a **measurable space** we mean a pair $(\Omega, \mathcal{S})$ consisting of a set Ω and a σ-algebra $\mathcal{S}$ of subsets of Ω; the members of $\mathcal{S}$ are then called **measurable sets**. Recall that measurable spaces are the objects for a category, with measurable mappings for the morphisms. A **measurable mapping** $f : (\Omega, \mathcal{S}) \to (\Omega', \mathcal{S}')$ is a function $f : \Omega \to \Omega'$ that makes the inverse image of each measurable set measurable — i.e., that satisfies $S' \in \mathcal{S}' \;\Rightarrow\; f^{-1}(S') \in \mathcal{S}$. Recall from 9.8 that a sufficient condition for measurability is that $G' \in \mathcal{G}' \;\Rightarrow\; f^{-1}(G') \in \mathcal{S}$, where $\mathcal{G}'$ is any collection of subsets of Ω' that generates the σ-algebra $\mathcal{S}'$.

Initial σ-algebras and *product σ-algebras* are defined as in 9.15 and 9.18.

In most cases of interest, the codomain Ω' is a topological space. When we speak of a measurable function from a measurable space to a topological space, the topological space Ω' will be understood to be equipped with its σ-algebra of Borel sets, unless some other arrangement is specified.

An analogous convention is not used for the domain Ω. For the most basic ideas of integration theory, developed later in this chapter, a topology is not needed on Ω. Even when a topology on Ω is present, several different σ-algebras on Ω may be useful (e.g., the Borel sets, the Lebesgue-measurable sets, or the almost open sets), and so we shall not assume any one of them is in use unless it is specified.

Most of our results about measurable functions $f : \Omega \to \Omega'$ require some sort of separability condition or small cardinality condition for the topological space Ω', as in 21.4 and 21.7. Without such assumptions, pathologies may arise, as in 21.8. An interesting exception is 21.3, which is valid for pseudometric spaces regardless of separability.

21.2. *Some exercises on measurability.* Assume $(\Omega, \mathcal{S})$ is a measurable space, X is a topological space, and $f : \Omega \to X$ is some mapping.

a. A sufficient condition for measurability of f is that the inverse image under f of each open set, or of each closed set, is measurable.

In particular, any continuous function is measurable, if the domain is a topological space equipped with its Borel σ-algebra.

b. If $f : \Omega \to X$ is measurable and $g : X \to Y$ is continuous, then the composition $g \circ f : \Omega \to Y$ is measurable.

c. A mapping $f : \Omega \to [-\infty, +\infty]$ is measurable if and only if the set $\{\omega \in \Omega : f(\omega) < r\}$ is measurable for each $r \in \mathbb{R}$ — or, equivalently, if and only if the set $\{\omega \in \Omega : f(\omega) \leq r\}$ is measurable for each $r \in \mathbb{R}$.

In particular, if Ω is a topological space equipped with its Borel σ-algebra, and $f : \Omega \to [-\infty, +\infty]$ is lower semicontinuous or upper semicontinuous, then f is measurable.

d. If $f : \Omega \to [-\infty, +\infty]$ is measurable, then so is the mapping $\omega \mapsto |f(\omega)|^p$, for any constant $p \in (0, +\infty)$.

21.3. Theorem. Let $(\Omega, \mathcal{S})$ be a measurable space, and let (X, d) be a pseudometric space (not necessarily separable). Let $f_1, f_2, f_3, \ldots$ be measurable functions from Ω into X, converging pointwise to a limit f. Then f is also measurable.

Proof (following Lang [1983]). Let any open set $T \subseteq X$ be given; we wish to show $f^{-1}(T) \in \mathcal{S}$. We may assume $\varnothing \subsetneqq T \subsetneqq X$. We have $x \in T$ if and only if $\operatorname{dist}(x, X \setminus T) > 0$. Consider the closed sets $F_p = \{x \in X \ : \ \operatorname{dist}(x, X \setminus T) \geq \frac{1}{p}\}$ and the open sets $G_p = \{x \in X \ : \ \operatorname{dist}(x, X \setminus T) > \frac{1}{p}\}$, for positive integers p. We have

$$x \in F_p \text{ for some } p \qquad \Rightarrow \qquad x \in T \qquad \Rightarrow \qquad x \in G_p \text{ for some } p.$$

(Actually, both of those implications are reversible, but we won't need that fact for the argument below.) For fixed p, use the facts that F_p is closed, G_p is open, and $f(\omega) = \lim_{n\to\infty} f_n(\omega)$, to show that

$$\begin{array}{lcl} f_n(\omega) \in F_p \text{ for all } n \text{ sufficiently large} & \Rightarrow & f(\omega) \in F_p, \text{ and} \\ f_n(\omega) \in G_p \text{ for all } n \text{ sufficiently large} & \Leftarrow & f(\omega) \in G_p. \end{array}$$

From this conclude that

$$\bigcup_{p=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{n=j}^{\infty} f_n^{-1}(F_p) \quad \subseteq \quad f^{-1}(T) \quad \subseteq \quad \bigcup_{p=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} f_n^{-1}(G_p). \qquad (*)$$

However, note that $G_p \subseteq F_p$. Hence the inclusions in $(*)$ are actually equalities, and therefore $f^{-1}(T)$ is measurable.

21.4. *Definitions and proposition.* Let $(\Omega, \mathcal{S})$ be a measurable space, and let (X, d) be a pseudometric space equipped with : [illegible] ebra of Borel sets.

We shall say that a mapping j [illegible] Y is **finitely valued** or **countably valued** or **separably valued** if the range c [illegible] nite set, a countable set, or a separable set, respectively.

Show that the following conditions are equivalent.

(A) f is separably valued and measurable.

(B) f is the uniform limit of a sequence (g_n) of countably valued, measurable functions.

(C) f is the pointwise limit of a sequence (g_n) of finitely valued, measurable functions.

(D) f is the pointwise limit of a sequence of separably valued, measurable functions.

Any function satisfying one, hence all, of these conditions will be called a **strongly measurable function**; the collection of all such functions will be denoted $SM(\mathcal{S}, X)$. Of course, if Y is separable, then $SM(\mathcal{S}, Y)$ is just the set of all measurable functions from $(\Omega, \mathcal{S})$ to the Borel subsets of Y. (The use of the term "strongly" is explained in part by comparison with the notion in 23.25.)

Further properties. If f is a strongly measurable function and the range of f is contained in a compact subset of X, then

(E) f is the uniform limit of a sequence (g_n) of finitely valued, measurable functions.

Proof of proposition. Obviously (B) $\Rightarrow$ (D) and (C) $\Rightarrow$ (D). For (D) $\Rightarrow$ (A), use 21.3; also show that $\text{Range}(f) \subseteq \text{cl}\left(\bigcup_{n=1}^{\infty} \text{Range}(g_n)\right)$, which is separable.

It remains to show that (A) implies both (B) and (C), and also that (A) implies (E) when the range of f is relatively compact. Let $(x_k : k = 1, 2, 3, \ldots)$ be a dense sequence in the range of f. For the compact case, choose the sequence (x_k) so that it has the further property that for each $\varepsilon > 0$, the range of f is covered by $\bigcup_{k=1}^{N} B(x_k, \varepsilon)$ for some $N \in \mathbb{N}$. Now, for (B) and (E), let $g_n(\omega)$ be the first term in the sequence (x_k) that satisfies $d(f(\omega), x_k) \leq 1/n$. For (C), let $g_n(\omega)$ be the closest member of $\{x_1, x_2, \ldots, x_n\}$ to $f(\omega)$ — or more precisely (since there may be a tie), let $g_n(\omega)$ be the first x_k in the finite sequence $(x_1, x_2, \ldots, x_n)$ that satisfies $d(f(\omega), x_k) = \text{dist}(f(\omega), \{x_1, x_2, \ldots, x_n\})$. We leave it as an exercise to prove that, in each case, g_n is measurable. (The exercise makes use of the fact that we took $g_n(\omega)$ to be the *first* x_k with a specified property, rather than simply *some* x_k with that property.)

21.5. *A related result.* If $f : \Omega \to [0, +\infty]$ is a measurable function, then there exist measurable, finitely valued functions $g_n : \Omega \to [0, +\infty)$ such that $g_n \uparrow f$ pointwise — i.e., such that $g_1 \leq g_2 \leq g_3 \leq \cdots \leq f$ and $g_n(\omega) \to f(\omega)$ for each $\omega \in \Omega$.

Hint: Let $g_n(\omega) = \max\{r : 2^n r \text{ is an integer}, r \leq n, \text{ and } r \leq f(\omega)\}$.

Joint Measurability

21.6. *Definition.* Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be measurable spaces. Recall that the **product σ-algebra** on $X \times Y$ is the product structure, as defined in 9.18, for the category of measurable spaces and measurable mappings. Thus, the product σ-algebra is the smallest σ-algebra on $X \times Y$ that makes both of the coordinate projections $(x, y) \mapsto x$ and $(x, y) \mapsto y$ into measurable mappings. Equivalently, it is the smallest σ-algebra that contains the collection

$\{S \times T : S \in \mathcal{S}, T \in \mathcal{T}\}$. We shall denote it by $\mathcal{S} \otimes \mathcal{T}$, to emphasize that in general it is not equal to $\{S \times T : S \in \mathcal{S}, T \in \mathcal{T}\}$.

Proposition on joint measurability. Let $(X, \mathcal{S})$, $(Y, \mathcal{T})$, and $(Z, \mathcal{U})$ be measurable spaces. Let $\mathcal{S} \otimes \mathcal{T}$ be the product σ-algebra on $X \times Y$. If $f : X \times Y \to Z$ is measurable from $\mathcal{S} \otimes \mathcal{T}$ to $\mathcal{U}$, then

(i) $f(x, \cdot) : Y \to Z$ is measurable from $\mathcal{T}$ to $\mathcal{U}$ (for each fixed x), and

(ii) $f(\cdot, y) : X \to Z$ is measurable from $\mathcal{S}$ to $\mathcal{U}$ (for each fixed y).

In particular, taking $Z = \{0, 1\}$, we obtain these important special cases: If $A \in \mathcal{S} \otimes \mathcal{T}$, then

(i′) the set $A_x = \{y \in Y : (x, y) \in A\}$ belongs to $\mathcal{T}$, for each fixed x, and

(ii′) the set $A^y = \{x \in X : (x, y) \in A\}$ belongs to $\mathcal{S}$, for each fixed y.

Proof. We first prove (i′). Fix any $x \in X$, and define $\mathcal{F}_x = \{A \subseteq X \times Y : A_x \in \mathcal{T}\}$. It is easy to show that $\mathcal{F}_x$ is a σ-algebra on $X \times Y$, and that $\mathcal{F}_x$ contains all sets of the form $S \times T$ for $S \in \mathcal{S}, T \in \mathcal{T}$. Hence $\mathcal{F}_x \supseteq \mathcal{S} \otimes \mathcal{T}$; this proves (i′). The proof of (ii′) is similar.

Now to prove (i), let any measurable $f : X \times Y \to Z$ and any measurable set $Q \subseteq Z$ be given. Then

$$f(x, \cdot)^{-1}(Q) \;=\; \{y \in Y : f(x, y) \in Q\} \;=\; \{y \in Y : (x, y) \in f^{-1}(Q)\} \;=\; \left[f^{-1}(Q)\right]_x$$

is measurable by (i′).

21.7. *Measurability in separable spaces.* Let (X, d) be a separable metric space. Then:

a. The product σ-algebra on $X \times X$ formed using the Borel σ-algebras on both X's is equal to the Borel σ-algebra determined by the product topology on X.

Hint: Let $\mathcal{B}$ be a countable base for the topology on X; then $\{B_1 \times B_2 \,:\, B_1, B_2 \in \mathcal{B}\}$ is a countable base for the product topology on $X \times X$.

b. The metric $d : X \times X \to [0, +\infty)$ is jointly measurable.

c. Let $(\Omega, \mathcal{S})$ be a measurable space. Consider mappings from Ω into $\mathbb{R}$ or $\mathbb{C}$ or $[0, +\infty]$ or $[-\infty, +\infty]$ (all of which are separable metric spaces). Then the maximum, minimum, sum, difference, product, or quotient of two measurable mappings is measurable, if it is defined — i.e., if it does not involve $\infty - \infty$ or division by 0 or other illegal operations. The set of all measurable real- or complex-valued functions is a unital algebra (see 11.3).

21.8. Nedoma's pathology (optional). Let $\mathcal{S}$ be a σ-algebra on a set X, with $\operatorname{card}(X) > \operatorname{card}(\mathbb{R})$. Then the diagonal set $I = \{(x, x) : x \in X\}$ is not a member of the product σ-algebra $\mathcal{S} \otimes \mathcal{S}$.

Corollaries. Let X be a set with $\operatorname{card}(X) > \operatorname{card}(\mathbb{R})$. Let d be a metric on X, let $\mathcal{S}$ be the resulting Borel σ-algebra on X, and let $\mathcal{S} \otimes \mathcal{S}$ be the product σ-algebra. Then the diagonal set I does not belong to $\mathcal{S} \otimes \mathcal{S}$. Hence

(i) $d : X \times X \to [0, +\infty)$ is not a measurable mapping, if we equip $X \times X$ with the product σ-algebra.

(ii) In view of 21.7.b, X cannot be equipped with a metric that makes X separable. (We already established this by other means in 15.37.a.)

(iii) On the other hand, the diagonal set I does belong to the σ-algebra determined by the product topology. Thus the functor that maps topologies to their Borel σ-algebras does not preserve product structures (as discussed in 9.35).

Proof of proposition. This proof is from Nedoma [1957]. Assume that $I \in \mathcal{S} \otimes \mathcal{S}$; we shall obtain a contradiction.

By definition in 9.18, $\mathcal{S} \otimes \mathcal{S}$ is the smallest σ-algebra on $X \times X$ that makes both of the coordinate projections measurable; this is the same as the smallest σ-algebra on $X \times X$ that contains all the sets of the form $E \times F$ for $E, F \in \mathcal{S}$.

By 5.26.h, we know that $\mathcal{S} \otimes \mathcal{S}$ is the union of the σ-algebras generated by countable subcollections of $\{E \times F : E, F \in \mathcal{S}\}$. In particular, since $I \in \mathcal{S} \otimes \mathcal{S}$, we know that I is a member of the σ-algebra $\mathcal{T}$ on $X \times X$ generated by $\{E \times F : E, F \in \mathcal{E}\}$ for some countable set $\mathcal{E} \subseteq \mathcal{S}$. Say $\mathcal{E} = \{E_1, E_2, E_3, \ldots\}$.

For each sequence $(\beta(1), \beta(2), \beta(3), \ldots)$ of 0s and 1s, form the set $D_\beta = \bigcap_{n=1}^{\infty} \complement^{\beta(n)} E_n$, where $\complement^1 E = \complement E = X \setminus E$ and $\complement^0 E = E$. Then the sets D_β (for $\beta \in 2^{\mathbb{N}}$) form a partition of X — that is, $D_\beta \cap D_{\beta'} = \varnothing$ for $\beta \neq \beta'$, and the union of the D_β's is X. (Some of the D_β's may be empty.) Each member of $\mathcal{E}$ is a union of D_β's.

The collection of sets $\{D_\beta \times D_\gamma : \beta, \gamma \in 2^{\mathbb{N}}\}$ is a partition of $X \times X$. Hence the sets of the form

$$M(A) \quad = \quad \bigcup_{(\beta,\gamma) \in A} (D_\beta \times D_\gamma) \qquad \text{for } A \subseteq 2^{\mathbb{N}} \times 2^{\mathbb{N}}$$

form a σ-algebra $\mathcal{T}'$ of subsets of $X \times X$. Note that $\{E \times F : E, F \in \mathcal{E}\} \subseteq \mathcal{T}'$, and therefore $\mathcal{T} \subseteq \mathcal{T}'$.

In particular, $I \in \mathcal{T}'$, and so I is of the form $M(A) = \bigcup_{(\beta,\gamma)\in A}(D_\beta \times D_\gamma)$ for some set $A \subseteq 2^{\mathbb{N}} \times 2^{\mathbb{N}}$. Moreover, any nonempty $D_\beta \times D_\gamma$ contained in I is actually of the form $D_\beta \times D_\beta = D_\beta^2$. Since

$$\mathrm{card}\left(\bigcup_{(\beta,\beta)\in A} D_\beta^2\right) = \mathrm{card}(I) = \mathrm{card}(X) > \mathrm{card}(\mathbb{R}) = \mathrm{card}(2^{\mathbb{N}} \times 2^{\mathbb{N}}) \geq \mathrm{card}(A)$$

at least one of the sets D_β^2 must contain more than one point of I. That is, there exists $(\beta, \beta) \in A$ and $u, v \in X$ with $u \neq v$ and

$$(u,u),\ (v,v) \quad \in \quad D_\beta^2 \quad = \quad \left(\bigcap_{n=1}^{\infty} \complement^{\beta(n)} E_n\right)^2.$$

Thus, for every n we have $u, v \in \complement^{\beta(n)} E_n$. We have $(u,v) \notin I$ since $u \neq v$. But

$$(u,v) \quad \in \quad \left(\bigcap_{n=1}^{\infty} \complement^{\beta(n)} E_n\right)^2 \quad = \quad D_\beta^2 \quad \subseteq \quad M(A) \quad = \quad I,$$

a contradiction. This completes the proof.

Positive Measures and Charges

21.9. *Remarks about positive charges.* Recall from 11.37 that a **positive charge** is a finitely additive mapping from an algebra of sets to $[0, +\infty]$, and a **positive measure** is a countably additive mapping from a σ-algebra of sets to $[0, +\infty]$. A *measurable space* is pair $(\Omega, \mathcal{S})$ consisting of a set Ω and a σ-algebra $\mathcal{S}$ of subsets of Ω; a **measure space** is a triple $(\Omega, \mathcal{S}, \mu)$, where the third component is a positive measure μ on $\mathcal{S}$.

Let $\mathcal{S}$ be an algebra of subsets of a set Ω. Note that if μ is a positive charge on $\mathcal{S}$, then

$$A \subseteq B, \qquad A, B \in \mathcal{S} \qquad \Rightarrow \qquad \mu(A) \leq \mu(B);$$

thus $\mu(A)$ is a measurement of how "big" A is. The largest value taken by μ is $\mu(\Omega)$. If $\mu(\Omega) < \infty$, the charge μ is said to be **finite** or **bounded**.

We can also use μ to measure how "big" is the difference between two sets. If μ is finite, verify that $d(A, B) = \mu(A \bigtriangleup B)$ is a pseudometric on $\mathcal{S}$; here $\bigtriangleup$ denotes symmetric difference. More generally, if μ is any positive charge (not necessarily bounded), then $d(A, B) = \arctan\left[\mu(A \bigtriangleup B)\right]$ defines a pseudometric d on $\mathcal{S}$. In place of the arctangent function, we could use any other bounded remetrization function; see 18.14.

Remark. We may define an equivalence relation $\approx$ on the algebra $\mathcal{S}$ using the pseudometric d as follows: $A \approx B \iff d(A, B) = 0$. For many purposes in analysis, equivalent sets can be used interchangeably; what is important is not the particular set but the equivalence class to which it belongs. The collection of all equivalence classes — i.e., the quotient $\mathcal{S}/\mu$ — is a Boolean algebra, on which d acts as a metric. This can be verified directly or by showing that $\mathcal{I} = \{A \in \mathcal{S} : \mu(A) = 0\}$ is an ideal in the Boolean algebra $\mathcal{S}$; then $\mathcal{S}/\mu = \mathcal{S}/\mathcal{I}$. The quotient Boolean algebra is sometimes called the **measure algebra**.

21.10. *More definitions.* By a **probability charge** or **probability measure** on Ω we shall mean a positive charge or measure that satisfies $\mu(\Omega) = 1$. If μ is a positive charge with $0 < \mu(\Omega) < \infty$, then most ideas about μ are unaffected if we replace μ with the probability charge ν defined by $\nu(S) = \mu(S)/\mu(\Omega)$. Thus, in many contexts we may restrict our attention to probabilities; this restriction often simplifies our notation.

Note that a probability charge μ has at least the two numbers 0, 1 in its range. A **two-valued probability** will mean a probability charge that has *only* the two values 0, 1 for its range.

21.11. *Some elementary examples.* If Ω is any set and $p : \Omega \to [0, +\infty]$ is any function, then $\mu(S) = \sum_{s \in S} p(s)$ defines a positive measure μ on the measurable space $(\Omega, \mathcal{P}(\Omega))$. A measure of this type will be called a **discrete measure**. Note that if $\mu(S) < \infty$, then $p(s)$ is nonzero for at most countably many points s in S; see 10.40. A few special kinds of discrete measures deserve further note:

a. Counting measure is the measure $\mu : \mathcal{P}(\Omega) \to \{0, 1, 2, \ldots, \infty\}$ obtained by using $p(s) = 1$ for all s. Thus, it is the discrete measure defined by

$$\mu(S) = \begin{cases} n & \text{if } S \subseteq \Omega \text{ is a finite set with } n \text{ elements} \\ +\infty & \text{if } S \text{ is an infinite subset of } \Omega. \end{cases}$$

Note that counting measure does not distinguish between different kinds of infinities. For instance, when μ is counting measure on $\mathbb{R}$, then $\mu(\mathbb{R}) = \mu(\mathbb{Z})$, even though $\mathrm{card}(\mathbb{R}) > \mathrm{card}(\mathbb{Z})$ by results of 10.44.f.

b. A **discrete probability measure** is a discrete measure that satisfies $\mu(\Omega) = 1$; that is, $\sum_{s\in\Omega} p(s) = 1$. Clearly, the function p must vanish everywhere outside some countable set.

A discrete probability measure on $\mathbb{N}$ may be described as a sequence (p_j) of nonnegative numbers that have sum equal to 1.

c. Let Ω be any set, and let $\xi \in \Omega$. Then the **unit mass at** ξ is the two-valued probability measure $\mu : \mathcal{P}(\Omega) \to \{0,1\}$ defined by

$$\mu(S) \quad = \quad 1_S(\xi) \quad = \quad \begin{cases} 1 & \text{if } \xi \in S \\ 0 & \text{if } \xi \notin S; \end{cases}$$

here 1_S is the characteristic function of S. Thus, μ is the discrete probability obtained by letting p be the characteristic function of the singleton $\{\xi\}$.

Exercise. Let μ_1, μ_2 be the unit masses at two distinct points ξ_1, ξ_2. Let $\mu(S) = \max\{\mu_1(S), \mu_2(S)\}$ for all $S \subseteq \Omega$. Show that μ is not a charge. Thus, *the setwise maximum of two measures* (*or charges*) *need not be a measure* (*or a charge*).

21.12. *Ultrafilters as charges.* Let Ω be a set, and let $\mathcal{F}$ be a collection of subsets of Ω. Let $1_{\mathcal{F}} : \mathcal{P}(\Omega) \to \{0,1\}$ be the characteristic function of $\mathcal{F}$; that is,

$$1_{\mathcal{F}}(S) = \begin{cases} 1 & \text{if } S \in \mathcal{F} \\ 0 & \text{if } S \notin \mathcal{F}, \end{cases}$$

for $S \subseteq \Omega$. Show that

a. $\mathcal{F}$ is an ultrafilter on Ω if and only if the function $1_{\mathcal{F}} : \mathcal{P}(\Omega) \to \{0,1\}$ is a two-valued probability charge.

b. $\mathcal{F}$ is the fixed ultrafilter at a point $\xi \in \Omega$ if and only if $1_{\mathcal{F}}$ is the unit mass at ξ.

c. $\mathcal{F}$ is a free ultrafilter on Ω if and only if $1_{\mathcal{F}}$ is a two-valued probability charge that vanishes on finite subsets of Ω.

d. If $\mathcal{F}$ is a free ultrafilter on $\mathbb{N}$, then $1_{\mathcal{F}}$ is a charge but not a measure — i.e., it is finitely additive but not countably additive.

21.13. *Preview.* Integrals will be defined later in this chapter. In 21.38(i) we shall prove that if $(\Omega, \mathcal{S}, \mu)$ is a measure space and $h : \Omega \to [0, +\infty]$ is measurable, then another positive measure ν can be defined by $\nu(S) = \int_S h(\omega)d\mu(\omega)$. Here is a typical example that is important in applications: Let m be a real number, and let s be a positive number. For Lebesgue-measurable sets $T \subseteq \mathbb{R}$, we may define

$$\nu(T) \quad = \quad \int_{-\infty}^{+\infty} \frac{1_T(x)}{s\sqrt{2\pi}} \exp\left[\frac{-(x-m)^2}{2s^2}\right] dx$$

where $1_T(\cdot)$ is the characteristic function of the set T and the integration is with respect to Lebesgue measure (defined later in this chapter). Then ν is the **Gaussian** (or **normal**) probability measure with **mean m** and **standard deviation s**. We omit further details; the interested reader is referred to any book on probability and statistics.

21.14. *Remark.* Additional examples of positive measures are given in 21.19 and 21.20.

Null Sets

21.15. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space. A set $N \subseteq \Omega$ is a **null set** (also known as a *negligible set*) if N is a subset of some measurable set that has measure 0. (The set N itself is not required to be measurable.) Note that the null sets form a σ-ideal (defined as in 5.2). It is a proper σ-ideal, except in the rather uninteresting case where $\mu(\Omega) = 0$.

A condition on points $\omega \in \Omega$ is said to hold **μ-almost everywhere**, commonly abbreviated **μ-a.e.**, if the set where it fails to hold is a null set. When μ is a probability measure, then other terms for "almost everywhere" are **almost surely** or **presque partout** or **with probability 1**, abbreviated *a.s.* or *p.p.* or *w.p. 1.*

Note that if $C_1, C_2, C_3, \ldots$ is a sequence of conditions, each of which holds μ-almost everywhere, then the condition

$$C_1 \text{ and } C_2 \text{ and } C_3 \text{ and } \cdots$$

also holds μ-almost everywhere, since the union of countably many null sets is a null set. We emphasize that if $\{C_\lambda : \lambda \in \Lambda\}$ is an *uncountable* collection of conditions, each of which holds μ-almost everywhere, it does not follow that the "and" of all the C_λ's necessarily holds μ-almost everywhere.

21.16. A measure space $(\Omega, \mathcal{S}, \mu)$ is said to be **complete** if every null set is measurable — i.e., if

$$A \subseteq B \subseteq \Omega, \quad B \in \mathcal{S}, \quad \mu(B) = 0 \qquad \Rightarrow \qquad A \in \mathcal{S}.$$

Example. All of our elementary measures in 21.11.a through 21.12 are complete, since they are defined on $\mathcal{P}(\Omega)$.

Not every measure space is complete, but every measure space $(\Omega, \mathcal{S}, \mu)$ can be extended to a complete measure space $(\Omega, \mathcal{S}', \mu')$, called its **completion**, in a natural way: Let $\mathcal{N}$ be the σ-ideal of null sets. The smallest σ-algebra that includes both $\mathcal{S}$ and $\mathcal{N}$ is (as in 5.28)

$$\mathcal{S}' \quad = \quad \mathcal{S} \,\triangle\, \mathcal{N} \quad = \quad \{S \,\triangle\, N \quad : \quad S \in \mathcal{S} \text{ and } N \in \mathcal{N}\}.$$

Then $\mu : \mathcal{S} \to [0, +\infty]$ extends uniquely to a complete measure $\mu' : \mathcal{S} \,\triangle\, \mathcal{N} \to [0, +\infty]$, defined by $\mu'(S \,\triangle\, N) = \mu(S)$. This measure is a complete extension of μ. In fact, it is the smallest complete extension, if we order measures by inclusion of graphs.

Exercise. Verify all the assertions above. In particular, show that if $S_1 \,\triangle\, N_1 = S_2 \,\triangle\, N_2$, then $\mu(S_1) = \mu(S_2)$.

21.17. Let $(\Omega, \mathcal{S}, \mu)$ be a positive measure space, and let $\mathcal{I}$ be the σ-ideal of null sets. Let X be any set. Two functions $f, g : \Omega \to X$ are **equal μ-almost everywhere** if the set where they differ is a null set — i.e., if $\{\omega \in \Omega : f(\omega) \neq g(\omega)\}$ is contained in a measurable set that has measure 0. It is easy to see that equality μ-almost everywhere is an equivalence relation on $X^\Omega = \{\text{functions from } \Omega \text{ into } X\}$; two functions f, g that are equal μ-almost everywhere may also be called **μ-equivalent**.

The set of equivalence classes is the reduced power $^*X = X^\Omega/\mathcal{I}$, defined as in 9.41. (In general it is not an ultrapower, since $\mathcal{I}$ generally is not a maximal ideal — i.e., not every subset of Ω is either a null set or the complement of a null set.) The members of $^*X$ — that is, the equivalence classes — are sometimes called X-valued **random variables**, especially if μ is a probability measure. In particular, members of $^*\mathbb{R}$ are **real random variables**.

For many purposes, any function can be replaced with any μ-equivalent function. Consequently, we may often identify μ-equivalent functions. By a slight abuse of notation, sometimes we may discuss a μ-equivalence class of functions as if it were a function. In such a context, a function may be defined arbitrarily on any null set or even left undefined on a null set. In particular, if f and g are μ-equivalent, X is a metric space, and f is strongly measurable (defined in 21.4), then for many purposes we may treat g as if it were strongly measurable.

A function $f : \Omega \to X$ is called **almost separably valued** (with respect to a given measure μ) if it is μ-equivalent to a separably valued function — that is, if by altering f on a set of measure 0 we can make it into a separably valued function. For most purposes, an almost separably valued function is just as good as a separably valued function.

If X is a pseudometric space, then we shall abbreviate

$$SM(\mu, X) \quad = \quad \{\varphi \in {}^*X \ : \ \varphi \text{ meets } SM(\mathcal{S}, X)\}.$$

In other words, a member of $SM(\mu, X)$ is a μ-equivalence class of functions that contains at least one strongly measurable function from $\mathcal{S}$ into X.

Exercise. Suppose X is a separable metric space, the measure space $(\Omega, \mathcal{S}, \mu)$ is complete (as defined in 21.16), and two functions $f, g : \Omega \to X$ are μ-equivalent. Then one of those functions is strongly measurable if and only if the other one is.

21.18. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let X be a topological space. Let (f_α) be a net in X^Ω, and let $f \in X^\Omega$. We say that (f_α) converges **μ-almost everywhere** to f if the statement $f_\alpha(\cdot) \to f(\cdot)$ is valid almost everywhere — i.e., if the set $\{\omega \in \Omega : f_\alpha(\omega) \not\to f(\omega)\}$ is contained in a measurable set that has measure 0. This condition is also written in various other ways, such as $f_\alpha \to f \quad \mu\text{-a.e.}$, or $f_\alpha \xrightarrow{\mu\text{-a.e.}} f$. The "$\mu$" may be omitted from the notation if it is understood. This convergence is also called **pointwise almost everywhere**, **almost surely** (**a.s.**), or **presque partout** (**p.p.**).

It is easy to verify that this convergence is centered and isotone, as defined in 7.34. However, it is not necessarily topological, or even pretopological; we shall prove that in 21.33.c.

In most cases of interest, the net (f_α) is actually a sequence. Then almost everywhere convergence is invariant under μ-equivalence: If f_n is μ-equivalent to g_n and f is μ-equivalent to g, then

$$f_n \to f \quad \mu\text{-a.e.} \qquad \Longleftrightarrow \qquad g_n \to g \quad \mu\text{-a.e.}$$

Thus, convergence almost everywhere makes sense for sequences of random variables.

Lebesgue Measure

21.19. *Preview of Lebesgue measure.* If $I_1, I_2, \ldots, I_n$ are intervals in $\mathbb{R}$, then the n-dimensional **Borel-Lebesgue measure** of the "box"

$$B \quad = \quad I_1 \times I_2 \times \cdots \times I_n$$

is the product of the lengths of those intervals. Here it is understood that the empty set and a singleton are intervals of length 0, an unbounded interval has length $+\infty$, and 0 times ∞ equals 0.

The **Borel σ-algebra** in $\mathbb{R}^n$ is the smallest σ-algebra $\mathcal{S}$ that contains all such boxes B; it is also equal to the smallest σ-algebra $\mathcal{S}$ that contains all open sets. The volume function for boxes extends uniquely to a measure μ on $\mathcal{S}$; that measure is called **Borel-Lebesgue measure**. Thus, we may think of this measure as the "volume" of a subset of $\mathbb{R}^n$. The uniqueness of Borel-Lebesgue measure follows easily (*exercise*) from 21.28.

Existence of one-dimensional Borel-Lebesgue measure will be proved in 24.35. Then n-dimensional Borel-Lebesgue measure is the product of n copies of one-dimensional Borel-Lebesgue measure, using the product construction given in 21.40. Then we can take the completion of n-dimensional Borel-Lebesgue measure, as in 21.16; the resulting measure is called n-dimensional **Lebesgue measure** and the members of the resulting σ-algebra are called **Lebesgue-measurable sets**. (They should not be called "Lebesgue sets" — that term unfortunately has another meaning, given in 25.16.)

Further properties of Lebesgue measure. As we have indicated, the volume function extends in a natural way, from boxes to a much larger collection of sets. Surprisingly, the volume function *cannot* be extended in a natural way to *all* subsets of $\mathbb{R}^n$; we shall prove that fact in 21.22. Thus, we can discuss the volume of a Lebesgue-measurable set, but not the volume of an arbitrary subset of $\mathbb{R}^n$. That is our main reason for studying σ-algebras.

It is easy to show (*exercise*) that the n-dimensional Lebesgue measure of any countable subset of $\mathbb{R}^n$ is zero. Some uncountable sets also have Lebesgue measure 0; we give examples in 24.39 and 25.19. In fact, 24.39 is an example of a comeager set with measure 0. Thus, a set can be "large" in one sense and "small" in another sense.

In 23.16 we shall introduce integrals $\int_\Omega f d\mu$ with respect to positive measures. The Lebesgue integral $\int_\Omega f d\mu$ is equal to the Riemann integral $\int_{\mathbb{R}^n} f(x)dx$ when both are defined; we shall prove that fact in 24.36.

21.20. Lebesgue measure in n dimensions is (i) translation invariant, and (ii) positive on each ball of positive radius. Banach spaces are a natural generalization of $\mathbb{R}^n$ introduced in the next chapter, but it is easy to show (using 23.22) that no positive measure on the Borel subsets of an infinite-dimensional Banach space can satisfy both (i) and (ii).

If we do not insist on translation-invariance, some interesting and useful measures do exist on infinite-dimensional spaces. The most famous of these is **Wiener measure**, which we shall now introduce briefly. (The details omitted here are major ones, not intended as an exercise.)

Let $C[0,1]$ be the Banach space of continuous functions from $[0,1]$ into $\mathbb{R}$, with its usual sup norm (discussed in 22.15), and let $\Omega = \{\omega \in C[0,1] : \omega(0) = 0\}$; this is an infinite-dimensional vector space. Whenever $[a_1, b_1]$, $[a_2, b_2]$, ..., $[a_n, b_n]$ are subintervals of $\mathbb{R}$ and $0 = t_0 < t_1 < t_2 < \cdots < t_n = 1$, the measure of the "box"

$$B \quad = \quad \bigcap_{i=1}^{n} \Big\{ \omega \in \Omega \ : \ \omega(t_i) - \omega(t_{i-1}) \in [a_i, b_i] \Big\}$$

is defined to be

$$\mu(B) \quad = \quad \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \int_{a_i}^{b_i} \exp\left[\frac{-x^2}{2(t_i - t_{i-1})}\right] dx.$$

In the terminology of probability, this says that the increments $\omega(t_i) - \omega(t_{i-1})$ are independent Gaussian random variables with mean 0 and variance $t_i - t_{i-1}$. It can be shown that μ extends to a probability measure on the σ-algebra generated by these "boxes."

The functions $\omega : [0, 1] \to \mathbb{R}$ correspond to the continuous but erratic paths taken by a particle starting from the origin and exhibiting Brownian motion. For any measurable set $S \subseteq C[0, 1]$, the measure $\mu(S)$ represents the probability that the path ω taken by such a particle will be an element of S. Although each path ω is continuous, it can be shown that with probability 1 the continuous function ω is *nowhere-differentiable.* (It is interesting to compare this result with 25.14(ii).)

More about Wiener measure can be found in Freedman [1971] or Kuo [1975].

21.21. (*Optional.*) Although a translation-invariant Lebesgue measure does not generalize naturally to infinite-dimensional spaces, a translation-invariant notion of *sets of Lebesgue measure 0* can be extended to that setting.

Definition. Let X be an Abelian group topologized by a complete metric that is translation-invariant (i.e., satisfying $d(x + u, y + u) = d(x, y)$; such metrics will be studied further in the next chapter). A set $S \subseteq X$ is called **shy** if there exists a positive measure μ on the Borel sets of X, with these properties:

(i) $0 < \mu(K) < \infty$ for some compact set $K \subseteq X$; and

(ii) there exists a Borel set B with $S \subseteq B \subseteq X$, such that $\mu(B + x) = 0$ for every $x \in X$.

The complement of a shy set is a **prevalent** set.

We emphasize that different shy sets may be exhibited using different measures μ, which may be chosen with particular applications in mind. For instance, if X is an infinite-dimensional vector space, then μ could be a Lebesgue measure on some finite-dimensional subset of X.

Following are some basic properties of shy sets and prevalent sets. We omit the proofs, which can be found in Hunt, Sauer, and Yorke [1992].

a. The shy sets form a proper σ-ideal; the prevalent sets form the corresponding proper δ-filter. Thus the shy sets form a collection of "small" sets, and the prevalent sets form a collection of "large" sets, in the sense of 5.3.

b. If S is a shy set (respectively, a prevalent set), then each of its translates $S + x$ is shy (respectively, prevalent).

c. If $X = \mathbb{R}^n$, then a set is shy if and only if its n-dimensional Lebesgue measure is zero.

d. Every shy set has empty interior; every prevalent set is dense in X.

e. If X is an infinite-dimensional Banach space (or, more generally, an infinite-dimensional F-space), then every compact set is shy.

Further remarks. Although "shy" is a more complicated notion than "meager," it plays a similar role and may be more natural for some measure-theoretic questions about "small" subsets of a topological vector space.

Here is one particularly interesting application: By 25.16 and 29.36, we know that any Lipschitzian function from $\mathbb{R}$ into $\mathbb{R}$ is differentiable almost everywhere. A generalization due to Rademacher (not proved in this book) says that any Lipschitzian function from $\mathbb{R}^n$ into $\mathbb{R}^n$ is differentiable except on a set whose n-dimensional Lebesgue measure is 0. Generalizations due to Christensen and others extend Rademacher's result to infinite-dimensional Banach spaces, with differentiable replaced by slightly weaker notions and with Lebesgue measure 0 replaced by shyness or similar notions.

Many applications and further references are listed in Hunt, Sauer, and Yorke [1992, 1993].

21.22. Vitali's Theorem. There exist subsets of $\mathbb{R}$ that are not Lebesgue-measurable. Furthermore, one-dimensional Lebesgue measure cannot be extended to a translation-invariant measure on *all* the subsets of $\mathbb{R}$.

Proof. Suppose μ were such a measure; we shall obtain a contradiction. Consider $\Omega = [0,1)$ as the circle group, i.e., the reals modulo 1 (introduced in 8.10.e). Then μ also acts as a translation-invariant measure defined on all subsets of $\Omega = [0,1)$, with $\mu(\Omega) = 1$.

Define an equivalence relation on $[0,1)$ by $x \approx y$ if $x - y$ is a rational number. Let S be a set consisting of one element from each equivalence class. (The existence of such a set S follows from the Axiom of Choice or from the slightly weaker principle (ACR) given in 6.12.)

By 8.23.a, the rationals are countable. Let $r_1, r_2, r_3, \ldots$ be an enumeration of the rationals in $[0,1)$. Then the sets $r_j + S$, for $j = 1, 2, 3, \ldots$, form a partition of Ω, and they have the same measure by translation invariance. Hence

$$1 \quad = \quad \mu([0,1)) \quad = \quad \sum_{j=1}^{\infty} \mu(r_j + S) \quad = \quad \sum_{j=1}^{\infty} \mu(S)$$

— but there is no number $\mu(S)$ in $[0, +\infty]$ that can satisfy this condition.

21.23. *Further remarks on extensions of Lebesgue measure.* Vitali's result, above, shows that Lebesgue measure cannot be extended to a translation-invariant measure on all the subsets of $\mathbb{R}$. Actually, it also shows that

> n-dimensional Lebesgue measure cannot be extended to a translation-invariant measure on all the subsets of $\mathbb{R}^n$, for any $n \geq 1$.

(Indeed, if μ were such a measure on $\mathbb{R}^n$ for some $n > 1$, then $\nu(S) = \mu\left(S \times [0,1)^{n-1}\right)$ would define such a measure on $\mathbb{R}$.)

What about if we do not require countable additivity? For any $n \geq 1$, there *does* exist a positive charge μ on the collection of all subsets of $\mathbb{R}^n$, which agrees with Lebesgue measure on the Lebesgue-measurable sets; that fact will follow easily from 29.32. The existence proof will use the Hahn-Banach Theorem, a weak form of the Axiom of Choice. Of course, the charge μ "constructed" in this fashion cannot be countably additive, as we have noted in the preceding paragraph.

On the other hand, what about translation invariance? It can be proved that, on $\mathbb{R}^1$ or $\mathbb{R}^2$, there exists a positive charge μ that agrees with Lebesgue measure on the Lebesgue-measurable sets and is translation-invariant; in fact, μ can be chosen to be invariant under isometries. (This includes not only translations, but also reflections, and — in the case of two dimensions — rotations.) The proof is longer and will not be given here; it can be found in Wagon [1985].

In three or more dimensions, such an invariant charge does not exist. Indeed, in three dimensions, this is an obvious consequence of the Banach-Tarski Decomposition, which was described (but not proved) in 6.16. In $n \geq 4$ dimensions, we may reason as follows: If μ is a positive charge on defined on all the subsets of $\mathbb{R}^n$, invariant under isometry, and extending Lebesgue measure, then $\nu(S) = \mu\left(S \times [0,1]^{n-3}\right)$ defines such a charge ν on $\mathbb{R}^3$, contradicting our previous remark.

For more about pathological charges, see Moore [1983] and Wagon [1985].

Some Countability Arguments

21.24. Let $\mathcal{S}$ be an algebra of subsets of a set Ω, and let μ be a positive charge on $\mathcal{S}$. We say that μ is ***σ*-finite** if $\Omega = \bigcup_{j=1}^{\infty} \Omega_j$ for some sequence of sets $\Omega_1, \Omega_2, \Omega_3, \ldots$ in $\mathcal{S}$, each of which satisfies $\mu(\Omega_j) < \infty$. Two important examples are Lebesgue measure on $\mathbb{R}^m$ and counting measure on $\mathbb{N}$.

Some basic properties. Suppose μ is a σ-finite charge on an algebra $\mathcal{S}$ of subsets of Ω. Show that

a. We can choose (Ω_j) to be an increasing sequence — i.e., to satisfy $\Omega_1 \subseteq \Omega_2 \subseteq \Omega_3 \subseteq \cdots$.

Proof. Replace (Ω_j) with the sequence (S_j), where $S_j = \Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_j$.

b. Or, if we prefer, we can choose the Ω_j's to be disjoint.

Proof. Assume $(\Omega_j : j \in \mathbb{N})$ is an increasing sequence; let $\Omega_0 = \varnothing$; then use the sequence (S_j), where $S_j = \Omega_j \setminus \Omega_{j-1}$.

c. If μ is a σ-finite measure, then we can write $\mu = \sum_{j=1}^{\infty} \mu_j$ where each μ_j is a finite measure.

Proof. Let the Ω_j's be disjoint, and let $\mu_j(S) = \mu(S \cap \Omega_j)$.

d. If μ is a σ-finite measure, then we can construct a probability measure ν that is positive on the same sets as μ.

Proof. Write $\mu = \sum_{j=1}^{\infty} \mu_j$ as above, and then let

$$\nu(S) \quad = \quad \sum_{j=1}^{\infty} \frac{\mu_j(S)}{2^j \mu_j(\Omega)} \qquad\qquad (S \in \mathcal{S}).$$

Remarks. For many purposes in measure and integration theory, a σ-finite charge is "as good as" a finite one. To prove some result for Ω, we can partition Ω into disjoint sets Ω_j and prove the result on each of those; putting all the pieces back together is then generally easy. This procedure can be applied in a mechanical way in proofs of many theorems.

21.25. *Some properties of positive charges and measures.* Let $\mathcal{S}$ be an algebra of subsets of a set Ω.

a. Let $\mu : \mathcal{S} \to [0, +\infty]$ be a positive charge. Then μ is countably additive if and only if it satisfies this condition:

> **Convergence Property for Increasing Sequences.** Whenever (A_n) is a sequence in $\mathcal{S}$ with $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ and with $\bigcup_{n=1}^{\infty} A_n \in \mathcal{S}$, then $\mu(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n) = \sup_{n\in\mathbb{N}} \mu(A_n)$.

b. Let $\mu : \mathcal{S} \to [0, +\infty)$ be a positive, *bounded* charge. Then μ is countably additive if and only if it satisfies this condition:

> **Convergence Property for Decreasing Sequences.** Whenever (B_n) is a sequence in $\mathcal{S}$ with $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$ and with $\bigcap_{n=1}^{\infty} B_n \in \mathcal{S}$, then $\mu(\bigcap_{n=1}^{\infty} B_n) = \lim_{n\to\infty} \mu(B_n) = \inf_{n\in\mathbb{N}} \mu(B_n)$

or, equivalently, this condition:

> **Property for Decreasing Free Sequences.** Whenever (B_n) is a sequence in $\mathcal{S}$ with $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$ and with $\bigcap_{n=1}^{\infty} B_n = \varnothing$, then $0 = \lim_{n\to\infty} \mu(B_n) = \inf_{n\in\mathbb{N}} \mu(B_n)$.

We cannot omit the assumption that $\mu(\Omega) < \infty$. For example, take μ to be counting measure on $\mathbb{N}$, and let $B_n = \{n, n+1, n+2, \ldots\}$.

c. Let μ be a bounded positive measure on a measurable space $(\Omega, \mathcal{S})$. Let (S_n) be any sequence of sets in $\mathcal{S}$. Define $\limsup_{n\to\infty} S_n$ and $\liminf_{n\to\infty} S_n$ as in 7.48. Show that

$$\mu\left(\liminf_{n\to\infty} S_n\right) \;\leq\; \liminf_{n\to\infty} \mu(S_n) \;\leq\; \limsup_{n\to\infty} \mu(S_n) \;\leq\; \mu\left(\limsup_{n\to\infty} S_n\right).$$

Hence, if $S_n \to S$ (in the sense of 7.48), then $\mu(S_n) \to \mu(S)$.

d. Let $(\mu_\delta : \delta \in \Delta)$ be a net of positive charges on $\mathcal{S}$, and assume that for each $S \in \mathcal{S}$ the net $(\mu_\delta(S) : \delta \in \Delta)$ increases to a limit $\mu(S)$ in $[0, +\infty]$. Then μ is also a positive charge on $\mathcal{S}$. If each μ_δ is countably additive, then so is μ.

Hint: Use 21.25.a; observe that $\sup_\delta \sup_n \mu_\delta(A_n) = \sup_n \sup_\delta \mu_\delta(A_n)$.

21.26. Approximation Lemma. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let $\mathcal{A}$ be an algebra of subsets of Ω that generates the σ-algebra $\mathcal{S}$. Assume μ is σ-finite on $\mathcal{A}$ — that is, assume Ω can be written as the union of countably many members of $\mathcal{A}$, each of which has finite measure.

Then $\mathcal{A}$ is dense in $\mathcal{S}$, in this sense: If $S \in \mathcal{S}$ with $\mu(S) < \infty$, and any number $\varepsilon > 0$ is given, then there exists a set $A \in \mathcal{A}$ with $\mu(A \triangle S) < \varepsilon$.

Proof. We first prove the proposition under the additional assumption that $\mu(\Omega) < \infty$. The measure μ defines a pseudometric d on $\mathcal{S}$ by $d(S,T) = \mu(S \triangle T)$. Let $\mathrm{cl}(\mathcal{A})$ be the closure of the set $\mathcal{A}$ in the pseudometric space $(\mathcal{S}, d)$. Observe that if (S_n) is an increasing sequence in $\mathcal{S}$ with union S, or (S_n) is a decreasing sequence in $\mathcal{S}$ with intersection S, then $d(S_n, S) \to 0$. Hence any closed subset of $(\mathcal{S}, d)$ is a monotone class. In particular, $\mathrm{cl}(\mathcal{A})$ is a monotone class. By the Monotone Class Theorem 5.29, $\mathrm{cl}(\mathcal{A}) = \mathcal{S}$. Thus, given any set $S \in \mathcal{S}$ and any $\varepsilon > 0$, there exists some $A \in \mathcal{A}$ with $\mu(A \triangle S) < \varepsilon$.

We turn now to the σ-finite case. By assumption, $\Omega = \bigcup_{n=1}^{\infty} \Omega_n$, where each Ω_n belongs to $\mathcal{A}$ and has finite measure. We may assume the Ω_n's are disjoint (see 21.24.b). Fix any n. Then $\mathcal{A}_n = \{A \cap \Omega_n : A \in \mathcal{A}\}$ is an algebra of subsets of Ω_n, and the σ-algebra that it generates on Ω_n is $\mathcal{S}_n = \{S \cap \Omega_n : S \in \mathcal{S}\}$. Let any $S \in \mathcal{S}$ be given. Apply the results of the preceding paragraph to the set $S \cap \Omega_n \in \mathcal{S}_n$; thus there exists some set $A_n \in \mathcal{A}_n$ with $\mu(A_n \triangle (S \cap \Omega_n)) < 2^{-n-1}\varepsilon$. Since μ is countably additive and the sets $S \cap \Omega_n$ form a partition of S, we have $\mu\left(\bigcup_{n=N+1}^{\infty} S_n\right) < \varepsilon/2$ for sufficiently large N. Let $A = \bigcup_{n=1}^{N} A_n$. Verify that $\mu\left(A \triangle \left(\bigcup_{n=1}^{N} S_n\right)\right) < \varepsilon/2$ and $\mu\left(\left(\bigcup_{n=1}^{N} S_n\right) \triangle S\right) < \varepsilon/2$, hence $\mu(A \triangle S) < \varepsilon$.

21.27. Corollary. Let Ω be an interval in $\mathbb{R}$ (possibly all of $\mathbb{R}$), and let μ be a σ-finite measure on the Borel subsets of Ω. Let any positive number ε and any Borel set $S \subseteq \Omega$ with finite measure be given. Then there exists a set $T \subseteq \Omega$ that is a union of finitely many intervals, such that $\mu(S \triangle T) < \varepsilon$.

Hints: 15.37.e and 21.26.

21.28. Uniqueness Lemma. Let $\mathcal{S}$ be the σ-algebra generated on some set Ω by some algebra of sets $\mathcal{A}$. Let μ and ν be two positive, σ-finite measures on $(\Omega, \mathcal{S})$; suppose $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$. Then $\mu = \nu$ on $\mathcal{S}$.

Proof. First suppose both μ and ν are finite. Then the collection $\mathcal{M} = \{M \in \mathcal{S} : \mu(M) = \nu(M)\}$ is a monotone class; apply the Monotone Class Theorem (5.29). For the general case, we may partition Ω into countably many sets, on each of which both μ and ν are finite.

Convergence in Measure

21.29. *Definitions.* Let $(\Omega, \mathcal{S}, \mu)$ be a measure space. The **outer measure** determined by

μ is the function $\mu^* : \mathcal{P}(\Omega) \to [0, +\infty]$ defined by

$$\mu^*(A) \quad = \quad \inf\{\mu(S) \ : \ S \supseteq A, \ S \in \mathcal{S}\} \qquad \text{for every set } A \subseteq \Omega.$$

The outer measure has some, but not all, the properties of a measure — it is defined on *all* subsets of Ω, but it is not necessarily countably additive.

Note that $\mu^*(A) = \mu(A)$ if $A \in \mathcal{S}$. For a slightly abridged reading, the beginner may restrict his or her attention to measurable functions; then all the sets considered below are measurable, so the μ^*'s can all be written instead as μ's.

Let (X, d) be a metric space. (We are chiefly interested in the cases where X is either $[0, +\infty]$ or a normed vector space.) We now introduce two more types of convergences on X^Ω. Let (f_α) be a net in X^Ω, and let $f \in X^\Omega$.

- We say (f_α) converges to f **in measure** if

$$\mu^*\Big\{\omega \in \Omega \ : \ d\Big(f_\alpha(\omega), f(\omega)\Big) > \varepsilon\Big\} \ \to \ 0 \text{ for each } \varepsilon > 0$$

  or, equivalently,

$$\text{for each } \varepsilon > 0, \text{ eventually } \mu^*\Big\{\omega \in \Omega \ : \ d\Big(f_\alpha(\omega), f(\omega)\Big) > \varepsilon\Big\} \ < \ \varepsilon.$$

  (*Exercise.* Prove that equivalence.) This is also called convergence **in probability** if $\mu(\Omega) = 1$.

- We say $f_\alpha \to f$ **μ-almost uniformly** if

  for each $\varepsilon > 0$, there exists a measurable set $S \subseteq \Omega$ such that $\mu(\Omega \setminus S) < \varepsilon$ and $f_\alpha \to f$ uniformly on S.

It is easy to verify that each of these is a centered, isotone convergence, as defined in 7.34.

Convergence in measure actually has much better properties. We shall see in 21.34 that it is determined by a pseudometric — or by a metric, if we identify functions that are μ-equivalent.

Almost uniform convergence is not given by a metric. In fact, we shall show in 21.33.c that almost uniform convergence is not topological, or even pretopological.

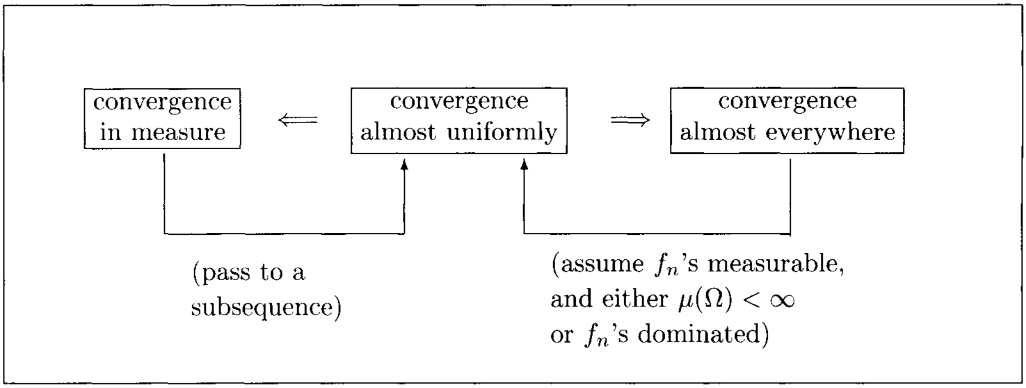
Observations. Convergence in measure is preserved if we replace functions with equivalent functions. That is, if f_α is μ-equivalent to g_α and f is μ-equivalent to g, then

$$f_\alpha \to f \text{ in measure} \qquad \Longleftrightarrow \qquad g_\alpha \to g \text{ in measure}.$$

Thus, convergence in measure makes sense for equivalence classes of functions.

Almost uniform convergence makes sense for *sequences* of equivalence classes of functions, just like almost everywhere convergence — see 21.18.

Preview. The following chart summarizes the relations that we shall establish between the three kinds of convergences.



21.30. *Proposition.* If $f_\alpha \to f$ μ-almost uniformly, then $f_\alpha \to f$ in measure and μ-almost everywhere. (*Proof.* Easy exercise.)

21.31. Theorem. If $g_n \to g$ in measure, then the sequence (g_n) has a subsequence that converges to g μ-almost uniformly (and therefore also converges pointwise μ-a.e.).

Hints: For each $\varepsilon > 0$, eventually $\mu^*\left\{\omega \in \Omega \,:\, d\Big(g_n(\omega), g(\omega)\Big) > \varepsilon\right\} < \varepsilon$ by assumption. Hence (g_n) has a subsequence (f_k) that satisfies, for some measurable sets S_k,

$$\mu(S_k) < 2^{-k-1} \qquad \text{and} \qquad S_k \supseteq \left\{\omega \in \Omega \,:\, d\Big(f_k(\omega), g(\omega)\Big) > 2^{-k-1}\right\}.$$

Let $T_k = S_k \cup S_{k+1} \cup S_{k+2} \cup \cdots$; then $\mu(T_k) < 2^{-k}$. Show that $f_j \to g$ uniformly on $\Omega \setminus T_k$ as $j \to \infty$.

21.32. Egorov's Theorem. Let (f_j) be a sequence in $SM(\mathcal{S}, X)$, converging pointwise to a limit $f : \Omega \to X$. Assume also $\mu(\Omega) < \infty$. Then $f_j \to f$ μ-almost uniformly (hence also $f_j \to f$ in measure).

Remarks. Note that we must assume the f_j's are strongly measurable. See also the related result in 26.12.f.

Hints: Let any $\varepsilon > 0$ be given. For positive integers k and m, let

$$B_{k,m} \quad = \quad \bigcup_{j=k}^{\infty} \left\{\omega \in \Omega \,:\, d(f_j(\omega), f(\omega)) > \frac{1}{m}\right\}.$$

Use the strong measurability of the f_j's (see 21.7) to show that $B_{k,m}$ is a measurable set. For fixed m, show that $\mu(\bigcap_{k=1}^{\infty} B_{k,m}) = 0$. Using the fact that $\mu(\Omega) < \infty$, show that $\mu(B_{k(m),m}) < 2^{-m}\varepsilon$ for some integer $k(m)$. Let $A_\varepsilon = \bigcup_{m=1}^{\infty} B_{k(m),m}$. Then $\mu(A_\varepsilon) < \varepsilon$, and $f_j \to f$ uniformly on $\Omega \setminus A_\varepsilon$.

21.33. *Examples.*

a. Let $(\Omega, \mathcal{S}, \mu)$ be the real line with Lebesgue subsets and Lebesgue measure. Let f_n be the characteristic function of the interval $[n, n+\frac{1}{n}]$. Then $f_n \to 0$ pointwise and in measure, but not μ-almost uniformly.

b. Let $(\Omega, \mathcal{S}, \mu)$ be the unit interval $[0,1]$, with Lebesgue subsets and Lebesgue measure; thus the measure of an interval is the length of that interval. Let $f_1, f_2, f_3, \ldots$ be the characteristic functions of the intervals

$$\left[\frac{0}{2}, \frac{1}{2}\right], \left[\frac{1}{2}, \frac{2}{2}\right], \quad \left[\frac{0}{3}, \frac{1}{3}\right], \left[\frac{1}{3}, \frac{2}{3}\right], \left[\frac{2}{3}, \frac{3}{3}\right], \quad \left[\frac{0}{4}, \frac{1}{4}\right], \left[\frac{1}{4}, \frac{2}{4}\right], \left[\frac{2}{4}, \frac{3}{4}\right], \left[\frac{3}{4}, \frac{4}{4}\right]$$

etc., in that order. Show that $f_n \to 0$ in measure, but not μ-almost uniformly or pointwise μ-a.e.

c. Example of non-pretopological convergence. Use the preceding example and the last few theorems to show that, in general, almost uniform convergence and almost everywhere convergence both lack the *sequential star property* introduced in 15.3.b. Hence, in general, those two convergences are not pretopological.

21.34. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let (X, d) be a pseudometric space. For $f, g \in X^\Omega$, define

$$D_\mu(f,g) \quad = \quad \inf_{\alpha>0} \arctan\left[\alpha \; + \; \mu^*\left\{\omega \in \Omega \; : \; d\Big(f(\omega), g(\omega)\Big) > \alpha\right\}\right].$$

(The arctan function can be replaced by any other bounded remetrization function; see 18.14.)

Admittedly, this formula is rather complicated. After we use it below to prove a few simple, basic properties, we will generally refer to those simple, basic properties, rather than the complicated formula for D_μ; we will very seldom want to make direct use of that formula. Still, the reader will probably find it conceptually helpful to see that there *is* some explicit formula for the pseudometric.

Show that

a. D_μ is a pseudometric on X^Ω.

b. $D_\mu(f,g) = 0$ if and only if $d(f(\cdot), g(\cdot)) = 0$ μ-almost everywhere. Thus, if d is a metric on X, then D_μ is a metric on the quotient space $^*X = X^\Omega/\mu$ — i.e., on the set of all μ-equivalence classes of functions.

c. The convergence determined by the pseudometric D_μ is the same as convergence in measure.

d. $SM(\mathcal{S}, X)$ (defined in 21.4) is a closed subset of the pseudometric space X^Ω; hence $SM(\mu, X)$ (defined in 21.17) is a closed subset of the metric space $^*X$. *Hint*: 21.3 and 21.31.

21.35. Theorem. If the pseudometric space (X, d) is complete, then the pseudometric space (X^Ω, D_μ) is complete.

Proof. Any D_μ-Cauchy sequence has a subsequence (f_k) satisfying

$$D_\mu(f_k, f_{k+1}) < \arctan(2^{-k-1}) \qquad \text{for} \quad k = 1, 2, 3, \ldots ;$$

it suffices to show that that subsequence is convergent in measure. By assumption, there is a measurable set $S_k \supseteq \{\omega \in \Omega : d(f_k(\omega), f_{k+1}(\omega)) > 2^{-k-1}\}$ with $\mu(S_k) < 2^{-k-1}$. Let $T_k = S_k \cup S_{k+1} \cup S_{k+2} \cup \cdots$; then $\mu(T_k) < 2^{-k}$. Since the T_k's form a decreasing sequence, we have

$$d(f_i(\omega), f_j(\omega)) \ < \ 2^{-j} \qquad \text{for } \omega \in \Omega \setminus T_j \text{ and } i \geq j. \tag{$**$}$$

Fix any k and any $\omega \in \Omega \setminus T_k$; consider $i \geq j \geq k$; the preceding estimate shows that the sequence $(f_k(\omega), f_{k+1}(\omega), f_{k+2}(\omega), \ldots)$ is Cauchy in (X, d). Since that space is complete, the sequence $(f_i(\omega) : i \in \mathbb{N})$ is convergent. Let $f(\omega)$ be any of its limits. (This is unique if d is a metric on X.) Take limits in $(**)$ as $i \to \infty$, to establish

$$d(f(\omega), f_j(\omega)) \quad \leq \quad 2^{-j} \qquad \text{for } \omega \in \Omega \setminus T_j.$$

Thus for $j \geq k$ we have

$$\mu^* \left\{\omega \in \Omega : d(f(\omega), f_j(\omega)) > 2^{-k}\right\} \quad \leq \quad \mu(T_j) \quad < \quad 2^{-j} \quad \leq \quad 2^{-k}$$

and therefore $f_j \to f$ in measure.

Integration of Positive Functions

21.36. *Definitions.* Let $(\Omega, \mathcal{S}, \mu)$ be a measure space. Let $f : \Omega \to [0, +\infty]$ be a measurable function, and let $S \in \mathcal{S}$. We shall define a number $\int_S f d\mu$ in $[0, +\infty]$, the **integral** of f over S with respect to μ, in two stages.

First, suppose f is a *simple function*, as defined in 11.42 — that is, f is measurable and its range is a finite set. Then we define the integral as in 11.42 — that is,

$$\int_S f(\omega)\, d\mu(\omega) \ = \ \sum_{x \in [0, +\infty]} \mu\left(S \cap f^{-1}(x)\right) x \ = \ \sum_{x \in \mathrm{Range}(f)} \mu\left(S \cap f^{-1}(x)\right) x.$$

These two summations are the same because when $x \in [0, +\infty] \setminus \mathrm{Range}(f)$ then $f^{-1}(x) = \varnothing$ and so $\mu(S \cap f^{-1}(x)) = 0$.

Second, when $f : \Omega \to [0, +\infty]$ is any measurable function, we define

$$\int_S f d\mu \quad = \quad \sup \left\{ \int_S g d\mu \ : \ g \text{ simple, } 0 \leq g \leq f \ \ \mu\text{-a.e.} \right\}.$$

Here the supremum is over all finitely valued measurable functions g that satisfy $0 \leq g \leq f$ μ-a.e. on Ω — or equivalently, that satisfy $0 \leq g \leq f$ μ-a.e. on S; this yields the same supremum (easy exercise). This supremum is well-defined — it is the supremum of a nonempty collection of numbers in $[0, +\infty]$, since trivially we could take $g = 0$; much better choices of g were noted in 21.5. Of course, this definition extends the one in the previous paragraph, for if f is finitely valued then we can take $g = f$.

We may sometimes refer to the integral defined in this fashion as the **positive integral**, to distinguish it from other kinds of integrals discussed in this book — see 11.41. The positive integral is also sometimes known as the **Lebesgue integral**, but that term has many other meanings as well.

A measurable function $f : \Omega \to [0, +\infty]$ is said to be **integrable** if $\int_\Omega f d\mu < \infty$. This terminology is unfortunately misleading: Some students may think that "integrable" means "capable of being integrated." But in fact, *any* measurable function taking values in $[0, +\infty]$ can be integrated to some value in $[0, +\infty]$; "integrable" means "yielding a *finite* value for its integral."

21.37. *Observations about integrals.* Let $(\Omega, \mathcal{S}, \mu)$ be a measure space. Let f and g be measurable functions from Ω into $[0, +\infty]$, and let $S \in \mathcal{S}$. Then:

a. $f \leq g$ μ-a.e. $\Rightarrow$ $\int_\Omega f \leq \int_\Omega g$.

b. $f = g$ μ-a.e. $\Rightarrow$ $\int_\Omega f = \int_\Omega g$.

c. $\int_S cf d\mu = c \int_S f d\mu$ for any constant $c \in [0, +\infty]$.

d. $\int_S f\, d\mu = \int_\Omega 1_S\, f\, d\mu$, where 1_S is the characteristic function of the set S.

e. $\int_S f d\mu = \int_S f d\mu_S$, where μ_S is the restriction of μ to subsets of S.

f. $\mu(S) = \int_\Omega 1_S d\mu$.

g. **(Chebyshev Inequality.)** For any $r \in (0, +\infty)$,

$$\mu\Big(\{\omega \in \Omega \ : \ f(\omega) \geq r\} \Big) \quad \leq \quad \frac{1}{r} \int_\Omega f d\mu.$$

h. $\int_\Omega f d\mu < \infty$ $\Rightarrow$ $f < \infty$ μ-a.e.

i. $\int_\Omega f d\mu = 0$ if and only if $f = 0$ μ-a.e.

j. Let $h : \Omega \to [0, +\infty]$ be a simple function. Then h is integrable (i.e., satisfies $\int_\Omega h d\mu < \infty$) if and only if the set $\{\omega \in \Omega \ : \ h(\omega) \neq 0\}$ has finite measure and the set $\{\omega \in \Omega \ : \ h(\omega) = \infty\}$ has measure 0.

k. An equivalent definition of the integral is

$$\int_S f d\mu \quad = \quad \sup \left\{ \int_S h d\mu \ : \ h \text{ simple, } 0 \leq h \leq f, \ \ h < f \text{ wherever } f \neq 0 \right\}.$$

Hints: To see this, first note that if g is a simple function with $0 \leq g \leq f$ on some measurable set M whose complement has measure 0, then we can replace g with the function $1_M(\cdot)g(\cdot)$. Thus we may assume $0 \leq g \leq f$ everywhere. Now we can approximate g from below by the sequence of functions $h_n = (1 - \frac{1}{n})g$; clearly $\int_S (1 - \frac{1}{n}) g\, d\mu \uparrow \int_S g\, d\mu$.

21.38. *A theorem in two parts.* Let $(\Omega, \mathcal{S}, \mu)$ be a measure space.

(i) **Integrals are measures.** Let $h : \Omega \to [0, +\infty]$ be measurable, and for each $S \in \mathcal{S}$ let $\nu(S) = \int_S h(\omega)d\mu(\omega)$. Then ν is a measure on $\mathcal{S}$.

(ii) **Lebesgue's Monotone Convergence Theorem.** Let $f_1, f_2, f_3, \ldots$ and f be measurable functions from Ω into $[0, +\infty]$. Suppose that $f_n \uparrow f$ pointwise μ-a.e. Then $\int_S f_n d\mu \uparrow \int_S f d\mu$ for each $S \in \mathcal{S}$.

Remark. An example of (i) is given in 21.13. A variant of (i) for vector-valued integrals is given in 29.10. A partial converse to (i) is given in 29.20.

Proof of theorem. Result (i) is easy when h is a simple function; we leave that easy case as an exercise. We shall now use that easy case as a step in our proof of (ii); then we shall use (ii) to prove (i) in its full generality.

Fix any $S \in \mathcal{S}$. Observe that the sequence $\int_S f_1 d\mu$, $\int_S f_2 d\mu$, $\int_S f_3 d\mu, \ldots$ is nondecreasing, hence converges to some limit $L \leq \int_S f d\mu$. Let $h : \Omega \to [0, +\infty)$ be any simple function that satisfies $0 \leq h \leq f$ and satisfies $h < f$ wherever $f \neq 0$. By 21.37.k, it suffices to show that $L \geq \int_S h\, d\mu$. Alter the f_n's on a set of measure 0, so that $f_n \uparrow f$ pointwise everywhere. From our choice of h it follows that for each ω we have $f_n(\omega) \geq h(\omega)$ for all n sufficiently large. Define $T(n) = \{\omega \in S : f_n(\omega) \geq h(\omega)\}$. Then $T(1) \subseteq T(2) \subseteq T(3) \subseteq \cdots$ and $\bigcup_{n=1}^{\infty} T(n) = S$. We have $\int_{T(n)} h d\mu \to \int_S h d\mu$ by 21.25.a and the special case of (i) already proved. Then

$$\int_S f_n d\mu \quad \geq \quad \int_{T(n)} f_n d\mu \quad \geq \quad \int_{T(n)} h d\mu.$$

Taking limits, we obtain $L \geq \int_S h d\mu$, which proves part (ii) of the theorem.

Finally, part (i) now follows easily in the general case, by 21.25.a.

21.39. *Corollaries.* Let $(\Omega, \mathcal{S}, \mu)$ be a measure space.

a. For any measurable functions $f, g : \Omega \to [0, +\infty]$ and any constant $c \geq 0$, we have $\int (f + g) = (\int f) + (\int g)$ and $\int (cf) = c \int f$. *Hint*: Prove this first for simple functions.

b. **Interchange of Limits (B. Levi).** Let $f_1, f_2, f_3, \ldots : \Omega \to [0, +\infty]$ be measurable functions. Then $\int (\sum_j f_j) = \sum_j (\int f_j)$; one side is finite if and only if the other side is.

c. **Fatou's Lemma.** Let $f_1, f_2, f_3, \ldots : \Omega \to [0, +\infty]$ be measurable functions. Then $\int \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int f_n$.

Hint: Let $g_k = \inf\{f_k, f_{k+1}, f_{k+2}, \ldots\}$; apply the Monotone Convergence Theorem to the sequence $g_1, g_2, g_3, \ldots$.

d. Suppose that each of $g, f_1, f_2, f_3, \ldots$ is an increasing continuous function from $[0, 1]$ into $[0, +\infty)$. Also suppose that

$$g(x) \quad = \quad f_1(x) + f_2(x) + f_3(x) + \cdots$$

for all x in $[0, 1)$. Then that equation is also valid for $x = 1$. This result will be used in 25.29.

Hint: Obtain a sequence of equations, by taking $x = 0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$. Then use the Monotone Convergence Theorem with counting measure on $\mathbb{N}$.

21.40. Product measures and Tonelli's Theorem. Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \nu)$ be σ-finite measures spaces. Define $\mathcal{S} \otimes \mathcal{T}$ as in 21.6. Then there exists a unique positive measure $\mu \otimes \nu$ on $\mathcal{S} \otimes \mathcal{T}$ that satisfies

$$(\mu \otimes \nu)(S \times T) \quad = \quad \mu(S)\nu(T) \qquad \text{for all} \quad S \in \mathcal{S}, \quad T \in \mathcal{T}.$$

That measure has this further property: If $f : X \times Y \to [0, +\infty]$ is jointly measurable, then both of the iterated positive integrals

$$\int_X \left[\int_Y f(x, y)\, d\nu(y) \right] d\mu(x), \qquad \int_Y \left[\int_X f(x, y)\, d\mu(x) \right] d\nu(y)$$

exist and are equal to the positive integral $\int_{X \times Y} f\, d(\mu \otimes \nu)$.

Proof. Our presentation is based on Cohn [1980]. By a **basic rectangle** we shall mean a set of the form $S \times T$, where $S \in \mathcal{S}$ and $T \in \mathcal{T}$. (These basic rectangles may also be called *basic measurable rectangles*, to distinguish them from another sort of "basic rectangle" introduced in 15.35.) Let $\mathcal{B}$ be the collection of unions of finitely many basic rectangles; verify that $\mathcal{B}$ is an algebra of sets. Clearly, the σ-algebra generated by $\mathcal{B}$ is $\mathcal{S} \otimes \mathcal{T}$.

Define sets A_x and A^y as in 21.6. We first show that the $[0, +\infty]$-valued mapping $x \mapsto \nu(A_x)$ is measurable. By 21.24.c, we may write $\nu = \sum_{j=1}^{\infty} \nu_j$ for some finite measures ν_j; it suffices to show that each of the mappings $x \mapsto \nu_j(A_x)$ is measurable. Fix any j. Let $\mathcal{N}_j$ be the collection of all sets $A \in \mathcal{S} \otimes \mathcal{T}$ for which the mapping $x \mapsto \nu_j(A_x)$ is measurable. Verify that $\mathcal{N}_j \supseteq \mathcal{B}$, and that $\mathcal{N}_j$ is a monotone class. Hence, by the Monotone Class Theorem 5.29, $\mathcal{N}_j = \mathcal{S} \otimes \mathcal{T}$.

Similarly, the $[0, +\infty]$-valued mapping $y \mapsto \mu(A^y)$ is measurable. Thus, for any set $A \in \mathcal{S} \otimes \mathcal{T}$, the integrals

$$I(A) = \int_X \nu(A_x)\, d\mu(x) \qquad \text{and} \qquad J(A) = \int_Y \mu(A^y)\, d\nu(y)$$

both exist. Verify that I and J are countably additive; thus they are measures on $\mathcal{S} \otimes \mathcal{T}$. Also, verify that $I(S \times T) = J(S \times T) = \mu(S)\nu(T)$ whenever $S \in \mathcal{S}$ and $T \in \mathcal{T}$. It follows from 21.28 that $I = J$ on $\mathcal{S} \otimes \mathcal{T}$. Thus, I and J are two different representations for the desired measure $\mu \otimes \nu$. From the equation

$$(\mu \otimes \nu)(A) \quad = \quad \int_X \nu(A_x)\, d\mu(x) \quad = \quad \int_Y \mu(A^y)\, d\nu(y)$$

we immediately obtain

$$\int_{X \times Y} f\, d(\mu \otimes \nu) \;=\; \int_X \left[\int_Y f(x, y)\, d\nu(y) \right] d\mu(x) \;=\; \int_Y \left[\int_X f(x, y)\, d\mu(x) \right] d\nu(y)$$

at least when f is finitely valued. For a general choice of f, approximate as in 21.5, and use the Monotone Convergence Theorem 21.38(ii).

21.41. *Exercise* (*optional*). Show that Tonelli's Theorem (21.40) implies Levi's Theorem (21.39.b), by using counting measure for one of the two factor measures. Then show that Levi's Theorem implies Lebesgue's Monotone Convergence Theorem 21.38(ii).

ESSENTIAL SUPREMA

21.42. Let $(\Omega, \mathcal{S}, \mu)$ be a σ-finite measure space. Let $SM(\mu, [-\infty, +\infty])$ be the collection of all μ-equivalence classes of measurable functions from Ω into $[-\infty, +\infty]$. Where no confusion will result, we shall use equivalence classes and members of those equivalence classes interchangeably. Let $SM(\mu, [-\infty, +\infty])$ be ordered in the obvious fashion:

$$f \preccurlyeq g \qquad \text{means that} \qquad f(\cdot) \leq g(\cdot) \quad \mu\text{-almost everywhere.}$$

Also, let $SM(\mu, \mathbb{R})$ and $SM(\mu, [0, 1])$ be the collection of equivalence classes of measurable functions with ranges in $\mathbb{R}$ or in $[0, 1]$, respectively.

Proposition and definition. The ordered set $SM(\mu, [-\infty, +\infty])$ is a complete lattice — i.e., it is a poset in which any nonempty set Φ has a supremum φ and an infimum ψ. Those functions are known as the **essential supremum** and **essential infimum** of Φ; they may be abbreviated ess-sup(Φ) and ess-inf(Φ). They have this further property: ess-sup(Φ) (respectively, ess-inf(Φ)) can be represented as the supremum (respectively, infimum) almost everywhere of some countable subcollection of Φ.

Similarly, $SM(\mu, [0, 1])$ is a complete lattice, and $SM(\mu, \mathbb{R})$ is a Dedekind complete vector lattice.

Caution: Another meaning for "essential supremum" is given in 22.28.

Proof of proposition. $SM(\mu, [-\infty, +\infty])$, $SM(\mu, \mathbb{R})$, and $SM(\mu, [0, 1])$ are certainly lattices, since the max or min of two measurable functions is measurable by 21.7.c. It is easy to verify that $SM(\mu, \mathbb{R})$ is a vector lattice. It remains to prove the assertions about order completeness. The Dedekind completeness of $SM(\mu, \mathbb{R})$ will follow from the completeness of $SM(\mu, [-\infty, +\infty])$. The transformation $\psi : t \mapsto \frac{1}{2} + \frac{1}{\pi} \arctan(t)$ is a strictly increasing bijection from $[-\infty, +\infty]$ onto $[0, 1]$, which is continuous (hence measurable) in both directions; thus it suffices to prove the order completeness of $SM(\mu, [0, 1])$.

By 21.24.d, we may assume $\mu(\Omega) < \infty$ (explain).

The supremum of countably many members of Φ is a measurable function. We may replace Φ with {sups of countably many members of Φ}, since that set has the same upper bounds as Φ does. Thus we may assume Φ is closed under countable sups.

Let $r = \sup\{\int_\Omega f d\mu : f \in \Phi\}$. Then r is some number in $[0, \mu(\Omega)]$. Choose functions $g_1, g_2, g_3, \ldots \in \Phi$ with $\int_\Omega g_n d\mu \to r$. Let $h_n = \max\{g_1, g_2, \ldots, g_n\}$; show that (h_n) is an increasing sequence in Φ and $\int_\Omega h_n d\mu \to r$. Let $\varphi = \sup\{h_1, h_2, h_3, \ldots\}$; show that $\varphi \in \Phi$ and $\int_\Omega \varphi d\mu = r$.

If some $f \in \Phi$ does not satisfy $f \leq \varphi$ almost everywhere, show that $g = \max\{f, \varphi\}$ is a member of Φ with integral strictly larger than r, a contradiction. Thus φ is an upper bound for Φ.

If some function β satisfies $\beta \geq f$ almost everywhere for each $f \in \Phi$, then in particular $\beta \geq h_n$ almost everywhere for each n; hence $\beta \geq \varphi$ almost everywhere. Thus φ is the least upper bound for Φ.

21.43. *Further related exercise.* Order convergence in a complete lattice was defined in

7.45. Show that order convergence in the complete lattice $SM(\mu, [-\infty, +\infty])$ is the same as convergence μ-almost everywhere (defined in 21.18).

Remark. We noted in 21.33.c that this convergence lacks the sequential star property, and thus is not topological or even pretopological.

21.44. The following result may be postponed; we shall use it in 22.36 and subsequently in 30.9.

Technical lemma on integration over compact sets. Let $(\Omega, \mathcal{S}, \mu)$ be a σ-finite measure space, and let (Z, d) be a metric space. Let $\Gamma : \Omega \times Z \to [0, +\infty]$ be a jointly measurable function — or, more generally, assume that the restriction of Γ to $\Omega \times Z_0$ is jointly measurable for each separable Borel set $Z_0 \subseteq Z$. For each compact set $K \subseteq Z$, assume that the integral

$$\int_\Omega \sup_{z \in K} \Gamma(\omega, z)\, d\mu(\omega)$$

exists and is finite.

Then for each compact set $K \subseteq Z$ there exist an open set $H \supseteq K$ and a function $\varphi \in L^1(\mu)$ with the property that

> whenever $v : \Omega \to H$ is a measurable function with relatively compact range, then $\Gamma(\omega, v(\omega)) \leq \varphi(\omega)$ for almost all ω.

Remarks. We emphasize that the set of almost all ω may depend on the particular choice of v; we do not assert that $\sup_v \Gamma(\omega, v(\omega)) \leq \varphi(\omega)$ for almost all ω.

The lemma's conclusion is only slightly weaker than the assertion that the integral $\int_\Omega \sup_{z\in H} \Gamma(\omega, z)\, d\mu(\omega)$ exists and is finite. In other words, if Γ is bounded (in a generalized sense) on compact subsets of Z, then Γ is also bounded on slightly larger subsets of Z. The lemma has the advantage that it is applicable even in some situations where $\int_\Omega \sup_{z\in H} \Gamma(\omega, z)\, d\mu(\omega)$ may not exist or may be infinite.

Proof of lemma. Fix K. For each number $r > 0$, define the open set $H(r) = \{z \in Z : \mathrm{dist}(z, K) < r\}$, and let $Q(r)$ be the set of all measurable functions $v : \Omega \to H(r)$ that have relatively compact ranges. We know $Q(r)$ is nonempty, since any constant function with value in K is a member.

We first claim that there is some $r > 0$ such that the number

$$\beta \quad = \quad \sup_{v \in Q(r)} \int_\Omega \Gamma(\omega, v(\omega))\, d\mu(\omega)$$

is finite. Indeed, suppose not. Then for $n = 1, 2, 3, \ldots$ there exist numbers $r_n \downarrow 0$ and functions $v_n \in Q(r_n)$ such that $\int_\Omega \Gamma(\omega, v_n(\omega))\, d\mu(\omega) > n$. However, since the v_n's have relatively compact ranges and are converging uniformly to K, the union of the ranges of the v_n's is contained in a single compact set L. Then $\Gamma(\omega, v_n(\omega)) \leq \sup_{z\in L} \Gamma(\omega, z)$, which yields a contradiction as soon as n is larger than $\int_\Omega \sup_{z \in L} \Gamma(\omega, z)\, d\mu$. This proves our claim.

Fix r and β as above; we shall take $H = H(r)$. As explained in 21.42, let φ be the supremum in $Meas(\mu, [-\infty, +\infty])$ of the set of functions $\omega \mapsto \Gamma(\omega, v(\omega))$ for $v \in Q(r)$. It remains only to show that $\int_\Omega \varphi\, d\mu$ is finite.

By the proposition in 21.42, there is some sequence (v_n) in $Q(r)$ such that φ is the pointwise supremum almost everywhere of the functions $\Gamma(\omega, v_n(\omega))$. For fixed $\omega \in \Omega$ and $n \in \mathbb{N}$, define $u_n(\omega)$ to be that member of the finite sequence $(v_1(\omega), v_2(\omega), \ldots, v_n(\omega))$ that maximizes $\Gamma(\omega, \cdot)$; if there is a tie, choose the first of the $v_j(\omega)$'s that yields the maximum value. Verify that the function u_n is a member of $Q(r)$, hence $\int_\Omega \Gamma(\omega, u_n(\omega))\, d\mu(\omega) \leq \beta$. Since $\Gamma(\omega, u_n(\omega)) = \max\{\Gamma(\omega, v_j(\omega)) : 1 \leq j \leq n\}$, the functions $\Gamma(\omega, u_n(\omega))$ increase pointwise almost everywhere to φ. By Lebesgue's Monotone Convergence Theorem 21.38(ii), the numbers $\int_\Omega \Gamma(\omega, u_n(\omega))\, d\mu(\omega)$ increase to $\int_\Omega \varphi\, d\mu$. Therefore $\int_\Omega \varphi\, d\mu \leq \beta$.

This Page Intentionally Left Blank

Part D

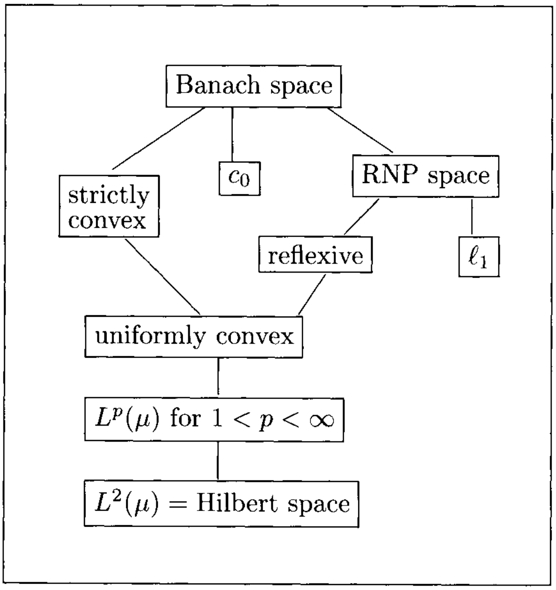
TOPOLOGICAL VECTOR SPACES

This Page Intentionally Left Blank

Chapter 22

Norms

22.1. *Preview.* A Banach space is a complete normed vector space. The following ch shows the relations between several types of Banach spaces that will be studied in this and later chapters.



(G-)(Semi)Norms

22.2. *Definitions.* Let X be an additive group. A **G-seminorm** on X is a mapping $\rho : X \to [0, +\infty)$ satisfying $\rho(0) = 0$ and these two conditions:

$$\begin{aligned} \rho(x+y) &\leq \rho(x) + \rho(y) && \text{(subadditive)} \\ \rho(-x) &= \rho(x) && \text{(symmetric)} \end{aligned}$$

for all $x, y \in X$. It is a **G-norm** if it also satisfies

$$x \neq 0 \quad \Rightarrow \quad \rho(x) > 0. \qquad \text{(positive-definite)}$$

In applications, we are usually concerned with either a single G-norm or a collection of infinitely many G-seminorms; see 5.15.h.

Most of the additive groups X considered in this book are actually vector spaces. In that setting, a G-seminorm $\rho : X \to [0, +\infty)$ is a **seminorm** if it also satisfies:

$$\rho(cx) = |c|\rho(x) \qquad \text{(homogeneous)}$$

for all scalars c and vectors $x, y \in X$. It is a **norm** if it is also positive-definite. Clearly, any seminorm is a G-seminorm; any norm is a G-norm. In Chapter 26 we shall study some G-(semi)norms that are not (semi)norms.

Below are some basic properties of G-(semi)norms and (semi)norms. Some readers may wish to skip ahead to our extensive collection of examples, which begins in 22.9.

22.3. *Notation.* A norm is usually denoted by $\| \ \|$; that is, we write $\|x\|$ instead of $\rho(x)$. Different norms on different vector spaces may both be denoted by the same symbol $\| \ \|$, when no confusion will result — e.g., we may write $\|x\|$ and $\|y\|$ when it is clear that $x \in X$ and $y \in Y$. When clarification is necessary we may use subscripts — e.g., let vector spaces X and Y have norms $\| \ \|_X$ and $\| \ \|_Y$.

Norms may also be denoted by $| \ |$ or $||| \ |||$. These symbols are used less often in the wider literature, but they will be used freely in this book to make it easier to distinguish between different norms. Most often, we shall use more bars to represent a "higher-order" norm — i.e., if $| \ |$ is used for a norm on spaces X or Y, then $\| \ \|$ will be used for a norm on some subspace of $Y^X = \{\text{functions from } X \text{ into } Y\}$; likewise, if $\| \ \|$ is used for a lower-order norm, then $||| \ |||$ will be used for the higher-order norm. In particular, $||| \ |||$ will be used for an *operator norm*, introduced in 23.1. Also, we shall use $| \ |$ for a norm especially when we do not wish to distinguish between a one-dimensional normed space (i.e., the scalar field) and higher-dimensional normed spaces — for instance, in 24.8.

22.4. If ρ is a G-seminorm on X, then $d(x, y) = \rho(x - y)$ defines a pseudometric d on X. Moreover, this pseudometric is **translation-invariant**; i.e., it satisfies

$$d(x + z,\ y + z) = d(x, y) \qquad \text{for all } x, y, z \in X.$$

Conversely, if d is a translation-invariant pseudometric on an additive group X, then $\rho(x) = d(x, 0)$ defines a G-seminorm ρ on X. For any additive group X, this correspondence $\rho \leftrightarrow d$ is a bijection between the G-seminorms on X and the translation-invariant pseudometrics on X. Positive-definiteness of ρ corresponds to that of d — that is, ρ is a G-norm if and only if d is a metric.

> *Hereafter, each G-seminorm $\rho : X \to [0, +\infty)$ will be identified with the corresponding translation-invariant pseudometric $d : X \times X \to [0, +\infty)$;*

we will use the two objects interchangeably. This convention will also apply in more specialized cases; e.g., a seminorm will be identified with its corresponding pseudometric, and in later chapters an F-seminorm will be identified with its corresponding pseudometric.

Topologies, uniformities, compactness, completeness, and other notions defined for pseudometrics will be transferred to G-seminorms in the obvious fashion.

22.5. *Exercise.* Let ρ and r be G-seminorms on an additive group X. Show that the following are equivalent.

(A) ρ is **stronger** than r — i.e., it yields a larger topology.

(B) ρ is uniformly stronger than r — i.e., it yields a larger uniformity.

(C) For each number $\varepsilon > 0$, there exists a number $\delta > 0$ such that $\rho(x) < \delta \Rightarrow r(x) < \varepsilon$.

(D) For each sequence (x_n) in X, if $\rho(x_n) \to 0$ then $r(x_n) \to 0$.

(E) For each net (x_α) in X, if $\rho(x_\alpha) \to 0$ then $r(x_\alpha) \to 0$.

If X is a linear space and ρ and r are seminorms, then the preceding conditions are also equivalent to:

(F) there exists a constant k such that $r(x) \leq k\rho(x)$ for all $x \in X$.

Further definitions. We say ρ is **strictly stronger** than r if ρ is stronger than r and r is not stronger than ρ. The two G-seminorms are **equivalent** if they determine the same topology — i.e., if each is stronger than the other. Note that they then determine the same uniform structure also. Equivalent (G-)(semi)norms can be used interchangeably for most purposes, but not for all purposes.

Further exercise. Two equivalent seminorms on a vector space yield the same collection of metrically bounded sets.

22.6. *Remarks.* The term "isomorphic" has different meanings in different parts of mathematics. Usually (but not always), an isomorphism of normed vector spaces is a linear homeomorphism. Occasionally it has the stronger meaning of a linear isometry.

> **Mazur-Ulam Theorem.** Let X and Y be real normed spaces, and let $f : X \to Y$ be a bijection that is distance-preserving. Then f is also affine — i.e., the map $f - f(0)$ is linear.

(The proof is long and will not be given here; it can be found in Banach [1932/1987].) Thus, the metric of a normed space is inextricably tied to its linear structure. Contrast this with the remarks about isometries in 22.9.d.

22.7. *Observations.* If (X, ρ) is a G-seminormed space, then the operation

$$x \mapsto -x \qquad \text{from } X \text{ into } X$$

is continuous, and the operation

$$(x, y) \mapsto x + y \qquad \text{from } X \times X \text{ into } X$$

is jointly continuous. If (X, ρ) is a seminormed space, then the operation

$$(c, x) \mapsto cx \qquad \text{from } \mathbb{F} \times X \text{ into } X$$

is jointly continuous, where $\mathbb{F}$ is the scalar field.

22.8. *Definition.* A complete normed vector space is called a **Banach space**.

Remarks. For most applications, a normed space can be viewed as a subspace of its completion; thus little is lost by restricting our attention to complete spaces. Much is gained: Completeness can be used to prove many theorems that make the spaces more useful or easier to understand.

In most of our examples of norms in this chapter, a vector space X is given, and a complete norm $\| \quad \|$ is given on X. That norm (or any norm equivalent to it) is called the "**usual norm**" for X: It is the norm used most frequently for X in applications.

The reader may wonder, though, why each space has only one "usual" norm, up to equivalence — i.e., why one particular norm (or equivalence class of norms) is preferred and singled out as *the* usual norm. Is this just a matter of custom and tradition? Or would some other complete norm be just as useful?

It turns out that among the examples one can find in applied analysis, there *isn't* any other complete norm. There are no explicitly constructible examples of inequivalent complete norms on a vector space. Thus the "usual norm" is determined uniquely (up to equivalence). In 27.18(iii) we shall prove the existence of inequivalent complete norms on a vector space, but this can only be accomplished by nonconstructive arguments — e.g., using the Axiom of Choice or some weakened form of AC. Applied mathematics generally is constructive and does not use the Axiom of Choice; thus it cannot produce inequivalent complete norms on X.

Analogous remarks apply to F-norms, a generalization of norms introduced in 26.2 for metrizable topological vector spaces; there are no explicitly constructible examples of inequivalent complete F-norms. Hence we may refer to "the usual F-norm" on a vector space. The proofs of these uniqueness results are rather deep; we postpone them until 27.47.b.

These uniqueness results do *not* generalize still further to G-norms. An Abelian group may have two explicitly constructible, inequivalent, complete G-norms; see the elementary example in 22.9.d. (But the uniqueness results do apply to *separable* complete G-normed groups. That follows from a version of the Closed Graph Theorem for separable groups, which can be found in Banach [1931] or Pettis [1950].)

BASIC EXAMPLES

22.9. *Elementary examples.*

a. Each of the scalar fields $\mathbb{R}$, $\mathbb{C}$ is a vector space over itself, and also $\mathbb{C}$ is a vector space over the scalar field $\mathbb{R}$. In each of these cases, the absolute value function is a norm, and the metric that it determines is complete.

b. If X is a vector space over the field $\mathbb{F}$ (either $\mathbb{R}$ or $\mathbb{C}$) and $f : X \to \mathbb{F}$ is a linear map, then $\rho(\cdot) = |f(\cdot)|$ is a seminorm on X. If $\dim(X) > 1$ then $\rho(\cdot) = |f(\cdot)|$ is not a norm. *Hint*: 11.9.j.

In particular, the constant function 0 is a seminorm.

c. If $\rho_1, \rho_2, \ldots, \rho_n$ are (semi)norms on a vector space X, then $\rho_1 + \rho_2 + \cdots + \rho_n$ and $\max\{\rho_1, \rho_2, \ldots, \rho_n\}$ are (semi)norms.

d. The Kronecker metric (defined in 2.12.b) is translation-invariant on any additive group, and thus it yields the **Kronecker G-norm**:

$$\rho(x) \quad = \quad 1 - \delta_{x0} \quad = \quad \begin{cases} 0 & \text{when } x = 0 \\ 1 & \text{when } x \neq 0. \end{cases}$$

It is complete. It is not a norm on any vector space X other than $\{0\}$ — in fact, it is not equivalent to a norm, since (*exercise*) scalar multiplication is not jointly continuous when this G-norm is used.

Note that the usual metric on $\mathbb{R}$ and the Kronecker metric on $\mathbb{R}$ are two inequivalent complete metrics, both determined by G-norms. Contrast this with the results for norms and F-norms discussed in 22.8.

The Kronecker G-norm of a group is not closely tied to its group structure — e.g., to its group homomorphisms. For instance, if X and Y are two groups equipped with their Kronecker G-norms, then *any injective map* from X into Y is an isometry (i.e., distance-preserving map) from X onto a subset of Y. Contrast this with the Mazur-Ulam Theorem, discussed in 22.6, which shows that any norm on a normed space is closely tied to its linear structure.

e. Suppose Ω is a set, $\mathcal{S}$ is an algebra of subsets of Ω, and μ is a positive charge on $\mathcal{S}$. As we noted in 21.9, we can define a pseudometric on $\mathcal{S}$ by $d(A, B) = \arctan \mu(A \triangle B)$, or more simply by $d(A, B) = \mu(A \triangle B)$ if μ is finite. Show that d is a translation-invariant pseudometric on the commutative group $(\mathcal{S}, \triangle, i_{\mathcal{S}}, \varnothing)$ discussed in 8.10.g, hence $\arctan \mu(\cdot)$ (or μ, if it is finite) is a G-pseudonorm on $\mathcal{S}$.

22.10. *Exercise.* Let $(X, \| \enspace \|)$ be a normed space. Show that X has a natural **completion.** That is, show that X is a dense linear subspace of a complete normed space Y, such that $\| \enspace \|$ is the restriction of the norm of Y. Show that this completion is unique, up to isomorphism (where an isomorphism preserves both linear and metric structure).

22.11. Finite-dimensional spaces. Let n be a positive integer. For any vector $x = (x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$ or $\mathbb{C}^n$, define

$$\|x\|_p \quad = \quad \{|x_1|^p + |x_2|^p + \cdots + |x_n|^p\}^{1/p} \qquad (0 < p < \infty)$$

and $\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$. Show that the functions $\| \enspace \|_p$, for $1 \leq p \leq \infty$, are complete norms on $\mathbb{R}^n$ or $\mathbb{C}^n$, all equivalent to one another (as defined in 22.5). (*Hint*: Show that $\| \enspace \|_p$ is the Minkowski functional of the convex, balanced, absorbing set $\{x : \|x\|_p \leq 1\}$; see 12.29.g.) If $0 < p < 1$, then $\| \enspace \|_p$ is not a norm, but $\| \enspace \|_p^p$ is a G-norm on $\mathbb{R}^n$ or $\mathbb{C}^n$ (*hint*: 12.25.e); in fact, we shall see in Chapter 26 that it is a special kind of G-norm, which we call an *F-norm.*

The functions $\| \quad \|_p$ $(1 \leq p \leq \infty)$ are sometimes referred to as the **usual norms** on $\mathbb{R}^n$ or $\mathbb{C}^n$. They yield the usual product topology and usual product uniform structure (studied in later chapters). The norm $\| \quad \|_2$ is often called the **Euclidean norm** because it gives $\mathbb{R}^n$ the metric of classical Euclidean geometry. Some textbooks use the notation E^n to refer to the vector space $\mathbb{R}^n$ equipped with the Euclidean norm.

The normed spaces $(\mathbb{F}^n, \| \quad \|_p)$ are a special case of the normed spaces $(L^p(\mu), \| \quad \|_p)$, which we shall study in 22.28 and thereafter.

Additional exercise. Draw graphs of the "unit circle"

$$C_p \quad = \quad \{(x_1, x_2) \in \mathbb{R}^2 \ : \ |x_1|^p + |x_2|^p = 1\}$$

for a few values of p — for instance, for $p = \frac{1}{5}, \frac{1}{2}, \frac{4}{5}, 1, \frac{5}{4}, 2, 5, +\infty$. (The values for $p < 1$ correspond not to norms but to F-norms, which will be studied in 26.4.d.) Observe that the "unit circle" is not circular when $p \neq 2$. In fact, when $p = \infty$ or $0 < p \leq 1$, the "unit circle" is not even round; it has corners.

Hints: First draw for $p = 1, 2, +\infty$; they'll be easiest and they may give you some idea of what to expect for the other sketches. For the remaining sketches, use the parametric representation

$$x_1 \quad = \quad |\cos\theta|^{2/p}, \qquad\qquad x_2 \quad = \quad |\sin\theta|^{2/p} \qquad\qquad (0 \leq \theta < 2\pi).$$

This is a good problem to do on a computer. If you have only a nongraphing calculator, just plot some points for $0 \leq \theta \leq \pi/2$ and then use symmetry to finish the sketch.

22.12. Product norms. We now generalize slightly the computations in 22.11. Let n be a positive integer, and let $(X_1, | \quad |), (X_2, | \quad |), \ldots, (X_n, | \quad |)$ be normed spaces. (The norms on the n spaces generally are different, but for simplicity we shall denote them all by the same symbol $| \quad |$.) For any vector $x = (x_1, x_2, \ldots, x_n)$ in the product space $X = \prod_{j=1}^n X_j$, define

$$\|x\|_p \quad = \quad \left\{|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right\}^{1/p} \qquad (0 < p < \infty)$$

and $\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$. Show that the functions $\| \quad \|_p^{\min\{1,p\}}$, for $0 < p \leq \infty$, are G-norms on X, all equivalent to one another, and the topology they determine on X is the product topology. When $1 \leq p \leq \infty$, then $\| \quad \|_p$ is a norm on X.

22.13. Quotient norms. Let X be an Abelian group (written as an additive group); let K be a subgroup; let $Q = X/K$ be the quotient group; let $\pi : X \to Q$ be the quotient map. For each G-seminorm $\rho : X \to [0, +\infty)$, we may define an associated function $\widehat{\rho} : Q \to [0, +\infty)$ by

$$\widehat{\rho}(q) \quad = \quad \inf\left\{\rho(x) \ : \ x \in \pi^{-1}(q)\right\}.$$

Show that

a. $\widehat{\rho}$ is a G-seminorm on Q. In fact, it is the largest G-seminorm on Q that satisfies $\widehat{\rho}(\pi(x)) \leq \rho(x)$ for all $x \in X$. *Hint*: 4.42.

b. If X is a vector space, ρ is a seminorm, and K is a linear subspace of X, then $\widehat{\rho}$ is a seminorm on the quotient space X/K. In some cases it is a norm; then it is called the **quotient norm**.

c. π preserves open balls:

$$\pi\Big(\{x \in X \ : \ \rho(x) < \varepsilon\}\Big) \quad = \quad \{q \in Q \ : \ \widehat{\rho}(q) < \varepsilon\}.$$

d. If $\rho^{-1}(0) \supseteq K$, then ρ is constant on each set of the form $\pi^{-1}(q)$, and so our definition of $\widehat{\rho}$ simplifies to $\widehat{\rho}(\pi(x)) = \rho(x)$.

e. The map between pseudometric spaces, $\pi : (X, \rho) \to (Q, \widehat{\rho})$, is a topological quotient map (defined as in 15.30).

More generally, let X be topologized by a gauge D consisting of G-seminorms, and let Q be topologized by the corresponding gauge $\widehat{D} = \{\widehat{\rho} : \rho \in D\}$. Suppose that D is *directed*, in the sense of 4.4.c. Then $\pi : X \to Q$ is an open mapping (by 22.13.c), hence it is a topological quotient map (by 15.31.e).

Sup Norms

22.14. As usual, let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$. If Λ is any nonempty set, then

$$B(\Lambda) \quad = \quad \{\text{bounded functions from } \Lambda \text{ into } \mathbb{F}\}$$

is a linear space. The usual norm on this space is $\|f\|_\infty = \sup\{|f(\lambda)| \ : \ \lambda \in \Lambda\}$; this is sometimes called the *sup norm*. It is complete. We have already seen one example of sup norms in 22.11.

The metric on $B(\Lambda)$ obtained from this norm is the same as the metric given in 4.41.f. The results in 4.41.f show that

every metric space (Λ, d) may be viewed as a subset of a Banach space.

Thus, in principle, metric spaces are not really "more general" than subsets of normed spaces.

However, this embedding is seldom used in applications. The additional linear structure of $B(\Lambda)$ may be merely distracting and not particularly relevant to the properties of the metric space (Λ, d) that one may be studying. For instance, we may gain some understanding of the "numbers" $\pm\infty$ by viewing them as elements of the metric space $[-\infty, +\infty]$ introduced in 18.24, but that understanding is not necessarily increased if we study the larger and more complicated space $B([-\infty, +\infty])$.

22.15. More sup-normed spaces. Let Ω be a topological space; then $B(\Omega) = \{$bounded functions from Ω into $\mathbb{F}\}$ is a Banach space when equipped with the sup norm. We now consider some interesting subspaces. First of all,

$$BC(\Omega) = \{\text{bounded, continuous functions from } \Omega \text{ into } \mathbb{F}\}.$$

is a closed linear subspace of $B(\Omega)$ — hence a Banach space, when equipped with the sup norm. If Ω is a uniform space, show that

$$BUC(\Omega) = \{\text{bounded, uniformly continuous functions from } \Omega \text{ into } \mathbb{F}\}$$

is a closed subspace of $BC(\Omega)$.

Suppose Ω is a locally compact Hausdorff space (such as $\mathbb{R}^n$, for instance). Then a function $f : \Omega \to \mathbb{F}$ is said to **vanish at infinity** if for each $\varepsilon > 0$ the set $\{x \in \Omega : |f(x)| > \varepsilon\}$ is relatively compact. In this setting,

$$C_0(\Omega) = \{\text{continuous functions from } \Omega \text{ into } \mathbb{F} \text{ that vanish at infinity}\}$$

is a closed linear subspace of $BC(\Omega)$, hence another Banach space. If Ω is also equipped with a uniform structure, show that $C_0(\Omega) \subseteq BUC(\Omega) \subseteq BC(\Omega)$. Of course, all these spaces are the same if Ω is a compact Hausdorff space.

A generalization. For any Banach space $(X, |\ \ |_X)$, we can define $BC(\Omega, X)$, $C_0(\Omega, X)$, $BUC(\Omega, X)$ in an analogous fashion; they are closed linear subspaces of the Banach space

$$B(\Omega, X) \quad = \quad \{\text{bounded functions from } \Omega \text{ into } X\}$$

with sup norm $\|f\|_\infty = \sup\{|f(\omega)|_X : \omega \in \Omega\}$.

A specialization. Let Ω be the set $\mathbb{N} = \{\text{positive integers}\}$, equipped with this metric: $d(m,n) = |\arctan(m) - \arctan(n)|$. This gives $\mathbb{N}$ its usual topology (i.e., the discrete topology), but gives $\mathbb{N}$ the uniform structure of a subset of the compact space $[0, +\infty]$. Then $B(\mathbb{N}) = BC(\mathbb{N})$ (since the topology on $\mathbb{N}$ is discrete), and the three Banach spaces $BC(\mathbb{N})$, $BUC(\mathbb{N})$, $C_0(\mathbb{N})$ can be rewritten respectively as

$$\begin{aligned} \ell_\infty &= \{\text{bounded sequences of scalars}\}, \\ c &= \{\text{convergent sequences of scalars}\}, \\ c_0 &= \{\text{sequences of scalars that converge to } 0\}, \end{aligned}$$

all equipped with the sup norm.

22.16. *Exercises.*

a. The sup-normed spaces $C[0,1]$ and $C_0(\mathbb{R})$ are separable.

Hint: We prove this for real scalars; the proof for complex scalars is similar. By a "rational piecewise affine function" we shall mean a continuous function whose graph consists of finitely many line segments, each of which has endpoints with rational coordinates; in the case of $C_0(\mathbb{R})$ we extend such a function by making it equal to 0 for all sufficiently large or small arguments. Show that there are only countably many rational piecewise affine functions. Show that members of $C[0,1]$ or $C_0(\mathbb{R})$ are uniformly continuous; use that fact to show that the rational piecewise affine functions are dense.

b. The sup-normed space $BUC(\mathbb{R})$ is not separable.

Hints: This is easiest in the case where the scalar field is $\mathbb{C}$ — that is, in the case where $BUC(\mathbb{R})$ represents the space of all bounded, uniformly continuous functions from $\mathbb{R}$ into $\mathbb{C}$. In that case, define $f_r(u) = \mathrm{cis}(ru) = \cos(ru) + i\sin(ru)$ for real numbers r and u. Show that $\|f_r - f_s\|_\infty = 2$ whenever $r \neq s$.

We can use that result to prove the nonseparability of $BUC(\mathbb{R})$ in the case of real scalars as well. Indeed, from the equation $\|f_r - f_s\|_\infty = 2$ we can prove that (r, s) must lie in at least one of the two sets

$$A = \left\{(r,s) \in \mathbb{R}^2 \ : \ \sup_{u\in\mathbb{R}} |\cos(ru) - \cos(su)| \geq 1\right\},$$

$$B = \left\{(r,s) \in \mathbb{R}^2 \ : \ \sup_{u\in\mathbb{R}} |\sin(ru) - \sin(su)| \geq 1\right\}.$$

Since $\{(r,s) \in \mathbb{R}^2 : r \neq s\}$ is uncountable, at least one of the two sets A, B is uncountable, and therefore $BUC(\mathbb{R})$ is not separable.

c. A Dominated Convergence Theorem for c_o. A set $S \subseteq c_0$ is relatively compact if and only if it is *dominated* in c_0 — i.e., if and only if there exists some sequence $r = (r_n) \in c_0$ such that $|s_n| \leq r_n$ for every sequence $s = (s_n) \in S$ and every $n \in \mathbb{N}$.

A sequence of members of c_0 converges in norm if and only if it is dominated and converges coordinatewise.

22.17. Lemma on spaces of vector-valued functions. Let Γ be a set, and let X be a Banach space. Let Φ be a collection of seminorms on the vector space X^Γ. Assume that the gauge topology given by Φ is equal to the topology of pointwise convergence on Γ — that is, for any net (f_α) in X^Γ,

$$\varphi(f_\alpha - f) \to 0 \ \text{ for each } \varphi \in \Phi \qquad \Longleftrightarrow \qquad f_\alpha(\gamma) \to f(\gamma) \ \text{ for each } \gamma \in \Gamma.$$

Now define $\|f\| = \sup\{\varphi(f) \ : \ \varphi \in \Phi\}$ for each $f \in X^\Gamma$, and let

$$V = \{f \in X^\Gamma \ : \ \|f\| < \infty\}.$$

Then $(V, \| \ \ \|)$ is a Banach space.

Remarks. This somewhat technical lemma will be used several times to show that certain linear spaces V are Banach spaces; see 22.18.c, 22.19.c, 29.6.c, and 29.29.f.

We emphasize that Γ need not be a complete metric space. In fact, Γ doesn't have to be equipped with any metric structure at all. Also, we emphasize that the convergence given by $\| \ \ \|$ or given by the gauge topology from Φ are, respectively, the convergences $\varphi(f_\alpha - f) \to 0$ uniformly for all φ or separately for each φ.

Proof of lemma. It is easy to verify that $(V, \| \ \ \|)$ is a seminormed linear space; that verification is left as an exercise. We shall show that the seminorm $\| \ \ \|$ is a norm and that it is complete.

To see that $\| \ \ \|$ is positive definite, note that if $f \in X^\Gamma \setminus \{0\}$, then $\varphi(f) > 0$ for at least one $\varphi \in \Phi$ (since the product topology on X^Γ is a Hausdorff topology), so $\|f\| > 0$.

It remains to show that $(V, \| \ \|)$ is complete. Let (f_n) be a $\| \ \|$-Cauchy sequence in V. Thus, for each number $\varepsilon > 0$ there is some integer N_ε such that $m, n \geq N_\varepsilon \Rightarrow \|f_m - f_n\| \leq \varepsilon$. Therefore, for each $\varphi \in \Phi$ we have

$$m, n \geq N_\varepsilon \qquad \Rightarrow \qquad \varphi(f_m - f_n) \leq \varepsilon. \tag{$*$}$$

Thus the net $(f_m - f_n : (m, n) \in \mathbb{N} \times \mathbb{N})$ converges to 0 pointwise on Γ. Since X is complete, for each $\gamma \in \Gamma$ there exists $f(\gamma) = \lim_{n\to\infty} f_n(\gamma)$. Since $f_n \to f$ pointwise, we have $\varphi(f_n - f) \to 0$ for each $\varphi \in \Phi$. Hold φ and m fixed and take limits in $(*)$ as $n \to \infty$; thus

$$m \geq N_\varepsilon \qquad \Rightarrow \qquad \varphi(f_m - f) \leq \varepsilon.$$

In other words, $m \geq N_\varepsilon \Rightarrow \|f_m - f\| \leq \varepsilon$. This proves that $f \in V$ and that (f_m) converges to f in $(V, \| \ \|)$.

22.18. The space of Hölder continuous functions. The definition of Hölder continuity, given in 18.4, simplifies slightly when the metric space Y is a normed space, with norm $| \ |_Y$. Let (X, d) be any metric space, and let $\alpha > 0$. For functions $f : X \to Y$, we obtain

$$\langle f \rangle_\alpha \quad = \quad \sup \left\{ \frac{|f(x_1) - f(x_2)|_Y}{d(x_1, x_2)^\alpha} \ : \ x_1, x_2 \in X, \ x_1 \neq x_2 \right\}.$$

Exercises. Show that

a. $\text{Höl}^\alpha(X, Y) = \{f \in Y^X : \langle f \rangle_\alpha < \infty\}$ is a linear space. The function $\langle \cdot \rangle_\alpha$ is not a norm, but rather a seminorm on $\text{Höl}^\alpha(X, Y)$. Indeed, we have $\langle f \rangle_\alpha = 0$ if and only if f is a constant function.

b. To get a norm, select any point in X; let us call that point "0" (although we shall not use any additive structure in X). Then $\|f\|_\alpha = \langle f \rangle_\alpha + |f(0)|_Y$ defines a norm $\| \ \|_\alpha$ on $\text{Höl}^\alpha(X, Y)$.

c. If Y is complete, then $\text{Höl}^\alpha(X, Y)$ (normed as above) is complete, regardless of whether X is complete. *Hint*: This is a special case of 22.17.

d. If $0 < \alpha < \beta \leq 1$, show that

$$\text{Höl}^\beta([0,1], Y) \quad \subseteq \quad \text{Höl}^\alpha([0,1], Y) \quad \subseteq \quad C([0,1], Y),$$

where the last space is the space of continuous functions from $[0, 1]$ into Y, equipped with the sup norm. The inclusions are continuous. If $Y = \mathbb{R}^n$ for some positive integer n, then the inclusions are compact — i.e., a bounded subset of one normed space is a relatively compact subset of the next space.

Hint: Use the Arzela-Ascoli Theorem 18.35.

e. *A related exercise.* This time we take the domain, rather than the codomain, to be a subset of a normed space.

Let C be a convex subset of a normed space $(X, | \ |_X)$, and let (Y, e) be any metric space. Show that if $\alpha > 1$, then $\text{Höl}^\alpha(C, Y)$ contains only constant functions.

Hint: Suppose $\langle p\rangle_\alpha = k$, and let any $u, v \in C$ be given. Let n be a large positive integer. Define $x_j = (1-\frac{j}{n})u + \frac{j}{n}v$ for $j = 0, 1, 2, \dots, n$. Then

$$e\Big(p(u), p(v)\Big) \;\le\; \sum_{j=1}^{n} e\Big(p(x_j), p(x_{j-1})\Big) \;\le\; \sum_{j=1}^{n} k|x_j - x_{j-1}|_X^\alpha \;=\; \frac{k|u-v|_X}{n^{\alpha-1}}.$$

22.19. Let $(X, |\ |)$ be a Banach space. Let $BV([a,b],X)$ be the set of all functions from $[a,b]$ into X that have bounded variation (as defined in 19.21). Show that

a. $BV([a,b],X)$ is a linear space, and $\mathrm{Var}(\cdot,[a,b])$ is a seminorm on $BV([a,b],X)$.

b. $\|\varphi\|_{\mathrm{BV}} = |\varphi(a)|_X + \mathrm{Var}(\cdot,[a,b])$ is a norm on $BV([a,b],X)$. Moreover, $\|\varphi\|_\infty \le \|\varphi\|_{BV}$.

c. $BV([a,b],X)$, normed as above, is complete. *Hint*: This is a special case of 22.17.

d. We say $f : [a,b] \to X$ is a **normalized function of bounded variation** on $[a,b]$ if f has bounded variation on $[a,b]$, f is right continuous on (a,b), and $f(0) = 0$. The collection of such functions will be denoted by $NBV([a,b],X)$; it will play an important role in 29.34. Note that $NBV([a,b],X)$ is a linear subspace of $BV([a,b],X)$, and $\mathrm{Var}(\cdot,[a,b])$ acts as a norm on $NBV([a,b],X)$.

Convergent Series

22.20. By a **series** in a normed space $(X, \|\ \|)$ we mean an expression of the form $\sum_{j=1}^\infty x_j = x_1 + x_2 + x_3 + \cdots$, where the x_j's are members of X. The **sum** of the series is the vector $v = \lim_{N\to\infty} \sum_{j=1}^N x_j$, if this limit exists. If it exists, we say the series is **convergent**; we may also write $v = \sum_{j=1}^\infty x_j$.

A series $\sum_{j=1}^\infty x_j$ is **absolutely convergent** if $\sum_{j=1}^\infty \|x_j\| < \infty$. Any absolutely convergent series in a Banach space is convergent; that follows from the completeness of X. In fact, (*exercise*) a normed space is complete if and only if every absolutely convergent series in the space is convergent.

See also the related results in 10.41, 23.26, and 23.27.

22.21. Dirichlet's test. Let V be a Banach space. Let $\sum_{k=1}^\infty v_k$ be a series in V whose partial sums $s_n = \sum_{k=1}^n v_k$ form a bounded sequence. Let (b_k) be a sequence of real numbers decreasing to 0. Then the series $\sum_{k=1}^\infty b_k v_k$ is convergent.

(A corollary is the Alternating Series Test, given in 10.41.g.)

Proof of Dirichlet's test. For any positive integers m, n with $n \ge m$, verify that

$$\sum_{k=m}^{n} b_k v_k \quad = \quad b_{n+1}s_n - b_m s_{m-1} - \sum_{k=m}^{n}(b_{k+1} - b_k)s_k.$$

By assumption, $S = \sup_n |s_n|$ is finite. Hence

$$\left|\sum_{k=m}^{n} b_k v_k\right| \;\le\; b_{n+1}S + b_m S + \sum_{k=m}^{n}(b_k - b_{k+1})S \;=\; 2b_m S.$$

It follows that the partial sums of the series $\sum_{k=1}^{\infty} b_k v_k$ form a Cauchy sequence.

22.22. *Example.* If (b_k) is any sequence of positive numbers decreasing to 0, then $\sum_{k=1}^{\infty} b_k \sin(kx)$ converges for each real number r. In particular, the series $\sum_{k=1}^{\infty} \frac{\sin(kx)}{k}$ and $\sum_{k=1}^{\infty} \frac{\sin(kx)}{\ln(k+1)}$ both converge. (Contrast this result with 10.43.)

Proof. First show that

$$2\left(\sin\frac{x}{2}\right)\left(\sin x + \sin 2x + \cdots + \sin nx\right) \;=\; \cos\left(\frac{1}{2}x\right) - \cos\left((n+\frac{1}{2})x\right),$$

either directly (using trigonometric identities), or by using the formulas $\sin\theta = (e^{i\theta} - e^{-i\theta})/2i$, $\cos\theta = (e^{i\theta} + e^{i\theta})/2$. Use that formula to show that the partial sums of the series $\sum_{k=1}^{\infty} \sin kx$ form a bounded sequence, for each fixed x. Now apply Dirichlet's Test.

22.23. Let $(X, \|\ \ \|)$ be a complex Banach space. Let $c_0, c_1, c_2, \ldots$ be some sequence in X, and let a be a complex number. (In the simplest case we take $a = 0$.) Then the expression $\sum_{n=0}^{\infty} c_n(\lambda - a)^n$ is called a **power series centered at** $\boldsymbol{a}$; the c_n's are called its **coefficients**. Associated with the power series is a number $R \in [0, +\infty]$ defined by

$$\frac{1}{R} \;=\; \limsup_{n\to\infty} \sqrt[n]{\|c_n\|}\,.$$

This number R is called the **radius of convergence** of the power series, the set $\{\lambda \in \mathbb{C} : |\lambda - a| < R\}$ is called the **disk of convergence**, and the set $\{\lambda \in \mathbb{C} : |\lambda - a| = R\}$ is called the **circle of convergence**. (The following results are also valid with real scalars, with intervals for "disks," but for simplicity of notation we shall only consider complex scalars.) The series, radius, and disk have these properties:

a. If only finitely many of the c_n's are nonzero, and $\lim_{n\to\infty} \|c_n\|/\|c_{n+1}\|$ exists in $[0, +\infty]$, then that limit is equal to R.

Remark. The expression $\lim_{n\to\infty} \|c_n\|/\|c_{n+1}\|$ is simpler, and thus is preferable in those cases where it is applicable. On the other hand, the more complicated expression $1/\limsup_{n\to\infty} \sqrt[n]{\|c_n\|}$ has the advantage that it is always applicable.

b. For each complex number λ with $|\lambda - a| < R$, the series $\sum_{n=0}^{\infty} c_n(\lambda - a)^n$ converges to a limit — that is, $\lim_{N\to\infty} \sum_{n=0}^{N} c_n(\lambda - a)^n$ exists in X. The series is absolutely convergent, and the convergence is uniform on compact subsets of the disk of convergence. Thus the power series defines a function on that disk; we summarize this by writing

$$f(\lambda) \;=\; \sum_{n=0}^{\infty} c_n(\lambda - a)^n \qquad (|\lambda - a| < R).$$

Hints: Any compact subset is contained in a set of the form $\{\lambda \in \mathbb{C} : |\lambda - a| \leq r\}$ for some number $r < R$. See 10.41.d and 22.20.

c. The series $\sum_n c_n(\lambda - a)^n$ is divergent (i.e., nonconvergent) for every $\lambda \in \mathbb{C}$ with $|\lambda - a| > R$.

Hint: If the series is convergent for some value of λ, then $c_n(\lambda - a)^n \to 0$ as $n \to \infty$. Then for all n sufficiently large, we have $\|c_n(\lambda - a)^n\| < 1$. If $\lambda \neq a$, then $\sqrt[n]{\|c_n\|} < 1/|\lambda - a|$.

Further properties of power series are described in 23.29(iii) and 25.27.

22.24. *Elementary examples.* A power series $\sum_{n=0}^{\infty} c_n(\lambda - a)^n$ converges inside the circle of convergence, and diverges outside that circle. The behavior is more complicated *on* the circle of convergence — i.e., at points λ satisfying $|\lambda - a| = R$. A series may converge at all, some, or none of these points. Following are a few simple examples with center $a = 0$ and with coefficients in $X = \mathbb{C}$.

a. Any polynomial (of a complex variable, with complex coefficients) is a power series with infinite radius of convergence. It has only finitely many nonzero coefficients.

b. The power series $f(\lambda) = \sum_{n=0}^{\infty} \lambda^n = 1 + \lambda + \lambda^2 + \lambda^3 + \cdots$ has radius of convergence equal to 1. Since

$$1 + \lambda + \lambda^2 + \cdots + \lambda^N = \frac{1 - \lambda^{N+1}}{1 - \lambda} \qquad \text{when } \lambda \neq 1,$$

we easily see that the power series $\sum_{n=0}^{\infty} \lambda^n$ converges to $\frac{1}{1-\lambda}$ when $|\lambda| < 1$ and diverges for every λ such that $|\lambda| \geq 1$.

c. The power series $f(\lambda) = \sum_{n=1}^{\infty} n^{-2}\lambda^n = \frac{\lambda}{1} + \frac{\lambda^2}{4} + \frac{\lambda^3}{9} + \cdots$ has radius of convergence equal to 1. Show that this series converges absolutely at every point on the circle of convergence.

d. Hardy gave an example of a power series that converges uniformly, but not absolutely, on its circle of convergence. Lusin gave an example of a series $\sum_{n=1}^{\infty} a_n\lambda^n$ such that $a_n \to 0$, but such that the series diverges at every point of the circle of convergence. These examples are much more complicated and will not be given here; they can be found in Landau [1929, pages 68–71].

22.25. Sequence spaces. Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$). For any sequence of scalars $x = (x_1, x_2, x_3, \ldots)$, define $\|x\|_\infty = \sup\Big\{|x_1|,\ |x_2|,\ |x_3|,\ \ldots\Big\}$ and

$$\|x\|_p = \Big\{|x_1|^p + |x_2|^p + |x_3|^p + \cdots\Big\}^{1/p} \qquad (0 < p < \infty).$$

Then define

$$\ell_p = \Big\{x \in \mathbb{F}^{\mathbb{N}} : \|x\|_p < \infty\Big\} \qquad (0 < p \leq \infty).$$

Then ℓ_p is a linear subspace of $\mathbb{F}^{\mathbb{N}}$. If $1 \leq p \leq \infty$, then $\| \ \|_p$ is a norm on ℓ_p (*hint*: 12.29.g); hence sequences of scalars satisfy **Minkowski's Inequality**:

$$\left\{\sum_{j=1}^{\infty} |x_j + y_j|^p\right\}^{1/p} \leq \left\{\sum_{j=1}^{\infty} |x_j|^p\right\}^{1/p} + \left\{\sum_{j=1}^{\infty} |y_j|^p\right\}^{1/p} \qquad (1 \leq p < \infty).$$

If $0 < p < 1$, then $\| \ \|_p$ generally is not a norm, but $\| \ \|_p^p$ is a G-norm on ℓ_p (*hint*: 12.25.e); in fact, we shall see in Chapter 26 that it is a special kind of G-norm that we call an *F-norm*.

The spaces ℓ_p are a simple but important special case of the spaces $L^p(\mu, X)$, introduced in 22.28. The completeness of the spaces ℓ_p and $L^p(\mu, X)$ will be proved in 22.31(i).

The **Cauchy-Schwarz Inequality** states that $\|xy\|_1 \leq \|x\|_2\|y\|_2$, where xy is the sequence whose nth term is $x_n y_n$. For a proof, take limits in 2.10.

Exercise. Let $0 < p < \infty$. Show that a subset S is relatively compact in ℓ_p if and only if it is metrically bounded and satisfies $\lim_{N\to\infty} \sup_{x\in S} \sum_{k=N}^{\infty} |x_k|^p = 0$.

A generalization. Let $\mathbb{J}$ be any nonempty set. For any function $x : \mathbb{J} \to \mathbb{F}$, define $\|x\|_\infty = \sup_{j\in\mathbb{J}} |x_j|$ and

$$\|x\|_p = \left\{\sum_{j\in\mathbb{J}} |x_j|^p\right\}^{1/p} \qquad (0 < p < \infty).$$

(Positive sums over arbitrary index sets are defined as in 10.40.) Then define

$$\ell_p(\mathbb{J}) = \{x \in \mathbb{F}^{\mathbb{J}} : \|x\|_p < \infty\} \qquad (0 < p \leq \infty).$$

For $1 \leq p \leq \infty$, $\ell_p(\mathbb{J})$ is a linear subspace of $\mathbb{F}^{\mathbb{J}}$ and $\| \ \|_p$ is a norm on that space. This generalization will be particularly useful in 22.56.

22.26. The James space J (optional). For sequences $x = (x_1, x_2, x_3, \ldots)$ of scalars, let

$$\begin{aligned}\|x\|_J = \sup\Big\{&\left|x_{k(1)} - x_{k(2)}\right|^2 + \left|x_{k(2)} - x_{k(3)}\right|^2 + \left|x_{k(3)} - x_{k(4)}\right|^2 \\ &+ \cdots + \left|x_{k(n-2)} - x_{k(n-1)}\right|^2 + \left|x_{k(n-1)} - x_{k(n)}\right|^2 + \left|x_{k(n)} - x_{k(1)}\right|^2\Big\}^{1/2}\end{aligned}$$

where the supremum is over all positive integers n and all finite increasing sequences $k(1) < k(2) < \cdots < k(n)$ of positive integers. Let $J = \{x \in c_0 : \|x\|_J < \infty\}$. This space was devised by James [1951] to answer several questions about normed spaces; one of those questions will be mentioned in the remarks in 28.41. The space J is discussed further by James [1982]. Show

a. $(J, \| \ \|_J)$ is a Banach space.

b. $\ell_2 \subsetneqq J$, and $\| \ \|_2$ is strictly stronger than $\| \ \|_J$ on ℓ_2. *Hint*: Use the Cauchy-Schwarz inequality.

c. $J \subsetneqq c_0$, and $\| \ \|_J$ is strictly stronger than the sup norm on J.

d. For $2 < p < \infty$, neither of ℓ_p or J includes the other, and neither of $\| \quad \|_p$ or $\| \quad \|_J$ is stronger than the other on $\ell_p \cap J$. *Hints*: To show $\|x\|_J / \|x\|_p$ is unbounded, consider

$$x = (1^{-r}, 0, 2^{-r}, 0, 3^{-r}, \ldots, 0, n^{-r}, 0, 0, 0, 0, 0, \ldots)$$

with $r \in (\frac{1}{p}, \frac{1}{2})$. To show $\|x\|_p / \|x\|_J$ is unbounded, consider a sequence of n 1s followed by infinitely many 0s.

Bochner-Lebesgue Spaces

22.27. Let $(\Omega, \mathcal{S})$ be a measurable space and $(X, | \quad |)$ be a Banach space. Let X be equipped with its σ-algebra $\mathcal{B}$ of Borel sets. Show that

a. The space

$$SM(\mathcal{S}, X) = \{\text{strongly measurable functions from } (\Omega, \mathcal{S}) \text{ to } (X, \mathcal{B})\}$$

is a linear subspace of X^Ω. *Hint*: Use 21.4(C).

b. The space

$$\mathcal{L}^\infty(\mathcal{S}, X) \quad = \quad \{f \in SM(\mathcal{S}, X) \ : \ f \text{ is bounded}\}$$

is a closed linear subspace of

$$B(\Omega, X) \quad = \quad \{\text{bounded functions from } \Omega \text{ into } X\}$$

when that space is equipped with the sup norm; hence $\mathcal{L}^\infty(\mathcal{S}, X)$ is a Banach space.

When $\mathbb{F}$ is the scalar field, the space $\mathcal{L}^\infty(\mathcal{S}, \mathbb{F})$ may be written more briefly as $\mathcal{L}^\infty(\mathcal{S})$. It follows from 21.4(E) that a dense subset of $\mathcal{L}^\infty(\mathcal{S})$ is given by the set of simple functions from $(\Omega, \mathcal{S})$ into $\mathbb{F}$, defined as in 11.42 — i.e., the measurable functions with finite ranges.

c. If X is separable, then the set $M(\mathcal{S}, X) = \{\text{measurable functions from } \Omega \text{ to } X\}$ is equal to $SM(\mathcal{S}, X)$; thus it is a linear space subspace of X^Ω.

d. If $\mathrm{card}(X) > \mathrm{card}(\mathbb{R})$, then there exists a measurable space $(\Omega, \mathcal{S})$ such that $M(\mathcal{S}, X) = \{\text{measurable functions from } \Omega \text{ to } X\}$ is *not* a linear space.

Proof. Let $\mathcal{B}$ be the σ-algebra of Borel subsets of X. Let $\mathcal{S} = \mathcal{B} \otimes \mathcal{B}$ denote the product σ-algebra on $\Omega = X \times X$. By 21.8, this is *not* the same as the σ-algebra of Borel subsets of the product topology on $X \times X$; in fact, the diagonal set belongs to that product topology but not to $\mathcal{S}$. Let $f(x, y) = x$ and $g(x, y) = y$. Then $f, g : \Omega \to X$ are measurable but $h = f - g$ is not, since $h^{-1}(0)$ is the diagonal set. This result is from Nedoma [1957].

e. *Remarks.* The results above show why we impose separability requirements throughout the theory of measure and integration.

Besides the cases described above, there is still one more case to consider: There exist some nonseparable Banach spaces X satisfying $\mathrm{card}(X) = \mathrm{card}(\mathbb{R})$. (*Exercise.*

Show that ℓ_∞ is such a space.) It is not presently known whether, whenever X is such a space and $(\Omega, \mathcal{S})$ is a measurable space, then $M(\mathcal{S}, X)$ is necessarily a linear space. Some related questions are considered by Stone [1976].

22.28. *Definitions.* Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let $(X, |\ \ |)$ be a Banach space. For each $f \in SM(\mathcal{S}, X)$, the function $|f(\cdot)| : \Omega \to [0, +\infty]$ is measurable. Hence, using the type of integral defined in 21.36, we can define the quantities

$$\begin{aligned} \|f\|_p &= \left\{ \int_\Omega |f(\omega)|^p d\mu(\omega) \right\}^{1/p} && (0 < p < \infty), \\ \|f\|_\infty &= \inf\{r > 0 \ : \ |f(\cdot)| \leq r \ \ \mu\text{-a.e.}\} \end{aligned}$$

— they are numbers in $[0, +\infty]$. (The case of $p = 1$ is particularly simple and important, so we shall restate it separately: $\|f\|_1 = \int_\Omega |f(\cdot)| d\mu$.) We can also define the set of functions

$$\mathcal{L}^p(\mu, X) \quad = \quad \{f \in SM(\mathcal{S}, X) \ : \ \|f\|_p < \infty\} \qquad\qquad (0 < p \leq \infty).$$

Then $\mathcal{L}^p(\mu, X)$ is a linear subspace of $SM(\mathcal{S}, X)$, for each $p \in (0, \infty]$. When $1 \leq p < \infty$, then $\|\ \ \|_p$ is a seminorm on that space. (*Hint*: 12.29.g.) When $0 < p \leq 1$, then $\|\ \ \|_p^p$ is a G-seminorm on that space (*hint*: 12.25.e); in fact, we shall see in Chapter 26 that it is an F-seminorm.

Note that $\mathcal{L}^\infty(\mathcal{S}, X)$ (defined in 22.27.b) includes only functions that are bounded, but $\mathcal{L}^\infty(\mu, X)$ consists of functions that are bounded almost everywhere. In fact, a function belongs to $\mathcal{L}^\infty(\mu, X)$ if and only if it agrees almost everywhere with some member of $\mathcal{L}^\infty(\mathcal{S}, X)$.

Remarks on membership in the Lebesgue spaces. Some mathematicians define the spaces $\mathcal{L}^p(\mu, X)$ a little differently, but in most cases their definitions are equivalent to the one given above. Note that f belongs to $\mathcal{L}^p(\mu, X)$ if and only if

1. f is "regular," in the sense that f belongs to $SM(\mathcal{S}, X)$, and
2. f is "not too big," in the sense that there exists some function $g \in \mathcal{L}^p(\mu, \mathbb{R})$ such that $|f(\cdot)| \leq g(\cdot)$.

These two conditions are entirely different in nature and can be studied separately from one another.

Associated metric spaces. For $0 < p \leq \infty$, in general the spaces $\mathcal{L}^p(\mu, X)$ are merely pseudometric spaces; we can make them into metric spaces by taking quotients in the usual fashion: Observe that $\|f - g\|_p = 0$ if and only if $f = g$ μ-a.e. This defines an equivalence relation $f \approx g$ on the pseudometric space $\mathcal{L}^p(\mu, X)$. The resulting metric space is denoted $L^p(\mu, X)$; we may call it the **Bochner-Lebesgue space** of order p. The seminorm $\|\ \ \|_p$ or G-seminorm $\|\ \ \|_p^p$ on $\mathcal{L}^p(\mu, X)$ (for $1 \leq p \leq \infty$ or $0 < p < 1$, respectively) acts as a norm or G-norm on $L^p(\mu, X)$.

In general, the spaces $\mathcal{L}^p(\mu, X)$ and $L^p(\mu, X)$ are different. Members of $\mathcal{L}^p(\mu, X)$ are functions, whereas members of $L^p(\mu, X)$ are equivalence classes of functions. In some contexts, members of $L^p(\mu, X)$ are discussed as if they were functions — i.e., the distinction between a function and its equivalence class is ignored. In certain contexts this abuse of language is convenient and does not cause confusion.

Although the spaces $\mathcal{L}^p(\mu, X)$ and $L^p(\mu, X)$ are different in general, they are the same in some special cases — for instance, when μ is counting measure, for then each equivalence class in $\mathcal{L}^p(\mu, X)$ contains only one function.

Notation for scalar-valued functions. When X is the scalar field $\mathbb{F}$, then we abbreviate $\mathcal{L}^p(\mu, X)$ as $\mathcal{L}^p(\mu)$ and abbreviate $L^p(\mu, X)$ as $L^p(\mu)$. The spaces $L^p(\mu)$ are called **Lebesgue spaces**.

When μ is counting measure on the finite set $\{1, 2, \ldots, n\}$, then $L^p(\mu) = \mathcal{L}^p(\mu)$ is just the finite dimensional space $\mathbb{F}^n$, normed as in 22.11. When μ is counting measure on $\mathbb{N}$, then $L^p(\mu) = \mathcal{L}^p(\mu)$ is just the sequence space ℓ_p; thus all the results proved below for integrals have corollaries about sums. More generally, when μ is counting measure on some set $\mathbb{J}$, then $L^p(\mu) = \mathcal{L}^p(\mu)$ is the generalized sequence space $\ell_p(\mathbb{J})$ introduced in 22.25.

When Ω is a subset of $\mathbb{R}^n$ and μ is n-dimensional Lebesgue measure, then $L^p(\mu)$ is usually written as $L^p(\Omega)$. For instance, if μ is one-dimensional Lebesgue measure on the interval $[0, 1]$, then $L^p(\mu)$ is usually written as $L^p(0, 1)$ or $L^p[0, 1]$. There is no substantial difference between $L^p(0, 1)$ and $L^p[0, 1]$ since a single point has Lebesgue measure 0.

Further notation. The number $\|f\|_\infty$ is sometimes called the **essential supremum** of the function f. *Caution*: That term has another meaning; see 21.42.

An **integrable** function is a member of $L^1(\mu, X)$ or $\mathcal{L}^1(\mu, X)$; this terminology is explained in 23.16.

22.29. Lebesgue's Dominated Convergence Theorem. Let $0 < p < \infty$. Let (f_n) be a sequence in $\mathcal{L}^p(\mu; X)$, converging pointwise to a limit f. Assume that the f_n's are dominated by some member of $\mathcal{L}^p(\mu; \mathbb{R})$ — i.e., assume that $|f_n(\omega)| \leq g(\omega)$ for some function $g \in \mathcal{L}^p(\mu; \mathbb{R})$. Then $f \in \mathcal{L}^p(\mu; X)$ and $\|f_n - f\|_p \to 0$.

Remarks. This theorem can be proved for Riemann integrals by more elementary methods — i.e., not involving σ-algebras and abstract measure theory. See Luxemburg [1971] and Simons [1995], and other papers cited therein.

Proof of theorem. We first prove this in the case of $p = 1$. Observe that $|f(\omega)| \leq g(\omega)$. Apply Fatou's Lemma (see 21.39.c) to the functions

$$h_n(\omega) \quad = \quad 2g(\omega) - |f_n(\omega) - f(\omega)|.$$

The remaining details for $p = 1$ are left as an exercise.

For other values of p, observe that the functions $F_n(\omega) = |f_n(\omega) - f(\omega)|^p$ converge pointwise to 0, and they are dominated by the function $G(\omega) = 2^p g(\omega)^p$, which lies in $\mathcal{L}^1(\mu; \mathbb{R})$. Hence $F_n \to 0$ in $\mathcal{L}^1(\mu; \mathbb{R})$ by the case of $p = 1$, and therefore $f_n \to f$ in $\mathcal{L}^p(\mu; X)$.

22.30. *Results about dense subsets.* Recall from 11.42 that a **simple function** is a measurable function whose range is a finite set. Let X be a Banach space, and let $0 < p < \infty$. Then:

a. A simple function f belongs to $\mathcal{L}^p(\mu, X)$ if and only if $\{\omega \in \Omega : f(\omega) \neq 0\}$ has finite measure.

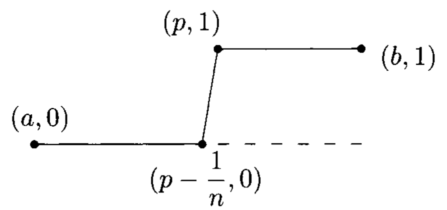
b. The simple functions that belong to $\mathcal{L}^p(\mu, X)$ are a dense subset of $\mathcal{L}^p(\mu, X)$. *Hint*: 21.4 and 22.29.

c. Let Ω be an interval in $\mathbb{R}$ (possibly all of $\mathbb{R}$), and let μ be a positive measure on the Borel subsets of Ω. Let X be a Banach space, and let $0 < p < \infty$. Then the integrable step functions are dense in $\mathcal{L}^p(\mu, X)$.

Hint: Use 21.27 to show that any integrable simple function can be approximated arbitrarily closely by an integrable step function.

d. Let μ be a positive finite measure on the Borel subsets of $[a, b]$. Then $C([a, b], X) = \{$continuous functions from $[a, b]$ into $X\}$ is a dense subset of $L^1(\mu, X)$.

Hints: In view of 22.30.c, it suffices to show that any step function can be approximated arbitrarily closely by a continuous function. Let us first consider how to approximate step functions of the form $1_{[p,b]}(\cdot)$, for any $p \in (a, b)$. Define f_n as in the following diagram. Show that $\|f_n - 1_{[p,b]}\|_1 \leq \mu\left([p - \frac{1}{n}, p)\right)$. That last quantity tends to 0 as $n \to \infty$, by 21.25.b. By left-right symmetry, we may approximate step functions of the form $1_{[a,p]}(\cdot)$ in an analogous fashion. Finally, show that any step function on $[a, b]$ is a linear combination of 1 and functions of the forms $1_{[a,p]}$ and $1_{[p,b]}$.



A continuous, piecewise-affine function f_n approximating $1_{[p,b]}$ in $L^1(\mu)$.

22.31. Let $0 < p < \infty$ and let X be a Banach space. Then:

(i) $L^p(\mu, X)$ is complete.

(ii) **(Converse to Dominated Convergence Theorem.)** Any convergent sequence in $L^p(\mu, X)$ has a subsequence that is convergent pointwise almost everywhere and is dominated by some member of $L^p(\mu, \mathbb{R})$.

Proof. To prove both statements, we shall show that any *Cauchy* sequence has a subsequence that is convergent pointwise almost everywhere and is dominated; then completeness follows from 22.29.

Let $\rho(f) = \|f\|_p^{\min\{1,p\}}$; then ρ is a G-norm on $L^p(\mu, X)$. Choose some subsequence (g_k) satisfying $\rho(g_k - g_{k+1}) < 2^{-k}$. Let $h_N(\omega) = |g_1(\omega)| + \sum_{k=1}^N |g_{k+1}(\omega) - g_k(\omega)|$. Since ρ is subadditive, we have $\rho(h_N) \leq \rho(g_1) + \sum_{k=1}^N \rho(g_{k+1} - g_k) \leq \rho(g_1) + 2^{-N+1}$. The functions h_N take values in $[0, +\infty)$, and they increase pointwise to the function $h(\omega) = |g_1(\omega)| + \sum_{k=1}^\infty |g_{k+1}(\omega) - g_k(\omega)|$. Use the Monotone Convergence Theorem to show that h is a member of $L^p(\mu, \mathbb{R})$, with $\rho(h) \leq \rho(g_1) + 2$; clearly h dominates the g_k's. Since

$\rho(h) < \infty$, it follows from 21.37.h that $h(\omega) < \infty$ for μ-almost every ω. From this it follows that the sequence $(g_k(\omega) : k \in \mathbb{N})$ is Cauchy in X for μ-almost every ω. By assumption, X is complete, so its Cauchy sequences converge.

22.32. Reverse Minkowski Inequality. If f and g are measurable functions taking values in $[0, +\infty)$ and $0 < s \le 1$, then $\|f + g\|_s \ge \|f\|_s + \|g\|_s$.

Proof. By 21.5 and the Dominated Convergence Theorem, it suffices to prove the present result for finitely valued functions. Let $r = 1/s \in [1, +\infty)$. We are to show that

$$\left\{\sum_{i=1}^{M}(x_i + y_i)^{1/r}\mu_i\right\}^r \quad \ge \quad \left\{\sum_{i=1}^{M} x_i^{1/r}\mu_i\right\}^r + \left\{\sum_{i=1}^{M} y_i^{1/r}\mu_i\right\}^r$$

for nonnegative numbers x_i, y_i, μ_i.

By induction on M, Minkowski's inequality yields $\left\|\sum_{i=1}^M v_i\right\|_r \le \sum_{i=1}^M \|v_i\|_r$ for any vectors $v_1, v_2, \ldots, v_M$. When those vectors are members of $\mathbb{R}^2$, that inequality tells us

$$\left(\sum_{j=1}^{2}\left|\sum_{i=1}^{M} v_{ij}\right|^r\right)^{1/r} \quad \le \quad \sum_{i=1}^{M}\left(\sum_{j=1}^{2}|v_{ij}|^r\right)^{1/r}$$

for any positive integer M and real numbers v_{ij}. Now take the rth power on both sides, and then substitute $v_{i1} = x_i^{1/r}\mu_i$ and $v_{i2} = y_i^{1/r}\mu_i$ for $i = 1, 2, \ldots, M$; this yields the desired inequality.

22.33. Let $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$. (Numbers p and q related in this fashion are called **conjugate exponents**.) Let f, g be measurable scalar-valued functions, and let fg be the pointwise product — i.e., the function whose value at ω is $f(\omega)g(\omega)$. Then we have **Hölder's Inequality**:

$$\|fg\|_1 \le \|f\|_p \|g\|_q$$

(whether those quantities are finite or not). Moreover, if $p, q \in (1, \infty)$ and $|f(\omega)|^p = c|g(\omega)|^q$ for all ω and some constant c, then we have **Hölder's Equality**: $\|fg\|_1 = \|f\|_p\|g\|_q$.

The special case of Hölder's inequality with $p = q = 2$ is important enough to deserve separate mention; it is the **Cauchy-Schwarz Inequality**:

$$\int_\Omega |f(\omega)g(\omega)|\, d\mu(\omega) \quad \le \quad \sqrt{\int_\Omega |f(\omega)|^2\, d\mu(\omega)}\,\sqrt{\int_\Omega |g(\omega)|^2\, d\mu(\omega)}\,.$$

Proof. The case of $p = 1$ and $q = \infty$, or vice versa, is easy; we omit the details. Assume that $p, q \in (1, \infty)$. We may assume both of the numbers $\|f\|_p$ and $\|g\|_q$ are nonzero. (Why?) Using homogeneity, we may replace the functions f and g with the functions $f(\cdot)/\|f\|_p$ and $g(\cdot)/\|g\|_q$, respectively; hence we may assume $\|f\|_p = \|g\|_q = 1$ (explain). If $|f(\omega)|^p = c|g(\omega)|^q$ for some constant c, then $c = 1$. By 12.20.a, we have $|f(\omega)g(\omega)| \le \frac{1}{p}|f(\omega)|^p + \frac{1}{q}|g(\omega)|^q$, with equality if $|f(\omega)|^p = |g(\omega|^q$. Now integrate to obtain $\|fg\|_1 \le \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q = 1$.

22.34. Suppose that $0 < \alpha < \beta \leq \infty$. Then:

a. Inequalities for sequence spaces. $\|x\|_\beta \leq \|x\|_\alpha$ for sequences x, and $\ell_\alpha \subsetneqq \ell_\beta$. In particular, $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$ and $\ell_1 \subsetneqq \ell_2 \subsetneqq \ell_\infty$.

Hints: Show that if $\|x\|_\alpha \leq 1$ then $|x_j| \leq 1$ for all j, hence $|x_j|^r \leq |x_j| \leq |x_j|^s$ if $r \geq 1 \geq s$. Also, find some constant t such that the sequence $f = (1^t, 2^t, 3^t, \ldots)$ belongs to ℓ_β and not to ℓ_α.

b. Inequalities for probability spaces. If μ is a probability measure, then $\|h\|_\alpha \leq \|h\|_\beta$ and $L^\beta(\mu, X) \subseteq L^\alpha(\mu, X)$. In particular, $\|h\|_1 \leq \|h\|_2 \leq \|h\|_\infty$ and $L^\infty(\mu, X) \subseteq L^2(\mu, X) \subseteq L^1(\mu, X)$.

Hints: Use 22.33 with $f = |h|^\alpha$, $g = 1$, $p = \beta/\alpha$, $q = \beta/(\beta - \alpha)$. Prove separately for $\beta = \infty$.

Remark. For a more general result, see the remarks in 27.29.

22.35. Clarkson's Inequality. Let $p, q \in (1, +\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, and let $\alpha = \min\{p, q\}$ and $\beta = \max\{p, q\}$. Then for any measurable scalar-valued functions f and g,

$$\|f + g\|_p^\beta + \|f - g\|_p^\beta \quad \leq \quad 2\left(\|f\|_p^\alpha + \|g\|_p^\alpha\right)^{\beta/\alpha}.$$

Note. There are several other inequalities also known as "Clarkson's Inequalities" — and in fact, some of them will appear in the proof below — but the inequality given above is the most important one for our applications in 22.41.a and thereafter.

Proof. This proof follows the presentation of Weir [1974]. For most steps of this proof, we shall give inequalities only for $p \geq 2 \geq q$; the reversed inequalities are then valid when $p \leq 2 \leq q$.

It follows from 10.35 that

$$|f(\omega) + g(\omega)|^p + |f(\omega) - g(\omega)|^p \quad \leq \quad 2\left(|f(\omega)|^q + |g(\omega)|^q\right)^{p/q} \qquad \text{if } p \geq 2 \geq q,$$

with inequality reversed if $p \leq 2 \leq q$. Now integrate; this yields

$$\|f + g\|_p^p + \|f - g\|_p^p \;\leq\; 2\int_\Omega \left(|f(\omega)|^q + |g(\omega)|^q\right)^{p/q} d\mu(\omega),$$

still assuming $p \geq 2 \geq q$, and we have the reverse of this inequality if $p \leq 2 \leq q$.

For any nonnegative scalar-valued measurable functions F and G, we have

$$\left[\int_\Omega (F(\omega) + G(\omega))^{p/q}\, d\mu(\omega)\right]^{q/p} \quad \leq \quad \left[\int_\Omega F(\cdot)^{p/q}\, d\mu\right]^{q/p} + \left[\int_\Omega G(\cdot)^{p/q}\, d\mu\right]^{q/p}$$

if $p \geq 2 \geq q$ (by Minkowski's Inequality), or the reverse of this inequality if $p \leq 2 \leq q$ (by the Reverse Minkowski Inequality). Raise both sides of this inequality to the power $\frac{p}{q}$, and then multiply by 2; thus

$$2\int_\Omega (F(\omega) + G(\omega))^{p/q}\, d\mu(\omega) \;\leq\; 2\left(\left[\int_\Omega F(\cdot)^{p/q}\, d\mu\right]^{q/p} + \left[\int_\Omega G(\cdot)^{p/q}\, d\mu\right]^{q/p}\right)^{p/q}$$

(or the reverse of this inequality). Apply this result with $F(\omega) = |f(\omega)|^q$ and $G(\omega) = |g(\omega)|^q$, and simplify the right side. We obtain

$$2\int_\Omega \Big(|f(\omega)|^q + |g(\omega)|^q\Big)^{p/q}\, d\mu(\omega) \;\le\; 2\Big(\|f\|_p^q + \|g\|_p^q\Big)^{p/q}$$

if $p \ge 2 \ge q$, or the reverse of this inequality if $p \le 2 \le q$.

Combine the conclusions of the last two paragraphs. We have established that

$$\begin{array}{lcll} \|f+g\|_p^p + \|f-g\|_p^p & \le & 2\left(\|f\|_p^q + \|g\|_p^q\right)^{p/q} & \text{if } p \ge 2 \ge q, \\ \|f+g\|_p^p + \|f-g\|_p^p & \ge & 2\left(\|f\|_p^q + \|g\|_p^q\right)^{p/q} & \text{if } p \le 2 \le q. \end{array}$$

Substituting $f = u + v$ and $g = u - v$ and then rearranging a bit yields

$$\begin{array}{lcll} 2\left(\|u\|_p^p + \|v\|_p^p\right)^{q/p} & \le & \|u+v\|_p^q + \|u-v\|_p^q & \text{if } p \ge 2 \ge q, \\ 2\left(\|u\|_p^p + \|v\|_p^p\right)^{q/p} & \ge & \|u+v\|_p^q + \|u-v\|_p^q & \text{if } p \le 2 \le q. \end{array}$$

The first and last of these four inequalities give us the result stated in the theorem.

22.36. *Definition.* Let $[a, b]$ be some interval in $\mathbb{R}$ equipped with its σ-algebra of Lebesgue-measurable sets, and let X and Y be Banach spaces equipped with their σ-algebras of Borel sets. Let G be an open subset of X. Let $f : [a,b] \times G \to Y$ be jointly measurable, and suppose f takes separable sets to separable sets. Also assume that for each fixed $x_0 \in G$, the mapping $f(\cdot, x_0) : [a,b] \to Y$ is integrable. We shall say that f is

> **integrably Lipschitz** if there exists a function $\lambda \in L^1[a,b]$ such that $\|f(t,x_1) - f(t,x_2)\| \le \lambda(t)\|x_1 - x_2\|$ for all $t \in [a,b]$ and $x_1, x_2 \in G$; or
>
> **integrably locally Lipschitz** if for each compact $K \subseteq G$ there exists a function $\lambda_K \in L^1[a,b]$ with $\|f(t,x_1) - f(t,x_2)\| \le \lambda_K(t)\|x_1 - x_2\|$ for all $t \in [a,b]$ and $x_1, x_2 \in K$.

Proposition. Let $f : [a,b] \times G \to Y$ be integrably locally Lipschitz, as above. Then for each compact set $K \subseteq G$ there exist an open set H with $K \subseteq H \subseteq G$ and a function $\varphi \in L^1[a,b]$ with this property: Whenever $u, v : [a,b] \to H$ are continuous functions, then

$$\|f(t,u(t)) - f(t,v(t))\| \quad \le \quad \varphi(t)\|u(t) - v(t)\|$$

for almost all $t \in [a,b]$. (This result will be used in 30.9.)

Proof. Note that any continuous function defined on $[a,b]$ is measurable and has compact range. It suffices to apply 21.44 with $Z = G \times G$, $\Omega = [a,b]$, and

$$\Gamma\Big(t,\ (x,y)\Big) \quad = \quad \begin{cases} \dfrac{\|f(t,x_1) - f(t,x_2)\|}{\|x_1 - x_2\|} & \text{when } x_1 \ne x_2 \\ \\ 0 & \text{when } x_1 = x_2. \end{cases}$$

(A more complicated but more general argument of this sort by Schechter [1981] dealt with not only Lipschitzness, but also uniform continuity.)

Strict Convexity and Uniform Convexity

22.37. *Observation.* Any norm is a convex function.

22.38. *Definitions.* Let $(X, \| \quad \|)$ be a normed linear space. We say X (or its norm) is

strictly convex if, whenever $\|x\| = \|y\| = 1$ and $\|x + y\| = 2$, then $x = y$;

locally uniformly convex if, whenever $\|x\| = 1$ and (y_n) is a sequence with $\|y_n\| = 1$ and $\|x + y_n\| \to 2$, then $\|x - y_n\| \to 0$;

uniformly convex if, whenever (x_n) and (y_n) are sequences satisfying $\|x_n\| = \|y_n\| = 1$ and $\|x_n + y_n\| \to 2$, then $\|x_n - y_n\| \to 0$.

Clearly, X is uniformly convex $\Rightarrow$ X is locally uniformly convex $\Rightarrow$ X is strictly convex.

We remark that strict, locally uniform, and uniform convexity are *not* topological properties. If we replace a norm with an equivalent norm — i.e., one that yields the same topology — then the convexity properties described above are not necessarily preserved; that will be clear from examples in 22.41. Nevertheless, uniform convexity of a norm does *imply* certain topological properties; see 28.46.

22.39. *Reformulations of the definition of strict convexity.* Let $(X, \| \quad \|)$ be a normed space, and let S be the unit sphere — i.e., let $S = \{x \in X \ : \ \|x\| = 1\}$. Then the following conditions are equivalent. If any (hence all) of them is satisfied, then the normed space $(X, \| \quad \|)$ is *strictly convex.*

(A) If $\|u + v\| = \|u\| + \|v\|$, then u, v, and 0 lie on one straight line.

(B) Any straight line intersects S in at most two points.

(C) Any convex subset of S contains at most one point.

(D) If C is a nonempty convex subset of X and $u \in X$, then at most one point of C is closest to u. That is, at most one point $c \in C$ satisfies $\|u - c\| = \mathrm{dist}(u, C)$.

(E) Any convex subset of X contains at most one point of minimum norm. (Compare with 28.41(F).)

(F) If $\|u\| = 1$ for all u in some line segment $[x, y]$ (with notation as in 12.5.i), then $x = y$.

(G) If x, y, z are distinct points satisfying $\|x\| = \|y\| = \|z\|$, then x, y, z are not all on one straight line.

(H) If x and y are points in X satisfying $\|x\| = \|y\| = \left\|\frac{1}{2}x + \frac{1}{2}y\right\|$, then $x = y$. (This condition is easily seen to be equivalent to the definition of strict convexity given in 22.38.)

(I) $\| \quad \|$ is a strictly convex function (as defined in 12.17.c) on each straight line that does not pass through 0. In other words, if u and v are points satisfying

$\|\lambda u + (1-\lambda)v\| = \lambda\|u\| + (1-\lambda)\|v\|$ for at least one $\lambda \in (0,1)$, then u, v, and 0 are all on one straight line.

Proof of (A) $\Rightarrow$ (B). Suppose that x, y, z are three distinct points on one straight line, with $\|x\| = \|y\| = \|z\|$. By relabeling we may assume that z lies between x and y; thus $z = \lambda x + (1-\lambda)y$ for some $\lambda \in (0,1)$. Let $u = \lambda x$ and $v = (1-\lambda)y$. Then

$$\|u+v\| \quad = \quad \|z\| \quad = \quad \lambda\|z\| + (1-\lambda)\|z\| \quad = \quad \|u\| + \|v\|.$$

Apply (A); this tells us that the line through x and y passes through 0; it also passes through z. This leads to a contradiction.

Proof of (B) $\Rightarrow$ (C). Trivial.

Proof of (C) $\Rightarrow$ (D). By translation (i.e., replacing C with $C - u$), we may assume $u = 0$. We may assume $\mathrm{dist}(0, C) > 0$ (explain). By homothety (i.e., replacing C with kC for a suitable positive number k), we may assume $\mathrm{dist}(0, C) = 1$. Then $\{c \in C : \|c\| = 1\} = \{c \in C : \|c\| \leq 1\}$ is a convex subset of S, so it contains at most one point.

Proof of (D) $\Rightarrow$ (E) $\Rightarrow$ (F). Obvious.

Proof of (F) $\Rightarrow$ (G). Assume that x, y, z are all on one line. We may assume that $\|x\| = \|y\| = \|z\| = 1$. One of these points is between the other two; by relabeling we may assume z is between x and y. We have $\|u\| \leq 1$ for all $u \in [x, y]$, by convexity of $\|\ \ \|$. If $\|u_1\| < 1$ for some $u_1 \in [x, y]$, then z lies between u_1 and x or between u_1 and y, hence $\|z\| < 1$ by convexity of $\|\ \ \|$.

Proof of (G) $\Rightarrow$ (H). Trivial.

Proof of (H) $\Rightarrow$ (I). The function $t \mapsto c(t) = \|tu + (1-t)v\|$ is convex. By 12.17.c, if $\|c(\lambda)\| = \lambda\|u\| + (1-\lambda)\|v\|$ for at least one $\lambda \in (0,1)$, then in fact $\|c(\lambda)\| = \lambda\|u\| + (1-\lambda)\|v\|$ for every $\lambda \in (0,1)$. We may assume that u and v are both nonzero. Choose $\lambda \in (0,1)$ to satisfy $\lambda\|u\| = (1-\lambda)\|v\|$; then let $x = \lambda u$ and $y = (1-\lambda)v$. Then apply (H).

Proof of (I) $\Rightarrow$ (A). Use $\lambda = 1/2$.

22.40. *Reformulations of the definition of uniform convexity.* Let $(X, \|\ \ \|)$ be a normed space. Then the following conditions are equivalent.

(A) X is *uniformly convex*, as defined in 22.38. That is, whenever (x_n) and (y_n) are sequences with $\|x_n\| = \|y_n\| = 1$ and $\|x_n + y_n\| \to 2$, then $\|x_n - y_n\| \to 0$.

(B) Whenever (u_n) and (v_n) are sequences with $\|u_n\|, \|v_n\| \to 1$ and $\|u_n + v_n\| \to 2$, then $\|u_n - v_n\| \to 0$.

(C) Whenever (p_n) is a sequence with $\|p_n\| \to 1$ and $\lim_{m,n\to\infty} \|p_m + p_n\| \to 2$, then (p_n) is Cauchy.

(D) For each $\varepsilon > 0$, there exists some $\delta = \delta(\varepsilon) > 0$ such that $\|u\|, \|v\| \leq 1$ and $\frac{1}{2}\|u + v\| \geq 1 - \delta$ imply $\|u - v\| \leq \varepsilon$.

In condition (D), the largest δ that will work is clearly

$$\delta(\varepsilon) \quad = \quad \inf\left\{1 - \left\|\frac{u+v}{2}\right\| \ : \ \|u\|, \|v\| \leq 1 \text{ and } \|u-v\| \geq \varepsilon\right\}.$$

This formula defines an increasing function $\delta : (0,2) \to (0,1)$, called the **modulus of convexity** of the space. By obvious substitutions, we obtain the following inequality, which may be more convenient in some applications:

$$\|p-x\|, \|p-y\| \leq r, \qquad \|x-y\| \geq \varepsilon r \qquad \Rightarrow \qquad \left\|p - \frac{x+y}{2}\right\| \leq (1-\delta(\varepsilon))\, r.$$

Hints for the equivalence proof: For (C) $\Rightarrow$ (B), let (p_n) be the sequence $(u_1, v_1, u_2, v_2, \ldots)$. For (A) $\Rightarrow$ (D), let $x_n = u_n/\|u_n\|$ and $y_n = v_n/\|v_n\|$.

22.41. *Examples.*

a. Let $1 < p < \infty$. When $\|f\| = \|g\| = 1$, then Clarkson's Inequality (proved in 22.35) yields modulus of convexity less than or equal to the function $\delta(\varepsilon) = 1 - \left[1 - \left(\frac{\varepsilon}{2}\right)^\beta\right]^{1/\beta}$; thus $L^p(\mu)$ is uniformly convex.

Optional remarks. When $p > 2 > q$ then this estimate is the best possible, and so the function δ defined above (with $\beta = p$) is actually equal to the modulus of convexity of $L^p(\mu)$; this is shown by Hanner [1955]. However, when $p < 2 < q$, then the estimate can be improved slightly; Hanner shows that the modulus of convexity $\delta(\varepsilon)$ is the slightly smaller function defined implicitly by the equation $\left(1 - \delta + \frac{\varepsilon}{2}\right)^p + \left(1 - \delta - \frac{\varepsilon}{2}\right)^p = 2$.

b. In general, norms of type $\| \ \|_1$ are not strictly convex. For instance, when $\mathbb{R}^2$ is equipped with the norm $\|(x_1,x_2)\|_1 = |x_1| + |x_2|$, then the unit sphere contains the line segment $\{(x_1,x_2) : x_1, x_2 \geq 0, x_1 + x_2 = 1\}$.

c. In general, norms of type $\| \ \|_\infty$ are not strictly convex. For instance, when $\mathbb{R}^2$ is equipped with the norm $\|(x_1,x_2)\|_\infty = \max\{|x_1|, |x_2|\}$, then the unit sphere contains the line segment $\{(x_1,x_2) : x_1 = 1 \text{ and } -1 \leq x_2 \leq 1\}$.

d. (*A renorming example due to Clarkson.*) Let $\mathbb{F}$ be the scalar field, and let $C[0,1] =$ {continuous functions from $[0,1]$ into $\mathbb{F}$}. Let $(t_n : n = 1,2,3,\ldots)$ be a dense sequence in (0,1) — e.g., the rationals in (0,1) or the dyadic rationals. For continuous $f : [0,1] \to \mathbb{F}$, let

$$\|f\|_C \quad = \quad \left[\|f\|_\infty^2 + \sum_{n=1}^{\infty} 4^{-n}|f(t_n)|^2\right]^{1/2}.$$

Show that $\| \ \|_C$ is a strictly convex norm on $C[0,1]$ that is equivalent to $\| \ \|_\infty$. *Hint*: Use the strict convexity of ℓ_2.

e. (*Lovaglia's example.*) Show that Clarkson's norm $\| \ \|_C$, given in 22.41.d, is not locally uniformly convex, by letting $x(t)$ be the constant $(3/4)^{1/2}$ and $y_n(t) = x(t)\min\{1, nt\}$.

22.42. If X is a uniformly convex normed space, then the completion of X is also uniformly convex; it has the same modulus of convexity.

(*Optional.*) The completion of a strictly convex space need not be strictly convex, as the following example shows. For sequences of scalars $y = (y_0, y_1, y_2, \ldots)$, let $\|y\| = |y_0| + \sqrt{\sum_{j=1}^{\infty} 4^{-j}|y_j|^2}$. Let Y be the set of all sequences for which $\|y\| < \infty$. Let X be the subspace consisting of those sequences y that also satisfy $\lim_{j\to\infty} y_j = 0$. Show that

- **a.** $(Y, \|\ \ \|)$ is a Banach space, and X is a dense linear subspace.
- **b.** X is strictly convex. *Hint*: Use the fact that ℓ_2 is complete and strictly convex.
- **c.** Y is not strictly convex. *Hint*: 22.41.b.

22.43. Clarkson's Renorming Theorem. Let $(X, \|\ \ \|)$ be a separable normed space. Then $\|\ \ \|$ is equivalent to a strictly convex norm.

Proof. This short proof is from Riley [1981]. Let $\mathbb{F}$ be the scalar field. Let (x_n) be a sequence in X with the property that every point in X is a limit of some subsequence of (x_n) (see 15.13.g). For $n = 1, 2, 3, \ldots$, let

$$f_n(y) \quad = \quad \mathrm{dist}(y, \mathbb{F}x_n) \quad = \quad \inf_{\lambda\in\mathbb{F}} \|y - \lambda x_n\|.$$

Define $\gamma(y) = \|y\| + \sum_{n=1}^{\infty} 2^{-n} f_n(y)$. Then show

- **a.** Each f_n is a seminorm on X, with $f_n(\cdot) \le \|\cdot\|$.
- **b.** γ is a norm on X that is equivalent to $\|\ \ \|$.
- **c.** Now let y and z be nonzero vectors in X, with $\gamma(y+z) = \gamma(y) + \gamma(z)$. It suffices to show that $y = tz$ for some $t > 0$. Show, first of all, that $\|y+z\| = \|y\| + \|z\|$, and $f_n(y+z) = f_n(y) + f_n(z)$ for all n.
- **d.** Since (x_n) is dense in X, there is some subsequence $(x_{n(j)})$ that is $\|\ \ \|$-convergent to $y+z$. That is, $\|y+z-x_{n(j)}\| \to 0$ as $j \to \infty$. Hence $f_{n(j)}(y+z) \to 0$, and therefore $f_{n(j)}(y) \to 0$. Thus there exist scalars λ_j with $\|y - \lambda_j x_{n(j)}\| \to 0$.
- **e.** We consider two cases now: First, suppose the sequence (λ_j) is unbounded. Replacing it with a subsequence (explain), we may assume that $1/\lambda_j \to 0$. Using the joint continuity of multiplication (noted in 22.7), show that $\|x_{n(j)}\| \to 0$, hence $\|y+z\| = 0$, hence $y = z = 0$, a contradiction.
- **f.** Thus, the sequence (λ_j) is bounded. Replacing it with a subsequence (explain), we may assume that (λ_j) converges to some finite scalar λ. In that case, again using the joint continuity of multiplication, show $y = \lambda(y+z)$; hence $\lambda \ne 0$.
- **g.** Similarly, $z = \mu(y+z)$ for some nonzero scalar μ, so $y = tz$ for some nonzero scalar t.
- **h.** Since also $\|y+z\| = \|y\| + \|z\|$, show that $|1+t| = 1 + |t|$, and therefore $t > 0$.

22.44. *Remarks.* The theorem above was originally proved for norms by Clarkson. The proof given above can also be applied to F-norms, if interpreted appropriately.

Still more is true, at least for norms. We have in fact

Kadec's Renorming Theorem. Every separable normed space has an equivalent norm that is locally uniformly convex.

The proof of Kadec's theorem is longer and deeper, and will not be given here.

Some other, related results: Any separable normed space has an equivalent norm that makes both X and its dual strictly convex (Klee, 1959). If both X and its dual are separable (Kadec, Klee, Asplund) or if X is reflexive (Troyanski), then X has an equivalent norm that makes both X and its dual locally uniformly convex. For further, related reading and references, see Diestel [1975], Istrăţescu [1984], and Lindenstrauss [1988].

22.45. Theorem on closest points. Let Q be a convex subset of a Banach space X. Assume either

(i) X is strictly convex and Q is compact, or

(ii) X is uniformly convex and Q is closed.

Then for each point $x \in X$ there is a unique point $\pi(x) \in Q$ that is closest to x — i.e., that satisfies $\|x - \pi(x)\| = \mathrm{dist}(x, Q)$. Furthermore, this function $\pi : X \to Q$ is continuous. It is called the **closest point projection** onto Q.

Proof. Uniqueness follows from 22.39(D).

For any $x \in X$, there exists a sequence (q_n) in Q that satisfies $\|x - q_n\| \to \mathrm{dist}(x, Q)$; any such sequence will be called a *minimizing sequence for* x in this proof. Note that any subsequence of a minimizing sequence is a minimizing sequence. To prove the existence of $\pi(x)$ it suffices to show that

(!) any minimizing sequence for x has a convergent subsequence.

That is easy in case (i), since any sequence in a compact metric space has a convergent subsequence. The proof of (!) will take slightly longer for case (ii). Let (q_n) be a minimizing sequence, and let $r = \mathrm{dist}(x, Q)$. The result is trivial if $r = 0$; we shall assume $r > 0$. By rescaling, we may assume $r = 1$. Thus $\|x - q_m\| \to 1$ and $\|x - q_n\| \to 1$ as $m, n \to \infty$. On the other hand, $\frac{1}{2}(q_m + q_n) \in Q$ since Q is convex; thus $\|x - \frac{1}{2}(q_m + q_n)\| \geq \mathrm{dist}(x, Q) = 1$. Therefore $\|(x - q_m) + (x - q_n)\| \to 2$. By 22.40(C) the sequence (q_n) is Cauchy. This completes the proof of (!). Thus π is defined everywhere on X.

To show π is continuous, suppose (x_n) is a sequence converging in X to some limit x_∞; we must show that $\pi(x_n)$ converges to $\pi(x_\infty)$. Suppose not. Replacing (x_n) with a subsequence, we may assume $\|\pi(x_n) - \pi(x_\infty)\| > \kappa$ for some constant $\kappa > 0$. We know that $\mathrm{dist}(x_n, Q) \to \mathrm{dist}(x_\infty, Q)$ by 4.41.b; hence $(\pi(x_n))$ is a minimizing sequence for x_∞. Replacing (x_n) with a subsequence, by (!) we know that $(\pi(x_n))$ converges to some limit $q \in Q$. Then $\|q - \pi(x_\infty)\| \geq \kappa > 0$, so $q \neq \pi(x_\infty)$. Thus, q is not the member of Q closest to x_∞, so $\|q - x_\infty\| > \mathrm{dist}(x_\infty, Q)$. Hence $\|q - x_\infty\| > r > \mathrm{dist}(x_\infty, Q)$ for some real number r. Then for all n sufficiently large we have $\|\pi(x_n) - x_n\| > r > \mathrm{dist}(x_n, Q)$, a contradiction. Thus π is continuous.

Hilbert Spaces

22.46. *Definition.* Let X be a linear space over $\mathbb{F}$. An **inner product** on X is a mapping $\langle\ ,\ \rangle : X \times X \to \mathbb{F}$ that satisfies:

$$\langle \cdot, y\rangle : X \to \mathbb{F} \text{ is linear, for each } y \in X \qquad \text{(linear in first component)}$$

$$x \neq 0 \quad \Rightarrow \quad \langle x, x\rangle > 0 \qquad \text{(positive-definiteness)}$$

$$\langle x, y\rangle = \overline{\langle y, x\rangle}, \qquad \text{(conjugate symmetry)}$$

where the bar denotes complex conjugation. The conjugate symmetry condition is sometimes called "antisymmetry." If the scalar field $\mathbb{F}$ is $\mathbb{R}$, then the complex conjugate of a scalar is equal to that scalar, and so the conjugate symmetry condition becomes

$$\langle x, y\rangle = \langle y, x\rangle, \qquad \text{(symmetry)}$$

and it also implies that $\langle\ ,\ \rangle$ is bilinear — i.e., linear in each of its two arguments.

An **inner product space** is a linear space equipped with an inner product. As we shall see in an exercise below, if $\langle\quad,\quad\rangle$ is an inner product then $\|x\| = \langle x, x\rangle^{1/2}$ is a norm on X. An inner product space will always be understood to be equipped with this norm, unless some other arrangement is specified. If the norm is complete, then the inner product space is called a **Hilbert space**.

22.47. *Examples.* If $(\Omega, \mathcal{S}, \mu)$ is any measure space, then $L^2(\mu)$ is a Hilbert space, with inner product defined by

$$\langle f, g\rangle \quad = \quad \int_\Omega f(\omega)\overline{g(\omega)}\, d\mu(\omega).$$

The convergence of the integral is guaranteed by Hölder's inequality. If the scalar field is $\mathbb{R}$, then the bar over the $g(j)$ may be omitted. We note some important special cases.

a. Let $(\Omega, \mathcal{S}, \mu)$ be some set $\mathbb{J}$ equipped with counting measure. Then we obtain the normed space $\ell_2(\mathbb{J})$ introduced in 22.25. It has inner product

$$\langle f, g\rangle \quad = \quad \sum_{j \in \mathbb{J}} f(j)\overline{g(j)}.$$

In 22.56 we shall prove that *every* Hilbert space can be expressed in this form — i.e., every Hilbert space is isomorphic to some $\ell_2(\mathbb{J})$. However, other representations of Hilbert spaces are often useful.

b. When $\mathbb{J}$ is a finite set containing n elements, we find that $\mathbb{F}^n$ is a Hilbert space when equipped with the inner product

$$\langle x, y\rangle \quad = \quad x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}\,.$$

If the scalar field is $\mathbb{R}$, then the bar over the y_j's may be omitted. In $\mathbb{R}^n$, the inner product is also known as the **dot product**; it is used in analytic geometry to give algebraic formulas for much of Euclidean geometry.

22.48. *Some elementary properties.* Let $\langle\ ,\ \rangle$ be an inner product on some vector space, and let $\|x\| = \langle x, x\rangle^{1/2}$. (We do not yet assert that $\|\ \|$ is a norm; that fact is shown below.) Show that

- **a.** $\|x+y\|^2 = \|x\|^2 + 2\,\mathrm{Re}\langle x,y\rangle + \|y\|^2$.
  (Hence $\mathrm{Re}\langle x,y\rangle$ is uniquely determined by $\|\ \|$. It follows easily that $\langle x,y\rangle$ is uniquely determined by $\|\ \|$.)
- **b. Schwarz Inequality.** $|\langle x,y\rangle| \le \|x\|\,\|y\|$.
  Hint: Substitute $c = \langle x,y\rangle/\|y\|^2$, and use $0 \le \|x - cy\|^2$.
- **c.** $\|\ \|$ is a norm on X.
- **d.** The mapping $(x,y) \mapsto \langle x,y\rangle$ is a continuous map from $X \times X$ (with the product topology) into $\mathbb{F}$.
- **e. Parallelogram Equation.** $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$.
  Remark. Clarkson's Inequality may be viewed as a generalization of the Parallelogram Equation. Clarkson's Inequality tells us that $\|\ \|_p$ norms, for $1 < p < \infty$, are "almost as good as" the norms of inner product spaces.
- **f.** Any inner product space is uniformly convex.

22.49. *Converse results (optional).* Let $(X, \|\ \|)$ be a normed space whose norm satisfies the Parallelogram Equality. Then $\|\ \|$ arises from an inner product $\langle\ ,\ \rangle$, which is uniquely determined by $\|\ \|$.

(i) If $\mathbb{F} = \mathbb{R}$, then $\frac{1}{4}\left[\|x+y\|^2 - \|x-y\|^2\right] = \langle x,y\rangle$.

(ii) If $\mathbb{F} = \mathbb{C}$, then $\frac{1}{4}\left[\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right] = \langle x,y\rangle$.

Hint: 22.48.a.

22.50. Let X be a linear space, and let $\langle\ ,\ \rangle$ be an inner product on X. In this context, we say that two elements $x, y \in X$ are **orthogonal** to each other, denoted $x \perp y$, if $\langle x,y\rangle = 0$. For any set $S \subseteq X$, the **orthogonal complement** of S is the set

$$S^{\perp} \quad = \quad \{x \in X \quad : \quad x \perp s \text{ for all } s \in S\}.$$

This definition is a special case of 4.12, with

$$\Gamma \quad = \quad \{(x,y)\ :\ x \perp y\} \quad = \quad \{(x,y)\ :\ \langle x,y\rangle = 0\},$$

and so the conclusions of 4.12 are applicable. The mapping $S \mapsto S^{\perp\perp}$ is then a Moore closure on X. (That closure is characterized further in 22.52.)

22.51. Theorem on closest points. Let C be a nonempty closed convex subset of a Hilbert space X. Then for each $u \in X$, there is a among the members of C a unique point $\pi(u)$ that is closest to u. It can be characterized as follows: It is the only point $\xi \in C$ that satisfies

$$\mathrm{Re}\langle u - \xi, x - \xi\rangle \ \le\ 0 \qquad \text{for all } x \in C.$$

(In terms of Euclidean geometry, this inequality says that the directed line segment from ξ to u and the directed line segment from ξ to x are separated by an **obtuse angle** — i.e., an angle greater than a right angle.) The mapping $\pi : X \to C$ is also nonexpansive — i.e., it satisfies $\langle\pi\rangle_{\text{Lip}} \le 1$.

If C is a closed linear subspace of X, then $\pi(u)$ can also be characterized as follows: It is the unique point $\xi \in C$ that satisfies $u - \xi \in C^{\perp}$.

Proof. Let C be closed, convex, and nonempty, and let $u \in X$. It follows from 22.45 that there is a closest point and that it is unique.

Now let ξ be a point in C. Then

$$\begin{array}{ll}
 & \xi \text{ is the point in } C \text{ that is closest to } u \\
\iff & \|\xi - u\| < \|x - u\| \text{ for all } x \in C \setminus \{\xi\} \\
\iff & \|\xi - u\| < \|\lambda x + (1-\lambda)\xi - u\| \text{ for all } x \in C \setminus \{\xi\} \text{ and } \lambda \in (0,1] \\
\iff & \|\xi - u\|^2 < \|(\xi - u) + \lambda(x - \xi)\|^2 \text{ for all } x \in C \setminus \{\xi\} \text{ and } \lambda \in (0,1] \\
\iff & \|\xi - u\|^2 < \|\xi - u\|^2 + 2\lambda \mathrm{Re}\langle \xi - u, x - \xi\rangle + \lambda^2 \|x - \xi\|^2 \\
 & \qquad\qquad \text{for all } x \in C \setminus \{\xi\} \text{ and } \lambda \in (0,1] \\
\iff & 0 < 2\lambda \mathrm{Re}\langle \xi - u, x - \xi\rangle + \lambda^2 \|x - \xi\|^2 \\
 & \qquad\qquad \text{for all } x \in C \setminus \{\xi\} \text{ and } \lambda \in (0,1] \\
\iff & 0 < 2\mathrm{Re}\langle \xi - u, x - \xi\rangle + \lambda \|x - \xi\|^2 \text{ for all } x \in C \setminus \{\xi\} \text{ and } \lambda \in (0,1] \\
\iff & 0 \le \mathrm{Re}\langle \xi - u, x - \xi\rangle \text{ for all } x \in C \setminus \{\xi\} \\
\iff & 0 \le \mathrm{Re}\langle \xi - u, x - \xi\rangle \text{ for all } x \in C.
\end{array}$$

This proves the first characterization.

Thus $0 \le \mathrm{Re}\langle \pi(u) - u, x - \pi(u)\rangle$ for all $u \in X$ and $x \in C$. Apply that result with $x = \pi(v)$ to obtain $0 \le \mathrm{Re}\langle \pi(u) - u, \pi(v) - \pi(u)\rangle$ for any $u, v \in X$. Reversing the roles of u and v yields $0 \le \mathrm{Re}\langle \pi(v) - v, \pi(u) - \pi(v)\rangle$. Combine that inequality with (1) and rearrange the results to obtain

$$\begin{aligned}
\|\pi(v) - \pi(u)\|^2 &= \mathrm{Re}\langle \pi(v) - \pi(u), \pi(v) - \pi(u)\rangle \le \mathrm{Re}\langle v - u, \pi(v) - \pi_{\mathbf{C}}(u)\rangle \\
&\le |\langle v - u, \pi(v) - \pi(u)\rangle| \le \|v - u\| \, \|\pi(v) - \pi(u)\|
\end{aligned}$$

and therefore $\|\pi(v) - \pi(u)\| \le \|v - u\|$. Thus π is nonexpansive.

Now suppose C is a linear subspace of X, and $\xi \in C$. Then as x varies over all members of C, $x - \xi$ also varies over all members of C. Hence

$$\begin{array}{ll}
 & \xi \text{ is the point in } C \text{ that is closest to } u \\
\iff & 0 \le \mathrm{Re}\langle \xi - u, x - \xi\rangle \text{ for all } x \in C \\
\iff & 0 \le \mathrm{Re}\langle \xi - u, y\rangle \text{ for all } y \in C \\
\iff & 0 \le \mathrm{Re}\langle \xi - u, ry\rangle \text{ for all } y \in C \text{ and all scalars } c \\
\iff & 0 = \langle \xi - u, y\rangle \text{ for all } y \in C \\
\iff & \xi - u \in C^{\perp}.
\end{array}$$

22.52. Theorem on orthogonal complements. Let X be a Hilbert space and $S \subseteq X$. Then $S^{\perp\perp}$ is the closed linear span of S. Thus S is an orthogonal complement if and only if S is a closed linear subspace of X.

Furthermore, if $S, T \subseteq X$ with $S^{\perp} = T$ and $T^{\perp} = S$, then S and T form an internal direct sum decomposition of X; that is, $S + T = X$ and $S \cap T = \{0\}$. The projections of X onto S and T are the **closest point** mappings; i.e., for any $x \in X$ the unique decomposition

$$x = s + t \qquad \text{with } s \in S,\ t \in T$$

is given by s and t being the points in S and T that are closest to x. These are continuous linear maps.

Remark. Compare this theorem with 11.61.

Proof of theorem. It is easy to see that any orthogonal complement is a closed linear subspace of X. Let $\mathrm{clsp}(S)$ denote the closed linear span of S; then $S^{\perp\perp} \subseteq \mathrm{clsp}(S)$. We wish to show equality here. Suppose that $x \in \mathrm{clsp}(S) \setminus S^{\perp\perp}$. Since $x \notin S^{\perp\perp}$, there is some $y \in S^{\perp}$ such that $\langle x, y\rangle \neq 0$. Since $y \in S^{\perp}$, we have

$$\langle s, y\rangle = 0$$

for every $s \in S$, hence (by the linearity of $\langle\cdot, y\rangle$) also for every $s \in \mathrm{span}(S)$, hence (by the continuity of $\langle\cdot, y\rangle$) also for every $s \in \mathrm{cl}(\mathrm{span}(S)) = \mathrm{clsp}(S)$. But this contradicts $\langle x, y\rangle \neq 0$. Thus, we must have $S^{\perp\perp} = \mathrm{clsp}(S)$.

Now suppose that $S^{\perp} = T$ and $T^{\perp} = S$. Let s be the point in S that is closest to x. By 22.51, $x - s$ is a member of $S^{\perp} = T$. This shows that x can be represented as the sum of an element of S and an element of T. Since S and T are linear subspaces of X and $S \cap T = \{0\}$, the representation is necessarily unique (see 8.13). Thus, in such a representation, the S component must be the member of S closest to x. By symmetric reasoning, the T component must be the member of T closest to x.

22.53. *Remarks.* The preceding theorem has a converse: If X is a normed space in which every closed linear subspace has an additive complement that is also a closed linear subspace, then X is isomorphic to a Hilbert space. This was proved in Lindenstrauss and Tzafriri [1971]; the proof is too long to give here.

22.54. *Definitions.* Let X be a Hilbert space. An **orthonormal set** in X is a set $S \subseteq X$ with the property that $\langle s, t\rangle = \delta_{st}$, where δ is the Kronecker delta — i.e.,

$$\langle s, t\rangle = \begin{cases} 0 & \text{if } s \neq t \\ 1 & \text{if } s = t. \end{cases}$$

Some easy observations. Suppose $\{e_1, e_2, \ldots, e_n\}$ is an orthonormal set. Then:

a. $\|e_j\| = 1$ for each j.

b. If $x = r_1e_1 + r_2e_2 + \cdots + r_ne_n$ and $y = s_1e_1 + s_2e_2 + \cdots + s_ne_n$ for some scalars r_j and s_j, then $\langle x, y\rangle = r_1\overline{s_1} + r_2\overline{s_2} + \cdots + r_n\overline{s_n}$ and $\|x\|^2 = |r_1|^2 + |r_2|^2 + \cdots + |r_n|^2$.

c. The e_j's are linearly independent — i.e., if $r_1e_1 + r_2e_2 + \cdots r_ne_n = 0$ then $r_1 = r_2 = \cdots = r_n = 0$.

d. If $x = r_1e_1 + r_2e_2 + \cdots + r_ne_n$ and $u \in X$, then

$$\|u - x\|^2 \quad = \quad \|u\|^2 + \sum_{j=1}^{n} |c_j - \langle u, e_j\rangle|^2 - \sum_{j=1}^{n} |\langle u, e_j\rangle|^2 .$$

Hence the member of $\operatorname{span}\{e_1, e_2, \ldots, e_n\}$ that is closest to u is the vector $x = r_1e_1 + r_2e_2 + \cdots + r_ne_n$ obtained by taking $r_j = \langle u, e_j\rangle$ for all j. Its distance from u is

$$\operatorname{dist}\Big(u, \operatorname{span}\{e_1, e_2, \ldots, e_n\}\Big) \quad = \quad \sqrt{\|u\|^2 - \sum_{j=1}^{n} |\langle u, e_j\rangle|^2}.$$

e. $\|u\|^2 \geq \sum_{j=1}^{n} |\langle u, e_j\rangle|^2$ for any $u \in X$.

22.55. Theorem. Let X be a Hilbert space, and let $\{e_j : j \in \mathbb{J}\}$ be an orthonormal subset of X. Then the following conditions are equivalent. If one (hence all three) of them are satisfied, we say $\{e_j : j \in \mathbb{J}\}$ is an **orthonormal basis** for X.

(A) $\{e_j : j \in \mathbb{J}\}$ is a maximal orthonormal set — i.e., an orthonormal set that is not contained in any other orthonormal set.

(B) The span of $\{e_j : j \in \mathbb{J}\}$ is dense in X.

(C) **Parseval's Identity.** $\|u\|^2 = \sum_{j\in\mathbb{J}} |\langle u, e_j\rangle|^2$ for every $u \in X$.

Remark. By Zorn's Lemma, any orthonormal set can be extended to a maximal orthonormal set. However, some Hilbert spaces have natural orthonormal bases that can be constructed without the Axiom of Choice. For instance, the space ℓ_2 has orthonormal basis consisting of the vectors $\xi_1, \xi_2, \xi_3, \ldots$, where $\xi_j = (0, 0, \ldots, 0, 1, 0, \ldots)$ has a 1 in the jth place and 0s elsewhere.

Proof of (A) $\Rightarrow$ (B). Suppose the span of $\{e_j\}$ is not dense. Then the closed span of $\{e_j\}$ — which we shall denote by Y — is not equal to X. Let $u \in X \setminus Y$. Let y be the point in Y that is closest to u. Then $z = y - u$ is nonzero, and z is orthogonal to all of Y — hence to all of $\{e_j\}$. Let $\xi = z/\|z\|$. Then $\{e_j : j \in \mathbb{J}\} \cup \{\xi\}$ is an orthonormal set; thus $\{e_j : j \in \mathbb{J}\}$ is not maximal.

Proof of (B) $\Rightarrow$ (C). Let any $u \in X$ and $\varepsilon > 0$ be given. Since the span of the e_j's is dense in X, there is some finite set $J_0 \subseteq \mathbb{J}$ such that some vector x in the span of $\{e_j : j \in J_0\}$ satisfies $\|x - u\| < \varepsilon$. Thus, by 22.54.d we obtain

$$0 \;\leq\; \sqrt{\|u\|^2 - \sum_{j\in J_0} |\langle u, e_j\rangle|^2} \;=\; \operatorname{dist}\Big(u, \operatorname{span}\{e_j : j \in J_0\}\Big) \;\leq\; \|x - u\| \;<\; \varepsilon.$$

That is, $\|u\|^2 - \varepsilon^2 \leq \sum_{j\in J_0} |\langle u, e_j\rangle|^2 \leq \|u\|^2$. Now let $\varepsilon \downarrow 0$ as J_0 increases.

Proof of (C) $\Rightarrow$ (A). Suppose $\{e_j : j \in \mathbb{J}\}$ satisfies (B), but is not maximal. Then there is some $\xi \in X$ such that $\{e_j : j \in \mathbb{J}\} \cup \{\xi\}$ is orthonormal. Then $|\xi| = 1$ and $\langle \xi, e_j \rangle = 0$ for all j, but from (C) we obtain $|\xi|^2 = \sum_{j \in \mathbb{J}} |\langle \xi, e_j \rangle|^2$, a contradiction.

22.56. Theorem. Every Hilbert space is isomorphic to some $\ell_2(\mathbb{J})$. More specifically:

Let $\{e_j : j \in \mathbb{J}\}$ be an orthonormal basis of a Hilbert space X, with scalar field $\mathbb{F}$. For each $x \in X$, define a mapping $\varphi_x : \mathbb{J} \to \mathbb{F}$ by $\varphi_x(j) = \langle x, e_j \rangle$. Then the mapping φ_x is a member of $\ell_2(\mathbb{J})$ (defined in 22.25). Furthermore, the mapping $\Phi : X \to \ell_2(\mathbb{J})$ given by $x \mapsto \varphi_x$ is an isomorphism — i.e., it is a bijection that preserves all relevant structures. It is linear and norm-preserving, and it even preserves the inner product:

$$\|x\|^2 = \sum_{j \in \mathbb{J}} |\varphi_x(j)|^2, \qquad \langle x, y \rangle = \sum_{j \in \mathbb{J}} \varphi_x(j)\overline{\varphi_y(j)}.$$

It maps members of the orthonormal basis of X to corresponding members of the usual orthonormal basis of $\ell_2(\mathbb{J})$ — that is, it maps e_j to the function $\xi_j : \mathbb{J} \to \mathbb{F}$, which is the characteristic function of the singleton $\{j\}$.

Hints: The map Φ is norm-preserving by Parseval's Identity (22.55(C)). It is obviously linear. It maps the span of the e_j's to the simple functions — i.e., the functions $f : \mathbb{J} \to \mathbb{F}$, which vanish outside a finite set.

Chapter 23

Normed Operators

Norms of Operators

23.1. Let $(X, \| \ \|_X)$ and $(Y, \| \ \|_Y)$ be normed vector spaces. Let $f : X \to Y$ be a linear map. Show that the following conditions are equivalent:

(A) f is continuous.

(B) f is uniformly continuous.

(C) f is Lipschitzian; i.e., the Lipschitz constant

$$|||f||| \quad = \quad \sup\left\{\frac{\|f(x) - f(x')\|_Y}{\|x - x'\|_X} \quad : \quad x, x' \in X,\ x \neq x'\right\}$$

is finite.

(D) f is a **bounded linear operator** — i.e., whenever $S \subseteq X$ is a bounded set, then $f(S) \subseteq Y$ is also a bounded set. (A generalization of this terminology will be given in 27.4.)

(E) The number $|||f||| = \sup\{\|f(x)\|_Y / \|x\|_X \ : \ x \in X \setminus \{0\}\}$ is finite.

(F) The number $|||f||| = \sup\{\|f(x)\|_Y \ : \ x \in X,\ \|x\|_X = 1\}$ is finite.

(G) The number $|||f||| = \sup\{\|f(x)\|_Y \ : \ x \in X,\ \|x\|_X \leq 1\}$ is finite.

(H) If (x_n) is a sequence in X with $\|x_n\|_X \to 0$, then $\|f(x_n)\|_Y \to 0$.

(I) If (u_n) is a sequence in X with $\|u_n\|_X \to 0$, then $\sup_n \|f(u_n)\|_Y < \infty$.

Moreover, if these conditions are satisfied, then all the numbers $|||f|||$ defined above are equal to each other.

Hint for 23.1(I) $\Rightarrow$ 23.1(H): Let $u_n = x_n / \sqrt{\|x_n\|_X}$.

Further notations. The set of all bounded linear operators from X into Y is a linear subspace of $Y^X = \{\text{maps from } X \text{ into } Y\}$, which we shall often denote by $BL(X, Y)$. It is a normed space, with $|||f|||$ (defined as above) for the norm of f. A norm obtained in this fashion is the **operator norm** determined by $\| \quad \|_X$ and $\| \quad \|_Y$. A bounded linear operator will generally be given this norm, unless some other norm is specified. In most of the literature the operator norm is denoted by $\| \quad \|$, but in this textbook we shall frequently denote it by $||| \quad |||$ to aid the beginner in distinguishing this norm from the "lower-level" norms of X and Y.

23.2. *Exercises and examples.*

a. If $f : X \to Y$ and $g : Y \to Z$ are bounded linear maps, then the composition $g \circ f : X \to Z$ is also a bounded linear map, with $|||g \circ f||| \leq |||f|||\ |||g|||$. Also, the identity map $i_X : X \to X$ has operator norm equal to 1. Thus, we could take the bounded linear maps as the morphisms of a category, with normed linear spaces for the objects.

b. Let Δ be a set, and let $B(\Delta) = \{$bounded functions from Δ into $\mathbb{R}\}$; then $B(\Delta)$ is a Banach space when equipped with the sup norm. Let $T : B(\Delta) \to \mathbb{R}$ be a positive linear map — i.e., assume $f \geq 0 \;\Rightarrow\; T(f) \geq 0$. Then T is also a bounded linear map; in fact, $|||T||| \leq |T(1)|$. In particular, any Banach limit (defined as in 12.33) is a bounded linear operator.

Similarly, if Δ is a topological space, then $BC(\Delta) = \{$bounded continuous functions from Δ into $\mathbb{R}\}$ is a Banach space when equipped with the sup norm, and any positive linear map from $BC(\Delta)$ into $\mathbb{R}$ is a bounded linear map.

c. The definitions of the vector space $BL(X, Y)$ and its operator norm $||| \quad |||$ depend on the norms $\| \quad \|_X$ and $\| \quad \|_Y$ of the spaces X and Y. Show that if $\| \quad \|_X$ and $\| \quad \|_Y$ are replaced with equivalent norms $\| \quad \|'_X$ and $\| \quad \|'_Y$, then the vector space $BL(X, Y)$ remains the same, and its norm $||| \quad |||$ is replaced with an equivalent norm $||| \quad |||'$. See also the related result in 23.29(iv).

d. If Y is complete, then the normed space $BL(X, Y)$ is complete — regardless of whether X is complete.

In particular, $BL(X, \mathbb{F})$ is complete, since the only scalar fields $\mathbb{F}$ that we are considering for normed spaces in this book are $\mathbb{R}$ and $\mathbb{C}$, both of which are complete.

e. Elementary Extension Theorem. Let X_0 be a dense linear subspace of a normed space X; let X_0 be normed with the restriction of the norm of X. Let Y be a Banach space. If $f_0 : X_0 \to Y$ is a continuous linear map, then f_0 extends uniquely to a continuous linear map $f : X \to Y$. Furthermore, f_0 and f have the same operator norm.

Proof. This is a special case of 19.27. (However, some readers may prefer to prove it directly.)

23.3. Example: matrix norms. Let T be an m-by-n matrix, with scalar t_{ij} in row i, column j. Consider elements $x \in \mathbb{F}^n$ as n-by-1 column vectors and elements $y \in \mathbb{F}^m$ as m-by-1 column vectors. Then T acts as a linear map from $\mathbb{F}^n$ into $\mathbb{F}^m$, with $y = Tx$ given as usual by $y_i = \sum_{j=1}^{n} t_{ij} x_j$ $(1 \leq i \leq m)$. The choice of the norms on $\mathbb{F}^m$ and $\mathbb{F}^n$ will affect the value of the operator norm $|||T|||$. For most choices, the value of $|||T|||$ is complicated and difficult to compute. But for the two following choices, the value of $|||T|||$ is fairly simple.

a. Let $\mathbb{F}^m$ and $\mathbb{F}^n$ both be normed by their respective $\| \quad \|_1$-norms, as defined in 22.11.

Then

$$|||T||| \quad = \quad \max_{1 \le j \le n} \sum_{i=1}^{m} |t_{ij}|.$$

Hints: Let $S_j = \sum_i |t_{ij}|$. Choose k so that $S_k = \max_j S_j$. To show that $|||T||| \ge S_k$, consider $\|Tx\|/\|x\|$ where $x_j = \delta_{jk}$.

b. Let $\mathbb{F}^m$ and $\mathbb{F}^n$ both be normed by their respective $\|\ \ \|_\infty$-norms, as defined in 22.11. Then

$$|||T||| \quad = \quad \max_{1 \le i \le m} \sum_{j=1}^{n} |t_{ij}|.$$

Hints: Let $S_i = \sum_j |t_{ij}|$. Choose k so that $S_k = \max_i S_i$. To show that $|||T||| \ge S_k$, consider $\|Tx\|/\|x\|$ where $x_j = |t_{jk}|/t_{jk}$ if $t_{jk} \ne 0$.

c. (The following observations use results developed later in this chapter, so beginners may wish to postpone reading this paragraph.) The similarity between the two results above is not just coincidental; either of those formulas can be obtained from the other as follows. The normed spaces $(\mathbb{F}^m, \|\ \ \|_1)$ and $(\mathbb{F}^m, \|\ \ \|_\infty)$ are each other's duals, as we shall see in 23.10. For any mapping $T : X \to Y$, the dual map $T^* : Y^* \to X^*$ satisfies $|||T^*||| = |||T|||$, by 23.20. In the case of an operator given by a matrix T, the dual operator T^* is given by the transpose matrix.

23.4. Example: the norm of an integral transform. (This example requires some familiarity with advanced calculus.) Let $\mathbb{F}$ be the scalar field. Let $[\alpha, \beta]$ and $[\gamma, \delta]$ be two closed bounded intervals in $\mathbb{R}$. Let $C[\alpha, \beta]$ and $C[\gamma, \delta]$ be the linear spaces of all continuous functions from $[\alpha, \beta]$ into $\mathbb{F}$, respectively from $[\gamma, \delta]$ into $\mathbb{F}$.

Let k be a continuous function from $[\alpha, \beta] \times [\gamma, \delta]$ into $\mathbb{F}$. For each $f \in C[\gamma, \delta]$, let

$$(Tf)(s) \quad = \quad \int_\gamma^\delta k(s,t) f(t)\, dt \qquad (\alpha \le s \le \beta).$$

Using uniform continuity arguments, show that $(Tf)(\cdot) : [\alpha, \beta] \to \mathbb{F}$ is a continuous function; hence T is a linear map from $C[\gamma, \delta]$ into $C[\alpha, \beta]$.

The choice of the norms on $C[\alpha, \beta]$ and $C[\gamma, \delta]$ will affect the value of the operator norm $|||T|||$. For the two choices given below, the value of $|||T|||$ is fairly simple to compute. (*Hint*: Any continuous function can be approximated uniformly by step functions.)

a. When $C[\alpha, \beta]$ and $C[\gamma, \delta]$ are normed as subspaces of $L^1[\alpha, \beta]$ and $L^1[\gamma, \delta]$, show that T is a bounded linear map from $C[\gamma, \delta]$ into $C[\alpha, \beta]$ with operator norm

$$|||T||| \quad = \quad \max_{\gamma \le t \le \delta} \int_\alpha^\beta |k(s,t)|\, ds.$$

b. When $C[\alpha, \beta]$ and $C[\gamma, \delta]$ are normed as subspaces of $L^\infty[\alpha, \beta]$ and $L^\infty[\gamma, \delta]$, show that T is a bounded linear map from $C[\gamma, \delta]$ into $C[\alpha, \beta]$ with operator norm

$$|||T||| \quad = \quad \max_{\alpha \le s \le \beta} \int_\gamma^\delta |k(s,t)|\, ds.$$

23.5. Example: quotient maps. Let $(X, \| \|_X)$ be a normed space, and let V be a closed linear subspace. Let $Q = X/V$ be the quotient vector space, and let $\pi : X \to Q$ be the quotient map. Then:

a. $\|q\|_Q = \inf\{\|x\|_X : \pi(x) = q\}$ defines a norm on Q; it is called the **quotient norm**. The topology it determines is the same as the quotient topology (defined in 15.30). (This is a special case of a construction given in 22.13.e.)

b. $\|\pi(x)\|_Q = \inf\{\|x + v\|_X : v \in V\} = \inf\{\|x - v\|_X : v \in V\} = \mathrm{dist}(x, V)$.

c. The quotient map $\pi : X \to Q$ has operator norm equal to 1.

23.6. Existence of unbounded linear maps. Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$).

a. *Explicit example with incomplete domain.* Let $C[0,1]$ = {continuous functions from $[0,1]$ $\mathbb{F}$}; this is a Banach space when equipped with the sup norm. Let $C^1[0,1] = \{f \in C[0,1] : f$ has a continuous derivative$\}$; this is a dense linear subspace of $C[0,1]$. Equipped with the sup norm, $C^1[0,1]$ is a normed vector space but not a Banach space. Define $T : C^1[0,1] \to \mathbb{F}$ by $T(f) = f'(0)$. Show that T is discontinuous.

b. *Nonconstructive example with arbitrary domain.* Let X be any infinite-dimensional normed vector space, with scalar field $\mathbb{F}$. Then there exists an unbounded linear functional T on X — that is, a linear map from X into $\mathbb{F}$ that is not bounded.

Hints: Let $\{x_\alpha : \alpha \in A\}$ be a vector basis for X. By assumption, A is infinite; hence we may assume $\mathbb{N} \subseteq A$. Define $T(x_n) = n\|x_n\|$ for each $n \in \mathbb{N}$; define T arbitrarily on the rest of the basis; then use 11.30.b.

c. *Remarks.* There is no explicitly constructible example with complete domain; that will follow from 27.45(ii).

23.7. Let $\mathbb{F}$ be the scalar field. As a special case of the normed space $BL(X, Y)$ introduced in 23.1, we now consider the space

$$X^* \quad = \quad BL(X, \mathbb{F}) \quad = \quad \{\text{bounded linear maps from } X \text{ into } \mathbb{F}\ \}.$$

It has norm

$$\|f\|_{X^*} \quad = \quad \sup\{|f(x)| \ : \ x \in X,\ \|x\|_X = 1\}.$$

(We emphasize that this supremum is not necessarily a maximum; contrast that with 28.41(G).) Our notation X^*, used in the remainder of this chapter, reflects the ideas of 9.55; the set X^* will be called the **dual** of X.

Caution: We remind the reader that the symbol X^* and the term "dual" have different meanings in different branches of mathematics; a few of the meanings are indicated by the list in 9.55. Also, we remark that X' is another notation often used for the set of all bounded linear maps from X into $\mathbb{F}$. In fact, the notation X' is probably used a little more widely in the literature than our own notation X^*. We prefer the notation X^* because (i) it ties in neatly with the other notions of "dual" discussed in 9.55, and (ii) the mark $'$ on a blackboard can be mistaken for a smudge too easily.

Preview of examples. In 23.10 we shall prove $(c_0)^* = \ell_1$, $(\ell_1)^* = \ell_\infty$, and $(\ell_\infty)^* \supsetneqq \ell_1$. In 29.30 we prove that $(\mathcal{L}^\infty(\mathcal{S}))^* = ba(\mathcal{S}, \mathbb{F})$ and $(\mathcal{L}^\infty(\mu))^* = ba(\mu)$. In 28.50 and 28.51 we prove that $(\mathcal{L}^p(\mu))^* = \mathcal{L}^q(\mu)$ if $1 \le p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ (though for $p = 1$ we must assume μ is σ-finite). In Chapter 28 we generalize this notion of "dual" to topological vector spaces that are not necessarily normable.

23.8. *Observation.* Let X be any normed space. By 23.2.d, we know that X^* is complete (i.e., a Banach space), regardless of whether X is complete.

23.9. In many cases we can prove that a dual space X^* is isomorphic to some simpler, more familiar normed space Y. In this context "isomorphic" means that we must preserve the linear structure and the norm. Thus, to each $y \in Y$ there is associated some bounded linear map $T_y : X \to \mathbb{F}$, satisfying $|||T_y|||X^* = \|y\|_Y$, and the mapping $y \mapsto T_y$ must be a linear map from Y onto X^*. Examples of such a duality are given below.

Hints on how to prove such a representation: In most proofs of such a representation, it is trivial to show that the mapping $y \mapsto T_y$ is linear.

The inequality $|||T_y|||_{X^*} \le \|y\|_Y$ means that $|T_y(x)| \le \|x\|_X \, \|y\|_Y$ for all $x \in X$ and $y \in Y$. Generally the proof of this inequality is straightforward — e.g., it may follow immediately from some result such as Hölder's inequality.

The inequality $|||T_y|||_{X^*} \ge \|y\|_Y$ means that $\sup_{\|x\|_X=1} |T_y(x)| \ge \|y\|_Y$. This may be harder to verify, because, as we noted in 23.7, this supremum is not necessarily a maximum. (Conditions for it to be a maximum are considered in 28.41(G).) Instead, for each $y \in Y$ we must show that there exist x_n's in X satisfying $\|x_n\|_X = 1$ and $\limsup_{n\to\infty} |T_y(x_n)| \ge \|y\|_Y$. Finding these x_n's may take some effort; their choice depends on the choice of y.

Finally, showing that the mapping $y \mapsto T_y$ is surjective may be a nontrivial matter. Following is one technique that works in several contexts: Let X_0 be a dense subset of X consisting of particularly nice elements (e.g., the polynomials are dense in certain spaces of continuous functions; the finitely valued functions are dense in certain spaces of measurable functions). Let f be any given element of X^*. Study how f acts on each member of X_0; use that information to find a corresponding $y \in Y$ such that $f = T_y$ on X_0. Since f and T_y are continuous maps agreeing on a dense set, they must agree everywhere on X.

23.10. *Exercises/examples.* For sequences of scalars x and y, define

$$T_y(x) \quad = \quad \sum_{j=1}^{\infty} x_j y_j$$

when this series converges. With notation as in 22.15, 22.25, and 23.9, show that

$$(c_0)^* \; = \; \ell_1, \qquad\qquad (\ell_1)^* \; = \; \ell_\infty, \qquad\qquad (\ell_\infty)^* \supsetneqq \ell_1.$$

(On the other hand, show that the finite-dimensional normed spaces $(\mathbb{F}^m, \|\ \ \|_1)$ and $(\mathbb{F}^m, \|\ \ \|_\infty)$ are each other's duals.)

Hint and remarks. The only tricky part of the proof is to show that $(\ell_\infty)^* \ne \ell_1$. If the scalar field is $\mathbb{R}$, then any sequential Banach limit is a member of $(\ell_\infty)^* \setminus \ell_1$. (Sequential Banach limits were defined in 12.33; their existence was proved in 12.31.) For complex scalars, the proof of $(\ell_\infty)^* \ne \ell_1$ then follows from 11.12.

We remark that there are other members of $(\ell_\infty)^* \setminus \ell_1$ besides the sequential Banach limits; a complete characterization of $(\ell_\infty)^*$ is a special case of results in 29.30. We also remark that this proof of $(\ell_\infty)^* \supsetneqq \ell_1$, or any other proof, must be nonconstructive — it cannot produce a particular example of some $f \in (\ell_\infty)^* \setminus \ell_1$ — that is, the proof does not give an algorithm that takes a constructive description of a sequence $x \in \ell_\infty$ and produces a constructive description of the corresponding scalar $f(x)$. In fact, we *cannot* give such an explicit algorithm; the members of $(\ell_\infty)^* \setminus \ell_1$ form an intangible class. The unavailability of explicit examples follows from 14.77 and 29.38.

For the finite-dimensional case, refer to 11.22.

Equicontinuity and Joint Continuity

23.11. *Remark.* Several of the results below involve "an Abelian group equipped with a gauge consisting of G-seminorms (defined as in 2.13, 22.2, and 22.4) and with the resulting topology (defined as in 5.15.h) and uniformity (defined as in 5.32)." That complicated-seeming object will simplify. It is nothing more than "a topological Abelian group," as we shall see in 26.14, 26.29, and 26.37.

23.12. Additivity and uniform continuity. We now state two analogous theorems side by side. Actually, the result in the left column is a special case of the result in the right column, with Φ consisting of a singleton; but it is a special case important enough to deserve separate mention.

Let X and Z be Abelian groups, each equipped with the topology and uniform structure determined by a gauge consisting of G-seminorms. Let $x_0 \in X$.

Let $f : X \to Z$ be an additive map. Then the following are equivalent:	Let Φ be a collection of additive maps from X into Z. Then the following are equivalent:
(A) f is continuous at x_0. (B) f is continuous at 0. (C) f is continuous. (D) f is uniformly continuous.	(A) Φ is equicontinuous at x_0. (B) Φ is equicontinuous at 0. (C) Φ is equicontinuous. (D) Φ is uniformly equicontinuous.
If X and Z are normed vector spaces, f is linear, and $\| \;\; \|$ is the operator norm defined as in 23.1, then (A)–(D) are also equivalent to:	If X and Z are normed vector spaces, Φ is a collection of linear maps, and $\| \;\; \|$ is the operator norm defined as in 23.1, then (A)–(D) are also equivalent to:
(E) $\|f\| < \infty$.	(E) $\sup_{f\in\Phi} \|f\| < \infty$.

23.13. Baire-Osgood Equicontinuity Theorem for Groups. Let X and Y be groups, with topology and uniform structure given by gauges consisting of G-seminorms. Assume that X is a Baire space. (That last condition is satisfied, for instance, if X is topologized by a single G-seminorm that is complete.)

Let $f_1, f_2, f_3, \ldots$ be continuous additive functions from X into Y. Assume that $f(x) = \lim_{n\to\infty} f_n(x)$ exists for each $x \in X$.

Then $\{f_1, f_2, f_3, \ldots\}$ is equicontinuous, f is continuous, and $f_n \to f$ uniformly on compact subsets of X as $n \to \infty$.

Proof. It suffices to prove equicontinuity; then the other conclusions follow from 18.32.b. Let R be the gauge on Y. In view of 18.30.b, to prove equicontinuity from X to (Y, R) it suffices to prove equicontinuity from X to (Y, ρ) for each G-seminorm ρ in R. Fix any ρ. Since the f_n's and f are additive, it suffices to prove that the sequence (f_n) is equicontinuous at some point $x \in X$ (and we may use different x's for different ρ's). That fact follows from the nonlinear version of the Baire-Osgood Theorem in 20.8.

23.14. Uniform Boundedness Theorem (normed space version). Let X and Y be normed spaces; assume X is complete. Let Φ be a collection of continuous linear maps from X into Y. Then these conditions are equivalent:

(A) Φ is bounded pointwise; that is, $\Phi(x) = \{\varphi(x) : \varphi \in \Phi\}$ is a bounded subset of Y for each $x \in X$.

(B) Φ is *uniformly bounded*, i.e., equicontinuous; that is, $\sup_{\varphi\in\Phi} |||\varphi||| < \infty$.

Proof. Obviously (B) $\Rightarrow$ (A). For (A) $\Rightarrow$ (B), suppose on the contrary that Φ is not equicontinuous. Then we can choose a sequence (φ_k) in Φ with $|||\varphi_k||| > k$. For each k, we can choose some $u_k \in X$ with $\|u_k\|_X = 1$ and $\|\varphi_k(u_k)\|_Y > k$. (*Remark.* These choices do not require the Axiom of Choice, but only the Axiom of Countable Choice, discussed in 6.25.)

We offer two different methods for finishing the proof. The first method is shorter but relies on earlier results that are rather nonelementary: By the Baire Category Theorem (a form of Dependent Choice), the complete metric space X is a Baire space. Since the functions φ_k are bounded pointwise, the functions $k^{-1/2}\varphi_k$ converge pointwise to 0, and therefore are equicontinuous by 23.13. Since the vectors $k^{-1/2}u_k$ converge to 0 in X, it follows that $k^{-1/2}\varphi_k(k^{-1/2}u_k) \to 0$ in Y. But $\|k^{-1/2}\varphi_k(k^{-1/2}u_k)\|_Y = k^{-1}\|\varphi_k(u_k)\|_Y > 1$, a contradiction.

The second proof, though longer, may be preferable to some readers, because it is self-contained and does not rely on the Baire Category Theorem or other deep topological theorems. (In fact, it uses Countable Choice but not Dependent Choice.) It is based on Hennefeld [1980]. Recursively define a sequence (x_n) in X and a sequence (f_n) in Φ, as follows: Let $x_0 = 0$ and choose any $f_0 \in \Phi$. Having chosen $x_0, x_1, \ldots, x_{n-1} \in X$ and $f_0, f_1, \ldots, f_{n-1} \in \Phi$ (clear for $n = 1$), define the numbers

$$\begin{aligned} A_n &= \sum_{j=0}^{n-1} \sup_{f\in\Phi} \|f(x_j)\|_Y \qquad \text{and} \\ B_n &= 2^n \max\{1,\ |||f_0|||,\ |||f_1|||,\ \ldots,\ |||f_{n-1}|||\}\,; \end{aligned}$$

these are both finite by our hypotheses. Now let f_n be some member of Φ that satisfies the inequality $|||f_n||| > (A_n+n)B_n$; if a canonical choice is desired, we may take f_n to be the first member of our sequence (φ_k) that satisfies that inequality. Now multiply the corresponding vector u_k by a suitable scalar, to obtain a vector $x_n \in X$ satisfying $\|x_n\|_X < B_n^{-1}$ and $\|f_n(x_n)\|_Y > A_n + n$. This completes our recursive definition; we obtain sequences (x_n) and (f_n).

Since $\|x_n\|_X < B_n^{-1} \leq 2^{-n}$ and X is complete, the sum

$$x = x_0 + x_1 + x_2 + \cdots$$

is a well-defined member of X. For all integers $j > n$, we have

$$\|f_n(x_j)\|_Y \leq |||f_n||| \, \|x_j\|_X \leq 2^{-j}B_j \cdot \frac{1}{B_j} = 2^{-j};$$

hence $\|f_n(x_{n+1} + x_{n+2} + \cdots)\|_Y \leq 1$. Also, $\|f_n(x_0 + x_1 + x_2 + \cdots + x_{n-1})\|_Y \leq A_n$ by the definition of A_n. Use the fact that

$$x = (x_0 + x_1 + x_2 + \cdots + x_{n-1}) + x_n + (x_{n+1} + x_{n+2} + \cdots).$$

It follows that $\|f_n(x)\|_Y > n - 1$ for every n, which contradicts the assumption that $\sup_{f \in \Phi} \|f(x)\|_Y < \infty$.

23.15. Theorems on joint continuity. Let X, Y, Z be groups, each of which is topologized by a gauge consisting of a collection of G-seminorms. Let $h : X \times Y \to Z$ be a biadditive, separately continuous map — i.e., assume that $z = h(x,y)$ is a continuous, additive function of either of the variables x, y when the other variable is held fixed. Then:

a. h is jointly continuous if and only if h is jointly continuous at 0 — i.e., if and only if whenever $\big((x_\alpha, y_\alpha) : \alpha \in \mathbb{A}\big)$ is a net converging to $(0,0)$ in $X \times Y$, then $h(x_\alpha, y_\alpha) \to 0$ in Z.

Proof. The "only if" part is obvious. For the "if" part, suppose that $(u_\alpha, v_\alpha) \to (u,v)$ in $X \times Y$. Then $(u_\alpha - u, v_\alpha - v) \to (0,0)$ in $X \times Y$, hence

$$h(u_\alpha, v_\alpha) - h(u,v) = h(u_\alpha - u, v_\alpha - v) + h(u, v_\alpha - v) + h(u_\alpha - u, v) \to 0$$

by joint continuity at $(0,0)$ and separate continuity.

b. Suppose that the topologies on X and Y are each given by a single G-seminorm and at least one of X, Y is complete. Then h is jointly continuous.

Proof. Say X is complete. Let $((x_n, y_n))$ be a sequence converging to $(0,0)$ in $X \times Y$; we wish to show $h(x_n, y_n) \to 0$ in Z. Define mappings $f_n : X \to Z$ by $f_n(x) = h(x, y_n)$. Then each f_n is continuous and additive, and $f_n \to 0$ pointwise. By 23.13, the sequence (f_n) is equicontinuous. Since $x_n \to 0$, we have $f_n(x_n) \to 0$ as required.

c. Suppose X, Y, Z are normed spaces, and let $h : X \times Y \to Z$ be a bilinear mapping. Then h is jointly continuous if and only if the number

$$|||h||| = \sup \{\|h(x,y)\|_Z : \|x\|_X, \|y\|_Y \leq 1\}$$

is finite.

Further observations. The jointly continuous, bilinear maps from $X \times Y$ into Z form a normed space, when normed by $|||h|||$ defined in this fashion. It is complete if Z is complete.

THE BOCHNER INTEGRAL

23.16. *Definitions.* Let $(X, |\quad|)$ be a Banach space, let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let $f \in L^1(\mu, X)$. Then the **Bochner integral**, or **Bochner-Lebesgue integral**, of f with respect to μ, denoted by $\int_\Omega f d\mu$, is an element of X defined as follows.

When $f \in L^1(\mu, X)$ is finitely valued, then we can define the integral as in 11.42; that is, $\int_\Omega f d\mu = \sum_x \mu\left(f^{-1}(x)\right) x$. It is easy to verify that

$$\left|\int_\Omega f d\mu\right| \quad \leq \quad \int_\Omega |f(\cdot)| d\mu \quad = \quad \|f\|_1 \, . \tag{$*$}$$

Thus the mapping $f \mapsto \int_\Omega f d\mu$ is a continuous linear map from a dense subspace of $L^1(\mu, X)$ into X. By 23.2.e, it therefore extends uniquely to a continuous linear map from $L^1(\mu, X)$ into X, satisfying $(*)$.

By a "measurable set" we mean a member of $\mathcal{S}$. For any measurable set $S \subseteq \Omega$, we may define the Bochner integral $\int_S f\, d\mu$ by restricting f to the set S and restricting μ to measurable subsets of S. However, $\int_S f\, d\mu$ is also equal to $\int_\Omega 1_S f\, d\mu$, where 1_S is the characteristic function of S. Note also $\left|\int_S f\, d\mu\right| \leq \int_S |f|\, d\mu$.

A few more basic properties of the Bochner integral are given below; some additional properties can be found in 29.10 and in the consequences of that result. The Bochner integral should also be contrasted with the Bartle integral, introduced in 29.30.

Further remarks. Terminology varies. The integral defined above (which we shall call the *Bochner integral* in this book) is often known as the **Lebesgue integral** — particularly in the special cases where μ is Lebesgue measure and/or X is finite-dimensional.

When $\mu(\Omega) = 1$ — i.e., when μ is a probability measure — then $\int_\Omega f\, d\mu$ is also called the **expectation** of f.

Exercise. If $f \in L^1(\mu, \mathbb{R})$ and $f \geq 0$, then the Bochner integral $\int_\Omega f\, d\mu$ is equal to the positive integral $\int_\Omega f\, d\mu$ defined in 21.36.

23.17. Let Ω_1, Ω_2, X be any sets. Then any function $f : \Omega_1 \times \Omega_2 \to X$ can also be viewed as a map $\widehat{f} : \Omega_2 \to X^{\Omega_1}$, whose value at any ω_2 is the mapping $[\widehat{f}(\omega_2)](\cdot) = f(\cdot, \omega_2) : \Omega_1 \to X$. This obviously gives us a bijection between $X^{\Omega_1 \times \Omega_2}$ and $\left(X^{\Omega_1}\right)^{\Omega_2}$.

Fubini's Theorem. Let $(\Omega_1, \mathcal{S}_1, \mu_1)$ and $(\Omega_2, \mathcal{S}_2, \mu_2)$ be σ-finite measure spaces, and let $(\Omega_1 \times \Omega_2, \mathcal{S}_1 \times \mathcal{S}_2, \mu_1 \times \mu_2)$ be the product measure space, defined as in 21.40. Let X be a Banach space, and let $Y = L^1(\mu_1, X)$. Then the mapping $f \mapsto \widehat{f}$ defined above gives an isomorphism (i.e., a linear, norm-preserving bijection) from $L^1(\mu_1 \times \mu_2, X)$ onto $L^1(\mu_2, Y)$.

Furthermore, if $f \in L^1(\mu_1 \times \mu_2, X)$, then the following two iterated Bochner integrals exist and are equal to the noniterated Bochner integral $\int_{\Omega_1 \times \Omega_2} f\, d(\mu_1 \times \mu_2)$.

(i) For almost every ω_2 in Ω_2, the Bochner integral $\int_{\Omega_1} f(\cdot, \omega_2)\, d\mu_1$ exists in X, and the mapping $\omega_2 \mapsto \int_{\Omega_1} f(\cdot, \omega_2)\, d\mu_1$ (from Ω_2 into X) is integrable in X with respect to μ_2. The resulting iterated integral (in X over μ_1 and then in X over μ_2) may be denoted $\int_{\Omega_2} \left[\int_{\Omega_1} f(\cdot, \omega_2)\, d\mu_1\right] d\mu_2(\omega_2)$.

(ii) For almost every ω_2 in Ω_2, the function $[\widehat{f}(\omega_2)](\cdot) : \Omega_1 \to X$ is a member of Y. The mapping $\omega_2 \mapsto \widehat{f}(\omega_2)$ (from Ω_2 into Y) is integrable in Y with respect to μ_2. Integrate in Y to obtain $\varphi = \int_{\Omega_2} \widehat{f}\, d\mu_2$, a member of Y. Then $\varphi \in Y = L^1(\mu_1, X)$ may itself be integrated in X; thus we obtain the iterated integral $\int_{\Omega_1} \varphi\, d\mu_1 = \int_{\Omega_1} \left[\int_{\Omega_2} \widehat{f} d\mu_2\right] d\mu_1$ (in Y over μ_2 and then in X over μ_1).

Proof. The map $f \mapsto \widehat{f}$ is obviously linear. We shall show that it maps a dense subset of $L^1(\mu_1 \times \mu_2, X)$ onto a dense subset of $L^1(\mu_2, Y)$, in a norm-preserving fashion; then the linear map obviously extends to an isomorphism between the two spaces.

By a *basic function* we shall mean a function $f : \Omega_1 \times \Omega_2 \to X$ of the form $s(\cdot) = \sum_{j=1}^n 1_{A_j \times B_j}(\cdot) x_j$, where n is a positive integer, the x_j's are members of X, and $1_{A_j \times B_j}$ is the characteristic function of $A_j \times B_j$, where $A_j \subseteq \Omega_1$ and $B_j \subsetneq \Omega_2$ are measurable sets with finite measure. We shall show that the basic functions s are dense in $L^1(\mu_1 \times \mu_2, X)$, and their images $\widehat{s}$ are dense in $L^1(\mu_2, Y)$.

Let any $f \in L^1(\mu_1 \times \mu_2, X)$ be given. We know that the integrable, finitely valued functions are dense in $L^1(\mu_1 \times \mu_2, X)$. Hence f can be approximated arbitrarily closely by a function of the form $\sum_{k=1}^m 1_{S_k}(\cdot) x_k$, where each S_k is a member of $\mathcal{S}_1 \times \mathcal{S}_2$ with finite measure. By 21.26, each S_k can be approximated arbitrarily closely in measure by a set that is a union of finitely many measurable rectangles — i.e., for any $\varepsilon > 0$, there exist $A_{k,1}, A_{k,2}, \ldots, A_{k,N(k)} \in \mathcal{S}_1$ and $B_{k,1}, B_{k,2}, \ldots, B_{k,N(k)} \in \mathcal{S}_2$ such that the sets

$$A_{k,1} \times B_{k,1}, \qquad A_{k,2} \times B_{k,2}, \qquad \ldots, \qquad A_{k,N(k)} \times B_{k,N(k)}$$

are disjoint subsets of S_k and have union S_k' satisfying $(\mu_1 \times \mu_2)(S_k \setminus S_k') < \varepsilon$. Then f is approximated by the basic function $s(\cdot) = \sum_{k=1}^m \sum_{i=1}^{N(k)} 1_{A_{k,i} \times B_{k,i}}(\cdot) x_k$.

On the other hand, let $f : \Omega_1 \times \Omega_2 \to X$ be a function with $\widehat{f} \in L^1(\mu_2, Y)$. Since the finitely valued, integrable functions are dense in $L^1(\mu_2, Y)$, we can approximate $\widehat{f}$ arbitrarily closely by a function of the form $\sum_{k=1}^m 1_{B_k}(\cdot) y_k$, where B_k is a member of $\mathcal{S}_2$ with finite measure and $y_k \in Y = L^1(\mu_1, X)$. Each $y_k(\cdot)$ can, in turn, be approximated arbitrarily closely by a finitely valued, integrable function $\sum_{i=1}^{N(k)} 1_{A_{k,i}}(\cdot) x_i$. It follows easily that $\widehat{f}$ is approximated arbitrarily closely by functions of the form $\widehat{s}$, where $s(\cdot) = \sum_{k=1}^m \sum_{i=1}^{N(k)} 1_{A_{k,i} \times B_k}(\cdot) x_k$.

Now let any corresponding functions $f \in X^{\Omega_1 \times \Omega_2}$ and $\widehat{f} \in \left(X^{\Omega_1}\right)^{\Omega_2}$ be given. We claim that

> f belongs to $\mathcal{L}^1(\mu_1 \times \mu_2, X)$ if and only if $\widehat{f}$ belongs to $\mathcal{L}^1\left(\mu_2, \mathcal{L}^1(\mu_1, X)\right)$, in which case $\|f\|_{L^1(\mu_1 \times \mu_2, X)} = \left\| \|\widehat{f}(\cdot)\|_Y \right\|_{L^1(\mu_2, \mathbb{R})}$.

This is clear for basic functions; equality of the norms follows easily from Tonelli's Theorem (21.40). Thus we obtain a norm-preserving linear map between dense subsets of the two Banach spaces $L^1(\mu_1 \times \mu_2, X)$ and $L^1(\mu_2, Y)$. Taking limits proves this claim for all f.

It is easy to verify that the three kinds of Bochner integrals agree for any basic function. The basic functions are dense in the integrable functions, and the Bochner integrals are continuous linear maps; therefore the Bochner integrals agree for all integrable functions.

Hahn-Banach Theorems in Normed Spaces

23.18. Following are several principles, any one of which may be referred to as "the Hahn-Banach Theorem;" they are equivalent to each other and to the Hahn-Banach Theorems presented in 12.31, 23.19, 26.56, 28.4, 28.14.a, and 29.32. Most of the principles below refer to the dual space X^* defined in 23.7.

Each of the principles below asserts the existence of some object, but does not specify how to find that object. In general, we may not be *able* to find the object. The existence proof is not constructive, and in fact it cannot be made constructive. The Hahn-Banach Theorem implies the existence of certain known intangibles; see 14.77 and 29.38.

The norm-preserving extension Λ described in (HB7) is not necessarily uniquely determined by λ. In 23.21 we consider some conditions for uniqueness.

Note that (HB10) asserts the equality of an infimum (the distance) and a supremum (the maximum). The principle (HB10) can be found in Luenberger [1969] or Nirenberg [1975].

The charge described in (HB12) is closely related to another charge described in 29.32.

(HB7) Norm-Preserving Extensions. Let $(X, \| \ \ \|)$ be a normed space, and let Y be a linear subspace of X. Let $\lambda \in Y^*$ — that is, let λ be a bounded linear map from Y into the scalar field, where Y is normed by the restriction of $\| \ \ \|$. Then λ can be extended to some $\Lambda \in X^*$ satisfying $\|\Lambda\|_{X^*} = \|\lambda\|_{Y^*}$.

(HB8) Functionals for Given Vectors. Let $(X, \| \ \ \|)$ be a normed vector space other than the degenerate space $\{0\}$, and let $x_0 \in X$. Then there exists some $\Lambda \in X^*$ such that $\|\Lambda\| = 1$ and $\Lambda(x_0) = \|x_0\|$. Hence the norm of a vector in X can be characterized in terms of the values of members of X^*:

$$\|x_0\| \quad = \quad \max\left\{|f(x_0)| \ : \ f \in X^*, \ \|f\|_{X^*} = 1\right\}.$$

(We emphasize that this is a maximum, not just a supremum; contrast that with 28.41(G).) Therefore each $x \in X$ acts as a bounded linear functional $T_x : X^* \to \mathbb{F}$ by the rule $T_x(f) = f(x)$, with norm $\|T_x\|_{X^{**}} = \|x\|_X$.

(HB9) Separation of Points. If X is a normed space, then X^* separates the points of X. That is, if x and y are distinct points of X, then there exists some $\Lambda \in X^*$ such that $\Lambda(x) \neq \Lambda(y)$. Equivalently, if $u \in X \setminus \{0\}$, then there exists some $\Lambda \in X^*$ such that $\Lambda(u) \neq 0$.

(HB10) Variational Principle. Let X be a normed space. Let V be a closed linear subspace, let $V^\perp = \{\lambda \in X^* : \lambda \text{ vanishes on } V\}$, and let $x_0 \in X \setminus V$. Then $V^\perp \setminus \{0\}$ is nonempty, and

$$\operatorname{dist}(x_0, V) \quad = \quad \max\left\{\frac{|\lambda(x_0)|}{\|\lambda\|} \;:\; \lambda \in V^\perp \setminus \{0\}\right\}.$$

(HB11) Separation of Subspaces. Let B be a closed linear subspace of a Banach space X, and let $\eta \in X \setminus B$. Then there exists a member of X^* that vanishes on B but not on η.

(HB12) Luxemburg's Measure. For every nonempty set Ω and every proper filter $\mathcal{F}$ of subsets of Ω, there exists a probability charge μ on $\mathcal{P}(\Omega)$ that takes the value 1 on elements of $\mathcal{F}$.

Proof of (HB2) $\Rightarrow$ (HB7). If the scalar field is $\mathbb{R}$, use $p(x) = \|x\|\,\|\lambda\|_{Y^*}$. If the scalar field is $\mathbb{C}$, use 11.12.

Proof of (HB7) $\Rightarrow$ (HB8). Let Y be the linear subspace spanned by x_0; define $\lambda(rx_0) = r\|x_0\|$ for all scalars r.

Proof of (HB8) $\Rightarrow$ (HB9). $x - y \neq 0$; choose $\Lambda \in X^*$ with $\Lambda(x-y) = \|x-y\|$.

Proof of (HB8) $\Rightarrow$ (HB10). If $v \in V$ and $\lambda \in V^\perp \setminus \{0\}$, then

$$\|v - x_0\| \quad \geq \quad \frac{|\lambda(v - x_0)|}{\|\lambda\|} \quad = \quad \frac{|\lambda(x_0)|}{\|\lambda\|}.$$

Take the infimum on the left over all choices of v, and take the supremum on the right over all choices of λ; this proves that

$$\operatorname{dist}(x_0, V) \quad \geq \quad \sup\left\{\frac{|\lambda(x_0)|}{\|\lambda\|} \;:\; \lambda \in V^\perp \setminus \{0\}\right\}.$$

It now suffices to exhibit some particular $\lambda \in V^\perp \setminus \{0\}$ that satisfies

$$\operatorname{dist}(x_0, V) \quad = \quad \frac{|\lambda(x_0)|}{\|\lambda\|}\,. \tag{!}$$

Let $Q = X/V$ be the quotient space, equipped with quotient norm $\|\ \|_Q$ as in 23.5. As we noted in 23.5, the quotient map $\pi : X \to Q$ has norm 1. As in (HB8), we may choose some functional $\Lambda \in Q^*$ with $\|\Lambda\| = 1$ and $\Lambda(\pi(x_0)) = \|\pi(x_0)\|_Q = \operatorname{dist}(x_0, V)$. Let $\lambda = \Lambda \circ \pi : X \to \{\text{scalars}\}$. The function λ vanishes on V, since π does. Thus λ satisfies (!).

Proof of (HB9) $\Rightarrow$ (HB11). Let $Q = X/B$ be the quotient space, equipped with the quotient topology; let $\pi : X \to Q$ be the quotient map. Refer to results in 23.5. Then $\pi(\eta)$ is different from the 0 element of Q. By (HB9), there is some continuous linear map

$\Lambda : Q \to \{\text{scalars}\}$ that does not vanish on $\pi(\eta)$. Then $\Lambda \circ \pi : X \to \{\text{scalars}\}$ is a continuous linear map that vanishes on B but not on η.

Proof of (HB10) $\Rightarrow$ (HB11). Obvious.

Proof of (HB11) $\Rightarrow$ (HB12). Let $\mathcal{I} = \{\Omega \backslash S : S \in \mathcal{F}\}$; this is a proper ideal. For each $E \in \mathcal{I}$, let $1_E : \Omega \to [0,1]$ be the characteristic function of the set E; also let $1 : \Omega \to [0,1]$ be the constant function 1. Let X be the Banach space of bounded functions from Ω into $\mathbb{R}$, with the sup norm. Let B be the closed span of the functions 1_E, for $E \in \mathcal{I}$.

We claim that $1 \notin B$. Indeed, consider any function g in the span of the 1_E's. It is of the form

$$g \quad = \quad c_1 1_{E_1} + c_2 1_{E_2} + \cdots + c_n 1_{E_n}$$

for some nonnegative integer n, some real numbers c_i, and some E_i's in $\mathcal{I}$. Since $\mathcal{I}$ is a proper ideal, the set $E = E_1 \cup E_2 \cup \cdots \cup E_n$ is a member of $\mathcal{I}$, hence a proper subset of Ω; let ω be some point in $\Omega \setminus E$. Then $g(\omega) = 0$, while $1(\omega) = 1$. This shows that any element of the span of the 1_E's has distance at least 1 from the constant function 1, and therefore the constant function 1 does not belong to B.

Now, using (HB11), we find that there is some bounded linear functional $\Lambda : X \to \mathbb{R}$ that vanishes on B but not on the constant function 1. Take $\nu(S) = \Lambda(1_S)$; then ν is a bounded real-valued charge on $\mathcal{P}(\Omega)$ that vanishes on each $E \in \mathcal{I}$ but not on Ω. As we saw in 11.47, the positive part of this charge is given by

$$\nu^+(S) \quad = \quad \sup\{\nu(A) : A \subseteq S\}.$$

This function ν^+ is a positive charge on $\mathcal{P}(\Omega)$ that vanishes on each $E \in \mathcal{I}$ but not on Ω. Finally, $\mu(S) = \nu^+(S)/\nu^+(\Omega)$ is a probability charge with the required properties.

Proof of (HB12) $\Rightarrow$ (HB1). Let $(\Delta, \preccurlyeq)$ be a directed set, and let $B(\Delta) = \{\text{bounded functions from } \Delta \text{ into } \mathbb{R}\}$. Let $\mathcal{F}$ be the filter of tails of $(\Delta, \preccurlyeq)$, as defined in 7.9 — that is, $\mathcal{F}$ is the filter on Δ consisting of the supersets of sets of the form $\{\delta \in \Delta : \delta \succcurlyeq \delta_0\}$.

By assumption, there exists a probability charge μ on $\mathcal{P}(\Delta)$ that takes the value 1 on elements of $\mathcal{F}$. Define $\mathrm{LIM}(u) = \int_\Delta u(\delta)\, d\mu(\delta)$ in the obvious fashion for simple functions u. Since simple functions are dense, we can extend this definition to $u \in B(\Delta)$ by taking limits. (This construction is a special case of the Bartle integral construction described in 29.30.)

Note that if $F \in \mathcal{F}$, then $\mathrm{LIM}(1_F) = \mu(F) = 1$ (where $1_F : \Omega \to \{0,1\}$ is the characteristic function of F), and

$$\mathrm{LIM}(1_{\Delta \backslash F}) \quad = \quad \mu(\Delta \setminus F) \quad = \quad \mu(\Delta) - \mu(F) \quad = \quad 1 - 1 \quad = \quad 0.$$

If g is a bounded real-valued function on Δ that vanishes on F, then $-\|g\|_\infty 1_{\Delta\backslash F} \leq g \leq \|g\|_\infty 1_{\Delta \backslash F}$; hence

$$0 \quad = \quad \mathrm{LIM}\left(-\|g\|_\infty 1_{\Delta\backslash F}\right) \quad \leq \quad \mathrm{LIM}(g) \quad \leq \quad \mathrm{LIM}\left(\|g\|_\infty 1_{\Delta\backslash F}\right) \quad = \quad 0$$

and thus $\mathrm{LIM}(g) = 0$. If h is any bounded real-valued function on Δ, then $h - h1_F$ vanishes on F, so $\mathrm{LIM}(h) = \mathrm{LIM}(h1_F)$.

We shall show that LIM is a Banach limit with the desired properties. Clearly LIM is a positive linear functional; it suffices to show that $\mathrm{LIM}(u) \leq \limsup_{\delta\in\Delta} u(\delta)$ for each $u \in B(\Delta)$. Fix any number $r > \limsup_{\delta\in\Delta} u(\delta)$; it suffices to show that $r \geq \mathrm{LIM}(u)$. By our choice of r, we have $r > u(\delta)$ for all δ sufficiently large — say for all $\delta \succcurlyeq \delta_0$. Thus $r1_F \geq u1_F$. The set $F = \{\delta \in \Delta : \delta \succcurlyeq \delta_0\}$ belongs to $\mathcal{F}$, hence

$$\mathrm{LIM}(u) \quad = \quad \mathrm{LIM}(u1_F) \quad \leq \quad \mathrm{LIM}(r1_F) \quad = \quad \mathrm{LIM}(r) \quad = \quad r.$$

23.19. *Luxemburg's Boolean equivalents of HB* (*optional*). The principle (HB12) involved algebras of sets. We shall generalize that principle to Boolean algebras. Admittedly, Boolean algebras don't seem to be much more general — indeed, (UF7) in 13.22 tells us that every Boolean algebra is isomorphic to an algebra of sets. However, UF is stronger than HB, so we are not permitted to use UF in the next few paragraphs when we prove that certain principles are equivalent to HB.

First, we must generalize some notions of charges and measures to Boolean algebras.

a. In a Boolean lattice X, we say that two elements x, y are **disjoint** if $x \wedge y = 0$. Note that 0 is disjoint from any element; 0 is even disjoint from itself. A subset of X is **disjoint** (or, for emphasis, **pairwise disjoint**) if each pair of elements of that set is disjoint.

b. A **probability** (or **probability charge**) on X is a function $\mu : X \to [0,1]$ such that $\mu(1) = 1$ and

$$\mu(x \vee y) = \mu(x) + \mu(y) \text{ for disjoint } x, y \in X.$$

Of course, if μ is any probability on a Boolean lattice X, then $\mu(0)$ is equal to 0, since 0 is disjoint from itself. Thus, we have $\{0,1\} \subseteq \mathrm{Range}(\mu) \subseteq [0,1]$. A **two-valued probability** on X is a probability with range equal to $\{0,1\}$. *Exercise.* Show that a two-valued probability is the same thing as a two-valued homomorphism — (defined in 13.8).

We can now generalize (HB12) to Boolean algebras. The principle (HB13) was recently used by Pawlikowski [1991] to prove that the Hahn-Banach Theorem implies the Banach-Tarski Decomposition. It is interesting to compare (UF8) and (HB13), both of which assert the existence of charges. Also, in 29.37 are some even weaker assertions of the existence of charges.

(HB13) On every Boolean algebra there exists a probability charge.

(HB14) Let X be a Boolean algebra. Then for every proper ideal I in X there exists a probability μ on X that vanishes on I.

Proof of (HB12) $\Rightarrow$ (HB13). Let X be a Boolean algebra. By the Tarski-Scott-Luxemburg Lemma 13.12, there exists a surjective homomorphism $f : \mathcal{S} \to X$, where $\mathcal{S}$ is some algebra of subsets of some set Ω. Then $\mathrm{Ker}(f)$ is a proper ideal in $\mathcal{S}$, and so there is a probability μ on $\mathcal{S}$ that takes the value 0 on elements of $\mathrm{Ker}(f)$. Hence μ determines a probability ν on $\mathcal{S}/\mathrm{Ker}(f)$, which is isomorphic to X.

Proof of (HB13) $\Rightarrow$ (HB14). Let $\pi : X \to X/I$ be the quotient map, and let ν be a probability on the Boolean algebra X/I. Define $\mu(x) = \nu(\pi(x))$; verify that μ is a probability on X that vanishes on elements of I.

Proof of (HB14) $\Rightarrow$ (HB12). Obvious.

A Few Consequences of HB

23.20. The dual functor in normed spaces. Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$). Let $\mathcal{C}$ and $\mathcal{C}^*$ both be the category of normed spaces over $\mathbb{F}$, with continuous linear maps for morphisms. Then we may define a dual functor $X \mapsto X^*$ as in 9.55 through 9.59, using the scalar field $\mathbb{F}$ for the object Δ discussed in 9.55. The resulting dual space X^* defined in 9.55 is the same as the normed vector space X^* defined in 23.7. For any morphism $f : X \to Y$, the dual map $f^* : Y^* \to X^*$ is defined by $f^*(\lambda) = \lambda \circ f$.

By the Hahn-Banach Theorem (HB9), X^* separates points of X. Therefore, points in X may be viewed as distinct functions acting on X^*. Moreover, the embedding $X \xrightarrow{\subseteq} X^{**}$ is norm-preserving, as we noted in (HB8) in 23.18.

For any morphism $f : X \to Y$, the bidual function $f^{**} : X^{**} \to Y^{**}$ is an extension of the function $f : X \to Y$. All of these statements and all of hypotheses (H1) through (H5) in 9.55 through 9.57 are now easy to verify.

A Banach space X is *reflexive* if $X^{**} = X$. Some Banach spaces are reflexive, but others are not. For instance, ℓ_p is reflexive for $1 < p < \infty$, but ℓ_1, ℓ_∞, and c_0 are not. Reflexivity of Banach spaces will be investigated further in 28.41(A).

A slightly more subtle result: $\|f\| = \|f^*\| = \|f^{**}\|$ for any continuous linear map $f : X \to Y$ between normed spaces.

Hints: From the definition $f^*(\lambda) = \lambda \circ f$, prove that $\|f^*\| \leq \|f\|$. Similarly, $\|f^{**}\| \leq \|f^*\|$. On the other hand, use the fact that $f^{**} : X^{**} \to Y^{**}$ is an extension of $f : X \to Y$ to show that $\|f\| \leq \|f^{**}\|$. Finally, combine these results: $\|f^*\| \leq \|f\| \leq \|f^{**}\| \leq \|f^*\|$.

23.21. Taylor-Foguel Theorem (optional). Let $(X, \|\ \ \|)$ be a Banach space, and let $(X^*, \|\ \ \|)$ be its dual. Then X^* is strictly convex if and only if every bounded linear functional on a subspace of X has a *unique* norm-preserving linear extension.

Proof. First, suppose there exists some linear subspace $X_0 \subseteq X$ and some $f_0 \in (X_0)^*$ that has distinct extensions $f_1, f_2 \in X^*$ with $\|f_0\| = \|f_1\| = \|f_2\|$. Let $f = (f_1 + f_2)/2$. Then f is also an extension of f_0, so $\|f\| \geq \|f_0\|$. Now f_1, f_2, f are collinear, so X^* is not strictly convex.

Conversely, X^* is not strictly convex; we shall show that X does not have unique norm-preserving extensions. By assumption, there exist distinct $f, g \in X^*$ with $\|f\|_{X^*} = \|g\|_{X^*} = \|\frac{1}{2}(f + g)\|_{X^*} = 1$. Let $M = \{x \in X : f(x) = g(x)\}$, and let φ be the restriction of f or g to the linear subspace M. It suffices to show that $\|\varphi\|_{M^*} = 1$, for then f and g

are distinct norm-preserving extensions. We know that $\|\varphi\|_{M^*} \le 1$ by the definition of the operator norm; thus it suffices to show that $\|\varphi\|_{M^*} \ge 1$. Since $f \ne g$, we may choose some $\xi \in X$ with $f(\xi) - g(\xi) = 1$. Then each $x \in X$ may be expressed in one and only one way in the form $x = y + a\xi$, where $y \in M$ and a is a scalar. Since $\|\frac{1}{2}(f+g)\|_{X^*} = 1$, we may choose a sequence (x_n) in X with $\|x_n\|_X = 1$ and $\frac{1}{2}(f+g)(x_n) \to 1$. Since $\|f\|_{X^*} = \|g\|_{X^*} = 1$, it follows that $f(x_n) \to 1$ and $g(x_n) \to 1$. Write $x_n = y_n + a_n\xi$ with $y_n \in M$ and scalar a_n. Then $a_n = (f-g)(a_n\xi) = (f-g)(x_n - y_n) = (f-g)(x_n) \to 1 - 1 = 0$, hence $\|y_n\|_M = \|y_n\|_X \to 1$. At the same time, $\varphi(y_n) = f(y_n) = f(x_n) - a_n f(\xi) \to 1$. Thus $\|\varphi\|_{M^*} \ge 1$, completing the proof.

23.22. Kottman's Theorem. Let X be an infinite-dimensional normed space. Then X^* is infinite-dimensional. Furthermore, there exists a sequence (x_n) in X such that $\|x_n\| = 1$ for each n and $\|x_m - x_n\| > 1$ whenever $m \ne n$.

Remark. It follows easily from compactness considerations (see 27.17) that such a sequence cannot exist in a finite-dimensional normed space.

Outline of proof. This theorem was first proved by Kottman, but the proof given here is due to T. Starbird and was published by Diestel [1984].

a. Show there exist $x_1 \in X$ and $\lambda_1 \in X^*$ with $\|x_1\| = \|\lambda_1\| = \lambda_1(x_1) = 1$.

b. We now proceed by induction. Assume

$$x_1, x_2, \ldots, x_k \in X \qquad \text{and} \qquad \lambda_1, \lambda_2, \ldots, \lambda_k \in X^*$$

have been chosen, all with norm 1, and with $\lambda_1, \lambda_2, \ldots, \lambda_k$ linearly independent. Show there exists $y \in X$ such that $\lambda_1(y), \lambda_2(y), \ldots, \lambda_k(y) < 0$.

c. Show there exists a nonzero x in $\bigcap_{i=1}^k \operatorname{Ker}(\lambda_i)$; here "Ker" denotes kernel.

d. Show that for any sufficiently large positive number K, we have $\|y\| < \|y + Kx\|$. Fix some such K.

e. Using the linear independence of the λ_i's, show that if $a_1, a_2, \ldots, a_k$ are scalars, not all 0, then $|\sum_{i=1}^k a_i\lambda_i(y + Kx)| < \|\sum_{i=1}^k a_i\lambda_i\| \; \|y + Kx\|$.

f. Let $x_{k+1} = (y + Kx)/\|y + Kx\|$, and then by (HB8) choose some $\lambda_{k+1} \in X^*$ with $\|\lambda_{k+1}\| = \lambda_{k+1}(x_{k+1}) = 1$. Using 23.22.e, show that λ_{k+1} is not a linear combination of $\lambda_1, \lambda_2, \ldots, \lambda_k$, completing the induction.

g. Show that if $1 \le i \le k$, then $\lambda_i(x_{k+1}) < 0$, and hence $\|x_{k+1} - x_i\| \ge |\lambda_i(x_{k+1} - x_i)| > 1$.

Duality and Separability

23.23. If $(X, \| \; \|)$ is a Banach space and X^* is separable, then X is separable.

Proof (following the exposition of M. Schechter [1971]). Let (φ_n) be a dense sequence in X^*. For each n, choose some $v_n \in X$ that satisfies $\|v_n\| = 1$ and $\langle \varphi_n, v_n \rangle \ge \frac{1}{2}\|\varphi_n\|$. Let

V be the closed linear span of the v_n's. Then V is a separable, closed linear subspace of X; it suffices to show $V = X$. Suppose, on the contrary, that $w \in X \setminus V$. By (HB11) in 23.18, there is some $\psi \in X^*$ that vanishes on V but not on w. By rescaling we may assume $\|\psi\| = 1$. For each n, we have

$$\frac{1}{2}\|\varphi_n\| \;\le\; \langle \varphi_n, v_n\rangle \;=\; \langle \varphi_n - \psi, v_n\rangle \;\le\; \|\varphi_n - \psi\|\,\|v_n\| \;=\; \|\varphi_n - \psi\|.$$

Hence

$$1 \;=\; \|\psi\| \;\le\; \|\varphi_n - \psi\| + \|\varphi_n\| \;\le\; 3\|\varphi_n - \psi\|.$$

But the φ_n's are dense in X^*, so they should come arbitrarily close to ψ, a contradiction.

Remark. It is possible to have X separable and X^* not separable. For instance, ℓ_1 is separable, but (see 23.10) its dual is ℓ_∞, which is not separable.

23.24. *Proposition.* Let $(X, \|\ \|)$ and $(Y, \|\ \|)$ be real Banach spaces. Assume one is the dual of the other (i.e., assume either $X = Y^*$ or $Y = X^*$). Let S be a separable subset of X. Then there exists a sequence (y_n) in Y satisfying $\|y_n\| = 1$ for all n, and such that

$$\|s\| \;=\; \sup_n \langle y_n, s\rangle \qquad \text{for each } s \in S.$$

Proof. Note that if $\|y\| \le 1$ then $\langle y, s\rangle \le \|s\|$. Let (s_k) be a dense sequence in S. We proceed by two different arguments:

(i) If $Y = X^*$, we may apply (HB8). For each k, there exists some $y_k \in Y$ satisfying $\|y_k\| = 1$ and $\langle y_k, s_k\rangle = \|s_k\|$.

(ii) If $X = Y^*$, we may apply the definition of the operator norm — i.e., the norm of X. For each k, we have $\|s_k\| = \sup\{\langle y, s_k\rangle : y \in Y, \|y\| = 1\}$. Hence we may choose a sequence $(y_{k,j} : j = 1, 2, 3, \ldots)$ in Y satisfying $\|y_{k,j}\| = 1$ and $\lim_{k\to\infty} \langle y_{k,j}, s_k\rangle = \|s_k\|$. Now arrange the doubly indexed set $\{y_{k,j} : j, k \in \mathbb{N}\}$ into a sequence (y_n) (see 2.20.e).

In either case we obtain a sequence (y_n) in Y, satisfying $\|y_n\| = 1$ for each n and satisfying $\sup_n \langle y_n, s_k\rangle = \|s_k\|$ for each k. Now let any $s \in S$ be given and any number $\varepsilon > 0$. Since (s_k) is dense in S, we have $\|s - s_k\| < \varepsilon$ for some k. For each n, we have $\langle y_n, s\rangle > \langle y_n, s_k\rangle - \varepsilon$, and therefore $\sup_n \langle y_n, s\rangle \ge \|s_k\| - \varepsilon > \|s\| - 2\varepsilon$. Now let $\varepsilon \downarrow 0$.

23.25. *Definitions.* Let X be a Banach space, and let $(\Omega, \mathcal{S})$ be a measurable space. A function $f : \Omega \to X$ is **weakly measurable** if the scalar-valued function $\langle \varphi, f(\cdot)\rangle$ is measurable for each fixed $\varphi \in X^*$. A function $\psi : \Omega \to X^*$ is **weak-star measurable** if the scalar-valued function $\langle \psi(\cdot), x\rangle$ is measurable for each fixed $x \in X$. A function satisfying either of these conditions will be called **scalarly measurable**.

Proposition. Any scalarly measurable, separably valued function is strongly measurable (defined as in 21.4).

Proof (modified slightly from Hille and Phillips [1957]). Here we assume X and Y are Banach spaces, one is the dual of the other, $f : \Omega \to X$ is separably valued, and $\langle y, f(\cdot)\rangle$ is measurable for each fixed $y \in Y$. Replacing each y with $\mathrm{Re}\, y$, we may assume the scalar field is $\mathbb{R}$ (see 11.12); this simplifies our notation slightly. Let S be the closed span of the range of f; then S is separable. As in 23.24, choose a sequence (y_n) in Y with $\|y_n\| = 1$ for all n and $\|s\| = \sup_n \langle y_n, s\rangle$ for each $s \in S$.

Temporarily fix any $v \in S$. The function $f(\omega)-v$ takes its values in S. Moreover, for each $n \in \mathbb{N}$, the real-valued function $\omega \mapsto \langle y_n, f(\omega) - v\rangle = \langle y_n, f(\omega)\rangle - \langle y_n, v\rangle$ is measurable. Hence, for each $v \in S$ the real-valued function $\omega \mapsto \|f(\omega) - v\| = \sup_n \langle y_n, f(\omega) - v\rangle$ is measurable.

In particular, if (x_k) is a dense sequence in S, then each of the functions $\omega \mapsto \|f(\omega) - x_k\|$ (for $k = 1,2,3,\ldots$) is measurable. Now, for each $\omega \in \Omega$ and $j \in \mathbb{N}$, let $f_j(\omega)$ be the first term in the sequence $x_1, x_2, x_3, \ldots$ whose distance from $f(\omega)$ is less than $\frac{1}{j}$. Show that $f_j : \Omega \to X$ is a countably valued, measurable function. Since (f_j) converges uniformly to f as $j \to \infty$, it follows that f is strongly measurable.

Unconditionally Convergent Series

23.26. In 10.42 we gave an elementary example of a series whose sum is affected by a reordering of its terms. We now investigate that phenomenon further.

Definition and proposition. Let $\sum_{j=1}^{\infty} x_j$ be a series in a Banach space $(X, \|\ \ \|)$. Then the following conditions are equivalent; if any (hence all) are satisfied we say the series is **unconditionally convergent**. Furthermore, when those conditions are satisfied, then all the series in (A) have the same sum, and that sum is equal to the limit in (B).

(A) $\sum_{k=1}^{\infty} x_{\pi(k)}$ is convergent for every permutation π of the positive integers. (This is the most commonly used definition of *unconditionally convergent.*)

(B) The net $\left(\sum_{j\in F} x_j : F \in \mathcal{F}\right)$ is convergent, where $\mathcal{F} = \{\text{finite subsets of } \mathbb{N}\}$ is directed by inclusion.

(C) The series $\sum_{j=1}^{\infty} |u(x_j)|$ converges uniformly for all u in the closed unit ball of X^*. That is, let $U = \{u \in X^* : \|u\| \leq 1\}$; then

$$\lim_{N\to\infty} \sup_{u\in U} \sum_{j=N}^{\infty} |u(x_j)| \quad = \quad 0.$$

In other words, the set of sequences of scalars $\{(ux_1, ux_2, ux_3, \ldots) : u \in U\}$ is a relatively compact subset of ℓ_1 (see the characterization of compactness in 22.25).

(D) For each sequence (β_j) of scalars with $|\beta_j| \leq 1$, the series $\sum_{j=1}^{\infty} \beta_j x_j$ converges.

(E) For each sequence (ε_j) with $\varepsilon_j = \pm 1$, the series $\sum_{j=1}^{\infty} \varepsilon_j x_j$ converges.

(F) Each **subseries** $\sum_{k=1}^{\infty} x_{j_k}$ is convergent — i.e., the series $\sum_{k=1}^{\infty} x_{j_k}$ is convergent for each choice of positive integers $j_1 < j_2 < j_3 < \cdots$.

Remarks. A seventh characterization of unconditional convergence will be given in 28.31.

This result and proof are taken from Singer [1970]. Before plowing through the proof of equivalence, some readers may find it helpful to glance ahead to the examples in 23.27. Also, this concept should be compared with the one in 10.40.

Proof that (A) *implies* (B), *and that the sums in* (A) *all equal the limit in* (B). Fix some particular permutation γ of $\mathbb{N}$, and let $x = \sum_{k=1}^{\infty} x_{\gamma(k)}$; suppose that $\sum_{j\in F} x_j$ does not converge to x; we shall obtain a contradiction. By our assumption, there exists some $\varepsilon > 0$ such that every finite set $F \subseteq \mathbb{N}$ is contained in some finite set G such that $\left\| x - \sum_{j\in G} x_j \right\| > \varepsilon$. Since $x = \sum_{k=1}^{\infty} x_{\gamma(k)}$, there is some positive integer N_1 such that

$$N \geq N_1 \qquad \Rightarrow \qquad \left\| x - \sum_{k=1}^{N} x_{\gamma(k)} \right\| < \frac{\varepsilon}{2}.$$

Recursively choose finite sets $F_1 \subseteq G_1 \subseteq F_2 \subseteq G_2 \subseteq F_3 \subseteq G_3 \subseteq \cdots \subseteq \mathbb{N}$ as follows: Let $F_1 = \{\gamma(1), \gamma(2), \ldots, \gamma(N_1)\}$. Given F_m, choose $G_m \supseteq F_m$ such that $\left\| x - \sum_{j\in G_m} x_j \right\| > \varepsilon$. Given G_m, choose $F_{m+1} = \{\gamma(1), \gamma(2), \ldots, \gamma(N_{m+1})\}$ with N_{m+1} large enough so that $N_{m+1} \geq m+1$ and $F_{m+1} \supseteq G_m$. This completes the recursion. Since $N_m \geq m$, the union of the F_m's is equal to $\mathbb{N}$. Now define a sequence $\pi(1), \pi(2), \pi(3), \ldots$ by listing first the elements of F_1 in any order, then the elements of $G_1 \setminus F_1$ in any order, then the elements of $F_2 \setminus G_1$, then the elements of $G_2 \setminus F_2$, etc. The resulting series $\sum_{k=1}^{\infty} x_{\pi(k)}$ is not convergent, since

$$\left\| \sum_{j\in G_m \setminus F_m} x_j \right\| \geq \left\| x - \sum_{j\in G_m} x_j \right\| - \left\| x - \sum_{j\in F_m} x_j \right\| > \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Proof of (B) $\Rightarrow$ (C). Let $x = \lim_{F\in\mathcal{F}} \sum_{j\in F} x_j$. Let any $\varepsilon > 0$ be given. By (B), there is some positive integer N such that if G is any finite subset of $\mathbb{N}$ with $G \supseteq \{1, 2, \ldots, N\}$, then $\left\| x - \sum_{j\in G} x_j \right\| < \frac{\varepsilon}{8}$. Fix any $u \in U$; it suffices to show that $\sum_{j=N+1}^{\infty} |u(x_j)| \leq \varepsilon$. Temporarily fix any integer $p \geq 1$. Define the sets

$$\begin{aligned} A_1 &= \Big\{ j \in \{N+1, N+2, \ldots, N+p\} \ : \ \mathrm{Re}\, u(x_j) \geq 0 \Big\}, \\ A_2 &= \Big\{ j \in \{N+1, N+2, \ldots, N+p\} \ : \ \mathrm{Re}\, u(x_j) < 0 \Big\}. \end{aligned}$$

Also let $B = \{1, 2, \ldots, N\}$. Then

$$\sum_{j=N+1}^{N+p} |\mathrm{Re}\, u(x_j)| = \sum_{k=1}^{2} \left| \mathrm{Re}\, u\left(\sum_{j\in A_k} x_j \right) \right| \leq \sum_{k=1}^{2} \left\| \sum_{j\in A_k} x_j \right\|$$

$$\leq \sum_{k=1}^{2} \left\{ \left\| x - \sum_{j\in A_k\cup B} x_j \right\| + \left\| x - \sum_{j\in B} x_j \right\| \right\} < \sum_{k=1}^{2} \left\{ \frac{\varepsilon}{8} + \frac{\varepsilon}{8} \right\} = \frac{\varepsilon}{2}.$$

Similarly, $\sum_{j=N+1}^{N+p} |\mathrm{Im}\, u(x_j)| < \frac{\varepsilon}{2}$. Hence $\sum_{j=N+1}^{N+p} |u(x_j)| < \varepsilon$. Now let $p \to \infty$.

Proof of (C) $\Rightarrow$ (D). We must show that the partial sums of $\sum_{j=1}^{\infty} \beta_j x_j$ form a Cauchy sequence. Given any $\varepsilon > 0$, choose N by (C), so that $\sup_{u \in U} \sum_{j=N}^{\infty} |u(x_j)| < \varepsilon$. Now for any integers n, p with $p \geq n \geq N$, use the Hahn-Banach Theorem (HB8) in 23.18 to choose some $u \in U$ (which may depend on n, p) to satisfy the first equation below:

$$\left\| \sum_{j=n}^{p} \beta_j x_j \right\| = \left| u\left(\sum_{j=n}^{p} \beta_j x_j \right) \right| \leq \sum_{j=n}^{p} |u(x_j)| < \varepsilon.$$

Proof of (D) $\Rightarrow$ (E). Obvious.

Proof of (E) $\Rightarrow$ (F). Any subseries $\sum_{k=1}^{\infty} x_{j_k}$ can be written as the average of the two convergent series $\sum_{j=1}^{\infty} x_j$ and $\sum_{j=1}^{\infty} \varepsilon_j x_j$, where

$$\varepsilon_j = \begin{cases} 1 & \text{if } j \text{ is among the numbers } j_1, j_2, j_3, \ldots \\ -1 & \text{otherwise.} \end{cases}$$

Proof of (F) $\Rightarrow$ (A). Suppose π is a permutation of $\mathbb{N}$ for which $\sum_k x_{\pi(k)}$ does not converge. Then the partial sums are not a Cauchy sequence. Hence there exists some constant $\varepsilon > 0$ and some sequence of positive integers $m_1 < m_2 < m_3 < \cdots$ such that $\left\| \sum_{k=m_n+1}^{m_{n+1}} x_{\pi(k)} \right\| > \varepsilon$ for all n. That is, $\left\| \sum_{k \in S_n} x_{\pi(k)} \right\| > \varepsilon$, where $S_n = \{m_n + 1, m_n + 2, \ldots, m_{n+1}\}$. The sets $\pi(S_n) = \{\pi(k) : k \in S_n\}$ are disjoint, finite sets with union $\mathbb{N}$. Let $\min \pi(S_n)$ and $\max \pi(S_n)$ be the minimum and maximum elements of $\pi(S_n)$. Note that $\lim_{n \to \infty} \min \pi(S_n) = \infty$. Therefore we can recursively choose positive integers $n(1) < n(2) < n(3) < \cdots$ so that $\min \pi(S_{n(p+1)}) > \max \pi(S_{n(p)})$. Form a subseries $\sum_{i=1}^{\infty} x_{j_i}$ by taking the positive integers $j_1 < j_2 < j_3 < \cdots$ to be the members of $\bigcup_{p=1}^{\infty} \pi(S_{n(p)})$ arranged in increasing order. Then for each p, there exist i' and i'' such that $\left\| \sum_{i=i'}^{i''} x_{j_i} \right\| = \left\| \sum_{k \in S_{n(p)}} x_{\pi(k)} \right\| > \varepsilon$, and the numbers i' and i'' tend to ∞ when $p \to \infty$. This shows that the series $\sum_{i=1}^{\infty} x_{j_i}$ is not convergent.

23.27. *Further exercises, examples, and observations.*

a. A convergent series is not necessarily unconditionally convergent. For instance, in the one-dimensional Banach space $\mathbb{R}$, the series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots$ is convergent; to see that, rewrite it as $\left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \cdots$. However, not every subseries is convergent; for instance, $-\frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \cdots$ does not converge (see 10.41.f).

b. If $\sum_{j=1}^{\infty} x_j$ is absolutely convergent — that is, if $\sum_{j=1}^{\infty} \|x_j\| < \infty$ — then $\sum_{j=1}^{\infty} x_j$ is unconditionally convergent.

c. If the Banach space X is finite-dimensional, then any unconditionally convergent series in X is absolutely convergent.

d. In an infinite-dimensional space, an unconditionally convergent series is not necessarily absolutely convergent. An example is given by the series $\sum_{j=1}^{\infty} x_j$ in the Banach space

$c_0 = \{$sequences of scalars converging to $0\}$, with

$$x_1 = (1, 0, 0, 0, \ldots), \quad x_2 = (0, \frac{1}{2}, 0, 0, \ldots), \quad x_3 = (0, 0, \frac{1}{3}, 0, \ldots), \quad \ldots$$

Remark. Actually, in *every* infinite-dimensional Banach space there exists a series that is unconditionally convergent but not absolutely convergent. This theorem was proved by Dvoretsky and Rogers in 1950, but its proof is too long to include here. (It is given by Diestel [1984], for instance.)

NEUMANN SERIES AND SPECTRAL RADIUS (OPTIONAL)

23.28. Let $(X, |\quad|)$ and $(Y, |\quad|)$ be Banach spaces over the scalar field $\mathbb{F}$. Let $BL(X,Y) = \{$bounded linear operators from X into $Y\}$, with operator norm $\|\quad\|$. Let $Inv(X,Y)$ be the set of all invertible bounded linear operators — that is, bijections f from X onto Y such that both $f : X \to Y$ and $f^{-1} : Y \to X$ are bounded linear operators. Show that

a. Let $i_X : X \to X$ be the identity map. If $p \in BL(X,X)$ with $\|p\| < 1$, then $i_X - p \in Inv(X,X)$ with

$$(i_X - p)^{-1} \quad = \quad \sum_{n=0}^{\infty} p^n; \qquad\qquad \text{hence} \qquad \|(i_X - p)^{-1}\| \le (1 - \|p\|)^{-1}.$$

Hint: The series is absolutely convergent in $BL(X,X)$; let s be its sum. Show that $s(i_X - p) = (i_X - p)s = i_X$.

b. More generally, assume $f \in Inv(X,Y)$ and $u \in BL(X,Y)$ with $\|u\| < \|f^{-1}\|^{-1}$. Then $f - u \in Inv(X,Y)$ with

$$(f-u)^{-1} \quad = \quad f^{-1} + f^{-1}uf^{-1} + f^{-1}uf^{-1}uf^{-1} + \cdots .$$

(*Hint*: Use the preceding result with $p = f^{-1}u$ or $p = uf^{-1}$.) This series is sometimes called the **Neumann series** for $(f-u)^{-1}$. Conclude that $Inv(X,Y)$ is an open subset of $BL(X,Y)$. Also note that $\|(f-u)^{-1}\| \le \left(\|f^{-1}\|^{-1} - \|u\|\right)^{-1}$.

23.29. *Proposition and definition.* Let $(X, |\quad|)$ be a real or complex Banach space; let $\|\quad\|$ denote the operator norm on $BL(X,X)$; let $g \in BL(X,X)$. Then

(i) the number $rad(g) = \lim_{n\to\infty} \|g^n\|^{1/n}$ exists; it is called the **spectral radius** of the operator g. It is also equal to each of the following quantities:

(ii) $rad(g) = \inf_{n\in\mathbb{N}} \|g^n\|^{1/n}$.

(iii) $rad(g) = R^{-1}$, where R is the radius of convergence of the power series $\psi(t) = \sum_{n=0}^{\infty} t^n g^n$.

(iv) $rad(g)$ is the infimum of all the numbers we obtain as operator norms $\|g\|$ for the operator g, when the norm of $(X, |\ \ |)$ is replaced by an equivalent norm and $\|\ \ \|$ is replaced by the resulting operator norm. (Compare 23.2.c.)

Still another equivalent definition of the spectral radius is stated (without proof) at the end of 23.30.

Proof of (i) *and* (ii). Fix any positive integer m. If n is any positive integer, then $n = mp+b$ for some $b \in \{0, 1, 2, \ldots, m-1\}$. Show that

$$|||f^n|||^{1/n} \quad \leq \quad |||f^m|||^{p/(mp+b)}\ |||f|||^{b/(mp+b)}.$$

Holding m fixed, let $n \to \infty$ to show that $\limsup_{n\to\infty} |||f^n|||^{1/n} \leq |||f^m|||^{1/m}$.

Proof of (iii). Immediate from the formula for the radius of convergence of power series, given in 22.23.

Proof of (iv). Since $\|g\| = \|g^1\|^1 \geq \inf_{n\in\mathbb{N}} \|g^n\|^{1/n} = rad(g)$, one direction is obvious. For the opposite inequality, suppose r is some number greater than $rad(g)$. Define $/x/ = \sum_{n=0}^{\infty} r^{-n}|g^n(x)|$. Show that $/\ \ /$ is a norm on X that is equivalent to $|\ \ |$, and $/g(x)/ \leq r/x/$. Thus the resulting operator norm satisfies $//g// \leq r$.

23.30. Let X be a complex Banach space; let $i_X : X \to X$ be the identity map. Define $BL(X, X)$ and $Inv(X, X)$ as in 23.28. Let $g \in BL(X, X)$. We define

$$\begin{array}{lcll} \rho(g) & = & \{\lambda \in \mathbb{C} : \lambda i_X - g \in Inv(X,X)\}, & \text{the } \textbf{resolvent set} \text{ of } g, \text{ and} \\ \sigma(g) & = & \{\lambda \in \mathbb{C} : \lambda i_X - g \notin Inv(X,X)\}, & \text{the } \textbf{spectrum} \text{ of } g. \end{array}$$

These two sets form a partition of $\mathbb{C}$. The function $(\lambda i_X - g)^{-1}$ is sometimes called the **resolvent** of g at λ, particularly in the literature of spectral theory. *Caution*: In some parts of mathematics — e.g., in the literature of semigroups of nonlinear operators — the term "resolvent" sometimes refers to the operator $(i_X - \lambda g)^{-1}$.

From 23.28 it is easy to see that $\rho(g)$ is an open subset of $\mathbb{C}$. Furthermore,

$$\text{if } \lambda \in \mathbb{C} \text{ and } |\lambda| > rad(g), \text{ then } \lambda \in \rho(g) \text{ and } \|(\lambda i_X - g)^{-1}\| \leq \tfrac{1}{|\lambda| - rad(g)}.$$

(Here $rad(g)$ is the spectral radius of g, defined in 23.29. *Hint*: First show that if $|\lambda| \geq \|g\|$, then $\lambda \in \rho(g)$ with $\|(\lambda i_X - g)^{-1}\| \leq \frac{1}{|\lambda| - \|g\|}$.) Thus $\sigma(g)$ is closed and bounded; hence it is a compact set.

More advanced results. We now state without proof a couple of further results about the spectrum. The proofs (which can be found in more advanced or more specialized books) depend on some knowledge of analytic functions.

(i) The spectrum $\sigma(g)$ is nonempty.

(ii) The spectral radius $rad(g)$ is equal to $\sup\{|\lambda| : \lambda \in \sigma(g)\}$.

Chapter 24

Generalized Riemann Integrals

24.1. *Preview.* Presumably the reader is familiar with the Riemann integral, which is introduced in college calculus. In this chapter we study the Riemann integral for Banach-space-valued functions. We also study the Henstock integral, a slight generalization of the Riemann integral. It is conceptually similar to the Riemann integral, but in its power it is more like the Lebesgue integral. In fact, for functions $f : [a, b] \to [0, +\infty)$ we shall prove in 24.36 that the Henstock and Lebesgue integrals are the same.

Still more generally, we shall study the Henstock-Stieltjes integral $\int_a^b f(t)\, d\varphi(t)$. Its notation is slightly more cumbersome, but the greater generality does not make the proofs longer, and the Henstock-Stieltjes integral offers certain advantages — particularly in explaining certain aspects of measure theory (see 24.35) and path integration (used especially in complex analysis; see 25.26). Readers who are entirely unfamiliar with Stieltjes integrals may also wish to glance ahead to 25.17, which shows the relationship between Stieltjes integrals and "ordinary" integrals.

Definitions of the Integrals

24.2. *Assumptions.* In this chapter we shall consider integrals only over compact intervals $[a, b]$; that is, we assume $-\infty < a < b < +\infty$. We shall consider integrals of functions $f : [a, b] \to X$, where $(X, \| \quad \|)$ is a normed vector space. For many of the results, we must assume X is complete; see especially 24.27.

The most important cases to keep in mind are $X = \mathbb{R}$ and $X = \mathbb{C}$, but other Banach spaces are also of interest and the more general theory of Banach-space-valued functions is not significantly harder; moreover, the greater generality will be needed in Chapter 30. It is possible to replace $[a, b]$ with a more general domain — see for instance McLeod [1980] — but the notation then becomes appreciably more complicated.

24.3. *Definition.* Let $f : [a, b] \to X$ be some function, and let $v \in X$. We say v is a **Riemann integral** of f over $[a, b]$ if for each number $\varepsilon > 0$ there exists some number $\delta > 0$ such that

> if $n \in \mathbb{N}$ and $a = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_n = b$ and $\tau_j \in [t_{j-1}, t_j]$ with $t_j - t_{j-1} < \delta$ for all j, then $\|v - \sum_{j=1}^{n}(t_j - t_{j-1}) f(\tau_j)\| < \varepsilon$.

When such an integral exists, we say f is **Riemann integrable**. The Riemann integrability of certain kinds of functions will be established later in this chapter; a characterization of Riemann integrability among real-valued functions will be given in 24.46.

Any function f has at most one Riemann integral v; this can be proved directly by ad hoc methods now (easy exercise) or proved via a broader insight given in 24.7.a. Hence we are justified in calling this vector *the* Riemann integral of f; we shall write it as $v = \int_a^b f(t)dt$. For emphasis it may sometimes be called the **proper Riemann integral**, to distinguish it from "improper Riemann integrals" such as $\int_0^1 t^{-1/2}dt = \lim_{\varepsilon\downarrow 0} \int_\varepsilon^1 t^{-1/2}dt$, which will not be considered here.

Remarks. The definition of "Riemann integral" given above is essentially the same as the definition published by Riemann in 1868 — at least, for real-valued functions f. Some calculus books use a different definition, which is equivalent for real-valued functions f but does not generalize readily to Banach-space-valued f: Any bounded function $f : [a, b] \to \mathbb{R}$ can be approximated both above and below by step functions (defined in 24.22), and those step functions can be integrated in an obvious fashion. When the infimum of the upper integrals equals the supremum of the lower integrals, the common value is called the **Darboux integral** or the **Riemann-Darboux integral**. It was used by Darboux in 1875.

24.4. We now generalize slightly. Let $f : [a, b] \to X$ be some function, and let $v \in X$. We say v is a **Henstock integral** of f over $[a, b]$ if for each number $\varepsilon > 0$ there exists some function $\delta : [a, b] \to (0, +\infty)$ such that

> if $n \in \mathbb{N}$ and $a = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_n = b$ and $\tau_j \in [t_{j-1}, t_j]$ with $t_j - t_{j-1} < \delta(\tau_j)$ for all j, then $\|v - \sum_{j=1}^{n}(t_j - t_{j-1})f(\tau_j)\| < \varepsilon$.

When such an integral exists, we say f is **Henstock integrable**. The Henstock integrability of certain kinds of functions will be established later in this chapter.

Any function f has at most one Henstock integral v; this can be proved directly by ad hoc methods now (easy exercise) or proved via a broader insight given in 24.7.a. Hence we are justified in calling this vector *the* Henstock integral of f; we shall write it as $v = \int_a^b f(t)dt$.

Clearly, any Riemann integral of f is also a Henstock integral of f. The Henstock integral is more general. For instance, $\int_0^1 t^{-1/2}\, dt$ is a Henstock integral with value 2 (if the integrand is defined arbitrarily at $t = 0$), but it is not a proper Riemann integral.

The Henstock integral is sometimes known as the **generalized Riemann integral**. It is also known as the **Kurzweil integral** or the **Henstock-Kurzweil integral**, although that last term also has another meaning — see 24.9. It was introduced independently at about the same time by Kurzweil and Henstock. Kurzweil used it briefly as a tool in the study of certain kinds of differential equations; see particularly Kurzweil [1957]. Henstock developed it in greater detail as part of a wider study of integration theory. The Henstock integral is sometimes known as the **gauge integral**, but that term has also been applied to some other integrals.

The integral studied in this chapter is also known by other names — e.g., the **special Denjoy integral** or the **Denjoy-Perron integral**, since it is equivalent to a more complicated integral worked out earlier by Denjoy and Perron. Research continues on related integrals; some recent references are Bullen et al. [1990], Henstock [1991], and Gordon

[1994].

24.5. *An equivalent definition (optional).* One of the chief advantages of the Henstock integral is that it so greatly resembles the Riemann integral with which we are already somewhat familiar. Thus our intuition about the Riemann integral can be carried over to this new, more general integral. Our definition in 24.4, which follows Henstock [1988], emphasizes this resemblance. However, we note that certain other books (such as McLeod [1980] and DePree and Swartz [1988]) use a slightly different definition for the Henstock integral. In those books, $v = \int_a^b f(t)dt$ means that

> $(*)$ for each number $\varepsilon > 0$, there exists some function $U : [a,b] \to \{$open subintervals of $\mathbb{R}\}$ satisfying $t \in U(t)$ for each t and such that
>
> if $n \in \mathbb{N}$ and $a = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_n = b$ and $\tau_j \in [t_{j-1}, t_j] \subseteq U(\tau_j)$ for all j, then $\|v - \sum_{j=1}^n (t_j - t_{j-1}) f(\tau_j)\| < \varepsilon$.

It is easy to show that this definition $(*)$ is equivalent to our definition of the Henstock integral in 24.4. Indeed, let any $\varepsilon > 0$ be given. If v and f satisfy $(*)$ with some U, then we can satisfy the definition in 24.4 by taking $\delta(\tau) > 0$ small enough so that $(\tau - \delta(\tau), \tau + \delta(\tau)) \subseteq U(\tau)$. Conversely, if v and f satisfy the definition in 24.4 with some δ, then we can satisfy $(*)$ by taking $U(\tau) = (\tau - \frac{1}{2}\delta(\tau), \tau + \frac{1}{2}\delta(\tau))$. (*Exercise.* Fill in the details of this argument.) Hereafter we shall only use the definition in 24.4.

24.6. *Definitions.* We now introduce several auxiliary notations that will be helpful in our study of the Riemann and Henstock integrals.

By a **gauge** we shall mean any function $\delta : [a,b] \to (0, +\infty)$; a positive constant may be viewed as a constant function and thus as a particularly simple gauge. (*Caution*: This kind of "gauge" is unrelated to the other "gauge," a collection of pseudometrics, defined in 2.11.)

By a **tagged division** of the interval $[a,b]$ we shall mean a system of numbers

$$T: \qquad a = t_0 \leq \tau_1 \leq t_1 \leq \tau_2 \leq t_2 \leq \tau_3 \leq \cdots \leq t_{n-1} \leq \tau_n \leq t_n = b,$$

where n is some positive integer; we may sometimes abbreviate this as $T = (n, t_j, \tau_j)$.

Some mathematicians impose the further restriction that $t_{j-1} < t_j$ for each j, to exclude degenerate intervals of length 0. Although that restriction is satisfied in most interesting cases, it is has no real effect on the development of the theory, and omitting that restriction simplifies the notation in some of our proofs — for instance, see 24.12.

A tagged division $T = (n, t_j, \tau_j)$ is called **δ-fine** for some positive constant δ if $t_j - t_{j-1} < \delta$ for all j. More generally, a tagged division $T = (n, t_j, \tau_j)$ is δ-fine for a gauge δ if $t_j - t_{j-1} < \delta(\tau_j)$ for all j.

For any function $f : [a,b] \to X$, the **approximating Riemann sum** corresponding to a tagged division $T = (n, t_j, \tau_j)$ is defined to be the sum

$$\Sigma[f,T] \quad = \quad \sum_{j=1}^n (t_j - t_{j-1})\, f(\tau_j)\,.$$

It is an element of the normed space X.

We can now restate our definitions of the integrals. A vector $v \in X$ is a Riemann integral (respectively, a Henstock integral) of a function $f : [a,b] \to X$ if

> for each number $\varepsilon > 0$ there exists a number $\delta > 0$ (respectively, a gauge $\delta > 0$) such that whenever T is a δ-fine tagged division of $[a,b]$, then $\|v - \Sigma[f,T]\| < \varepsilon$.

24.7. The definitions given above are admittedly complicated: "For each ε there exists a δ such that for each T we have ... ." That grammatical construction contains more quantifiers than are commonly used in a nonmathematical sentence. It takes some getting used to.

The Riemann or Henstock integral may be viewed very naturally as the limit of a certain net. Let us define the sets

$$\begin{aligned} \mathcal{D} &= \{(T,\delta) \ : \ \delta \in (0,+\infty) \text{ and } T \text{ is a } \delta\text{-fine tagged division of } [a,b]\}, \\ \mathcal{E} &= \{(T,\delta) \ : \ \delta \text{ is a gauge on } [a,b] \text{ and } T \text{ is a } \delta\text{-fine tagged division of } [a,b]\}. \end{aligned}$$

Then $\mathcal{D} \subseteq \mathcal{E}$, since every positive constant is a gauge. Both $\mathcal{D}$ and $\mathcal{E}$ will be viewed as directed sets, with this ordering: $(T_1,\delta_1) \preccurlyeq (T_2,\delta_2)$ if $\delta_1 \geq \delta_2$. Unwinding the notation, verify that

> "v is a Riemann integral of f" means that the net $\Big(\Sigma[f,T] \ : \ (T,\delta) \in \mathcal{D}\Big)$ converges in X to v, and
>
> "v is a Henstock integral of f" means that the net $\Big(\Sigma[f,T] \ : \ (T,\delta) \in \mathcal{E}\Big)$ converges in X to v.

Here are two immediate applications of this viewpoint:

a. Since the normed space $(X, \| \ \|)$ is a Hausdorff topological space, each net in X has at most one limit. Thus we have an immediate proof that each X-valued function has at most one Riemann integral or Henstock integral.

b. Assume the normed space X is complete. Then a function $f : [a,b] \to X$ is Riemann- or Henstock-integrable, respectively, if and only if the net $\Big(\Sigma[f,T] \ : \ (T,\delta) \in \mathcal{D}\Big)$ or the net $\Big(\Sigma[f,T] \ : \ (T,\delta) \in \mathcal{E}\Big)$ is Cauchy in X, where Cauchy nets are defined as in 19.2. In other words, f is Riemann integrable (respectively, Henstock integrable) if and only if

> for each $\varepsilon > 0$ there exists some number $\delta > 0$ (respectively, some gauge δ on $[a,b]$) such that whenever T, T' are δ-fine tagged divisions of $[a,b]$, then $\|\Sigma[f,T] - \Sigma[f,T']\| < \varepsilon$.

24.8. *Definitions.* We generalize still further. Let X be a normed space over the scalar field $\mathbb{F}$. (In the simplest case we may take $X = \mathbb{F}$, but greater generality is sometimes useful.) Let f and φ be two functions defined on $[a,b]$ — one of them X-valued, the other scalar-valued.

(Throughout most of this chapter, whenever possible, we shall be intentionally ambiguous about which of f, φ is scalar-valued and which is vector-valued, in order to cover both cases at once.) Define the *approximating Riemann-Stieltjes sum*

$$\Sigma[f,T,\varphi] \quad = \quad \sum_{j=1}^{n} f(\tau_j)\Big[\varphi(t_j) - \varphi(t_{j-1})\Big].$$

The **Riemann-Stieltjes integral** and the **Henstock-Stieltjes integral** are, respectively, the limits of the nets

$$\Big(\Sigma[f,T,\varphi]\ :\ (T,\delta)\in\mathcal{D}\Big) \qquad \text{and} \qquad \Big(\Sigma[f,T,\varphi]\ :\ (T,\delta)\in\mathcal{E}\Big),$$

where $\mathcal{D}, \mathcal{E}$ are defined as in 24.7. The resulting integrals are denoted $\int f d\varphi$ or $\int_a^b f(t)d\varphi(t)$. In other words, the integral is a vector v with the property that for each number $\varepsilon > 0$ there exists a number $\delta > 0$, respectively a gauge $\delta > 0$, such that whenever T is a δ-fine tagged division, then $|v - \Sigma[f,T,\varphi]| < \varepsilon$.

Since most of this chapter concerns itself with Henstock-Stieltjes integrals, when the Henstock-Stieltjes integral $\int_a^b f\,d\varphi$ exists we shall simply say that f is **φ-integrable**.

The theory of Stieltjes integrals generalizes that of Riemann and Henstock integrals, since we can take $\varphi(t) = t$. Readers who are entirely unfamiliar with Stieltjes integrals may wish to glance ahead to 24.18, 24.35, 25.17, and 25.26 for motivation.

24.9. *Remarks on generalizations and variants (optional).* We mention some other integrals that will not be studied in this book.

In defining $\int f d\varphi$, we could let f and φ both be vector-valued. Say they take values in vector spaces X and Y, respectively; then form a product using some bilinear mapping $\langle\ ,\ \rangle : X \times Y \to Z$. The resulting integral would take values in Z.

For any mapping $U : [a,b] \times [a,b] \to X$, we may define the *generalized Perron integral* of U as the limit (if it exists) of sums of the form

$$\Sigma[U,T] \ = \ \sum_{j=1}^{n} [U(\tau_j,t_j) - U(\tau_j,t_{j-1})]$$

for tagged divisions $T = (n, t_j, \tau_j)$. This generalizes the Henstock-Stieltjes integral $\int_a^b f\,d\varphi$ since we can take $U(\tau,t) = f(\tau)\varphi(t)$. For an introduction to this generalized integral and its applications to generalized differential equations, see Schwabik [1992].

Still more generally, let $h = h(\tau, J)$ be a Banach-space-valued function defined for real numbers τ and compact intervals J. The limit of the sums

$$\Sigma[h,T] \quad = \quad \sum_{j=1}^{n} h\left(\tau_j,\ [t_{j-1},t_j]\right),$$

when it exists, is sometimes called the *Henstock-Kurzweil integral* of h. For details the reader may refer to papers and books by Henstock and Kurzweil.

The Lebesgue integral is an **absolute integral** — i.e., if f is Lebesgue integrable, then so is $|f(\cdot)|$; this fact is built into our definition of the Lebesgue integral. The Henstock and Henstock-Stieltjes integrals are not absolute integrals; the Henstock integral is slightly more general than the Lebesgue integral. McShane [1983] studies a gauge integral which is defined slightly differently from 24.4; McShane's integral turns out to be exactly equivalent to the Lebesgue integral. Further information on McShane's integral can be found in R. Výborný [1994/95] and in the appendices of McLeod [1980].

Another integral, due to Fréchet, is particularly simple and noteworthy: Forget about gauges. Let $\mathfrak{T}$ be the set of all tagged divisions of $[a,b]$, with this ordering (which ignores the placement of the tags): $T_1 \preccurlyeq T_2$ if the partition of T_1 is a refinement of the partition of T_2 — i.e., if {divison points of T_2} $\subseteq$ {division points of T_1}. The limit of the resulting net $(\Sigma[f,T])$ is sometimes called the **refinement integral**. It has the advantage that, although it resembles the Riemann integral, it does not depend as heavily on the specialized nature of subintervals of $\mathbb{R}$ — it generalizes very easily to integrals over any measure space. It is discussed further by Hildebrandt [1963, pages 320–325]. The refinement integral is slightly simpler than the gauge integrals studied in this chapter, and perhaps it is a better approach in some respects; that question deserves further study. We prefer the gauge integral chiefly because at present it is more compatible with the wider body of mathematical literature.

24.10. *Proposition.* If $\delta : [a,b] \to (0,+\infty)$ is any gauge, then there exists a δ-fine tagged division of $[a,b]$.

Proof. Let $S = \{s \in [a,b]$: there exists a δ-fine tagged division of $[a,s]\}$. We are to show that $b \in S$. Trivially, $a \in S$. Let $\sigma = \sup(S)$. There is some $s \in S$ such that $s > \sigma - \delta(\sigma)$. Any tagged division of $[a,s]$ can be extended to a tagged division of $[a,\sigma]$ by tacking on the additional interval $[s,\sigma]$ with tag σ. This proves $\sigma \in S$. If $\sigma < b$, then any tagged division of $[a,\sigma]$ can be extended to a larger interval $[a,\sigma']$ by tacking on an additional subinterval $[\sigma,\sigma']$ with tag σ — thereby contradicting the maximality of σ. Thus $b = \sigma$, so $b \in S$.

24.11. *A useful gauge.* The following construction will be used in a few proofs later in this chapter. Let any finite, nonempty set $Q \subseteq [a,b]$ be given.

Let $\rho = \min\{|q-q'| : q,q' \in Q, q \neq q'\}$, or let $\rho = 1$ if Q consists of just one point. Define a gauge $\gamma : [a,b] \to (0,+\infty)$ by:

$$\gamma(t) \quad = \quad \begin{cases} \min\{\rho, \mathrm{dist}(t,Q)\} & \text{when } t \notin Q \\ \rho & \text{when } t \in Q. \end{cases}$$

Then it is easy to see that any γ-fine tagged division $T = (n, t_j, \tau_j)$ will have the following properties:

(i) No subinterval $[t_{j-1}, t_j]$ contains more than one member of Q.

(ii) If $q \in Q \cap [t_{j-1}, t_j]$, then q is equal to τ_j (i.e., the tag of the subinterval).

24.12. *A useful tagged division.* The following construction will be useful in a few proofs later in this chapter. Let $S = (m, s_i, \sigma_i)$ be any tagged division of an interval $[a,b]$. We can form a related, new tagged division $T = (2m, t_j, \tau_j)$ by the following rule: We have

$s_{i-1} \leq \sigma_i \leq s_i$, so we may subdivide each interval $[s_{i-1}, s_i]$ into the two subintervals

$$[t_{2i-2},\, t_{2i-1}] \quad = \quad [s_{i-1},\, \sigma_i] \qquad \text{and} \qquad [t_{2i-1},\, t_{2i}] \quad = \quad [\sigma_i,\, s_i],$$

with both new tags τ_{2i-1} and τ_{2i} equal to the old tag σ_i. (Of course, some of the new subintervals may have length 0, but that is not a difficulty — see the remarks in 24.6.) This tagged division T has the following important properties.

(i) For any gauge δ, if S is δ-fine, then T is also δ-fine.

(ii) For any function $g : [a, b] \to X$, we have $\Sigma[g, S] = \Sigma[g, T]$.

(iii) Each subinterval $[t_{j-1}, t_j]$ in the tagged division $T = (2m, t_j, \tau_j)$ has for its tag τ_j one of the subinterval's endpoints, t_{j-1} or t_j.

Basic Properties of Gauge Integrals

24.13. *Some trivial integrals.*

a. In our definitions of the integrals, we permit $a = b$. Trivially, $\int_a^a f(t)dt$ and $\int_a^a f(t)d\varphi(t)$ always exist and are equal to 0.

b. Let f be a constant function: $f(t) = x$ for all $t \in [a, b]$. Then we have the Riemann integral $\int_a^b x dt = (b - a)x$ or, more generally, the Riemann-Stieltjes integral $\int_a^b x d\varphi = [\varphi(b) - \varphi(a)]x$ for any function φ.

24.14. *Integrals as linear maps.* If T is any tagged division of $[a, b]$, then $f \mapsto \Sigma[f, T]$ is a linear map from $X^{[a,b]}$ into X. The Riemann integrable functions form a linear subspace of $X^{[a,b]}$; it is the set of all f for which the net $\big(\Sigma[f, T] \;:\; (T, \delta) \in \mathcal{D}\big)$ is convergent. The Riemann integral is a linear map from that linear subspace into X; it is the pointwise limit of the net of functions $\Sigma[\cdot, T]$.

Analogous remarks apply for the Henstock integral, with $\mathcal{D}$ replaced by $\mathcal{E}$.

Analogous remarks apply for the Stieltjes integral $\int_a^b f d\varphi$, as a function of f (with φ fixed) or as a function of φ (with f fixed).

24.15. Negligibility of small sets.

a. If $p : [a, b] \to X$ is a function that is only nonzero on a finite subset of $[a, b]$, then the Riemann integral $\int_a^b p(t)dt$ exists and equals 0. If $f, g : [a, b] \to X$ are functions that only differ on a finite subset of $[a, b]$, then the Riemann integral $\int_a^b f(t)dt$ exists if and only if the Riemann integral $\int_a^b g(t)dt$ exists, in which case they are equal. Thus, if we change the value of a function at finitely many points, its Riemann integral is not affected.

In the preceding statements, we cannot replace "finite" with "countable." For example, show that $1_{\mathbb{Q}}$, the characteristic function of the rational numbers, is not Riemann integrable on any interval of positive length.

b. If $p : [a,b] \to X$ is a function that is only nonzero on a countable subset $C = \{c_j\}$ of $[a,b]$, then the Henstock integral $\int_a^b p(t)dt$ exists and equals 0. (*Hint*: Given any number $\varepsilon > 0$, choose a gauge δ so that $\|p(c_j)\|\delta(c_j) < 2^{-j}\varepsilon$ for all j; choose δ arbitrarily outside C.) For instance, $\int_0^1 1_{\mathbb{Q}}(t)dt = 0$. If $f, g : [a,b] \to X$ are functions that only differ on a countable subset of $[a,b]$, then the Henstock integral $\int_a^b f(t)dt$ exists if and only if the Henstock integral $\int_a^b g(t)dt$ exists, in which case they are equal. Thus, if we change the value of a function at countably many points, its Henstock integral is not affected.

c. *Remarks.* The value of a Henstock integral is not affected if we change the integrand on a set of Lebesgue measure 0; that fact will follow from 21.37.i and 24.36. However, for some purposes involving Henstock integrals, we cannot ignore uncountable sets, even if they have measure 0; for instance, see 25.19 and 25.25.

d. Let f and φ be two functions on $[a,b]$, at least one of them scalar-valued. If φ vanishes at a and b and at all but finitely many points of (a,b), then the Henstock-Stieltjes integral $\int_a^b f\,d\varphi$ exists and equals 0. (*Hint*: Use 24.11 and 24.12.) If ψ_1, ψ_2 agree at a and b, and differ only on a finite subset of (a,b), then $\int_a^b f\,d\psi_1$ exists if and only if $\int_a^b f\,d\psi_2$ exists, in which case they are equal.

24.16. *Some elementary estimates.*

a. If $f : [a,b] \to X$ and $h : [a,b] \to \mathbb{R}$ are Henstock integrable and $\|f(\cdot)\| \le h(\cdot)$ on $[a,b]$, then $\|\int_a^b f(t)\,dt\| \le \int_a^b h(t)\,dt$.

More generally, if $\varphi : [a,b] \to \mathbb{R}$ is an increasing function, $f : [a,b] \to X$ and $h : [a,b] \to \mathbb{R}$ are φ-integrable, and $\|f(\cdot)\| \le h(\cdot)$ on $[a,b]$, then $\|\int_a^b f\,d\varphi\| \le \int_a^b h\,d\varphi$.

Hint: First show that $\|\Sigma[f,T,\varphi]\| \le \Sigma[h,T,\varphi]$.

b. *A mean value theorem.* If $f : [a,b] \to X$ is Henstock integrable, then $\frac{1}{b-a}\int_a^b f(t)\,dt$ is in the closed convex hull of the range of f.

More generally, if $\varphi : [a,b] \to \mathbb{R}$ is an increasing function and $f : [a,b] \to X$ is φ-integrable, then $[\varphi(b) - \varphi(a)]^{-1}\int_a^b f\,d\varphi$ is a member of the closed convex hull of the range of f.

Hint: First show that $[\varphi(b) - \varphi(a)]^{-1}\Sigma[f,T,\varphi] \in \mathrm{co}(\mathrm{Ran}(f))$.

c. Let f, φ be functions defined on $[a,b]$, at least one of them scalar-valued. Suppose φ has bounded variation and f is bounded and φ-integrable. Then $|\int_a^b f\,d\varphi| \le \|f\|_\infty \mathrm{Var}(\varphi, [a,b])$.

Hint: First show that $|\Sigma[f,T,\varphi]| \le \|f\|_\infty \mathrm{Var}(\varphi, [a,b])$.

24.17. Theorem on uniform limits. Assume the normed space X is complete. Suppose that $f_1, f_2, f_3, \ldots : [a,b] \to X$ are functions converging uniformly on $[a,b]$ to a function $f : [a,b] \to X$.

(i) If the f_n's are Riemann integrable or Henstock integrable, then f is integrable in the same sense, and $\int_a^b f_n(t)dt \to \int_a^b f(t)dt$.

(ii) More generally, suppose $\varphi : [a,b] \to \mathbb{R}$ is an increasing function. If the f_n's have Riemann-Stieltjes or Henstock-Stieltjes integrals with respect to φ, then f is integrable in the same sense and $\int_a^b f_n(t)d\varphi(t) \ \to \ \int_a^b f(t)d\varphi(t)$.

Hints: It suffices to prove (ii). By assumption, $\varepsilon_j = \|f - f_j\|_\infty = \sup_t |f(t) - f_j(t)|$ tends to 0 as $j \to \infty$. We have $\|f_j(t) - f_k(t)\| \le \varepsilon_j + \varepsilon_k$ for all t. Hence

$$\|\int_a^b f_j d\varphi - \int_a^b f_k d\varphi\| = \|\int_a^b (f_j - f_k)d\varphi\| \le \int_a^b (\varepsilon_j + \varepsilon_k)d\varphi = (\varphi(b) - \varphi(a))(\varepsilon_j + \varepsilon_k),$$

which tends to 0 as $j, k \to \infty$. Thus the sequence $(\int f_j d\varphi)$ is Cauchy and converges to some limit v. Now estimate

$$\|\Sigma[f,T,\varphi] - v\| \le \|\Sigma[f - f_j, T, \varphi]\| + \|\Sigma[f_j, T, \varphi] - \int f_j d\varphi\| + \|\int f_j d\varphi - v\|.$$

24.18. Reparametrization Theorem. Let $\sigma : [\widehat{a}, \widehat{b}] \to [a,b]$ be an increasing bijection. Let f and φ be functions defined on $[a,b]$, at least one of them scalar-valued. Then the Henstock-Stieltjes integral $\int_a^b f\, d\varphi$ exists if and only if the Henstock-Stieltjes integral $\int_{\widehat{a}}^{\widehat{b}} (f \circ \sigma)\, d(\varphi \circ \sigma)$ exists, in which case they are equal.

Proof. Let $\widehat{f} = f \circ \sigma$ and $\widehat{\varphi} = \varphi \circ \sigma$. We are to prove that $\int_a^b f\, d\varphi$ exists if and only if $\int_{\widehat{a}}^{\widehat{b}} \widehat{f}\, d\widehat{\varphi}$ exists, in which case they are equal. There is a symmetry between the "hat" quantities and the "no-hat" quantities, since $\sigma^{-1} : [a,b] \to [\widehat{a}, \widehat{b}]$ is an increasing bijection and we have $f = \widehat{f} \circ \sigma^{-1}$ and $\varphi = \widehat{\varphi} \circ \sigma^{-1}$.

Corresponding to each tagged division

$$T \quad : \quad a = t_0 \le \tau_1 \le t_1 \le \tau_2 \le \cdots \le t_{n-1} \le \tau_n \le t_n = b$$

is another tagged division

$$\widehat{T} \quad : \quad \widehat{a} = \widehat{t}_0 \le \widehat{\tau}_1 \le \widehat{t}_1 \le \widehat{\tau}_2 \le \cdots \le \widehat{t}_{n-1} \le \widehat{\tau}_n \le \widehat{t}_n = \widehat{b}$$

defined by $\widehat{t}_j = \sigma^{-1}(t_j)$ and $\widehat{\tau}_j = \sigma^{-1}(\tau_j)$. It is easy to verify that $\Sigma[f,T,\varphi] = \Sigma[\widehat{f}, \widehat{T}, \widehat{\varphi}]$. We are to prove that $\lim_T \Sigma[f,T,\varphi]$ exists if and only if $\lim_{\widehat{T}} \Sigma[\widehat{f}, \widehat{T}, \widehat{\varphi}]$ exists, in which case the limits are equal. Thus, it suffices to prove that the tagged divisions T become fine when and only when the tagged divisions $\widehat{T}$ become fine. By symmetry, it suffices to prove half of this implication.

Thus, let any gauge $\widehat{\delta}$ on $[\widehat{a}, \widehat{b}]$ be given; it suffices to prove the existence of a gauge δ on $[a,b]$ with the property that

$$\text{whenever } T \text{ is } \delta\text{-fine, then } \widehat{T} \text{ is } \widehat{\delta}\text{-fine.}$$

(*Caution*: The most obvious choice is $\delta = \widehat{\delta} \circ \sigma^{-1}$, but that choice doesn't work; we need something slightly more sophisticated.) Since $\sigma^{-1} : [a,b] \to [\widehat{a}, \widehat{b}]$ is an increasing bijection, it is continuous; in particular it is continuous at τ. For each number τ in $[a,b)$, we can choose $\delta(\tau)$ to be a positive number small enough so that $\tau + \delta(\tau) \in [a,b]$ and

$$\sigma^{-1}\Big(\tau + \delta(\tau)\Big) \quad < \quad \sigma^{-1}(\tau) + \frac{1}{2}\widehat{\delta}\left(\sigma^{-1}(\tau)\right). \tag{*1}$$

Also, for each number τ in $(a,b]$, we can choose $\delta(\tau)$ to be a positive number small enough so that $\tau-\delta(\tau)\in[a,b]$ and

$$\sigma^{-1}\Big(\tau-\delta(\tau)\Big) \quad > \quad \sigma^{-1}(\tau)-\frac{1}{2}\widehat{\delta}\left(\sigma^{-1}(\tau)\right). \tag{$*2$}$$

These conditions can be satisfied simultaneously since $\frac{1}{2}\widehat{\delta}\left(\sigma^{-1}(\tau)\right)$ is a positive number.

Now suppose T is δ-fine. Then for each j, we have $\tau_j\in[t_{j-1},t_j]$ and $t_j-t_{j-1}<\delta(\tau_j)$; hence

$$t_j\leq\tau_j+\delta(\tau_j) \qquad \text{and} \qquad t_{j-1}\geq\tau_j-\delta(\tau_j).$$

We can now prove $\sigma^{-1}(t_j)<\sigma^{-1}(\tau_j)+\frac{1}{2}\widehat{\delta}\left(\sigma^{-1}(\tau_j)\right)$ by two different arguments. If $\tau_j\in[a,b)$, then this inequality follows from $(*1)$; if $\tau_j=b$, then we deduce that $t_j=\tau_j$. Similarly, we obtain $\sigma^{-1}(t_{j-1})>\sigma^{-1}(\tau_j)-\frac{1}{2}\widehat{\delta}\left(\sigma^{-1}(\tau_j)\right)$. Hence

$$\widehat{t}_j-\widehat{t}_{j-1} \quad = \quad \sigma^{-1}(t_j)-\sigma^{-1}(t_{j-1}) \quad < \quad \widehat{\delta}\left(\sigma^{-1}(\tau_j)\right) \quad = \quad \widehat{\delta}(\widehat{\tau}_j),$$

so $\widehat{T}$ is $\widehat{\delta}$-fine.

Additivity over Partitions

24.19. Theorem. Let $p_0<p_1<p_2<\cdots<p_m$. Let f and φ be functions defined on $[p_0,p_m]$ — one taking values in a Banach space X, the other in the scalar field $\mathbb{F}$. Then f is φ-integrable on $[p_0,p_m]$ if and only if its restrictions to the subintervals $[p_{j-1},p_j]$ are all φ-integrable, in which case

$$\int_{p_0}^{p_m} f\,d\varphi \quad = \quad \int_{p_0}^{p_1} f\,d\varphi+\int_{p_1}^{p_2} f\,d\varphi+\cdots+\int_{p_{m-1}}^{p_m} f\,d\varphi.$$

Proof. It suffices to show this for $m=2$; then apply induction. Let us denote $a=p_0$, $q=p_1$, $b=p_2$. Thus, it suffices to consider $[a,b]=[a,q]\cup[q,b]$, where $a<q<b$.

First suppose that f is φ-integrable on $[a,b]$; we shall prove that f is φ-integrable on $[a,q]$. (A similar argument works on $[q,b]$.) We shall use the fact that f satisfies the Cauchy criterion 24.7.b on the interval $[a,b]$ to show that this criterion is also satisfied on the subinterval $[a,q]$. Let any $\varepsilon>0$ be given. By assumption, there is some positive number δ (or some gauge δ) such that if S,S' are δ-fine tagged divisions of $[a,b]$, then $\|\Sigma[f,S,\varphi]-\Sigma[f,S',\varphi]\|<\varepsilon$. Let $\widehat{\delta}$ be the restriction of δ to $[a,p]$. Now let T,T' be any two tagged divisions of $[a,p]$ that are $\widehat{\delta}$-fine. Using 24.10, show that T,T' can be extended to δ-fine tagged divisions S,S' of $[a,b]$ that are identical on $[p,b]$. Hence $\|\Sigma[f,T,\varphi]-\Sigma[f,T',\varphi]\|=\|\Sigma[f,S,\varphi]-\Sigma[f,S',\varphi]\|<\varepsilon$.

Conversely, suppose f is φ-integrable on both $[a,q]$ and $[q,b]$; we shall show that $\int_a^b f\,d\varphi$ exists and equals $\int_a^q f\,d\varphi+\int_q^b f\,d\varphi$. Let any $\varepsilon>0$ be given. It suffices to construct a gauge δ on $[a,b]$ with the property that if

$$S: \qquad a=s_0\leq\sigma_1\leq s_1\leq\sigma_2\leq s_2\leq\cdots\leq s_{n-1}\leq\sigma_n\leq s_n=b$$

is any δ-fine tagged division of $[a,b]$, then

$$\left\| \Sigma[f,S,\varphi] - \int_a^q f\,d\varphi - \int_q^b f\,d\varphi \right\| < \varepsilon. \qquad (**)$$

By hypothesis, since the integral exists on $[a,q]$ and $[q,b]$, there exists some gauge γ_1 on $[a,b]$ such that if U is any γ_1-fine tagged division of either $J = [a,q]$ or $J = [q,b]$, then $\|\Sigma[f,U,\varphi] - \int_J f\,d\varphi\| < \varepsilon/2$. Define a gauge γ as in 24.11 with $Q = \{q\}$. We shall show that the gauge $\delta = \min\{\gamma_1, \gamma\}$ satisfies condition $(**)$

Indeed, let S be a δ-fine tagged division of $[a,b]$. We must have $q \in [s_{k-1}, s_k]$ for at least one value of k; fix such a value of k. (There may be two such values of k, if q is equal to one of the division points s_j. In that case, let k be either one of the applicable values; let k remain fixed throughout the remainder of this argument.) Then $q = \sigma_k$, by one of the consequences of 24.11.

Now split S into tagged divisions U_1, U_2 of $[a,q], [q,b]$, splitting the subinterval $[s_{k-1}, s_k]$ into two subintervals both with tag q. Thus U_1 and U_2 have these tags and subintervals:

$$\begin{array}{ll} U_1: & \sigma_1 \in [s_0,s_1],\ \sigma_2 \in [s_1,s_2],\ \ldots,\ \sigma_{k-1} \in [s_{k-2},s_{k-1}],\ q \in [s_{k-1},q], \\ U_2: & q \in [q,s_k],\ \sigma_{k+1} \in [s_k,s_{k+1}],\ \ldots,\ \sigma_{n-1} \in [s_{n-2},s_{n-1}],\ \sigma_n \in [s_{n-1},s_n]. \end{array}$$

It is easy to see that U_1 and U_2 are both γ-fine, and that $\Sigma[f,S,\varphi] = \Sigma[f,U_1,\varphi] + \Sigma[f,U_2,\varphi]$. This completes the proof.

24.20. *Notation and corollary.* It is convenient to define $\int_b^a f\,d\varphi = -\int_a^b f\,d\varphi$. Thus an expression of the form $\int_p^q f\,d\varphi$ may be defined regardless of whether $p < q$ or $p > q$. With that notation, we have this corollary:

Let $f : [a,b] \to X$ be φ-integrable. Let p, q, r be any three numbers in $[a,b]$ (not necessarily in increasing order). Then

$$\int_p^q f\,d\varphi + \int_q^r f\,d\varphi = \int_p^r f\,d\varphi.$$

24.21. *Remarks.* A theorem analogous to 24.19 is also valid for Riemann integrals, with a similar but slightly longer proof. We shall omit the proof, since that result is not needed later in this book.

An analogous theorem is *not* valid for Riemann-Stieltjes integrals. Indeed, let

$$\varphi(t) = \begin{cases} 0 & \text{when } -1 \le t < 0 \\ 1 & \text{when } 0 \le t \le 1, \end{cases} \qquad f(t) = \begin{cases} 0 & \text{when } -1 \le t \le 0 \\ 1 & \text{when } 0 < t \le 1. \end{cases}$$

Then it is easy to prove that the Riemann-Stieltjes integrals $\int_{-1}^0 f\,d\varphi$ and $\int_0^1 f\,d\varphi$ both exist, but the Riemann-Stieltjes integral $\int_{-1}^1 f\,d\varphi$ does not exist. *Hint*: 24.22.c.

24.22. Recall from 11.43 that a **step function** on $[a,b]$ is a function that takes a constant value x_j on each open subinterval (p_{j-1}, p_j), for some division

$$a = p_0 < p_1 < p_2 < \cdots < p_m = b.$$

Show that

a. If f is a step function, then f is Riemann integrable, with $\int_a^b f(t)dt = \sum_{j=1}^m (p_j - p_{j-1})x_j$. Note that the values of $f(p_0)$, $f(p_1)$, ..., $f(p_m)$ are irrelevant — i.e., they can be altered without any effect on the value of $\int_a^b f(t)dt$ — a particular instance of the principle observed in 24.15.

b. More generally, suppose f is a step function on $[a,b]$, and φ is a function on $[a,b]$ that has right- and left-hand limits

$$\varphi(t+) = \lim_{u \downarrow t} \varphi(t) \qquad \text{and} \qquad \varphi(t-) = \lim_{u \uparrow t} \varphi(u)$$

at every $t \in [a,b]$, with the convention that $\varphi(a-) = \varphi(a)$ and $\varphi(b+) = \varphi(b)$. (This hypothesis is satisfied, for instance, if φ has bounded variation; see 19.21.) Assume that at least one of f, φ is scalar-valued. Then f is φ-integrable, with

$$\int_a^b f d\varphi = \sum_{j=1}^m [\varphi(p_j-) - \varphi(p_{j-1}+)]\, x_j + \sum_{j=0}^m [\varphi(p_j+) - \varphi(p_j-)]\, f(p_j).$$

(This formula is taken from McLeod [1980]. The constructions 24.11 and 24.12 may be useful in the proof.) Note that this formula *does* depend on the values of f at $p_0, p_1, p_2, \ldots, p_m$, unless φ is continuous at those points.

c. If f and φ are any functions on $[a,b]$ — one vector-valued, the other scalar-valued — that are both discontinuous at some point $p \in [a,b]$, then the Riemann-Stieltjes integral $\int_a^b f\, d\varphi$ does not exist.

24.23. Henstock-Saks Lemma. Let f and φ be functions on an interval $[a,b]$; assume one of them takes values in a Banach space X and the other is scalar-valued. Let f be φ-integrable — i.e., assume the Henstock-Stieltjes integral $\int_a^b f\, d\varphi$ exists. Let ε be a positive number, and (as in the definition of the integral) let δ be a gauge with the property that every δ-fine tagged division $T = (n, t_j, \tau_j)$ satisfies $\|\Sigma[f, T, \varphi] - \int_a^b f\, d\varphi\| < \varepsilon$. Then:

(i) Suppose $T = (n, t_j, \tau_j)$ is a tagged division, not necessarily δ-fine, but satisfying $t_j - t_{j-1} < \delta(\tau_j)$ for all j in some set $J \subseteq \{1, 2, 3, \ldots n\}$. Then

$$\left\| \sum_{j \in J} \left\{ f(\tau_j)\, [\varphi(t_j) - \varphi(t_{j-1})] - \int_{t_{j-1}}^{t_j} f\, d\varphi \right\} \right\| \le \varepsilon.$$

(ii) If $T = (n, t_j, \tau_j)$ is a δ-fine tagged division and X is a finite-dimensional space equipped with any norm equivalent to any of the usual norms, then

$$\sum_{j=1}^n \left\| f(\tau_j)\, [\varphi(t_j) - \varphi(t_{j-1})] - \int_{t_{j-1}}^{t_j} f(t)dt \right\| \le \kappa\varepsilon$$

for some constant κ that depends only on the choice of the space X and its norm $\| \;\; \|$. In particular, if $X = \mathbb{R}^q$ with norm $\|(x_1, x_2, \ldots, x_q)\| = |x_1| + |x_2| + \cdots + |x_q|$, then we can take $\kappa = 2q$.

Proof of (i). Let $K = \{1, 2, 3, \ldots, n\} \setminus J$. Throughout this argument, the tagged division $T = (n, t_j, \tau_j)$ and the sets J and K will be held fixed, but we shall consider certain other, varying tagged divisions based on T, J, K.

Temporarily fix any $k \in K$, and consider tagged divisions U of the subinterval $[t_{k-1}, t_k]$. We know that f is φ-integrable on $[t_{k-1}, t_k]$ (see 24.19); hence the net $(\Sigma[f, U, \varphi])$ converges to $\int_{t_{k-1}}^{t_k} f\, d\varphi$ as the tagged division U becomes finer.

Now, for the entire interval $[a, b]$, consider a tagged division

$$S: \qquad a = s_0 \leq \sigma_1 \leq s_1 \leq \sigma_2 \leq s_2 \leq \cdots \leq s_{m-1} \leq \sigma_m \leq s_m = b$$

that is identical to T on each of the intervals $[t_{j-1}, t_j]$ for $j \in J$, but is much finer on each of the intervals $[t_{k-1}, t_k]$ for $k \in K$. Then

$$\begin{aligned} \Sigma[f, S, \varphi] \quad = \quad & \sum_{j \in J} f(\tau_j) \left[\varphi(t_j) - \varphi(t_{j-1})\right] \\ & + \sum_{k \in K} \sum_{\substack{\{i \,:\, [s_{i-1}, s_i] \\ \subseteq [t_{k-1}, t_k]\}}} f(\sigma_i) \left[\varphi(s_i) - \varphi(s_{i-1})\right]. \end{aligned}$$

As the divisions on the intervals $[t_{k-1}, t_k]$ become finer, we have this convergence:

$$\Sigma[f, S, \varphi] \quad \to \quad \sum_{j \in J} f(\tau_j) \left[\varphi(t_j) - \varphi(t_{j-1})\right] \quad + \quad \sum_{k \in K} \int_{t_{k-1}}^{t_k} f\, d\varphi.$$

On the other hand, by applying 24.10 on each subinterval $[t_{k-1}, t_k]$ for $k \in K$, we can choose S so that it is δ-fine, hence $\|\Sigma[f, S, \varphi] - \int_a^b f\, d\varphi\| \leq \varepsilon$. Taking limits in this inequality as S becomes progressively finer on the subintervals $[t_{k-1}, t_k]$, we obtain

$$\left\| \sum_{j \in J} f(\tau_j) \left[\varphi(t_j) - \varphi(t_{j-1})\right] \; + \; \sum_{k \in K} \int_{t_{k-1}}^{t_k} f\, d\varphi \; - \; \int_a^b f\, d\varphi \right\| \; \leq \; \varepsilon.$$

Rearranging terms, we obtain conclusion (i).

Proof of (ii). To prove the result for $X = \mathbb{R}$, use conclusion (i) twice — once with J consisting of those j's for which the number

$$f(\tau_j) \left[\varphi(t_j) - \varphi(t_{j-1})\right] \; - \; \int_{t_{j-1}}^{t_j} f\, d\varphi$$

is positive and once with those j's for which that number is negative; then add the results. For an arbitrary positive integer q, apply the one-dimensional result to each of the real-valued functions $\pi_i \circ f$, where $\pi_i : \mathbb{R}^q \to \mathbb{R}$ is the ith coordinate projection. For the $\|\ \ \|_1$ norm, we just add the coordinatewise errors. Any equivalent norm $\|\ \ \|$ satisfies $c_1 \|x\| \leq \|x\|_1 \leq c_2 \|x\|$ for some positive constants c_1, c_2.

24.24. *Definition.* Let $f : [a,b] \to X$ be φ-integrable. Then an **indefinite integral** of f is a function $F : [a,b] \to X$ of the form $F(t) = x + \int_c^t f\,d\varphi$, for any constants $c \in [a,b]$ and $x \in X$. Note that any two indefinite integrals of f differ by a constant. For simplicity, the most common choice of F is with $c = a$ and $x = 0$. For most applications, the particular choice of x and c does not matter, so we may refer to F as "*the* indefinite integral of f." We may write it as

$$F(t) \quad = \quad \int_a^t f\,d\varphi \; + \; \text{constant}.$$

This is actually a whole collection of functions — one for each choice of the constant — but any one of those functions will work equally well in most applications.

24.25. Continuity Theorem. If f is Henstock integrable on $[a,b]$, then the indefinite integral $F(t) = \int_a^t f(s)ds$ is continuous.

More generally, if f is φ-integrable, then the indefinite integral $F(t) = \int_a^t f\,d\varphi$ is right continuous (respectively, left continuous) at each point where φ is.

Proof. Fix any $p \in [a,b)$ where φ is right continuous; we shall show that F is right continuous at p. (A similar argument works for left continuity.) Let any $\varepsilon > 0$ be given; choose some corresponding gauge δ as in the definition of the Henstock integral or in the Henstock-Saks Lemma. Replacing δ with a smaller gauge if necessary, we may assume $p + \delta(p) < b$. Now consider any $q \in (p, p+\delta(p))$. By applying 24.10 on the intervals $[a,p]$ and $[q,b]$ we can obtain a δ-fine tagged division T that has as one of its subintervals $[t_{\widehat{\jmath}-1}, t_{\widehat{\jmath}}] = [p,q]$ with tag $\tau_{\widehat{\jmath}} = p$. Apply the Henstock-Saks Lemma 24.23 with J equal to the singleton $\{\widehat{\jmath}\}$; thus $|f(p)[\varphi(q) - \varphi(p)] - \int_p^q f\,d\varphi| < \varepsilon$. This proves that

$$q \in (p, p+\delta(p)) \qquad \Rightarrow \qquad |F(q) - F(p)| < \varepsilon + |\varphi(q) - \varphi(p)|\,|f(p)|.$$

Hence $\limsup_{q \downarrow p} |F(q) - F(p)| \leq \varepsilon$. Since ε was chosen arbitrarily, we have

$$\limsup_{q \downarrow p} |F(q) - F(p)| \;\leq\; 0, \qquad \text{hence} \qquad \lim_{q \downarrow p} F(q) \;=\; F(p).$$

Integrals of Continuous Functions

24.26. Advanced calculus theorem: existence of the integral. Assume the normed space X is complete, and let $f : [a,b] \to X$ be continuous — or more generally, piecewise continuous (defined in 19.28). Then:

(i) The Riemann integral $\int_a^b f(t)dt$ exists.

(ii) More generally, let $\varphi : [a,b] \to \mathbb{F}$ be any function of bounded variation. Then the Riemann-Stieltjes integral $\int_a^b f d\varphi$ exists.

Hints: f is a uniform limit of step functions; we shall apply 24.22 and 24.17. It suffices to consider the case of real-valued φ, since $\varphi = \mathrm{Re}(\varphi) + i\,\mathrm{Im}(\varphi)$. It suffices to consider φ increasing, since any real-valued function of bounded variation is the difference of two increasing functions.

Remark. Much weaker hypotheses imply the existence of integrals; see 24.45 and 29.33.b.

24.27. *Converse proposition* (*optional*). Let $(X, \|\ \ \|)$ be a normed vector space that is not complete. Then there exists a continuous function $f : [0,1] \to X$ that is not Henstock integrable.

Proof. By assumption, there exists some sequence (x_n) in X that is Cauchy but does not converge. Replacing (x_n) with a subsequence, we may assume $\|x_n - x_{n+1}\| < 4^{-n}$ for all $n \in \mathbb{N}$. Let $u_n = x_n - x_{n+1}$; then $\|u_n\| < 4^{-n}$ but $\sum_{n=1}^{\infty} u_n = \lim_{N\to\infty} \sum_{n=1}^{N} u_n$ does not exist in X. Note that $\sum_{n=1}^{\infty} u_n$ does exist in the completion of X, which we shall denote by Y.

Now let $\varphi : [0,1] \to [0,+\infty)$ be some continuous function satisfying $\varphi(0) = \varphi(1) = 0$ and $c = \int_0^1 \varphi(t)dt > 0$. (The particular choice of φ does not matter; three such functions are $\frac{1}{2} - |t - \frac{1}{2}|$, $\frac{1}{4} - (t - \frac{1}{2})^2$, and $\sin(\pi t)$.)

Define $f : [0,1] \to X$ as follows: Let $f(0) = 0$. For $n = 1, 2, 3, \ldots$, on the subinterval $[2^{-n}, 2^{-n+1}]$, let $f(t) = 2^n \varphi(2^n t - 1) u_n$. Then f is continuous on that subinterval and vanishes at each end of that subinterval, and $\|f(t)\| \le 2^n \|\varphi\|_\infty \|u_n\| < 2^{-n}\|\varphi\|_\infty$ everywhere on that subinterval; hence f is continuous everywhere on $[0,1]$. An easy computation shows $\int_{2^{-n}}^{2^{-n+1}} f(t)dt = cu_n$.

We may view f as a continuous function from $[0,1]$ into the completion space Y. Then f is Riemann integrable in Y, by 24.26(i). It is intuitively obvious (and an only moderately difficult exercise to prove) that $\int_0^1 f(t)dt = c\sum_{n=1}^{\infty} u_n$, which exists in Y but not in X. If f has a Riemann or Henstock integral in X, then that integral must coincide with the Riemann integral in Y; thus f does not have a Henstock integral in X.

24.28. *Proposition.* Let $\mathbb{F}$ be the scalar field, and let $C[a,b] = \{$continuous functions from $[a,b]$ to $\mathbb{F}\}$. Let X be a Banach space. Define

$$C[a,b]^{\perp} = \left\{ \psi \in BV([a,b],X) \ : \ \int_a^b f\,d\psi = 0 \text{ for every } f \in C[a,b] \right\}.$$

Then a second, equivalent definition is

$$\begin{aligned} C[a,b]^{\perp} \quad = \quad & \Big\{ \psi \in BV([a,b],X) \ : \ \psi(a) = \psi(b) \text{ and the set} \\ & \{t \in [a,b] : \psi(t) \ne \psi(a)\} \text{ is at most countable} \Big\}. \end{aligned}$$

The linear space $BV([a,b],X)$ can be expressed as a direct sum of two linear subspaces, as follows:

$$BV([a,b],X) \quad = \quad C[a,b]^{\perp} \oplus NBV([a,b],X),$$

where $NBV([a,b],X)$ is defined as in 22.19.d. That is, any $\psi \in BV([a,b],X)$ can be written in one and only one way as $\psi = \psi_1 + \psi_2$ with $\psi_1 \in C[a,b]^{\perp}$ and $\psi_2 \in NBV([a,b],X)$. Furthermore, $\mathrm{Var}(\psi_2) \le \mathrm{Var}(\psi)$.

Proof (following Limaye [1981]). To prove the equivalence of the two definitions of $C[a,b]^{\perp}$, let any $\psi \in BV([a,b],X)$ be given. For simplicity of notation, replace $\psi(\cdot)$ with the function $\psi(\cdot) - \psi(a)$; thus we may assume $\psi(a) = 0$.

First suppose $\psi \in C[a,b]^{\perp}$ using the first definition. If we take f to be the constant function 1, we find that $\psi(b) = \psi(a)$. Now let $v(t) = \mathrm{Var}(\psi, [a,t])$; then v is increasing and hence has at most countably many discontinuities. Fix any point t_0 where v is continuous; it suffices to show that $\psi(t_0) = 0$. For large integers n, define the continuous function

$$f_n(t) \quad = \quad \begin{cases} 1 & \text{if } a \le t \le t_0 \\ 1 - (t - t_0)n & \text{if } t_0 \le t \le t_0 + \frac{1}{n} \\ 0 & \text{if } t_0 + \frac{1}{n} \le t \le b. \end{cases}$$

Using 24.19, we can compute

$$0 \quad = \quad \int_a^b f_n \, dt \quad = \quad \psi(t_0) \; + \; \int_{t_0}^{t_0 + \frac{1}{n}} f_n(t)\, d\psi \; + \; 0.$$

Since $\|f_n\|_\infty \le 1$, 24.16.c shows that $|\psi(t_0)| \le \mathrm{Var}(\psi, [t_0, t_0 + \frac{1}{n}]) = v(t_0 + \frac{1}{n}) - v(t_0)$. Now take limits as $n \to \infty$.

On the other hand, suppose that $\psi \in C[a,b]^{\perp}$ using the second definition. We use the fact that $\int_a^b f\, d\psi$ is a Riemann-Stieltjes integral, not just a Henstock-Stieltjes integral. We have $\int_a^b f\, d\psi = \lim_T \Sigma[f, T, \psi]$ for any choice of tagged divisions T that have subinterval lengths tending to 0. By our hypothesis on ψ, we can choose the tagged divisions T so that the subintervals $[t_{j-1}, t_j]$ satisfy $\psi(t_{j-1}) = \psi(t_j) = 0$. Then $\Sigma[f, T, \psi] = 0$ for all such tagged divisions. This completes the proof of the equivalence of the two definitions.

From the second definition of $C[a,b]^{\perp}$ it is clear that $C[a,b]^{\perp} \cap NBV([a,b],X) = \{0\}$; hence any ψ can be written in at most one way as $\psi_1 + \psi_2$. Let us show that it can be written in at least one way. Let any $\psi \in BV([a,b],X)$ be given. Since any constant function belongs to $C[a,b]^{\perp}$, we may replace $\psi(\cdot)$ with the function $\psi(\cdot) - \psi(a)$; thus we may assume $\psi(a) = 0$ to simplify our notation. Now define

$$\psi_2(t) \quad = \quad \begin{cases} 0 & \text{when } t = a \\ \psi(t+) & \text{when } t \in (a,b) \\ \psi(b) & \text{when } t = b. \end{cases}$$

Then ψ_2 is right continuous on (a,b).

To show that $\mathrm{Var}(\psi_2) \le \mathrm{Var}(\psi)$, let any partition $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ be given and any number $\varepsilon > 0$. For $j = 1, 2, \ldots, n-1$ choose some point $s_j \in (t_j, t_{j+1})$ with $|\psi_2(t_j) - \psi(s_j)| < \frac{1}{n}\varepsilon$. That inequality is also satisfied for $j = 0$ and $j = n$ by taking $s_0 = a$ and $s_n = b$. Hence

$$\sum_{j=1}^{n} |\psi_2(t_j) - \psi_2(t_{j-1})| \quad < \quad 2\varepsilon + \sum_{j=1}^{n} |\psi(s_j) - \psi(s_{j-1})| \quad \le \quad 2\varepsilon + \mathrm{Var}(\psi).$$

Thus $\mathrm{Var}(\psi_2) \leq 2\varepsilon + \mathrm{Var}(\psi)$; now let $\varepsilon \downarrow 0$. This proves $\mathrm{Var}(\psi_2) \leq \mathrm{Var}(\psi)$, and therefore $\psi_2 \in NBV([a,b], X)$.

Since ψ is continuous except at at most countably many points, the function $\psi_1 = \psi - \psi_2$ belongs to $C[a,b]^{\perp}$; that is clear from the second definition. This completes the proof of the theorem.

Monotone Convergence Theorem

24.29. Monotone Convergence Theorem for Henstock-Stieltjes integrals. Let $f_1, f_2, f_3, \ldots$ be functions from $[a,b]$ into $[0,+\infty)$. For each $t \in [a,b]$ assume that $0 \leq f_1(t) \leq f_2(t) \leq f_3(t) \leq \cdots$ and that the sequence $(f_k(t))$ converges to a finite limit $f(t)$. Let $\varphi : [a,b] \to \mathbb{R}$ be an increasing function. Assume each f_k is φ-integrable, and $\sup_k \int_a^b f_k \, d\varphi < \infty$. Then f is φ-integrable, and $\int_a^b f_k \, d\varphi \;\to\; \int_a^b f \, d\varphi$ as $k \to \infty$.

Remark. We shall use both this theorem and the analogous theorem for integrals over σ-algebras (given in 21.38(ii)) when we prove in 24.36 that the two kinds of integrals are equivalent.

Proof of theorem (following DePree and Swartz [1988]). Let $A = \lim_{k\to\infty} \int_a^b f_k \, d\varphi$; we are to show that the Henstock-Stieltjes integral $\int_a^b f \, d\varphi$ exists and equals A. Let any $\varepsilon > 0$ be given; we are to find a gauge δ such that every δ-fine tagged division T satisfies $|A - \Sigma[f, T, \varphi]| < \varepsilon$.

From our hypotheses, we can easily see that:

- There is some integer ν such that $0 \leq A - \int_a^b f_\nu \, d\varphi < \varepsilon/4$.
- For each $t \in [a,b]$, there is some positive integer $i(t)$ such that $0 \leq f(t) - f_{i(t)}(t) < \varepsilon/4[\varphi(b) - \varphi(a)]$.
- For each positive integer k, there is some gauge γ_k such that whenever S is a γ_k-fine tagged division of $[a,b]$, then $|\Sigma[f_k, S, \varphi] - \int_a^b f_k \, d\varphi| < 2^{-k-1}\varepsilon$.

Now let $\mu(t) = \max\{\nu, i(t)\}$, and then define a gauge δ by taking $\delta(t) = \gamma_{\mu(t)}(t)$. Let $T = (n, t_j, \tau_j)$ be any δ-fine tagged division; we shall show that $|\Sigma[f, T, \varphi] - A| < \varepsilon$. We may write $\Sigma[f, T, \varphi] - A = e_1 + e_2 + e_3$, with the error decomposed into these three pieces:

$$\begin{aligned}
e_1 &= \sum_{j=1}^{n} \left[\varphi(t_j) - \varphi(t_{j-1})\right] \left[f(\tau_j) - f_{\mu(\tau_j)}(\tau_j)\right], \\
e_2 &= \sum_{j=1}^{n} \left[\left[\varphi(t_j) - \varphi(t_{j-1})\right] f_{\mu(\tau_j)}(\tau_j) - \int_{t_{j-1}}^{t_j} f_{\mu(\tau_j)} \, d\varphi \right], \\
e_3 &= \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} f_{\mu(\tau_j)} \, d\varphi \;-\; A.
\end{aligned}$$

To estimate e_1, observe that $\mu(\tau_j) \geq i(\tau_j)$, so $0 \leq f(\tau_j) - f_{\mu(\tau_j)}(\tau_j) < \varepsilon/4[\varphi(b) - \varphi(a)]$; hence $|e_1| < \varepsilon/4$. To estimate e_2, temporarily fix any positive integer k. Define the set $J_k = \{j : \mu(\tau_j) = k\}$. For each $j \in J_k$, we have $\delta(\tau_j) = \gamma_k(\tau_j)$; hence $t_j - t_{j-1} < \gamma_k(\tau_j)$. By the Henstock-Saks Lemma (24.23) and our choice of γ_k,

$$\left| \sum_{j \in J_k} \left\{ [\varphi(t_j) - \varphi(t_{j-1})]\, f_k(\tau_j) - \int_{t_{j-1}}^{t_j} f_k\, d\varphi \right\} \right| \quad < \quad 2^{-k-1}\varepsilon.$$

The sets $J_1, J_2, J_3, \ldots$ form a partition of the set $\{1, 2, 3, \ldots, n\}$, and therefore

$$|e_2| \;=\; \left| \sum_{k=1}^{\infty} \sum_{j \in J_k} \left\{ [\varphi(t_j) - \varphi(t_{j-1})]\, f_k(\tau_j) - \int_{t_{j-1}}^{t_j} f_k\, d\varphi \right\} \right| \;<\; \sum_{k=1}^{\infty} 2^{-k-1}\varepsilon \;=\; \frac{\varepsilon}{2}\,.$$

Finally, to estimate e_3, let $p = \max\{\mu(\tau_1), \mu(\tau_2), \ldots, \mu(\tau_n)\}$. Then $\nu \leq \mu(\tau_j) \leq p$ for all j, hence

$$\int_{t_{j-1}}^{t_j} f_\nu\, d\varphi \quad \leq \quad \int_{t_{j-1}}^{t_j} f_{\mu(\tau_j)}\, d\varphi \quad \leq \quad \int_{t_{j-1}}^{t_j} f_p\, d\varphi,$$

and summing over j yields

$$A - \frac{\varepsilon}{4} \quad < \quad \int_a^b f_\nu\, d\varphi \;\leq\; \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} f_{\mu(\tau_j)}\, d\varphi \;\leq\; \int_a^b f_p\, d\varphi \;\leq\; A.$$

Therefore $|e_3| < \varepsilon/4$, which completes the proof.

24.30. Corollaries. Let $\varphi : [a, b] \to \mathbb{R}$ be an increasing function. Then:

a. Interchange of limits (theorem of Levi type). Suppose $g_1, g_2, g_3, \ldots : [a, b] \to [0, +\infty)$ are φ-integrable, and $\sum_{j=1}^{\infty} g_j(t)$ is finite for each t. Then $\sum_j \int_a^b g_j\, d\varphi$ is finite if and only if $\sum_j g_j$ is φ-integrable, in which case

$$\sum_{j=1}^{\infty} \int_a^b g_j\, d\varphi \quad = \quad \int_a^b \left\{ \sum_{j=1}^{\infty} g_j(t) \right\} d\varphi(t).$$

Hint: This is just a reformulation of the Monotone Convergence Theorem (24.29), with $g_j = f_j - f_{j-1}$ and $f_0 = 0$.

b. If $S \subseteq [a, b]$ is a union of countably many intervals, then its characteristic function 1_S is φ-integrable. In particular, by 15.37.d, we see that the characteristic function of any open subset of $[a, b]$ is φ-integrable.

ABSOLUTE INTEGRABILITY

24.31. *Notation.* Let $(X, \| \ \|)$ be a Banach space, and let $\varphi : [a, b] \to \mathbb{R}$ be an increasing function. We shall say that a function $f : [a, b] \to X$ is **φ-integrable** if the Henstock-Stieltjes integral $\int_a^b f\, d\varphi$ exists. We shall say that $f : [a, b] \to X$ is **absolutely φ-integrable** if both the Henstock-Stieltjes integrals $\int_a^b f(t)\, d\varphi(t)$ and $\int_a^b \|f(t)\|\, d\varphi(t)$ exist.

Recall from 19.21 and 22.19 the definition of the variation of a function.

Theorem. Let $(X, \| \ \|)$ be a Banach space. Let $\varphi : [a, b] \to \mathbb{R}$ be an increasing function. Suppose $f : [a, b] \to X$ is φ-integrable and has separable range, and let $F(t) = \int_a^t f(s)\, d\varphi(s)$ be its indefinite integral. Then F has bounded variation if and only if f is absolutely φ-integrable, in which case $\mathrm{Var}(F, [a, b]) = \int_a^b \|f\|\, d\varphi$.

Proof. If f is absolutely φ-integrable, then

$$\sum_{j=1}^{n} \|F(t_j) - F(t_{j-1})\| \quad = \quad \sum_{j=1}^{n} \left\| \int_{t_{j-1}}^{t_j} f\, d\varphi \right\| \quad \leq \quad \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} \|f(s)\|\, d\varphi(s)$$

and thus $\mathrm{Var}(F, [a, b]) \leq \int_a^b \|f\| d\varphi$.

Conversely, suppose that $v = \mathrm{Var}(F, [a, b]) < \infty$; we shall show that the Henstock-Stieltjes integral $\int_a^b \|f\| d\varphi$ exists and equals v. We prove this first in the case where X is finite-dimensional; then we shall use that case to prove the general case.

Let any $\varepsilon > 0$ be given. We are to construct a gauge δ such that whenever $S = (m, s_i, \sigma_i)$ is a δ-fine tagged division of $[a, b]$, then $|v - \Sigma\,[\|f\|, S, \varphi]| \leq \varepsilon$. Let κ be the constant corresponding to the norm $\| \ \|$ in the Henstock-Saks Lemma (24.23(ii)). Since f is Henstock integrable, we may choose some gauge γ_1 such that whenever $T = (n, t_j, \tau_j)$ is a γ_1-fine tagged division of $[a, b]$, then $\left\| \int_a^b f d\varphi - \Sigma[f, T, \varphi] \right\| < \varepsilon/2\kappa$. By the definition of v, there is some division $a = q_0 < q_1 < q_2 < \cdots < q_p = b$ of the interval $[a, b]$ such that $\sum_{k=1}^{p} \left\| \int_{q_{k-1}}^{q_k} f d\varphi \right\| > v - \frac{1}{2}\varepsilon$; this division will remain fixed throughout the remainder of the proof. Let $Q = \{q_0, q_1, \ldots, q_p\}$. Define a gauge γ as in 24.11. We shall show that the gauge $\delta = \min\{\gamma, \gamma_1\}$ has the required properties.

Let $S = (m, s_i, \sigma_i)$ be any δ-fine tagged division of $[a, b]$; we are to show that $|v - \Sigma[|f|, S, \varphi]| < \varepsilon$. Construct an auxiliary tagged division $T = (2m, t_j, \tau_j)$ as in 24.12. By 24.11 we have $Q \subseteq \{\sigma_1, \sigma_2, \ldots, \sigma_m\} = \{\tau_1, \tau_2, \ldots, \tau_{2m}\} \subseteq \{t_0, t_1, t_2, \ldots, t_{2m}\}$; hence

$$v - \frac{1}{2}\varepsilon \quad < \quad \sum_{k=1}^{p} \left\| \int_{q_{k-1}}^{q_k} f\, d\varphi \right\| \quad \leq \quad \sum_{j=1}^{2m} \left\| \int_{t_{j-1}}^{t_j} f\, d\varphi \right\| \quad \leq \quad v$$

and therefore $\left| v - \sum_{j=1}^{2m} \| \int_{t_{j-1}}^{t_j} f\, d\varphi \| \right| < \frac{1}{2}\varepsilon$. The tagged division T is δ-fine, hence γ_1-fine, so we may apply the Henstock-Saks Lemma, which yields the first inequality in the following string of inequalities:

$$\begin{array}{rcl}\frac{\varepsilon}{2} & \geq & \displaystyle\sum_{j=1}^{2m}\left\|\int_{t_{j-1}}^{t_j} f\,d\varphi - f(\tau_j)\,[\varphi(t_j)-\varphi(t_{j-1})]\right\| \\ & \geq & \displaystyle\left|\sum_{j=1}^{2m}\left\{\left\|\int_{t_{j-1}}^{t_j} f\,d\varphi\right\| - \|f(\tau_j)\|\,[\varphi(t_j)-\varphi(t_{j-1})]\right\}\right| \\ & = & \displaystyle\left|\sum_{j=1}^{2m}\left\|\int_{t_{j-1}}^{t_j} f\,d\varphi\right\| - \Sigma[\|f\|,T,\varphi]\right| \quad > \quad \Big|v - \Sigma[\|f(\cdot)\|,T,\varphi]\Big| - \frac{1}{2}\varepsilon.\end{array}$$

This completes the proof for the finite-dimensional case.

For the general case, we may reason as follows: The Banach space X and its dual X^* will have norms both denoted by $\|\ \ \|$. By assumption, the range of f is separable. By 23.24, there is some sequence (λ_j) in X^* such that $\|\lambda_j\| = 1$ for all j and $\|u\| = \sup_j |\lambda_j(u)|$ for every $u \in \mathrm{Range}(f)$.

Temporarily fix any positive integer m, and let $\mathbb{R}^m$ be equipped by the sup norm $\|\ \ \|_\infty$, defined as in 22.11. Define a function $f_m : [a,b] \to \mathbb{R}^m$ by

$$f_m(t) \quad = \quad \Big(\lambda_1 f(t),\ \lambda_2 f(t),\ \ldots,\ \lambda_m f(t)\Big).$$

The function f_m is φ-integrable in $\mathbb{R}^m$, since we can compute the integrals componentwise. For any $s,t \in [a,b]$, we have $\|f_m(t) - f_m(s)\|_\infty = \max_{1\leq j\leq m} |\lambda_j(f(t)-f(s))| \leq \|f(t) - f(s)\|$; from this it follows that $\mathrm{Var}(f_m,[a,b]) \leq \mathrm{Var}(f,[a,b]) = v$. By the case we have already proved in finite dimensions, the integral $\int_a^b \|f_m\|_\infty d\varphi$ exists and equals $\mathrm{Var}(f_m,[a,b])$; thus it is bounded above by v.

As $m \to \infty$, the function $\|f_m(t)\|_\infty = \max\{|\lambda_1 f(t)|, |\lambda_2 f(t)|, \ldots, |\lambda_m f(t)|\}$ increases to $\|f(t)\|$, by our choice of the λ_j's. By the Monotone Convergence Theorem 24.29, the integral $\int_a^b \|f\| d\varphi$ exists and is the limit of the integrals $\int_a^b \|f_m\|_\infty d\varphi$; hence $\int_a^b \|f\| d\varphi \leq v$. On the other hand, $v \leq \int_a^b \|f\| d\varphi$ as we showed at the beginning of this proof. Thus $v = \int_a^b \|f\| d\varphi$, completing the proof.

24.32. Corollaries. Let $\varphi : [a,b] \to \mathbb{R}$ be an increasing function. Then:

a. Suppose that $f : [a,b] \to X$ and $g : [a,b] \to \mathbb{R}$ are φ-integrable, f has separable range, and $\|f(t)\| \leq g(t)$ for all t. Then $\|f(\cdot)\|$ is φ-integrable — i.e., f is absolutely integrable. Moreover, $\left\|\int_a^b f d\varphi\right\| \leq \int_a^b \|f\| d\varphi \leq \int_a^b g d\varphi$.

Hint: $F(t) = \int_a^t f d\varphi$ has bounded variation, since $\|F(t) - F(s)\| = \left\|\int_s^t f d\varphi\right\| \leq \int_s^t g d\varphi$ by 24.16.a.

b. The lattice of absolutely integrable functions. The real-valued absolutely φ-integrable functions on $[a,b]$ form a vector lattice, with $(f \vee g)(t) = \max\{f(t), g(t)\}$ and $(f \wedge g)(t) = \min\{f(t), g(t)\}$. *Hints*:

$$f \vee g \;=\; \frac{1}{2}\,[f + g + |f-g|] \qquad \text{and} \qquad f \wedge g \;=\; \frac{1}{2}\,[f + g - |f-g|]\,.$$

Also, $|f(t) + g(t)| \le |f(t)| + |g(t)|$.

c. Suppose $f_1, f_2, f_3, \ldots : [a, b] \to [0, +\infty)$ are φ-integrable, and $g(t) = \sup_n f_n(t)$ exists for each t. If $g \le h$ for some φ-integrable function $h : [a, b] \to [0, +\infty)$, then g is φ-integrable, too.

Hint: The functions $F_n = \max\{f_1, f_2, \ldots, f_n\}$ are φ-integrable, by 24.32.b. Apply the Monotone Convergence Theorem to the sequence $F_1, F_2, F_3, \ldots$.

d. Fatou's Lemma. Let $f_1, f_2, f_3, \ldots : [a, b] \to [0, +\infty)$ be φ-integrable functions. Suppose that $\liminf_{n\to\infty} \int_a^b f_n \, d\varphi < \infty$. Then $\liminf_{n\to\infty} f_n(t)$ is a φ-integrable function of t, and $\int \liminf_{n\to\infty} f_n \, d\varphi \le \liminf_{n\to\infty} \int f_n \, d\varphi$.

Hint: Let $g_k = \inf\{f_k, f_{k+1}, f_{k+2}, \ldots\}$. Apply the Monotone Convergence Theorem to the increasing sequence $g_1, g_2, g_3, \ldots$.

Henstock and Lebesgue Integrals

24.33. *Remark.* The proofs in this subchapter, particularly the first one, are rather long and technical. In a first reading, beginners may find it helpful to read the statements of results but skip the proofs.

24.34. Technical lemma on regularity. Suppose that $\varphi : [a, b] \to \mathbb{R}$ is an increasing function, and $f : [a, b] \to [0, +\infty)$ is a function for which the Henstock-Stieltjes integral $\int_a^b f \, d\varphi$ exists. Let any number $\varepsilon > 0$ be given, and let $E = \{t \in [a, b] : f(t) \ge 1\}$. Then there exists an open set $G \supseteq E$ such that $\int_a^b 1_G \, d\varphi \le \varepsilon + \int_a^b f \, d\varphi$.

(Here "open" refers to the relative topology on $[a, b]$; for instance, the set $[a, b]$ is open. We know that the Henstock-Stieltjes integral $\int_a^b 1_G \, d\varphi$ exists by 24.30.b.)

Proof (modified from McLeod [1980]). It will be simplest to treat the point a separately from the rest of the interval. Let $1_{\{a\}}$ be the characteristic function of the singleton $\{a\}$, and let $1_{(a,b]}$ be the characteristic function of the remainder of the interval. It suffices to prove the lemma for each of the two functions $f1_{\{a\}}$ and $f1_{(a,b]}$ — i.e., it suffices to prove the lemma in the two cases where f vanishes outside $\{a\}$ and where f vanishes at a.

First, suppose f vanishes outside $\{a\}$. The theorem is trivial if $E = \varnothing$, so we may assume $f(a) \ge 1$. It is easy to compute $\int_a^b f \, d\varphi = f(a) [\varphi(a+) - \varphi(a)]$. Choose $\alpha > a$ small enough so that $\varphi(\alpha) - \varphi(a+) < \varepsilon$. Then $G = [a, \alpha)$ is an open set containing $\{a\}$, and

$$\int_a^b 1_G \, d\varphi \;=\; \varphi(\alpha-) - \varphi(a) \;\le\; \varphi(\alpha) - \varphi(a) \;\le\; \varepsilon + \frac{1}{f(a)} \int_a^b f \, d\varphi \;\le\; \varepsilon + \int_a^b f \, d\varphi$$

as required.

For the remainder of the proof we may assume $f(a) = 0$; hence $a \notin E$. We may define $\varphi(t) = \varphi(b)$ for all $t > b$; thus φ is right continuous at b. By 15.21.c, we know that φ is right continuous at all but countably many points of $(a, b]$; hence φ is right continuous at all points in a dense subset of $(a, b]$. We now form partitions $P_1, P_2, P_3, \ldots$ of $[a, b]$ into subintervals,

$$P_j \quad : \quad a = p_j^0 < p_j^1 < p_j^2 < \cdots < p_j^{Q(j)-1} < p_j^{Q(j)} = b$$

satisfying these requirements:

- P_{j+1} is a refinement of P_j (that is, P_j is a subsequence of P_{j+1}),
- φ is right continuous at each of the points $p_j^1, p_j^2, \ldots, p_j^{Q(j)}$,
- $\max\left\{p_j^1 - p_j^0,\, p_j^2 - p_j^1,\, \ldots,\, p_j^{Q(j)} - p_j^{Q(j)-1}\right\} < 2^{-j}(b-a)$.

(Although the particular method used for satisfying these conditions is not important, here is one way that it can be done, with $Q(j) = 3^j$: Each subinterval in P_j can be subdivided into 3 subintervals in P_{j+1}, where each of those subintervals has length less than half the length of the P_j subinterval and φ is right continuous at the endpoints of each subinterval.)

Since the Henstock-Stieltjes integral $\int_a^b f\, d\varphi$ exists, there is some gauge δ such that whenever T is a δ-fine tagged division, then $|\Sigma[f, T, \varphi] - \int_a^b f\, d\varphi| < \varepsilon/2$. We now use δ to *select* a sequence $((a_k, b_k], \sigma_k)_{k\in\mathbb{N}}$ of tagged intervals, chosen in stages by the following procedure. For the first stage, there are no previously selected intervals. For $j \geq 1$, in the jth stage we select all intervals $(a_k, b_k]$ that meet the following criteria:

(1) $(a_k, b_k]$ is one of the jth stage intervals $(p_j^0, p_j^1]$, $\ldots$, $(p_j^{Q(j)-1}, p_j^{Q(j)}]$ that make up partition P_j;

(2) $(a_k, b_k]$ is not contained in any previously selected interval; and

(3) there is at least one point $\sigma_k \in (a_k, b_k] \cap E$ that satisfies $b_k - a_k < \delta(\sigma_k)$.

For each interval $(a_k, b_k]$ selected in this fashion, there may be more than one point that is suitable for use as σ_k, but we choose one particular value for σ_k. Note that the resulting intervals $(a_k, b_k]$ are disjoint; also note that $f(\sigma_k) \geq 1$.

We claim that $\bigcup_{k=1}^{\infty}(a_k, b_k] \supseteq E$. To see this, fix any $z \in E$. For some integer j sufficiently large, we have $2^{-j}(b-a) < \delta(z)$. Since $a \notin E$, we have $z \neq a$, so one of the intervals $(p_j^{u-1}, p_j^u]$ (for $u = 1, 2, \ldots, Q(j)$) must contain z, and that interval $(p_j^{u-1}, p_j^u]$ must have length less than $2^{-j}(b-a)$. That interval must be selected in the jth stage, if it is not contained in an interval that was already selected in an earlier stage. Thus $z \in \bigcup_{k=1}^{\infty}(a_k, b_k]$, proving our claim.

Next we show that $\sum_{k=1}^{\infty}[\varphi(b_k) - \varphi(a_k)] \leq \frac{1}{2}\varepsilon + \int_a^b f\, d\varphi$. To see that, fix any positive integer K; it suffices to show $\sum_{k=1}^{K}[\varphi(b_k) - \varphi(a_k)] \leq \frac{1}{2}\varepsilon + \int_a^b f\, d\varphi$. The intervals $(a_1, b_1), (a_2, b_2), \ldots, (a_K, b_K)$ are disjoint, and the complement of their union is equal to the union of finitely many subintervals of $[a, b]$. Apply 24.10 to obtain a δ-fine tagged division of each of those subintervals. Putting all the subintervals together, we obtain a δ-fine tagged division $T = (m, t_v, \tau_v)$ of $[a, b]$, in which $([a_k, b_k], \sigma_k)_{1 \leq k \leq K}$ comprise some of the subintervals. Since $f(\sigma_k) \geq 1$, we obtain

$$\sum_{k=1}^{K} [\varphi(b_k) - \varphi(a_k)] \quad \leq \quad \sum_{v=1}^{m} [\varphi(t_v) - \varphi(t_{v-1})] \, f(\tau_v) \quad \leq \quad \frac{\varepsilon}{2} + \int_a^b f \, d\varphi$$

by our choice of δ. This proves the desired inequality.

For each k, the function φ is right continuous at b_k. Hence we may choose some $c_k > b_k$ with c_k close enough to b_k so that $\varphi(c_k) < \varphi(b_k) + 2^{-k-1}\varepsilon$; then $\sum_{k=1}^{\infty} [\varphi(c_k) - \varphi(b_k)] < \frac{1}{2}\varepsilon$. The intervals $(a_1, c_1), (a_2, c_2), (a_3, c_3), \ldots$ are not necessarily disjoint, but their union is an open set G that contains E. It follows from 24.30.a that the Henstock-Stieltjes integral $\int_a^b 1_G \, d\varphi$ exists and is less than or equal to

$$\sum_{k=1}^{\infty} \int_a^b 1_{(a_k, c_k)} \, d\varphi \; = \; \sum_{k=1}^{\infty} [\varphi(c_k-) - \varphi(a_k+)] \; \leq \; \sum_{k=1}^{\infty} [\varphi(c_k) - \varphi(a_k)] \; \leq \; \varepsilon + \int_a^b f \, d\varphi.$$

This completes the proof of the lemma.

24.35. Theorem on measures. One-dimensional Borel-Lebesgue measure and Lebesgue measure (defined as in 21.19) exist. Furthermore, a set $E \subseteq [a, b]$ is Lebesgue measurable if and only if its characteristic function 1_E is Henstock integrable, in which case the Lebesgue measure $\mu(E)$ is equal to the Henstock integral $\int_a^b 1_E(t) \, dt$.

More generally, let $\varphi : [a, b] \to \mathbb{R}$ be an increasing function. Let $\mathcal{K}$ be the collection of all sets $S \subseteq [a, b]$ for which the Henstock-Stieltjes integral

$$\mu_\varphi(S) \quad = \quad \int_a^b 1_S(t) d\varphi(t)$$

exists. Then $\mathcal{K}$ is a σ-algebra that includes $\mathcal{B} = \{\text{Borel sets}\}$, and $([a, b], \mathcal{K}, \mu_\varphi)$ is a complete measure space (as defined in 21.16); in fact, it is the completion of $([a, b], \mathcal{B}, \mu_\varphi)$. Furthermore, the measure space $([a, b], \mathcal{K}, \mu_\varphi)$ is *regular*, in this sense: A set $S \subseteq [a, b]$ belongs to $\mathcal{K}$ if and only if there exist an F_σ set A and a G_δ set B such that $A \subseteq S \subseteq B$ and $\mu_\varphi(B \setminus A) = 0$.

Every positive finite measure μ on the Borel sets $\mathcal{B}$ in $[a, b]$ is of the form μ_φ for some increasing function φ.

Remarks. Here G_δ is with respect to the relative topology of $[a, b]$; thus the set $[a, b]$ itself is considered to be open. Note that any measurable set is the union of an F_σ and a set of measure 0; contrast that with 20.22.

Proof of theorem. By the definition of $\mathcal{K}$, it is clear that if $K_1, K_2 \in \mathcal{K}$ and $K_1 \supseteq K_2$, then $K_1 \setminus K_2 \in \mathcal{K}$. In particular, the complement in $[a, b]$ of any member of $\mathcal{K}$ also belongs to $\mathcal{K}$. By 24.30.a, we see that the union of countably many disjoint members of $\mathcal{K}$ is also a member of $\mathcal{K}$; hence $\mathcal{K}$ is a monotone class (defined in 5.29). Also by 24.30.a, the mapping $\mu_\varphi : S \mapsto \int_0^1 1_S d\varphi$ is countably additive on $\mathcal{K}$. (We do not yet assert that $\mathcal{K}$ is a σ-algebra; that fact will be established much later in this proof.)

Let $\mathcal{F}$ be the algebra of all unions of finitely many subintervals of $[0,1]$. (We shall count the empty set and any singleton as subintervals.) Then $\mathcal{B}$, the σ-algebra of Borel sets, is the σ-algebra generated by $\mathcal{F}$. It is clear that $\mathcal{K} \supseteq \mathcal{F}$. By the Monotone Class Theorem 5.29, $\mathcal{K} \supseteq \mathcal{B}$. The restriction of μ_φ to $\mathcal{B}$ is a measure since μ_φ is countably additive on $\mathcal{K}$.

If $E \in \mathcal{K}$ and $n \in \mathbb{N}$, then (by 24.34 with $f = 1_E$ and $\varepsilon = \frac{1}{n}$) there is some open set $G_n \supseteq E$ with $\mu_\varphi(G_n) \leq \frac{1}{n} + \mu_\varphi(E)$. Then $B = \bigcap_{n=1}^\infty G_n$ is a Borel set (in fact, a G_δ set) with $B \supseteq E$ and $\mu_\varphi(B) = \mu_\varphi(E)$, hence $\mu_\varphi(B \setminus E) = 0$.

Let $\mathcal{N} = \{N \subseteq [a,b] : N \subseteq B \text{ for some Borel set } B \text{ with } \mu_\varphi(B) = 0\}$. The completion of $\mathcal{B}$ is the σ-algebra $\mathcal{B} \triangle \mathcal{N}$, defined as in 21.16. The inclusion $\mathcal{B} \triangle \mathcal{N} \subseteq \mathcal{K}$ is an easy exercise, using the definition of $\mathcal{K}$; we omit the details. To prove $\mathcal{K} \subseteq \mathcal{B} \triangle \mathcal{N}$, let any $E \in \mathcal{K}$ be given. Form a Borel set B as in the preceding paragraph. Then $B \setminus E$ is not necessarily a Borel set, but $B \setminus E$ is a member of $\mathcal{K}$ that has $\mu_\varphi(B \setminus E) = 0$. By the results of the last paragraph, there is some Borel set B' containing $B \setminus E$ with $\mu_\varphi(B') = 0$. Thus $B \setminus E \in \mathcal{N}$; hence $E = B \setminus (B \setminus E) \in \mathcal{B} \triangle \mathcal{N}$. This proves $\mathcal{K}$ is equal to the σ-algebra $\mathcal{B} \triangle \mathcal{N}$.

The result about F_σ sets is obtained by passing to complements.

If μ is a positive finite measure on the Borel subsets of $[a,b]$, then define an increasing right continuous function $\varphi : [a,b] \to \mathbb{R}$ by $\varphi(t) = \mu([a,t])$, and use it to define a measure μ_φ as above. We obtain $\mu_\varphi([a,t]) = \varphi(t)$ and $\mu_\varphi([a,t)) = \varphi(t-) = \mu([a,t))$. Thus the measures μ_φ and μ agree on $\mathcal{F}$; by the Monotone Class Theorem they agree on $\mathcal{B}$.

24.36. Theorem on integrals. Let $(X, |\ \ |)$ be a Banach space, and let $f : [a,b] \to X$ be some function. Then $f \in L^1([a,b], X)$ if and only if f is absolutely Henstock integrable and almost separably valued. Moreover, when those two conditions are satisfied, then the Bochner-Lebesgue and Henstock integrals $\int_a^b f(t)dt$ are equal.

More generally, suppose $(X, |\ \ |)$ is a Banach space, $f : [a,b] \to X$ is some function, and $\varphi : [a,b] \to \mathbb{R}$ is an increasing function. Define a measure μ_φ on the σ-algebra $\mathcal{K}$ of subsets of $[a,b]$, as in 24.35. Then the following two conditions are equivalent.

(A) $f \in L^1(\mu_\varphi, X)$.

(B) f is absolutely φ-integrable (i.e., the Henstock-Stieltjes integrals $\int_a^b f\, d\varphi$ and $\int_a^b |f(\cdot)|\, d\varphi$ both exist, in X and in $\mathbb{R}$ respectively), and f is almost separably valued (defined as in 21.17).

Moreover, when conditions (A) and (B) are satisfied, then the Bochner integral $\int_{[a,b]} f\, d\mu_\varphi$ is equal to the Henstock-Stieltjes integral $\int_a^b f\, d\varphi$.

Remarks. Of course, the separability condition is satisfied trivially and can be omitted from mention if X itself is separable — in particular, if X is finite-dimensional. We emphasize that measurability of f (from $\mathcal{K}$ to the Borel subsets of X) is an explicit part of condition (A), but not of condition (B). In fact, most of the proof is devoted to showing that condition (B) implies the measurability of f.

Proof of (A) $\Rightarrow$ (B) *and equality of the integrals.*

(i) We first prove (A) $\Rightarrow$ (B) in the case where f is finitely valued. This case is easy; the details are left as an exercise.

(ii) We next prove (A) $\Rightarrow$ (B) in the case where $X = \mathbb{R}$ and $f \geq 0$. Then there exists a sequence of finitely valued measurable functions f_n, increasing pointwise to f, by 21.5. We have $\int f_n d\mu = \int f_n(t)d\varphi(t)$ for each finitely valued function f_n. The numbers $\int f_n d\mu$ increase to $\int f d\mu$ by the Monotone Convergence Theorem for Lebesgue Integrals (21.38(ii)). Hence the numbers $\int f_n(t)d\varphi(t)$ increase to $\int f(t)d\varphi(t)$ by the Monotone Convergence Theorem for Henstock-Stieltjes Integrals (24.29).

(iii) Finally, we prove (A) $\Rightarrow$ (B) in full generality: Let any $f \in L^1(\mu_\varphi, X)$ be given. By 22.30.b, there exists a sequence of finitely valued functions f_n converging in $L^1(\mu_\varphi, X)$ to f. By 22.31(ii), passing to a subsequence, we may assume that $f_n \to f$ pointwise and that the sequence is dominated by some nonnegative function $h \in L^1(\mu_\varphi, \mathbb{R})$. By the remarks of the preceding paragraph, h is φ-integrable. Define $\nu(S) = \int_S h\, d\mu_\varphi$; this is a measure on $\mathcal{K}$ by 21.38(i).

Let any $\varepsilon > 0$ be given, and let $\delta = \varepsilon/6[\varphi(b) - \varphi(a)]$. By Egorov's Theorem (21.32), we may partition $[a, b]$ into some disjoint sets $J, K \in \mathcal{K}$ such that $\nu(J) < \varepsilon/6$ and $f_n \to f$ uniformly on K. Choose n large enough so that $\max_{t\in K} \|f_n(t) - f(t)\| < \delta$. The function $f_n \cdot 1_K$ is also finitely valued, hence φ-integrable. Since $f_n \cdot 1_K$ and $h \cdot 1_J$ are φ-integrable, we may choose a gauge γ such that whenever S is a γ-fine tagged division of $[a, b]$, then

$$\left\|\Sigma[f_n 1_K, S, \varphi] - \int_K f_n d\varphi\right\| < \frac{\varepsilon}{6} \qquad \text{and} \qquad \left\|\Sigma[h1_J, S, \varphi] - \int_J hd\varphi\right\| < \frac{\varepsilon}{6}.$$

Now consider any γ-fine tagged division S. We estimate

$$\begin{array}{rl}
\|\Sigma[f, S, \varphi] - \Sigma[f1_K, S, \varphi]\| & \leq\ \Sigma[h1_J, S, \varphi]\ \leq\ \frac{\varepsilon}{6} + \int_J hd\varphi\ =\ \frac{\varepsilon}{6} + \nu(J)\ <\ \frac{2\varepsilon}{6}, \\
\|\Sigma[f1_K, S, \varphi] - \Sigma[f_n 1_K, S, \varphi]\| & \leq\ \Sigma[\delta, S, \varphi]\ =\ [\varphi(b) - \varphi(a)]\delta\ =\ \frac{\varepsilon}{6}, \\
\|\Sigma[f_n 1_K, S, \varphi] - \int_K f_n d\mu\| & <\ \frac{\varepsilon}{6}\quad \text{(by our choice of } \gamma), \\
\|\int_K f_n d\mu - \int_K f d\mu\| & \leq\ \int_K \delta d\mu\ \leq\ \mu([a, b])\delta\ =\ [\varphi(b) - \varphi(a)]\delta\ =\ \frac{\varepsilon}{6}, \\
\|\int_K f d\mu - \int_{[a,b]} f d\mu\| & =\ \|\int_J f d\mu\|\ \leq\ \int_J hd\mu\ =\ \nu(J)\ <\ \frac{\varepsilon}{6}.
\end{array}$$

Putting all these ingredients together, we arrive at $\|\Sigma[f, S, \varphi] - \int_{[a,b]} f d\mu\| < \varepsilon$. This proves that the Henstock-Stieltjes integral $\int f d\varphi$ exists and equals $\int f d\mu$.

Proof of (B) $\Rightarrow$ (A). Again we proceed through several cases.

(i) We first prove (B) $\Rightarrow$ (A) in the case where $X = \mathbb{R}$, $f \geq 0$, and $\int_a^b f\, d\varphi = 0$. Let any numbers $r, \varepsilon > 0$ be given. By Lemma 24.34, there is an open set G containing $\{t \in [a, b] : \frac{1}{r} f(t) \geq 1\}$, such that $\mu_\varphi(G) \leq \varepsilon$. It follows easily that $\{t \in [a, b] : f(t) \geq r\}$ is a member of $\mathcal{K}$ with measure 0. Hence f is measurable, $f = 0$ almost everywhere, and $f \in L^1(\mu_\varphi)$.

(ii) We next prove (B) $\Rightarrow$ (A) in the case where $X = \mathbb{R}$ and $f \geq 0$ (but $\int_a^b f\, d\varphi$ is not necessarily 0). Temporarily fix any number $\varepsilon > 0$. For $n = 0, 1, 2, \ldots$, define the functions

$$u_n(t) \quad = \quad (f(t) - n\varepsilon)^+ \wedge \varepsilon \quad = \quad \begin{cases} 0 & \text{if } f(t) \le n\varepsilon \\ f(t) - n\varepsilon & \text{if } n\varepsilon \le f(t) \le (n+1)\varepsilon \\ \varepsilon & \text{if } (n+1)\varepsilon \le f(t). \end{cases}$$

Then the functions u_n are all absolutely φ-integrable, since the absolutely φ-integrable functions form a vector lattice, as noted in 24.32.b. It is easy to verify that $\sum_{n=0}^{\infty} u_n(t) = f(t)$ for each t; hence $\sum_{n=0}^{\infty} \int_a^b u_n\, d\varphi = \int_a^b f\, d\varphi$ by 24.30.a.

By Lemma 24.34, for each $n \ge 0$ we may choose some open set $G_n \supseteq \{t \in [a,b] : \varepsilon^{-1} u_n(t) \ge 1\} = \{t \in [a,b] : u_n(t) = \varepsilon\}$, satisfying

$$\int_{[a,b]} 1_{G_n} d\mu_\varphi \quad = \quad \int_a^b 1_{G_n}\, d\varphi \quad \le \quad 2^{-n-1} + \frac{1}{\varepsilon} \int_a^b u_n\, d\varphi.$$

Then $g = \varepsilon \sum_{n=0}^{\infty} 1_{G_n}$ is Borel measurable, and $\int_{[a,b]} g\, d\mu_\varphi = \int_a^b g\, d\varphi \le \varepsilon + \int_a^b f\, d\varphi$ by the Levi Theorems 21.39.b and 24.30.a. Note that the sets $H_n = \{t \in [a,b] : n\varepsilon < f(t) < (n+1)\varepsilon\}$ (for $n = 0, 1, 2, 3, \ldots$) are disjoint, and $(1_{G_n} + 1_{H_n})\varepsilon \ge u_n$. Hence, summing over n, we obtain $g + \varepsilon \ge f$.

Our construction of g depended on the choice of ε. Now construct such a function $g = g_k$ for each of the values $\varepsilon = \frac{1}{k}$ (for $k = 1, 2, 3, \ldots$). Thus we obtain functions $g_k \in L^1(\mu_\varphi, \mathbb{R})$ with $g_k + \frac{1}{k} \ge f$ and $\int_{[a,b]} g_k\, d\mu_\varphi \le \frac{k}{+} \int_a^b f\, d\varphi$. Let $h = \liminf_{k\to\infty} g_k$. Then h is Borel measurable, $h \ge f$, and by Fatou's Lemma (21.39.c) we have $\int_a^b h\, d\varphi = \int_{[a,b]} h\, d\mu_\varphi \le \int_a^b f\, d\varphi$. Hence $f - h$ is nonnegative and $\int_a^b (f-h)\, d\varphi = 0$. By the special case discussed earlier in this proof, it follows that $f - h \in L^1(\mu_\varphi)$, and therefore $f \in L^1(\mu_\varphi)$. This completes the proof in the case where $X = \mathbb{R}$ and $f \ge 0$.

(iii) We next prove (B) $\Rightarrow$ (A) in the case where $X = \mathbb{R}$ (but f is not necessarily nonnegative). As we noted in 24.32.b, the absolutely φ-integrable functions form a vector lattice. Hence we may write the Jordan Decomposition $f = f^+ - f^-$ (see 8.42.f). The functions f^+, f^- are absolutely φ-integrable, so the problem is reduced to the previous case.

(iv) Finally, we prove (B) $\Rightarrow$ (A) in general — i.e., where X is any Banach space. Any complex Banach space may be viewed as a real Banach space, by "forgetting" how to multiply vectors by members of $\mathbb{C} \setminus \mathbb{R}$. This has no effect on conditions (A) and (B); hence we may assume the scalar field is $\mathbb{R}$. By assumption, f is almost separably valued, so by changing f on a set of μ_φ-measure 0 we may assume f is separably valued. By assumption, the vector-valued function f is absolutely φ-integrable; hence the real-valued function $|f(\cdot)|$ is also absolutely φ-integrable. By the previous case of this theorem, we know that $|f(\cdot)| \in L^1(\mu_\varphi, \mathbb{R})$.

Temporarily fix any $\lambda \in X^*$. We claim that the function $\lambda \circ f : [a,b] \to \mathbb{R}$ is φ-integrable, with $\int_a^b \lambda(f(\cdot))\, d\varphi = \lambda\left(\int_a^b f\, d\varphi\right)$; indeed, this is clear from the estimate

$$\left| \sum_{i=1}^{m} \lambda(f(\sigma_i))\, [\varphi(s_i) - \varphi(s_{i-1})] - \lambda\left(\int_a^b f\, d\varphi\right) \right|$$

$$\begin{aligned} &= \left|\lambda\left(\sum_{i=1}^{m} f(\sigma_i)\left[\varphi(s_i)-\varphi(s_{i-1})\right] - \int_a^b f\,d\varphi\right)\right| \\ &\le \|\lambda\| \left|\sum_{i=1}^{m} f(\sigma_i)\left[\varphi(s_i)-\varphi(s_{i-1})\right] - \int_a^b f\,d\varphi\right|. \end{aligned}$$

A similar estimate shows that $\int_p^q \lambda(f(\cdot))\,d\varphi = \lambda\left(\int_p^q f\,d\varphi\right)$ for any subinterval $[p,q] \subseteq [a,b]$. Hence for any partition $a = p_0 < p_1 < p_2 < \cdots < p_n = b$ we have

$$\begin{aligned} \sum_{j=1}^{n}\left|\int_{p_{j-1}}^{p_j} \lambda(f(\cdot))\,d\varphi\right| &= \sum_{j=1}^{n}\left|\lambda\left(\int_{p_{j-1}}^{p_j} f\,d\varphi\right)\right| \\ &\le \|\lambda\| \sum_{j=1}^{n}\left|\int_{p_{j-1}}^{p_j} f\,d\varphi\right| \le \|\lambda\| \int_a^b |f(\cdot)|\,d\varphi. \end{aligned}$$

By 24.31, therefore, $\lambda \circ f : [a,b] \to \mathbb{R}$ is absolutely φ-integrable. Apply the previous case (iii); thus $\lambda \circ f \in L^1(\mu_\varphi, \mathbb{R})$ with $\int_{[a,b]} \lambda \circ f\,d\mu_\varphi = \int_a^b \lambda(f(\cdot))d\varphi$.

In particular, $\lambda \circ f$ is measurable from $\mathcal{K}$ to the Borel sets of $\mathbb{R}$. The function f is separably valued and weakly measurable, hence (see 23.25) strongly measurable. Since $|f(\cdot)| \in L^1(\mu_\varphi, \mathbb{R})$ (established earlier in this proof), it follows that $f \in L^1(\mu_\varphi, X)$. The equation $\int_{[a,b]} f\,d\mu_\varphi = \int_a^b f\,d\varphi$ was established when we proved (A) $\Rightarrow$ (B).

24.37. Corollary. Let X be a Banach space with scalar field $\mathbb{F}$ (equal to $\mathbb{R}$ or $\mathbb{C}$). If $\varphi : [a,b] \to \mathbb{F}$ has bounded variation and $f : [a,b] \to X$ is bounded and strongly measurable (from the Borel sets to the Borel sets), then the Henstock-Stieltjes integral $\int_a^b f\,d\varphi$ exists.

24.38. *Remarks.* The Henstock integral can be generalized, though only with some difficulty, to domains more general than an interval $[a,b]$. The Bochner/Lebesgue approach is more powerful, in that it applies easily to a very wide collection of measure spaces $(\Omega, \mathcal{S}, \mu)$.

In certain other respects, however, the Henstock integral is actually more general. Built into the definition of the Bochner/Lebesgue integral are a separability condition and an absolute integrability condition (i.e., not only f but $\|f(\cdot)\|$ must be integrable). These restrictions are not imposed on the Henstock integral; hence we can devise functions that are Henstock integrable but not Bochner/Lebesgue integrable by violating either the separability condition or the absolute integrability condition.

Violations of the separability condition are perhaps contrived and artificial, since all of applied mathematics (all of "the real world") happens in separable Banach spaces, or in separable subspaces of Banach spaces. Violations of the absolute integrability condition are not so contrived, however. A study of the continuous dependence on parameters and asymptotic behavior for solutions to differential equations with rapidly oscillating terms leads to functions very much like the pathological function in 25.20, which is Henstock integrable but not Lebesgue integrable. In fact, it was the study of such solutions to differential equations that led Kurzweil to his independent discovery of the Henstock integral (also known as the Kurzweil integral); for instance, see Kurzweil [1957].

MORE ABOUT LEBESGUE MEASURE

24.39. *Example: meager but full.* The sets with Lebesgue measure 0 and the meager sets form two σ-ideals on $\mathbb{R}$ and thus two different notions of "small" sets. These notions are not directly related; a set may be small in one sense while large in the other sense. That is evident from the following example.

Let (r_j) be an enumeration of the rationals. For $i, j \in \mathbb{N}$, define the open interval $H_{i,j} = (r_i - 2^{-i-j}, r_i + 2^{-i-j})$. Then $G_j = \bigcup_{i=1}^{\infty} H_{i,j}$ is an open dense subset of $\mathbb{R}$, and so $C = \bigcap_{j=1}^{\infty} G_j$ is a comeager set with Lebesgue measure 0. Thus it is "small" with respect to Lebesgue measure, but "large" with respect to Baire category. Its complement has these properties reversed. Note also that C is uncountable since it is not meager.

24.40. *Proposition on regularity of Lebesgue measure.* Let μ denote Lebesgue measure on $\mathbb{R}$. If $S \subseteq \mathbb{R}$ is Lebesgue measurable, then

$$\mu(S) \quad = \quad \sup\{\mu(K) \,:\, K \subseteq S, \quad K \text{ compact}\}.$$

Proof. Let any $\varepsilon > 0$ be given. By 24.35, for each integer $n \in \mathbb{Z}$ we can find some compact set $K_n \subseteq S \cap [n, n+1]$ such that $\mu(K_n) > \mu(S \cap [n, n+1]) - 2^{-|n|-2}\varepsilon$. Any overlap among the K_n's or among the sets $S \cap [n, n+1]$ is contained in the set $\mathbb{Z}$, which has measure 0. Hence for any $N \in \mathbb{N}$, we have $\mu(\bigcup_{|n| \leq N} K_n) = \sum_{|n| \leq N} \mu(K_n) \geq \mu(S \cap [-N, N]) - \varepsilon$. The set $\bigcup_{|n| \leq N} K_n$ is compact, and the numbers $\mu(S \cap [-N, N])$ increase to $\mu(S)$ as $N \to \infty$.

24.41. *Further results for Lebesgue-integrable functions.* Let X be a Banach space. Recall that $L^1([a,b], X)$ means the space $L^1(\mu, X)$ where μ is Lebesgue measure on the Lebesgue measurable subsets of $[a, b]$. Show the following.

a. Continuity of the indefinite integral. Let $I_f(u) = \int_a^u f(t)dt$. Then the map $f \mapsto I_f$ is continuous from $L^1([a,b], X)$ into $L^{\infty}([a,b], X)$.

Hint: First show that the map $f \mapsto I_f$, from $L^1([a,b], X)$ into $L^{\infty}([a,b], X)$, is nonexpansive. Also show that when f is continuous, I_f is continuous. Recall from 22.30.d that the continuous functions are dense in $L^1([a,b], X)$.

b. Riemann-Lebesgue Lemma. If $f \in L^1([a,b], X)$, then $\int_a^b \sin(nt) f(t)dt$ converges to 0 as $n \to \infty$.

Hints: Let $R_n(f) = \int_a^b \sin(nt) f(t)\, dt$. Show that R_n, considered as a mapping from $L^1([a,b], X)$ to the scalar field, is nonexpansive. Prove $\lim_{n \to \infty} R_n(f) = 0$ first when f is a step function. Then recall from 22.30.c that step functions are dense in $L^1([a,b], X)$.

Remark. This result will be generalized in 26.47.

24.42. *Proposition (optional).* Suppose a function $f : \mathbb{R} \to \mathbb{R}$ is additive; i.e.,

$$f(x+y) \;=\; f(x) + f(y) \qquad \text{for all} \quad x, y \in \mathbb{R}.$$

Also suppose f is measurable (from the Lebesgue-measurable sets to the Borel sets). Then $f(x) = xf(1)$ for all x. Thus, any additive, Lebesgue measurable function from $\mathbb{R}$ to itself is continuous. (Compare this with 11.30.c.)

Proof. This proof is based on Hille and Phillips [1957].

It is easy to show that $f(rx) = rf(x)$ for all rational numbers r and all real numbers x; that does not require measurability. In particular, $f(0) = 0$ and $f(-x) = -f(x)$.

Let μ denote Lebesgue measure. We first claim that

$$\text{if} \quad a > 0 \quad \text{and} \quad S = \left\{ s \in [0,a] : f(s) \geq \frac{1}{2} f(a) \right\}, \quad \text{then} \quad \mu(S) \geq \frac{1}{2} a.$$

To see this, let $T = a - S = \{a - s : s \in S\}$; then $\mu(T) = \mu(S)$. Also, $S \cup T = [0,a]$ since $f(s) + f(a-s) = f(a)$. Hence $\mu(S) \geq \frac{1}{2}\mu([0,a]) = \frac{1}{2}a$.

Next we claim that f is bounded above on $[1,2]$. Indeed, suppose not. Then there is a sequence (x_n) in $[1,2]$ with $f(x_n) \to +\infty$. Passing to a subsequence, we may assume $f(x_n) > 2n$. For each n, the Lebesgue-measurable set $S_n = \{s \in [0, x_n] : f(s) \geq n\}$ has measure $\mu(S_n) \geq \frac{1}{2}x_n$; hence for each n the measurable set $M_n = \{s \in [0,2] : f(s) \geq n\}$ has measure $\mu(M_n) \geq \frac{1}{2}$. However, the sets M_n form a decreasing sequence with empty intersection, hence $\lim_{n\to\infty} \mu(M_n) = 0$, a contradiction.

Thus f is bounded above on $[1,2]$. Replacing f with $-f$ (which satisfies the same hypotheses), f is also bounded below on $[1,2]$. It follows easily that f is bounded on each bounded subinterval of $\mathbb{R}$.

Next we show that f is continuous at 0. Indeed, suppose not. Show that there exists a sequence (t_n) converging to 0 in $\mathbb{R}$ with $f(t_n) > \varepsilon$ for some constant $\varepsilon > 0$. Passing to a subsequence, we may assume the t_n's are all positive or all negative. Let us assume they are all positive; the proof is similar in the other case. Passing to a subsequence again, we may assume $t_n < 2^{-n}$. Then the numbers $s_n = t_1 + t_2 + \cdots + t_n$ all take their values in a bounded interval, but $f(s_n) > n\varepsilon$, a contradiction.

Since f is continuous at 0, it follows easily by translation that f is continuous everywhere on $\mathbb{R}$. Since $f(x) = xf(1)$ for all rational x, this equation is also valid for all real x.

24.43. Lemma on the maximal function for Lebesgue measure. Let μ be Lebesgue measure on $\mathbb{R}$, let X be a Banach space, and let $f \in L^1(\mathbb{R}, X)$. For each $t \in \mathbb{R}$, let

$$g(t) \quad = \quad \sup_B \frac{1}{\mu(B)} \int_B |f(s)|\, ds,$$

where the supremum is over all open intervals B that contain t. Then g is defined uniquely at each point of $\mathbb{R}$ (even if f is left ambiguous on a set of measure 0); g is lower semicontinuous (hence measurable); and

$$\mu\Big(\{t \in \mathbb{R} \ : g(t) > \alpha\} \Big) \quad \leq \quad \frac{3}{\alpha}\|f\|_1$$

for any number $\alpha > 0$. The function g is called the **maximal function** associated with f. (This lemma will be used in the proof of 25.16. It is somewhat comparable to 29.18.)

Proof of lemma. (This presentation is from Fefferman [1977].) The definition of g is not affected if we change f on a set of measure 0. To see that g is measurable, note that

$$g(t) = \sup_{a,b>0} g_{a,b}(t), \qquad \text{where} \qquad g_{a,b}(t) = \frac{1}{a+b}\int_{t-a}^{t+b} |f(s)|\,ds.$$

Thus g is a supremum of continuous functions, hence g is lower semicontinuous.

Fix any $\alpha > 0$, and let $S = \{t \in \mathbb{R} : g(t) > \alpha\}$. For each $t \in S$, there is some open interval B_t containing t such that $\int_{B_t} |f(s)|\,ds > \alpha\mu(B_t)$. Fix any finite number $r < \mu(S)$. The set $G = \bigcup_{t\in S} B_t$ is open, hence measurable; it contains S, hence $\mu(G) > r$. By the regularity property established in 24.40, there is some compact $K \subseteq G$ with $\mu(K) > r$.

Since it is compact, K can be covered by finitely many of the members of $\mathcal{B}$ — say by $B_{t_1}, B_{t_2}, \ldots, B_{t_n}$. We may assume that these are arranged in order of decreasing length — that is, $\mu(B_{t_1}) \geq \mu(B_{t_2}) \geq \cdots \mu(B_{t_n})$.

We shall choose a subsequence $B_{u_1}, B_{u_2}, \ldots, B_{u_m}$ with the property that the B_{u_j}'s are disjoint and $\sum_{j=1}^m \mu(B_{u_j}) > r/3$. Let $B_{u_1} = B_{t_1}$. Thereafter, let $B_{u_{j+1}}$ be the first one of the B_{t_i}'s that does not meet any of $B_{u_1}, B_{u_2}, \ldots, B_{u_j}$. The resulting collection $B_{u_1}, B_{u_2}, \ldots, B_{u_m}$ is clearly disjoint. To see that $\sum_{j=1}^m \mu(B_{u_j}) > r/3$, reason as follows: For each j, let L_j be the open interval that has the same midpoint as B_{u_j} but is three times as long. If some B_{t_i} is not among the B_{u_j}'s, then B_{t_i} meets some B_{u_j} that is at least as large as B_{t_i}. Then $B_{t_i} \subseteq L_j$. Hence $K \subseteq \bigcup_{i=1}^n B_{t_i} \subseteq \bigcup_{j=1}^m L_j$, so $r < \mu(K) \leq \sum_{j=1}^m \mu(L_j) = 3\sum_{j=1}^m \mu(B_{u_j})$.

Let C be the union of the B_{u_j}'s. Then

$$\frac{\alpha r}{3} \quad < \quad \alpha\sum_{j=1}^m \mu(B_{u_j}) \quad < \quad \sum_{j=1}^m \int_{B_{u_j}} |f(s)|\,ds \quad = \quad \int_C |f(s)|\,ds \quad \leq \quad \|f\|_1.$$

The desired inequality follows immediately, in view of our choice of r.

More about Riemann Integrals (Optional)

24.44. *Proposition.* If $f : [a,b] \to X$ is Riemann integrable, then f is bounded.

Proof of proposition. Suppose (p_k) is a sequence in $[a,b]$ with $\|f(p_k)\| \to \infty$. Choose some number $\delta > 0$ such that whenever T is a δ-fine tagged division, then $|\Sigma[f,T] - \int_a^b f(t)dt| < 1$. Then whenever T and T' are δ-fine tagged divisions, we have $|\Sigma[f,T] - \Sigma[f,T']| < 2$. Choose a partition $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ with $\max_j(t_j - t_{j-1}) < \delta$. Show that some subinterval $[t_{j-1}, t_j]$ contains infinitely many of the p_k's. Consider tagged divisions T^k, all of which are identical except for their tag τ_j in the subinterval $[t_{j-1}, t_j]$; let T^k use p_k for that tag. Then $\Sigma[f, T^k] - \Sigma[f, T^{k'}] = [f(p_k) - f(p_{k'})]\,(t_j - t_{j-1})$; from this obtain a contradiction.

24.45. Theorem. Let $(X, \|\ \ \|)$ be a Banach space. Let $f : [a,b] \to X$ be bounded, and continuous at almost every point of $[a,b]$. Then f is Riemann integrable.

Proof. For simplicity of notation we may assume $[a,b] = [0,1]$. Also, we may extend f to all of $\mathbb{R}$ by defining $f(t) = 0$ for all $t \in \mathbb{R} \setminus [0,1]$.

For each positive integer p, define a step function $w_p : \mathbb{R} \to [0, +\infty)$ by taking

$$w_p(t) = \sup\left\{\|f(r) - f(s)\| : r, s \in \left(\frac{k-1}{p}, \frac{k+1}{p}\right]\right\} \text{ for } t \in \left(\frac{k-\frac{1}{2}}{p}, \frac{k+\frac{1}{2}}{p}\right].$$

The functions w_p are bounded, since f is bounded. Also, at every point t where f is continuous, we have $w_p(t) \to 0$ as $p \to \infty$. Thus $\int_0^1 w_p(t)dt \to 0$ by the Dominated Convergence Theorem.

Now let us consider any two tagged divisions

$$\begin{aligned} S &: \quad 0 = s_0 \le \sigma_1 \le s_1 \le \sigma_2 \le s_2 \le \cdots \le s_{m-1} \le \sigma_m \le s_m = 1, \\ T &: \quad 0 = t_0 \le \tau_1 \le t_1 \le \tau_2 \le t_2 \le \cdots \le t_{n-1} \le \tau_n \le t_n = 1, \end{aligned}$$

of the interval $[a,b]$, such that $\max_i(s_i - s_{i-1}) < \frac{1}{2p}$ and $\max_j(t_j - t_{j-1}) < \frac{1}{2p}$; we wish to compare the approximating Riemann sums $\Sigma[f,S]$ and $\Sigma[f,T]$. For each pair i,j, the set $[s_{i-1}, s_i] \cap [t_{j-1}, t_j]$ is either an interval $[u_{ij}, v_{ij}]$ or the empty set. Let $I = \{(i,j) : [s_{i-1}, s_i] \cap [t_{j-1}, t_j] \neq \varnothing\}$; then

$$\Sigma[f,S] = \sum_{(i,j)\in I} f(\sigma_i)(v_{ij} - u_{ij}), \qquad \Sigma[f,T] = \sum_{(i,j)\in I} f(\tau_j)(v_{ij} - u_{ij}).$$

Therefore $\Sigma[f,S] = \int_0^1 f_{S,T}(t)dt$ and $\Sigma[f,T] = \int_0^1 f_{T,S}(t)dt$, where $f_{S,T}$ and $f_{T,S}$ are the step functions that take the values $f(\sigma_i)$ and $f(\tau_j)$, respectively, on the interval $(u_{ij}, v_{ij}]$.

For any $(i,j) \in I$, consider any point $t \in (u_{ij}, v_{ij}]$. Both σ_i and τ_j lie within distance $\frac{1}{2p}$ from the number t. Choose the integer k that satisfies $t \in \left(\frac{k-\frac{1}{2}}{p}, \frac{k+\frac{1}{2}}{p}\right]$; then $\sigma_i, \tau_j \in \left(\frac{k-1}{p}, \frac{k+1}{p}\right]$, hence $\|f_{S,T}(t) - f_{T,S}(t)\| = \|f(\sigma_i) - (\tau_j)\| \le w_p(t)$. Thus $\|\Sigma[f,S] - \Sigma[f,T]\| \le \int_0^1 w_p(t)dt$. This proves that the net $\Sigma[f,\cdot]$ is Cauchy, and therefore f is Riemann integrable.

24.46. Theorem (Lebesgue). A function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if it is measurable (from Lebesgue-measurable sets to Borel sets) and bounded and its discontinuities make up a subset of $[a,b]$ that has Lebesgue measure 0.

Proof. Parts of this theorem were proved in 24.44 and 24.45. It remains to show that if f is Riemann integrable, then its discontinuities make up a set of measure 0. For simplicity of notation, we may take $[a,b] = [0,1]$; also, we may define $f(t) = 0$ for all $t < 0$ and for all $t > 1$. For positive integers n, define lower and upper step functions by taking

$$\left.\begin{aligned} l_n(t) &= \inf\left\{f(s) : \tfrac{k-1}{2^n} < s \le \tfrac{k}{2^n}\right\} \\ u_n(t) &= \sup\left\{f(s) : \tfrac{k-1}{2^n} < s \le \tfrac{k}{2^n}\right\} \end{aligned}\right\} \text{ when } \frac{k-1}{2^n} < t \le \frac{k}{2^n},$$

and taking $l_n(0) = u_n(0) = f(0)$. Then $l_n \le f \le u_n$; the sequence (l_n) increases pointwise to a limit l; the sequence (u_n) decreases pointwise to a limit u. By the Monotone or

Dominated Convergence Theorems, $\int_0^1 l_n(t)dt \to \int_0^1 l(t)dt$ and $\int_0^1 u_n(t)dt \to \int_0^1 u(t)dt$. On the other hand, we are assuming f is Riemann integrable; since $\int_0^1 l_n(t)dt$ and $\int_0^1 u_n(t)dt$ are equal to approximating sums $\Sigma[f,T]$, both of these integrals converge to $\int_0^1 f(t)dt$ as $n \to \infty$. Thus $u - l$ is a nonnegative function with Lebesgue integral 0, so in fact $u - l = 0$ almost everywhere. Thus, $u(t) = l(t) = f(t)$ fails only on a set of measure 0. The set $\{2^{-n}k : n, k \in \mathbb{N}\}$ is countable, hence has measure 0. It is easy to see that f is continuous at any point t that is not in either of these sets of measure 0.

24.47. A pathological, Riemann-integrable function. The theorem in 24.46 applies to *real-valued* functions. It extends easily to functions taking values in a finite-dimensional Banach space. However, it is not valid for infinite-dimensional Banach spaces, as we shall show with an example of Gordon [1991].

Let ℓ_2 be the space of all square-summable sequences of real numbers, as in 22.25. That space is a separable Hilbert space; thus it is the most nonpathological of all infinite-dimensional Banach spaces. Nevertheless, we shall describe a Riemann-integrable function $f : [0,1] \to \ell_2$ that is discontinuous everywhere.

For each positive integer m, let e_m be the sequence with 1 in the mth place and 0s elsewhere. Let $(r_m : m \in \mathbb{N})$ be an enumeration of the rational numbers in $[0,1]$. Define $f(r_m) = e_m$ for all m and $f(t) = 0$ when t is irrational. Then f is discontinuous everywhere.

To prove that f is Riemann integrable, let $T = (n, t_j, \tau_j)$ be any tagged division with $\max_j(t_j - t_{j-1}) < \varepsilon$. We may merge two consecutive subintervals $[t_{j-1}, t_j]$ and $[t_j, t_{j+1}]$ if they have the same tag — i.e., if $t_j = \tau_{j-1} = \tau_j$, then we may replace the two subintervals with a single subinterval; this does not affect the value of the approximating Riemann sum $\Sigma[f,T]$. The resulting new tagged division still satisfies $\max_j(t_j - t_{j-1}) < 2\varepsilon$, and no two of its tags are identical. Hence $\langle f(\tau_j), f(\tau_k)\rangle = 0$ for $j \neq k$. Therefore

$$\|\Sigma[f,T]\|^2 \;=\; \left\| \sum_{j=1}^{n} f(\tau_j)(t_j - t_{j-1}) \right\|^2 \;=\; \sum_{j=1}^{n} \|f(\tau_j)\|^2 (t_j - t_{j-1})^2 \;\leq\; 2\varepsilon.$$

Thus $\int_0^1 f(t)dt = 0$.

Chapter 25

Fréchet Derivatives

Definitions and Basic Properties

25.1. *Definitions.* Let X and Y be normed spaces, and let $f : \Omega \to Y$ be some function with domain $\Omega \subseteq X$. We say that L is a **derivative** of f at a point $\xi \in \Omega$ if $L : X \to Y$ is a bounded linear map satisfying

$$\lim_{x \to \xi} \frac{\|f(x) - f(\xi) - L(x-\xi)\|}{\|x - \xi\|} \quad = \quad 0, \tag{$*$}$$

or, in greater detail,

$$\lim_{r \downarrow 0} \quad \sup_{\substack{x \in \Omega, \\ 0 < \|x-\xi\| < r}} \quad \frac{\|f(x) - f(\xi) - L(x-\xi)\|}{\|x - \xi\|} \quad = \quad 0.$$

Then f is said to be **differentiable** at ξ. This condition says, roughly, that f can be approximated closely by a continuous affine operator at points near ξ. The operator L may also be called the **Fréchet derivative** of f at ξ, to distinguish it from several other kinds of derivatives.

In most cases of interest, we can show that there is at most one operator L satisfying the conditions above — see 25.3. Thus we are justified in calling it *the* derivative of f at ξ. It may then be denoted by $f'(\xi)$; this is the **Lagrange notation** for the derivative. The **Cauchy notation** is $D_x f(\xi)$. Alternatively, the derivative can be written in **Leibniz notation**: df/dx or $\frac{df}{dx}$ or $\frac{df}{dx}(\xi)$; this can also be written as $\frac{dy}{dx}$ if $y = f(x)$. Each notation has its advantages: Lagrange notation is usually preferable if all of our functions are expressed in terms of the same independent variable x. Leibniz notation is usually preferable if we are working with several different choices of the independent variable (as in the Chain Rule, in 25.6).

Caution: Leibniz notation makes the derivative look like a quotient of two simpler, more elementary quantities. At least in the case where $X = Y = \mathbb{R}$, it is possible to explain the derivative as the quotient of two dependent variables, or even as the quotient of two infinitesimals. However, that explanation is not simple and is not recommended for beginners. The beginner is probably safer thinking of dy/dx as just one expression and ignoring the fact that the derivative appears to be the quotient of two simpler expressions (except when that appearance is helpful for mnemonic purposes, as in 25.6).

25.2. *Alternate definition for functions of a scalar variable.* When X is the scalar field ($\mathbb{R}$ or $\mathbb{C}$), then a linear operator $L : X \to Y$ can be represented by a vector $y_0 \in Y$, in this fashion: $L(t) = ty_0$ for all $t \in X$. (In fact, $y_0 = L(1)$.) Thus f' may be viewed as a mapping from Ω into Y. Since it is possible to divide by scalars, the definition in 25.1 can be restated in an equivalent but simpler form:

$$f'(\xi) \quad = \quad \lim_{x \to \xi} \frac{f(x) - f(\xi)}{x - \xi}.$$

In particular, when Y is also the scalar field, then $f'(\xi)$ is just a scalar, as in college calculus.

25.3. *Proposition: uniqueness of the derivative.* Let X and Y be normed spaces, let $\Omega \subseteq X$, let $f : \Omega \to Y$ be some mapping, and let $\xi \in \Omega$. Suppose that either (i) ξ is in the interior of Ω, or (ii) Ω is convex and has nonempty interior.

Then the derivative of f at ξ, if it exists, is unique — i.e., there is at most one bounded linear operator $L : X \to Y$ satisfying condition 25.1(*).

Remarks. These hypotheses can be weakened, but apparently not without making them more complicated. Note that hypothesis (ii) is satisfied if Ω is an interval in the real line.

Proof of proposition. By replacing f with the function $u \mapsto f(\xi + u)$, we may assume $0 \in \Omega$ and $\xi = 0$; this will simplify our notation. By assumption,

$$\lim_{r \downarrow 0} \quad \sup_{x \in \Omega,\, 0 < \|x\| < r} \frac{\|f(x) - f(0) - L(x)\|}{\|x\|} \quad = \quad 0.$$

Suppose that L_1 and L_2 are two bounded linear operators satisfying this condition; we must show that the bounded linear operator $M = L_1 - L_2$ is equal to 0. We know that M satisfies

$$\lim_{r \downarrow 0} \quad \sup_{x \in \Omega,\, 0 < \|x\| < r} \frac{\|M(x)\|}{\|x\|} \quad = \quad 0.$$

To show that the linear mapping M equals 0, it suffices to show that M vanishes on some nonempty open subset of X; in particular, it suffices to show that M vanishes on $\mathrm{int}(\Omega)$. Consider any nonzero point $v \in \mathrm{int}(\Omega)$. If either Ω is convex or $0 \in \mathrm{int}(\Omega)$, then $tv \in \Omega$ for all $t > 0$ sufficiently small. Then $0 = \lim_{t \downarrow 0} \|M(tv)\|/\|tv\| = \|M(v)\|/\|v\|$, so $\|M(v)\| = 0$.

25.4. *Definitions.* We say that f is **differentiable on the set** Ω if the Fréchet derivative $f'(\xi) = \frac{df}{dx}(\xi)$ exists for every point $\xi \in \Omega$. Thus we define a function $f' = \frac{df}{dx} : \Omega \to BL(X, Y)$, where $BL(X, Y)$ is the normed space of bounded linear operators from X into Y (introduced in 23.1). We say that f is **continuously differentiable** on Ω if it is differentiable and the mapping $f' : \Omega \to BL(X, Y)$ is continuous. The linear space of all continuously differentiable maps from Ω into Y is sometimes denoted $C^1(\Omega, Y)$.

If $f : \Omega \to Y$ is continuously differentiable, then the continuous mapping $f' : \Omega \to BL(X, Y)$ might also be differentiable at some point ξ. Then its derivative is the **second derivative** of f, denoted $f''(\xi) = \dfrac{d^2 f}{dx^2}(\xi)$. That operator is a member of $BL(X, BL(X, Y))$; it may be viewed as a map from $X \times X$ into Y. Similarly, we may define $f'''(\xi) = \dfrac{d^3 f}{dx^3}(\xi)$, etc., and in general $f^{(n)}(\xi) = \dfrac{d^n f}{dx^n}(\xi)$. We denote by $C^n(\Omega, Y)$ the class of functions f for which $f^{(n)}$ exists and is continuous. When Y is the scalar field ($\mathbb{R}$ or $\mathbb{C}$), then $C^n(\Omega, Y)$ may be written more briefly as $C^n(\Omega)$. A function is called **smooth** if it has derivatives of all orders — i.e., if it belongs to $C^\infty(\Omega, Y) = \bigcap_{n=1}^{\infty} C^n(\Omega, Y)$.

Note that $f^{(n)}(\xi) = [f^{(n)}(\xi)](\cdot)$ is a mapping from X^n into Y; we may write it as $[f^{(n)}(\xi)](x_1, x_2, \ldots, x_n)$. It is linear in each x_j if ξ and all the other x_i's are held fixed. In general it is *not* linear in ξ if all the x_i's are held fixed.

25.5. *Elementary examples and properties.*

a. If f is differentiable at ξ, then f is also continuous at ξ.

b. If f is a constant function, then $f'(\xi) = 0$ for all $\xi \in \Omega$.

c. Review (from a calculus text) the proof of the product rule:

$$\frac{d}{dt}\left(f(t)g(t)\right) \quad = \quad f'(t)g(t) + f(t)g'(t)$$

for scalar-valued functions f and g.

d. If $M : X \to Y$ is a bounded linear operator and $f(x) = M(x)$ for all $x \in \Omega$, then $f'(\xi) = M$ for each $\xi \in \Omega$. Thus the mapping $f' : \Omega \to BL(X, Y)$ is a constant mapping, since it takes the value M at each point of Ω. Hence $f'' = 0$.

However, f itself is *not* constant (unless $M = 0$). This is a subtle distinction that may confuse some beginners. We have $f(x) = M(x)$ and $f'(x) = M$, but these are two different things: $M(x)$ is a particular member of Y, whereas M is a mapping from X into Y.

e. If $f : X \to Y$ is some mapping and Y is the scalar field, then $f'(x)$ (if it exists) is a member of the dual space X^*. This situation should not be confused with the situation in 25.2.

f. When $X = \mathbb{R}$, then we can use one-sided limits (as in 15.21) to define the **one-sided derivatives**:

$$f^+(\xi) \;=\; \lim_{x \downarrow \xi} \frac{f(x) - f(\xi)}{x - \xi} \qquad \text{and} \qquad f^-(\xi) \;=\; \lim_{x \uparrow \xi} \frac{f(x) - f(\xi)}{x - \xi}.$$

When ξ lies in the interior of the set Ω, then the Fréchet derivative $f'(\xi)$ (defined as in 25.1) exists if and only if both the one-sided derivatives exist and are equal, in which case $f'(\xi)$ is equal to their common value. If Ω is an interval and ξ is the left endpoint of that interval, then the limits and derivatives from the left at ξ are meaningless and the Fréchet derivative $f'(\xi)$ is (by our definition in 25.1) the same as the right-handed derivative $f^+(\xi)$. Similarly, when ξ is the right endpoint of the interval, then the limits and derivatives from the right are meaningless and $f'(\xi) = f^-(\xi)$.

g. If we replace the norms of X and Y with equivalent norms, then the linear space $BL(X,Y) = \{$bounded linear operators from X into $Y\}$ remains unchanged and its norm also gets replaced by an equivalent norm. The existence and value of $f'(\xi)$ are unaffected by these replacements. Thus, our calculations are actually being performed, not in a normed space, but in a *normable space* — i.e., in a topological vector space whose topology can be given by various norms but that does not have one of those norms specified in particular.

h. (*Optional.*) Let X, Y be Banach spaces. With notation as in 23.28, recall that $Inv(X,Y)$ is an open subset of the Banach space $BL(X,Y)$. Define a mapping $\Xi : Inv(X,Y) \to BL(Y,X)$ by $\Xi(f) = f^{-1}$. Show that Ξ is continuously differentiable, with $[\Xi'(f)](h) = -f^{-1}hf^{-1}$. *Hint*: Use the series in 23.28.b.

i. (*Optional.*) Let X be a complex Banach space, and let $BL(X,X)$ be the Banach space of bounded linear operators from X into X. Let T be a member of $BL(X,X)$, and let $\rho(T)$ be its resolvent set (defined as in 23.30). Show that the mapping $\lambda \mapsto (\lambda I - T)^{-1}$ is a differentiable mapping from $\rho(T)$ into $BL(X,X)$. What is its derivative?

Another example of a derivative in infinite dimensions is given in 25.22.

25.6. Chain rule of differential calculus. Let X, Y, Z be normed spaces. Let $S \subseteq X$ and $T \subseteq Y$ be open sets. Let $f : S \to Y$ and $g : T \to Z$ be some functions. Suppose that $x_0 \in S$ and $y_0 = f(x_0) \in T$. Suppose that the derivatives at these points, $f'(x_0)$ and $g'(y_0)$, both exist. (We do not assume that f and g are differentiable anywhere else, though we do not prohibit that either.) Then the composition $g \circ f$ is differentiable at x_0, and we have

$$(g \circ f)'(x_0) \quad = \quad g'(y_0) \circ f'(x_0),$$

or, in other terms, $(g \circ f)'(\xi) \;=\; g'(f(\xi)) \circ f'(\xi)$. This formula is easier to remember in Leibniz notation: If z is a function of y and y is a function of x, then

$$\frac{dz}{dx} \quad = \quad \frac{dz}{dy} \circ \frac{dy}{dx}.$$

The dy's appear to "cancel out" in this formula. The proof of the Chain Rule is similar to that given in any calculus book for $X = Y = Z = \mathbb{R}$, with epsilons and deltas. We leave the details as an exercise. (It is interesting to compare this with 29.12.b.)

Cautionary remark. When $X = Y = Z = \mathbb{R}$, as in college calculus, the linear operators $\frac{dz}{dy}$ and $\frac{dy}{dx}$ are simply the operations of multiplication by a real number (see 25.2); hence the composition of those two operators is just the multiplication of those two real numbers. In that setting, it does not matter in what order we put the factors $\frac{dz}{dy}$ and $\frac{dy}{dx}$, since multiplication of real numbers is commutative. However, in the more general setting of three arbitrary normed spaces X, Y, Z, the order of the two factors is very important. The formula must be stated $\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$; it is *incorrect* if written $\frac{dz}{dx} = \frac{dy}{dx}\frac{dz}{dy}$. Indeed, the composition $\frac{dy}{dx}\frac{dz}{dy}$ may not even make sense, for $\frac{dy}{dx}$ is a bounded linear operator from X into Y while $\frac{dz}{dy}$ is a bounded linear operator from Y into Z. Even when $X = Y = Z$, the compositions $\frac{dz}{dy}\frac{dy}{dx}$ and $\frac{dy}{dx}\frac{dz}{dy}$ need not be equal, since the composition of bounded linear operators from X into itself generally is not commutative if $\dim(X) \geq 2$ — see 8.27.

PARTIAL DERIVATIVES

25.7. *The matrix of partial derivatives.* In some cases of interest, the normed spaces X and Y are products of finitely many normed spaces:

$$X = X_1 \times X_2 \times \cdots \times X_m, \qquad Y = Y_1 \times Y_2 \times \cdots \times Y_n$$

for some positive integers m, n. As we noted in 25.5.g, for our present purposes any norm can be replaced by any equivalent norm. Hence the product topology on X can be given by

$$\begin{aligned} \|(x_1, x_2, \ldots, x_m)\| &= \|x_1\| + \|x_2\| + \cdots + \|x_m\| \qquad \text{or} \\ \|(x_1, x_2, \ldots, x_m)\| &= \max\{\|x_1\|, \|x_2\|, \ldots, \|x_m\|\} \end{aligned}$$

or any other convenient product norm; similarly for Y. Let Ω be a subset of X; then a function $f : \Omega \to Y$ can be represented by a wide assortment of notations:

$$\begin{aligned} y = (y_1, y_2, \ldots, y_n) &= f(x) = f(x_1, x_2, \ldots, x_m) = \Big(f_1(x), f_2(x), \ldots, f_n(x)\Big) \\ &= \Big(f_1(x_1, x_2, \ldots, x_m),\ f_2(x_1, x_2, \ldots, x_m),\ \ldots,\ f_n(x_1, x_2, \ldots, x_m)\Big). \end{aligned}$$

We shall use these different expressions interchangeably, switching to whichever one is most convenient in any particular context — usually suppressing whatever information is not currently being used.

The **partial derivative** $\partial y_i/\partial x_j$ is the derivative of the mapping $x_j \mapsto y_i$ that we obtain when we consider x_j to be the only variable and view all the other x_p's as constants — i.e., hold their values fixed. With $j = 1$, for instance, $\frac{\partial y_i}{\partial x_1}(\xi)$ is a bounded linear operator $L : X_1 \to Y_i$ that satisfies

$$\lim_{u \to \xi_1} \frac{\|f_i(u, \xi_2, \xi_3, \ldots, \xi_m) - f_i(\xi_1, \xi_2, \xi_3, \ldots, \xi_m) - L(u - \xi_1)\|}{\|u - \xi_1\|} = 0$$

(if such an operator L exists). We define $\partial y_i/\partial x_j$ for other j's analogously.

Exercise. Let us represent vectors $x \in X$ and $y \in Y$ as column matrices; that is,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \qquad \text{and} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Suppose the Fréchet derivative $dy/dx = f'(x)$ (defined as in 25.1) exists. Then all the partial derivatives exist, and the Fréchet derivative is equal to the matrix of partial derivatives:

$$f'(\xi) \quad = \quad \frac{dy}{dx} \quad = \quad \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} & \cdots & \dfrac{\partial y_1}{\partial x_m} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} & \cdots & \dfrac{\partial y_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial y_n}{\partial x_1} & \dfrac{\partial y_n}{\partial x_2} & \cdots & \dfrac{\partial y_n}{\partial x_m} \end{bmatrix}.$$

The expression $L(x-\xi)$ in 25.1(*) is then evaluated by the usual method for multiplying a matrix times a vector, as in 8.28. Of course, here we must extend the meaning of the term "matrix:" the components $\partial y_i/\partial x_j$ of this matrix are not necessarily numbers, or even members of a single ring; rather, $\partial y_i/\partial x_j$ is a bounded linear operator from X_j into Y_i. The components of the matrix are numbers (i.e., scalars) when X_j and Y_i are both one-dimensional, as in 25.2.

Example. Here is a typical example in two dimensions. If $\Omega = X = Y = \mathbb{R}^2$ and $(y_1, y_2) = (x_1 \cos x_2,\ x_1 \sin x_2)$, then

$$f'(x_1, x_2) \quad = \quad D_x \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad = \quad \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} \end{bmatrix} \quad = \quad \begin{bmatrix} \cos x_2 & -x_1 \sin x_2 \\ \sin x_2 & x_1 \cos x_2 \end{bmatrix}.$$

Thus, for instance, $\partial y_1/\partial x_1$ is the function we obtain by viewing x_2 as a constant and x_1 as a variable, and differentiating the function $y_1 = x_1 \cos x_2$ with respect to that variable. The other partial derivatives $\partial y_i/\partial x_j$ are defined similarly. To say that the formula above gives the derivative is to say that this quotient

$$\frac{\left\| \begin{bmatrix} x_1 \cos x_2 \\ x_1 \sin x_2 \end{bmatrix} - \begin{bmatrix} \xi_1 \cos \xi_2 \\ \xi_1 \sin \xi_2 \end{bmatrix} - \begin{bmatrix} \cos \xi_2 & -\xi_1 \sin \xi_2 \\ \sin \xi_2 & \xi_1 \cos \xi_2 \end{bmatrix} \begin{bmatrix} x_1 - \xi_1 \\ x_2 - \xi_2 \end{bmatrix} \right\|}{\left\| \begin{bmatrix} x_1 - \xi_1 \\ x_2 - \xi_2 \end{bmatrix} \right\|}$$

converges to 0 when $x_1 \to \xi_1$ and $x_2 \to \xi_2$.

Further observations. If the Fréchet derivative exists, then the partial derivatives all exist. The converse is not valid: There are some functions that possess partial derivatives but do not possess Fréchet derivatives; one example is the function f given in 15.28.b. (*Exercise.* Prove this.) One convenient sufficient condition for the existence of a Fréchet derivative is given in the exercise below, but it is not a necessary and sufficient condition.

Exercise. Let $X = X_1 \times \cdots \times X_m$ and $Y = Y_1 \times \cdots \times Y_n$; let Ω be an open subset of X; let $f : \Omega \to Y$ be some mapping. Suppose that all the partial derivatives $g_{ij} = \partial y_i / \partial x_j$ exist. Also assume that each function $g_{ij}(x_1, x_2, \ldots, x_m)$ is a jointly continuous function from X into Y_i. Then f has a Fréchet derivative given by the matrix of partial derivatives.

Hints: We can consider each Y_i separately. (Why is that?) Therefore, in the computation below, we shall suppress the subscript i. For simplicity of notation we shall consider only the case of $m = 3$; the proof for arbitrary m is similar. Use the fact that

$$\begin{aligned} &\left\| f(x_1 + h_1,\ x_2 + h_2,\ x_3 + h_3) - f(x_1, x_2, x_3) - \sum_{j=1}^{3} g_j(x_1, x_2, x_3) h_j \right\| \\ &\leq \ \| f(x_1 + h_1, x_2 + h_2, x_3 + h_3) - f(x_1, x_2 + h_2, x_3 + h_3) - g_1(x_1, x_2, x_3) h_1 \| \\ &+ \| f(x_1, x_2 + h_2, x_3 + h_3) - f(x_1, x_2, x_3 + h_3) - g_2(x_1, x_2, x_3) h_2 \| \\ &+ \| f(x_1, x_2, x_3 + h_3) - f(x_1, x_2, x_3) - g_3(x_1, x_2, x_3) h_3 \|. \end{aligned}$$

25.8. *Real derivatives versus complex derivatives.* The spaces $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic when considered as *real* Banach spaces — i.e., as complete normed vector spaces over the scalar field $\mathbb{R}$. The obvious bijection preserves the linear structure and also the topology. However, the two-dimensional real vector space $\mathbb{C}$ and the one-dimensional complex vector space $\mathbb{C}$ have different differentiable structures, as we shall now show.

Proposition. Let $(W, \|\ \ \|)$ be any complex Banach space (for instance, $\mathbb{C}$ itself), and let $f : \mathbb{C} \to W$ be some function. We may view W also as a real Banach space and define a corresponding function $g : \mathbb{R}^2 \to W$ by $g(x, y) = f(x + iy)$ for real x and y. Suppose g has a Fréchet derivative $g'(x_0, y_0) = \left[\ \frac{\partial g}{\partial x}(x_0, y_0) \quad \frac{\partial g}{\partial y}(x_0, y_0)\ \right]$ at some point (x_0, y_0). Then f has a Fréchet derivative at the point $z_0 = x_0 + iy_0$ if and only if the partial derivatives of g satisfy

$$\frac{\partial g}{\partial y}(x_0, y_0) \quad = \quad i\,\frac{\partial g}{\partial x}(x_0, y_0), \tag{$**$}$$

in which case $f'(z_0) = \frac{\partial g}{\partial x}(x_0, y_0)$. Equation $(**)$ may be called the **vector version of the Cauchy-Riemann equations**.

Hints: We emphasize that it is assumed that g has a Fréchet derivative; this is stronger than the assumption that g has partial derivatives. By the definition of Fréchet derivative, a complex number λ is the Fréchet derivative of f at z_0 if and only if λ satisfies

$$\lambda \quad = \quad \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \quad = \quad \lim_{(h,k) \to (0,0)} \frac{g(x_0 + h, y_0 + k) - g(x_0, y_0)}{h + ik}.$$

If the limit exists, then we must get the same value for the limit no matter how (h, k) approaches $(0, 0)$. In particular, approach along the horizontal direction or along the vertical direction. Thus we get the values

$$\begin{aligned} \lim_{h \to 0} \frac{g(x_0 + h, y_0) - g(x_0, y_0)}{h} \quad &= \quad \frac{\partial g}{\partial x}(x_0, y_0), \\ \lim_{k \to 0} \frac{g(x_0, y_0 + k) - g(x_0, y_0)}{ik} \quad &= \quad \frac{1}{i}\frac{\partial g}{\partial y}(x_0, y_0), \end{aligned}$$

which must therefore be equal; this proves $(**)$. Conversely, if $(**)$ holds, apply the definition of the Fréchet derivative.

Further remarks. A particularly important special case is that of $W = \mathbb{C}$. In that case we may write $f = u + iv$, where u and v are real-valued functions (see 10.25). Then equation $(**)$ can be rewritten as $\frac{\partial}{\partial y}(u+iv) = i\frac{\partial}{\partial x}(u+iv)$ — that is, $u_y + iv_y = iu_x - v_x$. Since $u(x,y)$ and $v(x,y)$ are real functions of real variables, all their partial derivatives u_x, u_y, v_x, v_y are real. Hence we may equate the real parts of the preceding condition, as well as the imaginary parts. Thus, for $W = \mathbb{C}$, equation $(**)$ can be rewritten in the form

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

These are the classical **Cauchy-Riemann Equations**.

Example. Let $f(z) = \overline{z}$. This is the complex conjugate of z; it is a continuous function of z. We have $u(x,y) = x$ and $v(x,y) = -y$. The **real derivative** of this function exists:

$$\begin{bmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2ex] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

However, the **complex derivative**

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

does *not* exist for this function f; the Cauchy-Riemann Equations are not satisfied.

25.9. The Chain Rule takes a particularly interesting form when the spaces X, Y, Z can be factored into simpler spaces, as in 25.7. If

$$x = (x_1, x_2, \ldots, x_m), \qquad y = (y_1, y_2, \ldots, y_n), \qquad z = (z_1, z_2, \ldots, z_p),$$

then the Chain Rule says

$$\begin{bmatrix} \dfrac{\partial z_1}{\partial x_1} & \cdots & \dfrac{\partial z_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial z_p}{\partial x_1} & \cdots & \dfrac{\partial z_p}{\partial x_m} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial z_1}{\partial y_1} & \cdots & \dfrac{\partial z_1}{\partial y_n} \\ \vdots & & \vdots \\ \dfrac{\partial z_p}{\partial y_1} & \cdots & \dfrac{\partial z_p}{\partial y_n} \end{bmatrix} \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial y_n}{\partial x_1} & \cdots & \dfrac{\partial y_n}{\partial x_m} \end{bmatrix}$$

with usual multiplication of matrices. Thus, the entry in row i, column j of the product is

$$\frac{\partial z_i}{\partial x_j} = \frac{\partial z_i}{\partial y_1}\frac{\partial y_1}{\partial x_j} + \frac{\partial z_i}{\partial y_2}\frac{\partial y_2}{\partial x_j} + \cdots + \frac{\partial z_i}{\partial y_n}\frac{\partial y_n}{\partial x_j}.$$

This formula is sometimes taught in college calculus texts, especially in the special cases where one or two of the integers m, n, p are equal to 1.

STRONG DERIVATIVES

25.10. *Definition.* Let X and Y be normed spaces, and let $f : \Omega \to Y$ be some function with domain $\Omega \subseteq X$. We say that L is a **strong derivative** of f at a point $\xi \in \Omega$, denoted $L = f'(\xi)$, if $L : X \to Y$ is a bounded linear map satisfying

$$\lim_{x,u \to \xi} \frac{\|f(x) - f(u) - L(x-u)\|}{\|x-u\|} = 0,$$

or, in greater detail,

$$\lim_{r \downarrow 0} \sup_{\substack{x,u \in B_r(\xi), \\ x \neq u}} \frac{\|f(x) - f(u) - L(x-u)\|}{\|x-u\|} = 0,$$

where $B_r(\xi)$ is the ball of radius r centered at ξ. Then we say f is **strongly differentiable** at ξ. Clearly, this is a stronger property than Fréchet differentiability.

Note that strong differentiability is a condition at a single point. It is possible for f to be strongly differentiable at ξ and yet be nondifferentiable at every point in $\Omega \setminus \{\xi\}$. We shall see in 25.23 that if f is differentiable on an open set, then f is strongly differentiable if and only if f is continuously differentiable.

Our results on strong derivatives are taken from Behrens [1974] and Nijenhuis [1974 and 1976].

Some basic properties.

a. Suppose $f : \Omega \to Y$ is strongly differentiable at some point $\xi \in \Omega$. Then f is Lipschitzian on some neighborhood of ξ. In fact, if $r > |||f'(\xi)|||$, then for some neighborhood G of ξ we have $\langle f|_G \rangle_{\text{Lip}} \leq r$.

Proof. Let $\varepsilon = r - |||f'(\xi)|||$. Choose G small enough so that $u, v \in G \Rightarrow \|f(u) - f(v) - f'(\xi)(u-v)\| \leq \varepsilon \|u - v\|$.

b. If we merely assume that f is *differentiable* on a neighborhood of ξ, then f need not be strongly differentiable at ξ.

Example. Let $f(t) = t^2 \sin(t^{-4})$ with $f(0) = 0$. Then f is differentiable at every point of $\mathbb{R}$, but f is not strongly differentiable at 0.

Proof. $f'(0) = 0$ since $|f(t)| \leq t^2$ for all t. For $t \neq 0$, we easily compute $f'(t) = 2t \sin(t^{-4}) - 4t^{-3} \cos(t^{-4})$. Now consider $u_n = (2\pi n)^{-1/4}$ and $v_n = u_n + \delta_n$ for some very small positive number δ_n to be specified. Then $f(u_n) = 0$ and $f'(u_n) = -4(2\pi n)^{3/4}$, hence

$$\lim_{\delta \downarrow 0} \frac{f(u_n + \delta) - f(u_n)}{\delta} = f'(u_n) = -4(2\pi n)^{3/4}.$$

Thus, for δ_n positive but sufficiently small,

$$\frac{f(v_n) - f(u_n)}{v_n - u_n} < -3(2\pi n)^{3/4},$$

and therefore $[f(v_n) - f(u_n)]/(v_n - u_n)$ does not tend to some finite limit $L(0)$ as $u_n, v_n \to 0$.

25.11. If f is a (possibly vector-valued) function of a *single real variable* and f is merely assumed to be differentiable at a point ξ, then f still has a property that is nearly as useful as strong differentiability:

Straddle Lemma. If f is differentiable at ξ, then $[f(u) - f(v)]/(u - v)$ approaches $f'(\xi)$ when u and v approach ξ *from opposite directions* (i.e., when u and v "straddle" ξ). That is, $\lim_{u\uparrow\xi,\, v\downarrow\xi} (f(u) - f(v))/(u - v) = f'(\xi)$ or, more precisely,

$$\lim_{\delta\downarrow 0} \quad \sup\left\{\left\|\frac{f(u)-f(v)}{u-v} - f'(\xi)\right\| \quad : \quad \begin{matrix} \xi-\delta \le u \le \xi \le v \le \xi+\delta, \\ u \ne v \end{matrix}\right\} \quad = \quad 0.$$

(This result will be used in 25.14 and 25.17.)

Proof. By the definition of $f'(\xi)$, we have $f(x) - f(\xi) - f'(\xi)(x - \xi) = (x - \xi)\varepsilon(x)$, where $\varepsilon(x)$ is a function satisfying $\lim_{x\to\xi} \varepsilon(x) = 0$. Subtract one of the equations

$$f(u) - f(\xi) - f'(\xi)(u - \xi) = (u - \xi)\varepsilon(u), \qquad f(v) - f(\xi) - f'(\xi)(v - \xi) = (v - \xi)\varepsilon(v)$$

from the other and then divide through by $u - v$, to obtain the equation

$$\frac{f(u)-f(v)}{u-v} - f'(\xi) \quad = \quad \frac{u-\xi}{u-v}\varepsilon(u) + \frac{\xi-v}{u-v}\varepsilon(v).$$

Because ξ lies between u and v, the right side of the equation above is a convex combination of $\varepsilon(u)$ and $\varepsilon(v)$, so it tends to 0 as $u, v \to \xi$.

25.12. Inverse Function Theorem. Let X and Y be Banach spaces, let $\Omega \subseteq X$ be open, and let $p : \Omega \to Y$ be some function. Assume that p is strongly differentiable at some point $x_0 \in \Omega$. (We do not assume that p is differentiable anywhere else.) Assume that the linear mapping $p'(x_0) : X \to Y$ is an isomorphism — i.e., a continuous linear bijection with continuous inverse.

Then p is locally invertible, in the following sense: There exist open sets $U \subseteq \Omega \subseteq X$ and $V \subseteq Y$, with $x_0 \in U$, such that the restriction $p\big|_U$ gives a bijection from U onto V whose inverse is strongly differentiable at $p(x_0)$. The derivative of p^{-1} at that point is equal to $p'(x_0)^{-1}$.

Proof. Let $\varphi = p'(x_0)$. It suffices to show that the mapping $q = \varphi^{-1} \circ p : \Omega \to X$ is locally invertible, for then $p = \varphi \circ q$ and $p^{-1} = q^{-1} \circ \varphi^{-1}$ have the required properties. Thus, we may assume $Y = X$. Replacing p with $p(x_0 + \cdot) - p(x_0)$, we may assume $x_0 = p(x_0) = 0$.

Denote open and closed balls by B and K, as in 5.15.g. Let $g(x) = x - p(x)$; then $g(0) = 0$ and $g'(0) = 0$. By 25.10.a, there is some $r > 0$ such that g is Lipschitzian on the closed ball $K(0, r)$ with Lipschitz constant at most $\frac{1}{2}$. Since $g(0) = 0$, g maps $K(0, r)$ into $K(0, r/2)$. For any constant $y \in K(0, r/2)$, the mapping $g_y = y + g$ is Lipschitzian with the same Lipschitz constant and maps $K(0, r)$ into $K(0, r)$. By Banach's Fixed Point Theorem

(19.39), g_y has a unique fixed point x in $K(0,r)$. Unwinding the notation, that says that for each $y \in K(0,r/2)$, the problem

$$p(x) = y, \qquad x \in K(0,r)$$

has a unique solution x. The same conclusion is reached if we replace r with any slightly smaller value. Hence for each $y \in B(0,r/2)$, the problem $p(x) = y$, $x \in B(0,r)$ has a unique solution. Let $V = B(0,r/2)$ and $U = B(0,r) \cap p^{-1}(V)$; then U, V are open and the restriction of p is a bijection from U onto V.

Next we show that $p^{-1} : V \to U$ is Lipschitzian. Since $\langle g \rangle_{\text{Lip}} \leq \frac{1}{2}$, we have

$$\|x_1 - x_2\| - \|p(x_1) - p(x_2)\| \quad \leq \quad \|x_1 - p(x_1) - x_2 + p(x_2)\| \quad \leq \quad \frac{1}{2}\|x_1 - x_2\|$$

and therefore $\|x_1 - x_2\| \leq 2\|p(x_1) - p(x_2)\|$. Thus $\langle p^{-1} \rangle_{\text{Lip}} \leq 2$.

Finally, we show that p^{-1} is strongly differentiable at $p(x_0)$, with derivative equal to $p'(x_0)^{-1}$. Let $\varphi = p'(x_0)$. Let $y_1 = p(x_1)$ and $y_2 = p(x_2)$. Then $\frac{1}{\|y_1 - y_2\|} \leq \frac{2}{\|x_1 - x_2\|}$, since $\langle p^{-1} \rangle_{\text{Lip}} \leq 2$. Then

$$\begin{aligned}
& \frac{\|p^{-1}(y_1) - p^{-1}(y_2) - \varphi^{-1}(y_1 - y_2)\|}{\|y_1 - y_2\|} \\
= \quad & \frac{\big\|(x_1 - x_2) - \varphi^{-1}[p(x_1) - p(x_2)]\big\|}{\|y_1 - y_2\|} \\
= \quad & \frac{\big\|\varphi^{-1}\left\{\varphi(x_1 - x_2) - [p(x_1) - p(x_2)]\right\}\big\|}{\|y_1 - y_2\|} \\
\leq \quad & 2\,|||\varphi^{-1}|||\, \frac{\|\varphi(x_1 - x_2) - [p(x_1) - p(x_2)]\|}{\|x_1 - x_2\|}\,.
\end{aligned}$$

When $y_1, y_2 \to p(x_0)$, then $x_1, x_2 \to x_0$, and then that last fraction tends to 0 by definition of $\varphi = p'(x_0)$.

Remarks. This theorem has many generalizations and variants. For instance, Clarke [1976] gives an inverse function theorem for Lipschitzian functions that are not necessarily differentiable. The bounded linear operator $f'(x_0)$ is replaced by a collection of approximating operators.

25.13. In the next theorem we solve the following problem: Let $f(x,y)$ be some function of two variables. When x and z are known, then we may try to solve the equation $f(x,y) = z$ for y. Does this make y into a function of x and z?

Implicit Function Theorem. Let X, Y, Z be Banach spaces, let $\Omega \subseteq X \times Y$ be an open set, and let $f : \Omega \to Z$ be some mapping that is strongly differentiable at some point $(x_0, y_0) \in \Omega$. Also suppose that $L = \frac{\partial f}{\partial y}(x_0, y_0) : Y \to Z$ is an isomorphism — i.e., suppose L is a linear bijection from Y onto Z such that L and L^{-1} are continuous. Let $z_0 = f(x_0, y_0)$.

Then there exists a Lipschitzian mapping $q : N \to Y$, defined on a neighborhood N of (x_0, z_0) in $X \times Z$, such that

$$q(x_0, z_0) = y_0, \qquad \text{and} \qquad f\Big(x, q(x, z)\Big) = z \text{ for all } (x, z) \in N.$$

Moreover, this mapping q is locally unique, in the following sense: For all (x, y, z) in some neighborhood of (x_0, y_0, z_0) in $X \times Y \times Z$, we have $f(x, y) = z$ if and only if $y = q(x, z)$.

Proof. Let $z_0 = f(x_0, y_0)$. Define a mapping $g : \Omega \to X \times Z$ by $g(x, y) = (g_1, g_2) = (x, f(x, y))$. Then $g(x_0, y_0) = (x_0, z_0)$. Also, g is strongly differentiable at (x_0, y_0), with

$$g'(x,y) \quad = \quad \begin{bmatrix} \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} \\ \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} \end{bmatrix} \quad = \quad \begin{bmatrix} i_X & 0 \\ \\ \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{bmatrix}.$$

In particular,

$$g'(x_0, y_0) \;=\; \begin{bmatrix} i_X & 0 \\ \\ f_x(x_0, y_0) & L \end{bmatrix} \qquad \text{has inverse} \qquad \begin{bmatrix} i_X & 0 \\ \\ -L^{-1} f_x(x_0, y_0) & L^{-1} \end{bmatrix}.$$

Thus the mapping $g'(x_0, y_0) : X \times Y \to X \times Z$ is an isomorphism. By the Inverse Function Theorem (in 25.12), the restriction of g gives a bijection from some open neighborhood of (x_0, y_0) onto some open neighborhood of (x_0, z_0), and both g and g^{-1} are Lipschitzian. Let $g^{-1}(x, z) = (p(x, z), q(x, z))$.

When $z = f(x, y)$, then $g(x, y) = (x, z)$, hence $(p(x, z), q(x, z)) = (x, y)$. This proves that $p(x, y) = x$ for all (x, y) near (x_0, y_0). Now for (x, y, z) near (x_0, y_0, z_0), we have

$$f(x, y) = z \iff g(x, y) = (x, z) \iff (x, y) = g^{-1}(x, z) \iff y = q(x, z).$$

25.14. Existence of nowhere-differentiable functions. The functions studied in an undergraduate course in calculus are, for the most part, piecewise continuously differentiable. That is, they are functions from an interval of $\mathbb{R}$ to $\mathbb{R}$, which are differentiable and have continuous derivatives except at finitely many points. However, those functions are atypical if we change our viewpoint slightly. We now present two different proofs of the existence of continuous functions that are nowhere-differentiable.

(i) $f(t) = \sum_{n=1}^{\infty} 2^{-n} \sin(2^{2n} \pi t)$ is a particular example of a function that is continuous but nowhere-differentiable.

(ii) The nowhere-differentiable functions comprise a comeager subset of the sup-normed Banach space $C[0, 1] = \{$continuous, real-valued functions on $[0, 1]\}$. (In other words, *most* continuous functions are nowhere-differentiable. This explains Poincaré's remark, on page v at the front of this book. See also 21.20.)

Proof of (i). This is similar to a proof given by Billingsley [1982]. Since $\sin(\cdot)$ is bounded, $f(t)$ is a uniform limit of continuous functions; therefore it is continuous. Fix any $t \in \mathbb{R}$; we shall show that $[f(b) - f(a)]/(b-a)$ is unbounded as $a, b \to t$ with $a \le t < b$, and hence f is not differentiable at $t \in \mathbb{R}$ by the Straddle Lemma (25.11).

Temporarily fix any large positive integer m; then choose k, a, b so k is an integer and

$$a \quad = \quad \frac{k}{2^{2m+1}} \quad \le \quad t \quad < \quad \frac{k+1}{2^{2m+1}} \quad = \quad b\,.$$

We now analyze the number $2^{-n}|\sin(2^{2n}\pi a) - \sin(2^{2n}\pi b)|$ in three cases:

- When n is an integer greater than m, then $2^{2n}a$ and $2^{2n}b$ are integers, hence $\sin(2^{2n}\pi a) = \sin(2^{2n}\pi b) = 0$.
- When $n = m$, then one of $2^{2n}a, 2^{2n}b$ is an integer and the other differs from it by $\frac{1}{2}$, so $2^{-n}|\sin(2^{2n}\pi a) - \sin(2^{2n}\pi b)| = 2^{-n} = 2^{-m}$.
- Since $|\frac{d}{dx}\sin(x)| \le 1$, the function $\sin(\cdot)$ is nonexpansive; so when n is a positive integer less than m we can estimate $2^{-n}|\sin(2^{2n}a) - \sin(2^{2n}b)| \le 2^{-n}|(2^{2n}a) - (2^{2n}b)| = 2^{n-2m-1}$.

Combining these results shows that $|f(b) - f(a)| \ge 2^{-m} - \sum_{n=-\infty}^{m-1} 2^{n-2m-1} = 2^{-m-1}$. Thus $|[f(b) - f(a)]/(b-a)| \ge 2^m$, which is not bounded as $m \to \infty$.

Example of a zigzag line

Proof of (ii). For $n = 1, 2, 3, \ldots$, let M_n be the set of all functions $x \in C[0,1]$ that have the following property:

$$\text{There exists some point } t_0 \in [0,1] \text{ such that } \sup_{s \in [0,1] \setminus t_0} \left| \frac{x(s) - x(t_0)}{s - t_0} \right| \quad \le \quad n.$$

It is clear that if x is continuous on $[0,1]$ and differentiable at some point t_0, then x satisfies the inequality above for some n, and thus $x \in \bigcup_{n=1}^{\infty} M_n$. It suffices to show that each set M_n is nowhere-dense. It is easy to see that M_n is closed in $C[0,1]$; thus it suffices to show that M_n has no interior. Let y be any continuous function on $[0,1]$, and let any $\varepsilon > 0$ be given; it suffices to show that some member of $C[0,1] \setminus M_n$ is within distance ε from y. Since y is uniformly continuous, we can partition $[0,1]$ into finitely many subintervals, on each of which y changes less than $\frac{\varepsilon}{2}$. Then we approximate y by a polygonal function — i.e., a continuous function z whose graph is a "zig-zag line" consisting of finitely many line segments. We may change those line segments to be very numerous and short and to all have slope greater than n or less than $-n$. Then $z \notin M_n$.

Derivatives of Integrals

25.15. First Fundamental Theorem of Calculus. Let X be a normed vector space. Suppose $u : [a,b] \to X$ is Henstock integrable. Then "the derivative of the integral of u is equal to u." More precisely:

Define a function $F : [a,b] \to X$ by $F(t) = \int_a^t u(s)ds$. Then F is differentiable at every point t_0 where u is continuous, and $F'(t_0) = u(t_0)$. Likewise, F has a one-sided derivative (equal to u) at every point where u is continuous from one side.

Hints: Suppose u is continuous at t_0. Given any $\varepsilon > 0$, choose $\delta > 0$ small enough so that $s \in [t_0 - \delta, t_0 + \delta] \cap [a,b]$ implies $\|u(s) - u(t_0)\| < \varepsilon$. Then $t \in [t_0 - \delta, t_0 + \delta] \cap [a,b]$ implies

$$\|F(t) - F(t_0) - (t - t_0)u(t_0)\| \quad = \quad \left\| \int_{t_0}^t u(s)ds - \int_{t_0}^t u(t_0)ds \right\| \quad \leq \quad |t - t_0|\varepsilon$$

by 24.16.a. Hence $\lim_{t \to t_0} [F(t) - F(t_0)]/(t - t_0) = u(t_0)$. The same proof applies, with obvious modifications, to one-sided continuity and one-sided derivatives.

25.16. Lebesgue's Theorem on differentiation of the integral. Let $u \in \mathcal{L}^1(\mathbb{R}, X)$. Let $F(t) = \int_0^t u(s)ds$. Then $F'(t)$ exists and equals $u(t)$ for almost every $t \in \mathbb{R}$. In fact, we have a slightly stronger conclusion: There exists a set L_u whose complement has Lebesgue measure 0, such that for each $t \in L_u$ we have

$$\lim_{h \to 0} \frac{1}{h} \int_t^{t+h} |u(s) - u(t)|ds \quad = \quad 0.$$

(This is a two-sided limit — i.e., we permit h to approach 0 from either side, with the notational convention of 24.20.) The set L_u is sometimes called the **Lebesgue set** for u; its members are called the **Lebesgue points** of u.

Remark. In most of integration theory we would work with a member of $L^1(\mathbb{R}, X)$ — that is, an equivalence class of functions. However, in the present theorem we work with a member of $\mathcal{L}^1(\mathbb{R}, X)$ — that is, a particular function from $\mathbb{R}$ into X. Different functions in a single equivalence class may have different Lebesgue sets, but those sets will differ by a set with Lebesgue measure 0.

Proof of theorem (following Fefferman [1977]). We are to show that the set

$$\left\{ t \in \mathbb{R} \ : \ \lim_{r \downarrow 0} \ \sup_{0 < |h| \leq r} \left| \frac{1}{h} \int_t^{t+h} |u(s) - u(t)|ds \right| > 0 \right\}$$

has measure 0. Hence it suffices to show that for each positive integer n, the set

$$S_n \quad = \quad \left\{ t \in \mathbb{R} \ : \ \lim_{r \downarrow 0} \ \sup_{0 < |h| \leq r} \left| \frac{1}{h} \int_t^{t+h} |u(s) - u(t)|ds \right| > \frac{1}{n} \right\}$$

has measure less than $\frac{1}{n}$.

The continuous functions that belong to $\mathcal{L}^1(\mathbb{R}, X)$ are dense in that space. Thus we may write $u = v + f$, where $v \in C(\mathbb{R}, X) \cap \mathcal{L}^1(\mathbb{R}, X)$ and $f \in \mathcal{L}^1(\mathbb{R}, X)$ with $\|f\|_1 < 1/8n^2$. Let g be the maximal function of f (defined as in 24.43). Since $\lim_{h \to 0} \frac{1}{h} \int_t^{t+h} |v(s) - v(t)| ds = 0$ for every t, we have

$$\begin{array}{rcl} S_n & = & \left\{ t \in \mathbb{R} \ : \ \lim\limits_{r \downarrow 0} \sup\limits_{0 < |h| \leq r} \left| \dfrac{1}{h} \displaystyle\int_t^{t+h} |f(s) - f(t)| ds \right| > \dfrac{1}{n} \right\} \\ & \subseteq & \left\{ t \in \mathbb{R} \ : \ \sup\limits_{h \neq 0} \left| \dfrac{1}{h} \displaystyle\int_t^{t+h} |f(s) - f(t)| ds \right| > \dfrac{1}{n} \right\} \\ & \subseteq & \left\{ t \in \mathbb{R} \ : \ |f(t)| > \dfrac{1}{2n} \right\} \cup \left\{ t \in \mathbb{R} \ : \ \sup\limits_{h \neq 0} \left| \dfrac{1}{h} \displaystyle\int_t^{t+h} |f(s)| ds \right| > \dfrac{1}{2n} \right\} \\ & \subseteq & \left\{ t \in \mathbb{R} \ : \ |f(t)| > \dfrac{1}{2n} \right\} \cup \left\{ t \in \mathbb{R} \ : \ g(t) > \dfrac{1}{2n} \right\}. \end{array}$$

Hence (letting μ denote Lebesgue measure)

$$\begin{array}{rcl} \mu(S_n) & \leq & \mu\left(\left\{ t \in \mathbb{R} \ : \ |f(t)| > \dfrac{1}{2n} \right\} \right) + \mu\left(\left\{ t \in \mathbb{R} \ : \ g(t) > \dfrac{1}{2n} \right\} \right) \\ & \leq & 2n\|f\|_1 + 6n\|f\|_1 \quad = \quad 8n\|f\|_1 \quad \leq \quad \dfrac{1}{n} \end{array}$$

by Chebyshev's Inequality (21.37.g) and the Lemma on the Maximal Function (24.43).

Integrals of Derivatives

25.17. Theorem relating the Henstock and Stieltjes integrals. Let X be a normed vector space space with scalar field $\mathbb{F}$. Let f and φ be two functions defined on $[a, b]$ — one of them vector-valued, the other scalar-valued. Suppose that φ is continuous, and its derivative $\varphi'(t)$ exists except at at most countably many values of t. Then the Henstock-Stieltjes integral $\int_a^b f(t) d\varphi(t)$ exists if and only if the Henstock integral $\int_a^b f(t)\varphi'(t) dt$ exists, in which case they are equal.

Clarification. By $\int_a^b f(t)\varphi'(t)dt$ we mean $\int_a^b f(t)G(t)dt$, where G is any function that satisfies $G(t) = \varphi'(t)$ except at perhaps countably many points. Thus G may be defined arbitrarily on a countable set. It is interesting to compare this theorem with 29.12.a.

Proof of theorem. We wish to show that if either of the nets $\Sigma[f, T, \varphi]$, $\Sigma[fG, T]$ (for tagged divisions T) converges to a limit, then the other converges to the same limit. Thus it suffices to show that $\|\Sigma[f, T, \varphi] - \Sigma[fG, T]\|$ becomes small as T progresses.

Let $S = \{\sigma_1, \sigma_2, \sigma_2, \ldots\}$ be an enumeration of the points where φ' does not exist or

$G \neq \varphi'$. Let any $\varepsilon > 0$ be given. Define a gauge δ separately on S and on $[a,b] \setminus S$, as follows:

Define δ on S by this rule: Choose $\delta(\sigma_k) > 0$ small enough so that

$$t, t' \in [\sigma_k - \delta(\sigma_k), \sigma_k + \delta(\sigma_k)] \cap [a,b]$$
$$\Rightarrow \qquad \|f(\sigma_k)\| \Big\{ \|\varphi(t) - \varphi(t')\| + \|G(\sigma_k)\|(t - t') \Big\} \; < \; 2^{-k-2}\varepsilon.$$

(We can do this since $\|f(\sigma_k)\|$ and $\|G(\sigma_k)\|$ are finite numbers and φ is continuous.)

Define δ on $[a,b] \setminus S$ by this rule: At each point τ in $[a,b] \setminus S$, by the Straddle Lemma (25.11) there is some number $\delta(\tau) > 0$ with the property that

$$u, v \in \Big[\tau - \delta(\tau), \tau + \delta(\tau)\Big] \cap [a,b], \qquad u \leq \tau \leq v, \qquad u \neq v$$
$$\Rightarrow \qquad \|f(\tau)\| \left\| \frac{\varphi(u) - \varphi(v)}{u - v} - G(\tau) \right\| < \frac{\varepsilon}{2(b-a)}.$$

In this fashion we define a gauge $\delta : [a,b] \to (0, +\infty)$. Now let $T = (n, t_j, \tau_j)$ be any δ-fine tagged division of $[a,b]$; we shall show that $\|\Sigma[f,T,\varphi] - \Sigma[fG,T]\| \leq \varepsilon$. We may drop any degenerate subinterval (i.e., any subinterval with $t_{j-1} = t_j$) since such subintervals contribute 0 to $\Sigma[f,T,\varphi]$ and $\Sigma[fG,T]$. Thus we may assume that $t_{j-1} < t_j$ for each j. Let $\lambda_j = \|f(\tau_j)\,[\varphi(t_j) - \varphi(t_{j-1})] - f(\tau_j)G(\tau_j)(t_j - t_{j-1})\|$; then $\|\Sigma[f,T,\varphi] - \Sigma[fG,T]\| \leq \sum_{j=1}^n \lambda_j$. We shall estimate the λ_j's in two ways, according to whether τ_j does or does not belong to S.

If $\tau_j = \sigma_k$ for some k, then $\lambda_j < 2^{-k-2}\varepsilon$ by our choice of $\delta(\sigma_k)$. Any σ_k appears at most twice among the tags τ_j (since we have no degenerate subintervals). Hence the sum of all such λ_j's is less than $2\sum_{k=1}^{\infty} 2^{-k-2}\varepsilon = \frac{1}{2}\varepsilon$.

On the other hand, if $\tau_j \notin S$, then $\lambda_j < (t_j - t_{j-1})\varepsilon/2(b-a)$ by our choice of $\delta(\tau_j)$. The sum of all such λ_j's is less than $\varepsilon/2$. This completes the proof.

25.18. Second Fundamental Theorem of Calculus. Let $\varphi : [a,b] \to X$ be a mapping from some compact interval into a normed vector space. Then "the integral of the derivative of φ is φ." More precisely:

(i) (College calculus version.) Suppose φ is differentiable at every point of $[a,b]$, and assume φ' is Riemann integrable. Then $\int_a^b \varphi'(t)dt = \varphi(b) - \varphi(a)$.

(ii) (Henstock integral version.) More generally, just assume φ is differentiable at every point of $[a,b]$. Then φ' is Henstock integrable, and $\int_a^b \varphi'(t)dt = \varphi(b) - \varphi(a)$.

(iii) (Henstock integral with bad points.) Still more generally, let $G : [a,b] \to X$ be some function. Suppose that $\varphi : [a,b] \to X$ is continuous, and suppose that the derivative $\varphi'(t)$ exists and equals $G(t)$ for all but countably many points t in $[a,b]$. Then G is Henstock integrable, and $\int_a^b G(t)dt = \varphi(b) - \varphi(a)$.

Proof. It suffices to prove (iii). Apply 25.17 with $f(t) = 1$.

25.19. *Pathological example.* In 25.18(iii) we cannot replace "countable set" with "set with measure 0." We shall now exhibit a continuous function $f : [0,1] \to [0,1]$ that is *not constant, but nevertheless satisfies $f'(t) = 0$ for all t outside a set of measure 0.*

Our example f is known as the **Cantor function**. (In chaos theory it is also known sometimes as the **devil's staircase** — see Devaney [1989].) It is fairly complicated; we shall construct it as the uniform limit of a sequence of simpler functions f_n. These are most easily understood by their graphs; see the graphs of the first few f_n's on the next page.

In general, $f_n(t)$ has a graph consisting of horizontal line segments of varying width, alternating with diagonal line segments that have slope $(3/2)^n$, which go up 2^{-n} units while going to the right 3^{-n} units. Each f_{n+1} is formed from f_n by this procedure: Leave unchanged each of the horizontal line segments in the graph of f_n. Replace each non-horizontal line segment (which covers a horizontal distance of 3^{-n}) and replace it with three new line segments (each of which covers a horizontal distance of 3^{-n+1}); the middle one of these new line segments is horizontal.

It is easy to see that the functions f_n are continuous and converge uniformly to a limit f, which is therefore continuous. The function f is constant on each of the open intervals

$$\left(\frac{1}{3},\frac{2}{3}\right), \quad \left(\frac{1}{9},\frac{2}{9}\right),\left(\frac{7}{9},\frac{8}{9}\right), \quad \left(\frac{1}{27},\frac{2}{27}\right),\left(\frac{7}{27},\frac{8}{27}\right),\left(\frac{19}{27},\frac{20}{27}\right),\left(\frac{25}{27},\frac{26}{27}\right), \quad \cdots$$

and thus we have $f' = 0$ at every point of the set

$$\left(\frac{1}{3},\frac{2}{3}\right) \cup \left(\frac{1}{9},\frac{2}{9}\right)\cup\left(\frac{7}{9},\frac{8}{9}\right) \cup \left(\frac{1}{27},\frac{2}{27}\right)\cup\left(\frac{7}{27},\frac{8}{27}\right)\cup\left(\frac{19}{27},\frac{20}{27}\right)\cup\left(\frac{25}{27},\frac{26}{27}\right) \cup \cdots,$$

which has measure equal to

$$\left(\frac{1}{3}\right)+\left(\frac{1}{9}+\frac{1}{9}\right)+\left(\frac{1}{27}+\frac{1}{27}+\frac{1}{27}+\frac{1}{27}\right)+\cdots = \frac{1}{3}\left[\left(\frac{2}{3}\right)^0+\left(\frac{2}{3}\right)^1+\left(\frac{2}{3}\right)^2+\cdots\right]$$

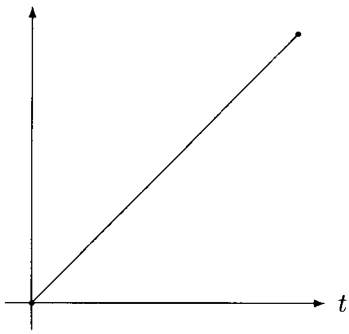
which is 1 (by 10.41.d). It follows that $f' = 0$ almost everywhere.

Some Applications of the Second Fundamental Theorem of Calculus

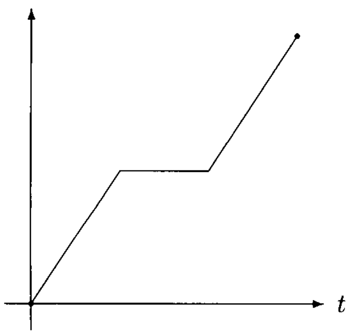
25.20. *Pathological example.* Let $f(t) = t^2\cos(\pi t^{-2})$ when $0 < t \le 1$, and let $f(0) = 0$. We shall show that this function's derivative, $f'(t)$, is *Henstock integrable but not Lebesgue integrable* on $[0,1]$. We easily compute $f'(t) = 2t\cos(\pi t^{-2}) + 2\pi t^{-1}\sin(\pi t^{-2})$ for $0 < t \le 1$, and $f'(0) = 0$. Then f' is Henstock integrable, by 25.18, and $f(b) - f(a) = \int_a^b f'(t)dt$ for any $[a,b] \subseteq [0,1]$. On the other hand, let

$$a_j = \sqrt{\frac{2}{4j+1}}, \quad b_j = \sqrt{\frac{2}{4j}}; \qquad \text{then} \qquad f(a_j) = 0, \quad f(b_j) = \frac{1}{2j}.$$

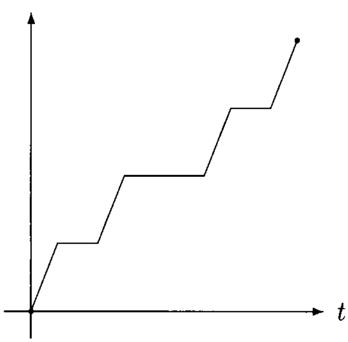
Approximations to
Cantor's function



$f_0(t) = t$ for all t in $[0, 1]$



$$f_1(t) = \begin{array}{rl} \frac{3}{2}t & \text{when } 0 \le t \le \frac{1}{3}, \\ \frac{1}{2} & \text{when } \frac{1}{3} \le t \le \frac{2}{3}, \\ \frac{3}{2}(t - \frac{1}{3}) & \text{when } \frac{2}{3} \le t \le 1. \end{array}$$



$$f_2(t) = \begin{array}{rl} \frac{9}{4}t & \text{when } 0 \le t \le \frac{1}{9}, \\ \frac{1}{4} & \text{when } \frac{1}{9} \le t \le \frac{2}{9}, \\ \frac{9}{4}(t - \frac{1}{9}) & \text{when } \frac{2}{9} \le t \le \frac{1}{3}, \\ \frac{1}{2} & \text{when } \frac{1}{3} \le t \le \frac{2}{3}, \\ \frac{9}{4}(t - \frac{4}{9}) & \text{when } \frac{2}{3} \le t \le \frac{7}{9}, \\ \frac{3}{4} & \text{when } \frac{7}{9} \le t \le \frac{8}{9}, \\ \frac{9}{4}(t - \frac{5}{9}) & \text{when } \frac{8}{9} \le t \le 1. \end{array}$$

Since $0 < a_n < b_n < a_{n-1} < b_{n-1} < \cdots < a_1 < b_1 < 1$, for any positive integer n we have $\int_0^1 |f'(t)|dt \geq \sum_{j=1}^n \left|\int_{a_j}^{b_j} f'(t)dt\right| = \sum_{j=1}^n \frac{1}{2j}$. Thus $\int_0^1 |f'(t)|dt = \infty$, so $f' \notin L^1[0,1]$, and we cannot define $\int_0^1 f'(t)dt$ as a Lebesgue integral.

25.21. Theorem on differentiation under the integral sign. Let X be a Banach space (equipped with its σ-algebra of Borel sets, as usual). Let $(\Omega, \mathcal{S}, \mu)$ be a measure space. Let $f : \mathbb{R} \times \Omega \to X$ be jointly measurable (where $\mathbb{R}$ is equipped with Lebesgue measure on the Lebesgue-measurable sets). Assume that

$$\frac{\partial f}{\partial s}(s,\omega) \quad = \quad \lim_{h\to 0} \frac{f(s+h,\omega) - f(s,\omega)}{h}$$

exists in X for all $(s,\omega) \in \mathbb{R} \times \Omega$, and assume $\frac{\partial f}{\partial s} \in L^1(\mathbb{R} \times \Omega, X)$ (where the product space is equipped with the product measure). Assume that $I(s_0) = \int_\Omega f(s_0,\cdot)d\mu$ exists for at least one real number s_0.

Then $I(s) = \int_\Omega f(s,\cdot)d\mu$ exists for every $s \in \mathbb{R}$. Also, $I'(s)$ and $\int_\Omega \frac{\partial f}{\partial s}(s,\cdot)d\mu$ exist and are equal for almost every $s \in \mathbb{R}$. Thus, we have $\frac{d}{ds}\int_\Omega f(s,\cdot)d\mu = \int_\Omega \frac{\partial}{\partial s} f(s,\cdot)d\mu$.

Proof. When we need to write Lebesgue measure explicitly, we shall denote it by λ.

Let us denote $g(s,\omega) = \frac{\partial f}{\partial s}(s,\omega)$. By the Second Fundamental Theorem of Calculus 25.18, we have $f(b,\omega) - f(a,\omega) = \int_a^b g(t,\omega)dt$. Hence

$$\int_\Omega \|f(b,\omega) - f(a,\omega)\|_X d\mu(\omega) \quad = \quad \int_\Omega \left\| \int_a^b g(t,\omega)dt \right\|_X d\mu(\omega) \quad \leq \quad \|g\|_{L^1(\lambda\times\mu, X)},$$

which is finite by hypothesis. It follows that $f(s,\cdot) \in L^1(\mu, X)$, and $I(s)$ exists for every $s \in \mathbb{R}$. Then $I(b) - I(a) = \int_\Omega \left[\int_a^b g(t,\omega)dt\right] d\mu(\omega)$.

Let $Y = L^1(\mu, X)$. By Fubini's Theorem (23.17), $L^1(\lambda\times\mu, X)$ is isomorphic to $L^1(\lambda, Y)$. For each $v \in L^1(\lambda \times \mu, X)$ let $\widehat{v}$ be the corresponding member of $L^1(\lambda, Y)$. The Bochner integrals $B_{[a,b]}u = \int_a^b u(t)dt$ and $B_\Omega v = \int_\Omega v(\omega)d\omega$ define continuous linear operators $B_{[a,b]} : L^1([a,b], Y) \to Y$ and $B_\Omega : L^1(\mu, X) \to X$; one of the conclusions of Fubini's Theorem 23.17 is that $B_\Omega B_{[a,b]}\widehat{v} = \int_\Omega \left[\int_a^b v(t,\omega)dt\right] d\mu(\omega)$.

By Lebesgue's Theorem on Differentiation (25.16) applied to members of $L^1(\lambda, Y)$, for almost every $s \in \mathbb{R}$ we have $\widehat{g}(s) = \lim_{h\to 0} \frac{1}{h}\int_s^{s+h} \widehat{g}(t)dt$; the limit and the integral are both in the Banach space Y. Fix any s for which that equation is valid. The equation can be restated $\widehat{g}(s) = \lim_{h\to 0}\frac{1}{h}B_{[s,s+h]}\widehat{g}$. Apply the operator B_Ω on both sides of that equation, keeping in mind that it is continuous and therefore preserves limits. We obtain

$$\int_\Omega g(s,\omega)d\mu(\omega) \;=\; B_\Omega\widehat{g}(s) \;=\; B_\Omega \lim_{h\to 0}\frac{1}{h}B_{[s,s+h]}\widehat{g} \;=\; \lim_{h\to 0}\frac{1}{h}B_\Omega B_{[s,s+h]}\widehat{g}$$

$$= \lim_{h\to 0}\frac{1}{h}\int_\Omega \left[\int_s^{s+h} g(t,\omega)dt\right] d\mu(\omega) \;=\; \lim_{h\to 0}\frac{1}{h}\left[I(s+h) - I(s)\right].$$

This completes the proof.

25.22. *Example: differentiation of an integral operator.* Let $X, (\Omega, \mathcal{S}, \mu), f$ be as in 25.21, and in addition assume that $\mu(\Omega) < \infty$, Ω is a compact metric space, and $\frac{\partial f}{\partial s} : \mathbb{R} \times \Omega \to X$ is jointly continuous. Let

$$BC(\mathbb{R}, X) \quad = \quad \{\text{bounded, continuous functions from } \mathbb{R} \text{ to } X\};$$

this is a Banach space when equipped with the sup norm. Define

$$\Big[F(\gamma)\Big](t) \quad = \quad \int_\Omega f\Big(\gamma(t), \omega\Big) d\mu(\omega) \qquad \text{for} \quad \gamma \in BC(\mathbb{R}, \mathbb{R}), \quad t \in \mathbb{R}.$$

We shall show that $[F(\gamma)](\cdot)$ belongs to $BC(\mathbb{R}, X)$, and the mapping

$$F \quad : \quad BC(\mathbb{R}, \mathbb{R}) \to BC(\mathbb{R}, X)$$

defined in this fashion is Fréchet differentiable. The derivative at any $\gamma \in BC(\mathbb{R}, \mathbb{R})$ is given by the continuous linear map $F'(\gamma) : BC(\mathbb{R}, \mathbb{R}) \to BC(\mathbb{R}, X)$ whose value at any $\psi \in BC(\mathbb{R}, \mathbb{R})$ is given by

$$\Big[F'(\gamma)\psi\Big](t) \quad = \quad \int_\Omega \frac{\partial f}{\partial s}\Big(\gamma(t), \omega\Big)\psi(t) d\mu(\omega).$$

Proof. Let $g = \frac{\partial f}{\partial s}$. Since Range($\gamma$) is a bounded subset of $\mathbb{R}$, it is contained in a compact set K. Then $f, g : K \times \Omega \to X$ are continuous functions on a compact set, hence they are bounded and uniformly continuous there. For fixed t, the function $\omega \mapsto f(\gamma(t), \omega)$ is measurable and bounded, hence integrable on the finite measure space Ω. For $t \in K$, the integrand is bounded; hence $[F(\gamma)](t)$ is a bounded function of t. That it is also a continuous function of t follows easily by Lebesgue's Dominated Convergence Theorem (22.29).

To show that $F'(\gamma)$ has the indicated value, replace K with a slightly larger compact subset of $\mathbb{R}$, so that $\Big(\gamma(t) + \psi(t), \omega\Big) \in K \times \Omega$ for all ψ sufficiently small. Let any $\varepsilon > 0$ be given. Since g is uniformly continuous on $K \times \Omega$, there is some $\delta > 0$ such that

$$\|\psi\|_\infty < \delta, \quad \alpha \in [0, 1] \qquad \Rightarrow \qquad \Big\| g\Big(\gamma(t) + \alpha\psi(t), \omega\Big) - g\Big(\gamma(t), \omega\Big) \Big\|_X < \varepsilon.$$

Let us denote $[E(\gamma)\psi](t) = \int_\Omega g\Big(\gamma(t), \omega\Big)\psi(t) d\mu(\omega)$. Observe that

$$\begin{aligned} &[F(\gamma + \psi) - F(\gamma) - E(\gamma)\psi](t) \\ &= \int_\Omega \Big\{ f\Big(\gamma(t) + \psi(t), \omega\Big) - f\Big(\gamma(t), \omega\Big) - g\Big(\gamma(t), \omega\Big)\psi(t) \Big\} d\mu(\omega) \\ &= \int_\Omega \Big\{ \int_0^1 g\left(\gamma(t) + \alpha\psi(t), \omega\right)\psi(t) dt - g\Big(\gamma(t), \omega\Big)\psi(t) \Big\} d\mu(\omega) \end{aligned}$$

and therefore $\|[F(\gamma+\psi) - F(\gamma) - E(\gamma)\psi](t)\|_X \le \varepsilon \|\psi\|_\infty \mu(\Omega)$. Finally, take the supremum over all choices of t; this yields $\|F(\gamma + \psi) - F(\gamma) - E(\gamma)\psi\|_\infty \le \varepsilon\|\psi\|_\infty \mu(\Omega)$. Since ε is arbitrarily small, this proves $F'(\gamma) = E(\gamma)$.

25.23. Theorem relating continuous differentiability to strong differentiability. Let X and Y be Banach spaces, let $\Omega \subseteq X$ be an open set, and let $f : \Omega \to Y$ be some differentiable function. Let $\xi \in \Omega$. Then f is strongly differentiable at ξ (as defined in 25.10) if and only if $f'(\cdot) : \Omega \to BL(X,Y)$ is continuous at ξ.

Proof. First suppose f is strongly differentiable at ξ. Let (x_n) be a sequence in Ω converging to ξ; we wish to show that $f'(x_n) \to f'(\xi)$. Since f is strongly differentiable at ξ, for any number $\varepsilon > 0$ there is some number $\delta > 0$ such that

$$\|u - \xi\|, \|v - \xi\| \le 2\delta \qquad \Rightarrow \qquad u, v \in \Omega \text{ and } \|f(u) - f(v) - f'(\xi)(u - v)\| \le \|u - v\|\varepsilon.$$

For n sufficiently large (say for $n \ge N_\varepsilon$) we have $\|x_n - \xi\| < \delta$; then

$$\|h\| \le \delta \qquad \Rightarrow \qquad \|f(x_n + h) - f(x_n) - f'(\xi)h\| \le \|h\|\varepsilon.$$

On the other hand, since f is differentiable at x_n, there is some $\delta_n > 0$ such that

$$\|h\| \le \delta_n \qquad \Rightarrow \qquad \|f(x_n + h) - f(x_n) - f'(x_n)h\| \le \|h\|\varepsilon.$$

Thus $\|h\| \le \min\{\delta, \delta_n\} \Rightarrow \|f'(\xi)h - f'(x_n)h\| \le 2\|h\|\varepsilon$. Therefore $n \ge N_\varepsilon \Rightarrow \|\|f'(\xi) - f'(x_n)\|\| \le 2\varepsilon$, so $f'(x_n) \to f'(\xi)$.

Conversely, suppose $f'(\cdot)$ is continuous at ξ. Temporarily fix any two points x_0, x_1 near ξ, and let $x_t = (1-t)x_0 + tx_1$ for $0 \le t \le 1$. Then, applying the Chain Rule (25.6) and the Second Fundamental Theorem of Calculus (25.18),

$$f(x_1) - f(x_0) \;=\; \int_0^1 \left[\frac{d}{dt} f(x_t)\right] dt \;=\; \int_0^1 f'(x_t)(x_1 - x_0)dt,$$

and therefore

$$f(x_1) - f(x_0) - f'(\xi)(x_1 - x_0) \quad = \quad \int_0^1 \left[f'(x_t) - f'(\xi)\right](x_1 - x_0)dt.$$

When x_0 and x_1 are near ξ, then all the x_t's are near ξ, hence $\|f'(x_t) - f'(\xi)\|$ stays small for all $t \in [0,1]$. This can be made precise with epsilons and deltas; we omit the details. It follows easily that $\|f(x_1) - f(x_0) - f'(\xi)(x_1 - x_0)\| \,/\, \|x_1 - x_0\| \to 0$ as $x_1, x_0 \to \xi$.

25.24. Theorem relating Lipschitzness to derivatives. Let Ω be a convex open subset of a Banach space X. Suppose that $f : \Omega \to Y$ is differentiable at every point of Ω. (We do not require that the derivative of f be continuous.) Then $\langle f \rangle_{\text{Lip}} = \|f'\|_\infty$. That is,

$$\sup_{x_1 \ne x_2} \frac{\|f(x_1) - f(x_2)\|_Y}{\|x_1 - x_2\|_X} \quad = \quad \sup_{x \in \Omega} \|f'(x)\|_{BL(X,Y)}$$

(with one side of this equation finite if and only if the other side is finite). Thus, f is Lipschitzian if and only if f' is bounded. This generalizes a result in 18.3.b.

Proof. First suppose $\langle f \rangle_{\text{Lip}} \leq r$, and let $L = f'(\xi)$ and $h \in X \setminus \{0\}$; we wish to show $\|Lh\| \leq r\|h\|$. Replacing h with ch for some small nonzero scalar c if necessary, we may assume $\xi + h \in \Omega$. Then also $\xi + th \in \Omega$ for all $t \in [0, 1]$, by convexity. Let any $\varepsilon > 0$ be given. By the definition of derivative we have

$$\frac{\|f(\xi + th) - f(\xi) - tLh\|}{t\|h\|} < \varepsilon \qquad \text{for all } t > 0 \text{ sufficiently small.}$$

For those t, we have $\|f(\xi + th) - f(\xi) - tLh\| < \varepsilon t\|h\|$, and therefore

$$t\|Lh\| = \|tLh\| \leq \|f(\xi + th) - f(\xi)\| + \varepsilon t\|h\| \leq rt\|h\| + \varepsilon t\|h\|.$$

Divide by t to obtain $\|Lh\| \leq (r + \varepsilon)\|h\|$; then let $\varepsilon \downarrow 0$.

Conversely, suppose $\|f'(\xi)\| < r$ for all $\xi \in \Omega$; we shall show $\langle f \rangle_{\text{Lip}} \leq r$. Let any $x_0, x_1 \in \Omega$ be given. Since Ω is convex, the points $x_t = tx_1 + (1 - t)x_0$ lie in Ω for all $t \in [0, 1]$. By the Chain Rule (25.6) and the Second Fundamental Theorem of Calculus (25.18), we have

$$f(x_1) - f(x_0) = \int_0^1 \left[\frac{d}{dt} f(x_t)\right] dt = \int_0^1 f'(x_t)(x_1 - x_0)dt.$$

Therefore $\|f(x_1) - f(x_0)\| \leq \int_0^1 r\|x_1 - x_0\|dt = r\|x_1 - x_0\|$ by 24.16.b or 24.16.a.

25.25. Theorem characterizing convex functions on an interval. Let $J \subseteq \mathbb{R}$ be an open interval, and let $f : J \to \mathbb{R}$ be some function. Then f is convex if and only if

- (i) f is continuous,
- (ii) the derivative $f'(t)$ exists except for at most countably many points $t \in J$, and
- (iii) there exists some increasing function $g : J \to \mathbb{R}$ such that $f'(t) = g(t)$ for all but at most countably many points $t \in J$.

Moreover, if f is convex, then both of the one-sided derivatives

$$f^+(t) = \lim_{s \downarrow t} \frac{f(s) - f(t)}{s - t}, \qquad f^-(t) = \lim_{s \uparrow t} \frac{f(s) - f(t)}{s - t}$$

exist for all $t \in J$, and either of these functions satisfies the requirements on g listed in (iii).

Proof. First assume $f : J \to \mathbb{R}$ is convex. Show that

$$\frac{f(y_1) - f(x_1)}{y_1 - x_1} \leq \frac{f(y_1) - f(x_2)}{y_1 - x_2} \leq \frac{f(y_2) - f(x_2)}{y_2 - x_2} \tag{!}$$

whenever $x_1 \leq x_2$, $y_1 \leq y_2$, and $x_j < y_j$ for $j = 1, 2$. The function $[f(y) - f(x_2)]/(y - x_2)$ is an increasing function of y for $y > x_2$, and it is bounded below by $[f(y_1) - f(x_1)]/(y_1 - x_1)$.

Hence $f^+(x_2)$ exists for every $x_2 \in J$, and therefore f is continuous from the right at every $x_2 \in J$. Take limits in (!) as $y_1 \downarrow x_1$ and $y_2 \downarrow x_2$, to prove $f^+(x_1) \leq f^+(x_2)$ for $x_1 < x_2$; thus f^+ is an increasing function. Similarly, f^- exists everywhere on J and is an increasing function, and f is continuous from the left. Combining these results, f is continuous on J. Since f^+ and f^- are increasing functions, they are continuous except at at most countably many points (see 15.21.c). Take limits in (!) as $x_2 \uparrow y_2$ and $y_1 \downarrow x_1$ to prove $f^-(y_2) \geq f^+(x_1)$ when $y_2 > x_1$; or take limits in (!) as $y_2 \downarrow x_2$ and $x_1 \uparrow y_1$ to prove $f^+(x_2) \geq f^-(y_1)$ when $x_2 > y_1$. Thus, at any point t where f^- and f^+ are both continuous, they must be equal, and there f' exists. Thus (i), (ii), (iii) are satisfied, with $g = f^+$ or $g = f^-$.

Conversely, suppose (i), (ii), (iii) are satisfied. By the Second Fundamental Theorem of Calculus (25.18), g is Henstock integrable on each closed interval $[a,b] \subseteq J$, with $\int_a^b g(t)dt = f(b) - f(a)$. We have $g(s) \leq g(b) \leq g(t)$ whenever $a \leq s \leq b \leq t \leq c$, hence

$$\begin{aligned}\frac{f(b)-f(a)}{b-a} &= \frac{1}{b-a}\int_a^b g(s)ds \;\leq\; \frac{1}{b-a}\int_a^b g(b)ds \;=\; g(b)\\ &= \frac{1}{c-b}\int_b^c g(b)dt \;\leq\; \frac{1}{c-b}\int_b^c g(t)dt \;=\; \frac{f(c)-f(b)}{c-b}\end{aligned}$$

when $a < b < c$. This inequality can be rearranged to yield

$$f\left(\frac{c-b}{c-a}a + \frac{b-a}{c-a}c\right) = f(b) \;\leq\; \frac{c-b}{c-a}f(a) \;+\; \frac{b-a}{c-a}f(c),$$

which proves f is convex.

Path Integrals and Analytic Functions (Optional)

25.26. *Remark.* This subchapter can be omitted. Its results will not be needed later in this book, except for a few brief examples.

Definitions. By a **path** in $\mathbb{C}$ we shall mean a function $\varphi : [a,b] \to \mathbb{C}$ that satisfies these conditions:

(i) φ is continuous,

(ii) φ is nondifferentiable at at most countably many points, and

(iii) φ has bounded variation.

We say that the path **begins** at $\varphi(a)$ and **ends** at $\varphi(b)$, or that $\varphi(a)$ and $\varphi(b)$ are the **initial** and **terminal points** of the path. A **closed path** is a path $\varphi : [a,b] \to \mathbb{C}$ that also satisfies $\varphi(a) = \varphi(b)$.

Let X be a complex Banach space. Let $\varphi : [a,b] \to \mathbb{C}$ be a path, and let $h : \mathrm{Range}(\varphi) \to X$ be a function that is measurable (from Borel sets to Borel sets) and bounded, with separable range. Then the **path integral** $\int_\varphi h(z)dz$ is defined to be the value of the Henstock-Stieltjes integral $\int_a^b (h \circ \varphi)\, d\varphi$ or, equivalently,

the Henstock integral $\int_a^b h(\varphi(t))\, \varphi'(t)dt$.

The existence of the former integral follows from 24.37; the existence of the latter integral and the equality of the two integrals follows from 25.17. If φ is a closed path (i.e., if $\varphi(a) = \varphi(b)$), then the path integral $\int_\varphi h(z)dz$ is sometimes written $\oint_\varphi h(z)dz$ for emphasis.

The terminology varies in the literature. Some mathematicians may prefer a less general or more general definition of "path." Also, some mathematicians will call φ_1 and φ_2 "different parametrizations of the same path" if $\varphi_1 = \varphi_2 \circ \sigma$ where σ is some continuous, strictly increasing function, because then φ_1 and φ_2 are interchangeable for the most important purposes: We have

$$\int_{\varphi_1} h(z)dz \quad = \quad \int_{\varphi_2} h(z)dz \tag{!!}$$

as an immediate consequence of 24.18. This equation says, roughly, that two motorists who drive from New York to Chicago along the same road will get the same value for any quantity computed as a path integral (with their driving routes as paths), even if they follow different timetables (different starting times, different speeds, etc.) when traversing that road.

In the preceding remarks we have established (!!) for any bounded measurable function h. We emphasize that we have established (!!) only when φ_1 and φ_2 are different parametrizations of the same path — i.e., the two motorists must follow the same road. However, for certain special functions h considered in 25.27, we obtain equality (!!) even if the motorists follow different roads — e.g., if one goes from New York to Chicago by way of Nashville, while the other goes from New York to Chicago by way of Buffalo.

25.27. We now state without proof a few results about analytic functions. Proofs can be found in books on functions of a complex variable. (Usually these basic results are presented for mappings from subsets of $\mathbb{C}$ into $\mathbb{C}$, but the proofs are not much different for mappings from subsets of $\mathbb{C}$ into a complex Banach space. However, we remark that the theory becomes much more complicated when one considers mappings from subsets of $\mathbb{C}^n$ into a complex Banach space, or even into $\mathbb{C}$. That theory will not be indicated here.)

Theorem. Let X be a complex Banach space, let Ω be an open subset of $\mathbb{C}$, and let $h : \Omega \to X$ be some function. Then the following conditions are equivalent.

(A) If γ is a closed path contained in an open convex subset of Ω, then $\oint_\gamma h(z)dz = 0$. (Cauchy called a function **holomorphic** if it satisfied a condition like this.)

(B) h has a complex derivative on Ω (as defined in 25.8). That is, $h'(\zeta) = \lim_{z\to\zeta} \frac{h(z)-h(\zeta)}{z-\zeta}$ exists in X for each point $\zeta \in \Omega$. (Riemann called such functions **complex differentiable**.)

(C) Locally, h is a sum of a power series. That is, for each $z \in \Omega$ there exist $R > 0$

and vectors $c_n \in X$ (which depend on z) such that

$$h(\zeta) \quad = \quad \sum_{n=0}^{\infty}(\zeta - z)^n c_n \qquad \text{for all } \zeta \text{ with } |\zeta - z| < R.$$

(Weierstrass called a function h **analytic** if it had this property.)

Furthermore, if the conditions above are satisfied, then the radius of convergence of the series in 25.27(C) is at least as large as the radius of the largest disk centered at z and contained in Ω. The function h has derivatives of all orders, and $c_n = h^{(n)}(z)/n!$ in that power series formula.

Any power series $h(\zeta) = \sum_{n=0}^{\infty} c_n(\zeta - z)^n$ is analytic inside its disk of convergence $\{\zeta : |\zeta - z| < R\}$. It can be differentiated term by term: $h'(\zeta) = \sum_{n=1}^{\infty} nc_n(\zeta - z)^{n-1}$ is a power series with the same radius of convergence. It can also be integrated term by term: $\int_\gamma h(\zeta)d\zeta = \sum_{n=0}^{\infty} \frac{c_n}{n+1}(q^{n+1} - p^{n+1})$ if γ is a path inside the disk of convergence with initial point p and terminal point q.

Montel's Theorem. Let Ω be an open subset of the complex plane, and let $h_1, h_2, h_3, \ldots : \Omega \to \mathbb{C}$ be a sequence of scalar-valued analytic functions. Suppose the sequence is uniformly bounded on compact sets — that is, assume

$$\sup_{k\in\mathbb{N}} \sup_{z\in K} |h_k(z)| \; < \; \infty \qquad \text{for each compact set } K \subseteq \Omega.$$

Then some subsequence (h_{k_j}) converges uniformly on compact sets, and the limit is also an analytic function.

Remark. This theorem takes an interesting form when restated in the terminology of topological vector spaces; see 26.10.

25.28. The following example illustrates the difference between smooth functions and analytic functions. Define a function $\psi : \mathbb{R} \to \mathbb{R}$ by

$$\psi(r) \quad = \quad \begin{cases} e^{-1/r} & \text{when } r > 0 \\ 0 & \text{when } r \le 0. \end{cases}$$

Verify the following:

a. $\psi(r) > 0$ for $r > 0$, and $\psi(r) = 0$ for $r \le 0$.

b. There exist some polynomials $p_0, p_1, p_2, \ldots$ such that the derivatives of ψ are of the form

$$\psi^{(n)}(r) \quad = \quad \begin{cases} p_n\left(\frac{1}{r}\right) e^{-1/r} & \text{when } r > 0 \\ 0 & \text{when } r \le 0. \end{cases}$$

Hint: Don't bother trying to *find* the polynomials explicitly; that is more work than is necessary. Just use induction on n to show that there *exist* such polynomials.

c. ψ is a smooth function — i.e., infinitely many times differentiable — with $\psi^{(n)}(0) = 0$ for all n.

Hint: Show that $\lim_{r\downarrow 0} e^{-1/r}\, r^k = 0$ for every integer k.

d. Conclude that generally $\psi(r) \neq \psi(0) + r\psi'(0) + \frac{r^2}{2}\psi''(0) + \cdots + \frac{r^n}{n!}\psi^{(n)}(0) + \cdots$.

e. *Related exercise.* Define a function $\beta : \mathbb{R}^m \to \mathbb{R}$ by $\beta(x) = \psi(1 - \|x\|_2^2)$. The function β could be called a *smooth unit bump function,* because it has these properties: β is smooth (i.e., infinitely many times differentiable), $\beta(x) > 0$ for $\|x\|_2 < 1$, and $\beta(x) = 0$ for $\|x\|_2 \geq 1$. It can be used to define *approximate identity* functions

$$\iota_\varepsilon(x) \quad = \quad \frac{\beta\left(\frac{x}{\varepsilon}\right)}{\varepsilon^m \int_{\mathbb{R}^m} \beta(u)du} \qquad\qquad (\varepsilon > 0,\ x \in \mathbb{R}^m).$$

These functions have the following properties: ι_ε is smooth, $\iota_\varepsilon(x) > 0$ for $\|x\|_2 < \varepsilon$, $\iota_\varepsilon(x) = 0$ for $\|x\|_2 \geq \varepsilon$, and $\int_{\mathbb{R}^m} \iota_\varepsilon(x)dx = 1$. Thus, the ι_ε's have unit weight, but that weight is concentrated near the origin when ε is small. Moreover, for many types of functions $g : \mathbb{R}^m \to \mathbb{R}$, the function $g * \iota_\varepsilon : \mathbb{R}^m \to \mathbb{R}$ defined by

$$(g * \iota_\varepsilon)(x) \quad = \quad \int_{\mathbb{R}^m} g(y)\,\iota_\varepsilon(x - y)dy$$

is smooth and converges to g as $\varepsilon \downarrow 0$. The type of convergence — pointwise, uniform, etc. — depends on the regularity assumptions about g — integrable, continuous, etc.; we omit the details.

25.29. *An example: series for* $\ln 2$ *and* $\pi/4$. The following example answers a question that some readers may have wondered about when they studied calculus.

Begin by forming the geometric series

$$\begin{aligned} \frac{1}{1+x} &= 1 - x + x^2 - x^3 + x^4 + \cdots && (|x| < 1), \\ \frac{1}{1+x^2} &= 1 - x^2 + x^4 - x^6 + x^8 - \cdots && (|x| < 1). \end{aligned}$$

Since a power series can be integrated term by term inside its radius of convergence, we obtain

$$\begin{aligned} \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots && (|x| < 1), \\ \arctan(x) &= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \cdots && (|x| < 1). \end{aligned}$$

The reader may wonder: Do these series formulas remain valid at $x = 1$? They do, but the proof of that fact is generally beyond the scope of courses in advanced calculus.

However, the proof is now quite easy, using a corollary of the Monotone Convergence Theorem. First, rewrite the series as

$$\begin{aligned} \ln(1+x) &= \left(x - \frac{x^2}{2}\right) + \left(\frac{x^3}{3} - \frac{x^4}{4}\right) + \cdots + \left(\frac{x^{2n-1}}{2n-1} - \frac{x^{2n}}{2n}\right) + \cdots && (|x| < 1), \\ \arctan(x) &= \left(x - \frac{x^3}{3}\right) + \left(\frac{x^5}{5} - \frac{x^7}{7}\right) + \cdots + \left(\frac{x^{4n-3}}{4n-3} - \frac{x^{4n-1}}{4n-1}\right) + \cdots && (|x| < 1). \end{aligned}$$

Observe that the functions $\frac{x^{2n-1}}{2n-1} - \frac{x^{2n}}{2n}$ and $\frac{x^{4n-3}}{4n-3} - \frac{x^{4n-1}}{4n-1}$ are increasing functions on $[0,1]$ since their derivatives are nonnegative on that interval. Now apply 21.39.d; we find that the formulas are valid for $x = 1$. Thus

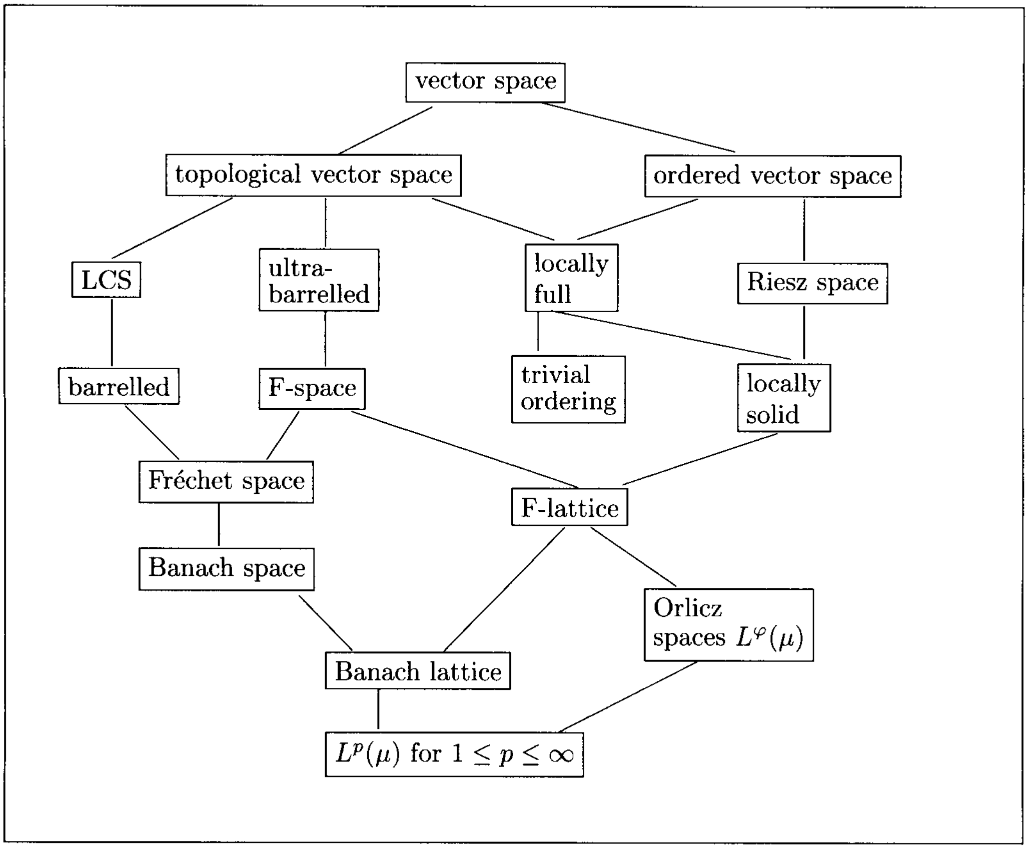
$$\ln 2 \;=\; 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots, \qquad\qquad \frac{\pi}{4} \;=\; 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots.$$

(The formula for $\ln 2$ was also proved by a different method in 10.42.)

Chapter 26

Metrization of Groups and Vector Spaces

26.1. *Preview.* The following chart shows the relations between several types of spaces that will be studied in this and later chapters.



A *topological vector space* (TVS) is a vector space equipped with a topology that makes the vector space operations continuous. More generally, a *topological Abelian group* (TAG) is an Abelian group equipped with a topology that makes the group operations continuous. This makes the topology translation-invariant, so there is a uniform structure naturally associated with that topology. Even when scalars are present (i.e., in a TVS), the most basic properties of the uniform structure do not involve the scalars; thus it is natural to first introduce uniform structure in the simpler and more general setting of TAG's.

In 5.32, 16.16, and Chapter 18 we saw that uniform spaces can be analyzed in terms of pseudometrics. Since the structure of a TAG is translation-invariant, it can be analyzed naturally in terms of translation-invariant pseudometrics. These are called *G-seminorms* (or *group seminorms*) in the theory developed below. In special kinds of TAG's we can use special kinds of G-seminorms: In TVS's we can use F-seminorms (the "F" stands for Fréchet), and in LCS's we can use seminorms.

The following table lists several kinds of distance functions in order of decreasing generality. Thus, every Riesz seminorm is also a seminorm, every seminorm is also an F-seminorm, every F-seminorm is also a paranorm, etc.

FUNCTIONAL	PROPERTIES	SUP TOPOLOGY
quasipseudometric	triangle inequality	topological space
pseudometric	symmetric	uniform space
G-seminorm	translation-invariant	topological additive group
paranorm	continuous multiplication	topological linear space
F-seminorm	balanced	topological linear space
seminorm	homogeneous	locally convex space
Riesz seminorm	isotone	locally solid LCS

Caution: Although the most basic theorems of this subject are fairly well standardized by now, the terminology is not; it varies slightly from one book or article to another. In particular, different mathematicians attach different meanings to the prefixes "quasi," "pseudo," or "semi." The reader is urged to check definitions whenever reading anything on this subject. In this text we have chosen an arrangement of terminology that is internally consistent, and is in agreement with the literature to the extent that this is possible. Here, both "pseudo" and "semi" mean "not necessarily positive-definite." When "pseudo" or "semi" appears in parentheses in a sentence, the sentence should be read once with the parenthesized term included and once with it omitted.

F-Seminorms

26.2. The reader will find it helpful to review the definitions of (semi)norm and G-(semi)norm given in 22.2. We shall now define some structures that are midway between (semi)norms and G-(semi)norms.

Definitions. Let X be a real or complex linear space; let $\mathbb{F}$ be the scalar field. A **paranorm** on X is a G-seminorm $\rho : X \to [0, +\infty)$ that makes scalar multiplication jointly continuous — i.e., that satisfies the following.

If $c_n \to c$ in $\mathbb{F}$ and $x_n \to x$ in the metric space (X, ρ), then $c_n x_n \to cx$ in (X, ρ).

Actually, it suffices to assume that scalar multiplication is separately continuous — i.e., that

> if $c_n \to c$ in $\mathbb{F}$ and $x_n \to x$ in the metric space (X, ρ), then $c_n x \to cx$ and $cx_n \to cx$ in (X, ρ)

— for in this context *separate continuity implies joint continuity*; that rather nontrivial fact follows (*exercise*) from 23.15.b and the fact that the scalar field $\mathbb{F}$ (always $\mathbb{R}$ or $\mathbb{C}$ in this book) is a complete metric space.

An **F-seminorm** is a paranorm that is also *balanced* — i.e., satisfying

$$x \in X, \quad c \in \mathbb{F}, \quad |c| \le 1 \qquad \Rightarrow \qquad \rho(cx) \le \rho(x).$$

If it is also positive-definite, then ρ is called an **F-norm**, and (X, ρ) is an **F-normed space**. We may refer to X itself as the (F-)normed space, if no confusion will result.

The definitions above are admittedly complicated, but their importance will become evident in 26.29.

An **F-space** is a vector space topologized by an F-norm that is complete. (Equivalently, it is a complete metrizable topological vector space; we shall prove that equivalence in 26.29 and 26.32.) *Caution*: In the modern literature, an "F-normed space," an "F-space," a "Fréchet space," and a "topological space with a Fréchet topology" are four different things; see 26.14.

Any seminorm is also an F-seminorm (easy exercise). Other examples of F-seminorms and paranorms will be given below.

Further remarks about terminology. This book's terminology conforms to the literature whenever possible, but it is not always possible; the literature varies greatly in its terminology for F-norms and related notions. Our definition of "F-norm" is equivalent to the definition used by Kalton, Peck, and Roberts [1984] and Köthe [1969]. Our definition of "paranorm" follows that of Wilansky [1978]; Swartz [1992] calls this object a **quasinorm**. If a paranorm is positive-definite, then Wilansky [1978] calls it a "**total paranorm**," Yosida [1964] calls it a "**quasinorm**," and Swartz [1992] calls it a "**total quasinorm**." Many other books and papers — including the classic work of Banach [1932/1987] — use positive-definite paranorms without attaching any name to them. A very extensive treatment of metric linear spaces is given by Rolewicz [1985].

26.3. *Relations between G-seminorms, paranorms, and F-seminorms.*

a. A function ρ is an F-seminorm on a vector space X if and only if ρ is a balanced G-seminorm that satisfies this **scalar continuity** condition:

$$\text{For any } x \in X, \text{ if } |c_n| \to 0, \text{ then } \rho(c_n x) \to 0.$$

Hint: 12.25.f.

b. Any paranorm ρ is equivalent to an F-seminorm σ.

Hint (modified from Rolewicz [1985]): The set $\{t \in \mathbb{F} : |t| \le 1\}$ is compact; hence the number $\sigma(x) = \max\{\rho(tx) : |t| \le 1\}$ is finite.

c. Let X be a linear space, and let $\rho : X \to [0, +\infty)$ be some mapping. Then the following are equivalent:

(A) ρ is a seminorm;

(B) ρ is a convex function and an F-seminorm;

(C) ρ is a convex, balanced G-seminorm.

Hints: We shall prove (C) $\Rightarrow$ (A); the other implications are easy. Show that $\rho(sx) \leq s\rho(x)$ for $x \in X$ — first for $s \in \mathbb{N}$, by the subadditivity of G-seminorms; then for $s \in (0, 1]$, by the assumed convexity of ρ; then for $s > 0$ by combining those two results. Then show $\rho(sx) \geq s\rho(x)$ for $s > 0$ by replacing s with $1/s$. Then what?

26.4. *Basic examples.*

a. *Open and closed balls.* Let (X, d) be a metric space. As in 5.15.g, define the open ball and closed ball

$$B_d(z, r) \;=\; \{x \in X \;:\; d(x, z) < r\}, \qquad K_d(z, r) \;=\; \{x \in X \;:\; d(x, z) \leq r\}.$$

As we noted in 5.18.c, $\mathrm{cl}[B_d(z, r)] \subseteq K_d(z, r)$ in any metric space. Show that

$$\mathrm{cl}[B_d(z, r)] \;=\; K_d(z, r) \qquad \text{in any normed space.}$$

That conclusion is false in some F-normed spaces; for instance, show that it is false in $\mathbb{R}$ equipped with the F-norm $\rho(x) = \min\{1, |x|\}$.

b. *Pathological example.* Consider $\mathbb{C}$ (the complex numbers) as a vector space with scalar field $\mathbb{R}$; then

$$\rho(z) \;=\; |\mathrm{Re}\; z| + |\mathrm{Im}\; z|$$

is a norm. On the other hand, if we consider $\mathbb{C}$ as a vector space with scalar field $\mathbb{C}$, then ρ is not a norm or an F-norm (since it is not balanced), but it is a paranorm on $\mathbb{C}$. Here we understand that the absolute value of a scalar is defined as usual: $|c| = \sqrt{(\mathrm{Re}\; c)^2 + (\mathrm{Im}\; c)^2}$.

c. *Another pathological example.* Using the identity $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$, show that the function $f(x) = |\sin(\pi x)|$ is subadditive on $\mathbb{R}$ — that is, $f(x + y) \leq f(x) + f(y)$ for $x, y \in \mathbb{R}$. Then show that the function

$$\rho(x) \;=\; |\sin(\pi x)| \;+\; \min\{2, |x|\}$$

is a paranorm on $\mathbb{R}$ that is equivalent to the usual norm on $\mathbb{R}$, but ρ is not balanced.

d. For $0 < p < 1$, the functions

$$\|x\|_p^p \;=\; |x_1|^p + |x_2|^p + \cdots + |x_n|^p$$

are F-norms on $\mathbb{R}^n$ or $\mathbb{C}^n$ that yield the product topology. This follows easily from 12.25.e. (Here $\|\ \ \|_p$ is defined as in 22.11.) Similarly, the functions

$$\|x\|_p^p \;=\; |x_1|^p + |x_2|^p + |x_3|^p + \cdots$$

are F-norms on ℓ_p, defined as in 22.25. We emphasize that the functions $\|\ \ \|_p$ are not necessarily norms when $p < 1$.

e. The sum of finitely many paranorms or (G-)(F-)(semi)norms is another object of the same type.

f. The pointwise maximum of finitely many paranorms or (G-)(F-)(semi)norms is another object of the same type.

g. The restriction of any paranorm or (F-)(semi)norm to a linear subspace is a paranorm or (F-)(semi)norm.

h. *Bounded equivalents.* Let β be a bounded remetrization function; in 18.14 we saw that if d is a pseudometric on a set X, then $\beta \circ d$ is a bounded, uniformly equivalent pseudometric. Show that if ρ is a G-(semi)norm or an F-(semi)norm, then $\beta \circ \rho$ is a bounded, uniformly equivalent G-(semi)norm or F-(semi)norm.

We cannot make an analogous assertion for seminorms, however. If ρ is a seminorm, then $\beta \circ \rho$ is an equivalent F-seminorm, but in general it is not a seminorm. Indeed, the only bounded seminorm is the constant function 0.

i. If X is a vector space, ρ is an F-seminorm on X, K is a linear subspace, and the quotient G-seminorm $\widehat{\rho}$ is defined as in 22.13, then $\widehat{\rho}$ is an F-seminorm too.

26.5. *Change of scalar field.* Let X be a complex vector space. Then X may also be viewed as a real vector space, if we "forget" how to multiply members of X by members of $\mathbb{C} \setminus \mathbb{R}$. Show that

a. If ρ is a paranorm, F-seminorm, or seminorm on the complex vector space X, then ρ is a paranorm, F-seminorm, or seminorm (respectively) on the real vector space X.

b. If ρ is a seminorm or F-seminorm on the real vector space X, show that

$$\gamma(x) \quad = \quad \sup\{\rho(tx) \ : \ t \in \mathbb{C},\ |t| \leq 1\}$$

defines a seminorm or F-seminorm γ on the complex vector space X, which is "semi-equivalent" to ρ in this sense: $\rho(x) \leq \gamma(x) \leq \rho(x) + \rho(ix)$ for all $x \in X$. (*Hint*: $\rho(tx) \leq \rho(\mathrm{Re}(t)x) + \rho(\mathrm{Im}(t)ix)$.)

Moreover, if X is equipped with some topology and $\rho : X \to [0, +\infty)$ is lower semicontinuous, then $\gamma : X \to [0, +\infty)$ is lower semicontinuous. (*Hint*: It is the sup of the lower semicontinuous functions $x \mapsto \rho(tx)$.)

(This construction is based on Rolewicz [1985].)

26.6. *Fréchet combinations.* Let $(\rho_j : j \in \mathbb{N})$ be a sequence of F-seminorms on a vector space X. Then $\varphi(x) = \sum_{j=1}^{\infty} 2^{-j} \min\{1, |\varphi_j(x)|\}$ defines an F-seminorm φ that is uniformly equivalent to the gauge $\{\rho_1, \rho_2, \rho_3, \ldots\}$.

More generally, let (α_j) be a sequence of positive numbers with finite sum. Let $\Gamma : [0, +\infty) \to [0, +\infty)$ be a bounded remetrization function (defined as in 18.14). For $x \in X$, let

$$\varphi(x) \quad = \quad \sum_{j \in \mathbb{N}} \alpha_j \, \Gamma\Big[\rho_j(x)\Big].$$

This sum is called a **Fréchet combination** of the ρ_j's; it is a special case of the sum developed in 18.17. Show that

a. For any sequence (x_n) in X (or, more generally, any net), we have

$$\varphi(x_n) \to 0 \qquad \text{if and only if} \qquad \rho_j(x_n) \to 0 \ \text{ for each } j. \tag{$*$}$$

b. φ is a bounded F-seminorm on X. (*Hint*: Use $(*)$ for an easy proof that φ is scalar continuous.)

c. φ is uniformly equivalent to the gauge $\{\rho_1, \rho_2, \rho_3, \ldots\}$. Thus, the topology that φ determines on X is the supremum of the topologies of the pseudometric spaces (X, ρ_j), and the uniform structure of (X, φ) is the supremum of the uniform structures of the (X, ρ_j)'s. (*Hint*: Again, use $(*)$.)

d. If we replace Γ with some other bounded remetrization function and/or replace (α_j) with some other sequence of positive numbers with finite sum, then condition $(*)$ remains valid; hence the resulting Fréchet combination is equivalent to φ.

e. φ is an F-norm if and only if the ρ_j's separate points of X — i.e., if and only if they have the further property that whenever $\rho_j(x) = 0$ for all j, then $x = 0$.

Remarks. In most applications of this formula, the F-seminorms ρ_j are actually seminorms — i.e., they are homogeneous. In that case X is locally convex; that will follow from 26.20.d. (Hence, if the metric determined by φ is complete, then X is a Fréchet space.) However, φ cannot be a seminorm, since φ is bounded. In many applications, φ is not even *equivalent* to a seminorm; see the examples in 27.7.c and 27.8.

26.7. Example: the space of all sequences. Let $\mathbb{F}$ be the scalar field — that is, $\mathbb{R}$ or $\mathbb{C}$; then $\mathbb{F}^{\mathbb{N}} = \{\text{sequences of scalars}\}$. The product topology and product uniform structure on $\mathbb{F}^{\mathbb{N}}$ are given by any of the following F-norms:

$$\sum_{j=1}^{\infty} \frac{\min\{1, |x_j|\}}{j!}, \qquad \sum_{j=1}^{\infty} \frac{\arctan |x_j|}{2^j}, \qquad \sum_{j=1}^{\infty} \frac{j^{-2}|x_j|}{1 + |x_j|},$$

applied to any sequence $x = (x_1, x_2, x_3, \ldots)$. Indeed, these formulas are all special cases of 26.6, with $\rho_j(x) = |x_j|$. The resulting F-norms are complete; any one of them may be referred to as the **usual F-norm** on $\mathbb{F}^{\mathbb{N}}$. Although all the ρ_j's are seminorms, none of the resulting F-norms is a norm or even *equivalent* to a norm — i.e., the product topology on $\mathbb{F}^{\mathbb{N}}$ cannot be given by a norm; that will be proved in 27.8.

26.8. Example: the space of all continuous functions. Let $C(\mathbb{R})$ be the set of all continuous functions from $\mathbb{R}$ into the scalar field $\mathbb{F}$ (which may be $\mathbb{R}$ or $\mathbb{C}$). For $f \in C(\mathbb{R})$, let

$$\varphi(f) \quad = \quad \sum_{j=1}^{\infty} 2^{-j} \arctan\left(\max_{t \in [-j,j]} |f(t)| \right).$$

Show that φ is an F-norm on $C(\mathbb{R})$ that is complete and that gives the topology of uniform convergence on compact sets (introduced in 18.26). This F-norm (or any other F-norm equivalent to it) is the **usual F-norm** on $C(\mathbb{R})$. It is not equivalent to a norm; see 27.8.

26.9. Example: the space of locally integrable functions. Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$). A function $f : \mathbb{R} \to \mathbb{F}$ is called **locally integrable** if its restriction to each compact interval $[a, b]$ is integrable. More generally, let Ω be an open subset of $\mathbb{R}^n$, equipped with Lebesgue measure; a function $f : \Omega \to \mathbb{F}$ is called **locally integrable** if its restriction to some open neighborhood of each point in Ω is integrable. The set of all (equivalence classes of) locally integrable functions on Ω is denoted by $L^1_{\text{loc}}(\Omega)$. Show that

a. $L^p(\Omega) \subseteq L^1_{\text{loc}}(\Omega)$ for every p in $[1, +\infty]$. (*Hint*: Hölder's Inequality.)

b. If $f : \Omega \to \mathbb{F}$ is locally bounded (i.e., bounded on compact sets) and f is measurable, then $f \in L^1_{\text{loc}}(\Omega)$. In particular, any continuous function from Ω to $\mathbb{F}$ is locally integrable.

Thus, the function $f(t) = \exp(t^2)$ is locally integrable on $\mathbb{R}$, even though it does not belong to $L^p(\mathbb{R})$ for any $p \in [1, +\infty]$.

c. $L^1_{\text{loc}}(\Omega)$ can be made into a Fréchet space in a natural fashion, using the sequence of seminorms $\rho_n(f) = \int_{G_n} |f(t)|\, dt$, where the G_n's form an open cover of Ω and each G_n is contained in some compact subset of Ω (see 17.18.a). The resulting F-norm is

$$\rho(f) \quad = \quad \sum_{n=1}^{\infty} 2^{-n} \max\left\{1, \int_{G_n} |f(t)|\, dt\right\}.$$

(In particular, $L^1_{\text{loc}}(\mathbb{R})$ can be made into a Fréchet space using the sequence of seminorms $\rho_n(f) = \int_{-n}^{n} |f(t)|\, dt$.)

Exercise. Different choices of the sequence (G_n) of relatively compact sets may yield different F-norms ρ. Show that any two such F-norms are equivalent. (*Hint*: 17.18.b.) In fact, the topology can be described as follows: A sequence (f_n) is ρ-convergent to a limit f if and only if, for each open set G that is contained in a compact subset of Ω, we have $\lim_{n\to\infty} \int_G |f_n(t) - f(t)|\, dt = 0$.

Further exercise. Prove the completeness of $L^1_{\text{loc}}(\Omega)$.

26.10. Example: the space of holomorphic functions. Let Ω be an open subset of the complex plane, and let $\text{Hol}(\Omega) = \{$holomorphic functions from Ω into $\mathbb{C}\}$ (defined as in 25.27).

The usual topology for $\text{Hol}(\Omega)$ is the topology of uniform convergence on compact subsets of Ω, introduced in 18.26. That topology makes $\text{Hol}(\Omega)$ a Fréchet space; it can be metrized as follows.

Let $G_1, G_2, G_3, \ldots$ be an open cover of Ω, where each G_n is contained in some compact subset of Ω. (See 17.18.a.) Define

$$\rho_n(f) \quad = \quad \max_{\omega \in G_n} |f(\omega)| \qquad\qquad (n = 1, 2, 3, \ldots).$$

Then each ρ_n is a seminorm on $\text{Hol}(\Omega)$, and the seminorms $\rho_1, \rho_2, \rho_3, \ldots$ determine the topology of uniform convergence on compact sets. The particular choice of the sequence (G_n) does not matter — see 17.18.b.

A further property of $\text{Hol}(\Omega)$ is noted in 27.10.c.

26.11. Example: convergence in measure. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let $(X, |\ \ |)$ be a Banach space. Then the pseudometric D_μ defined in 21.34 on $SM(\mathcal{S}, X)$ takes a slightly simpler form: It can be rewritten $D_\mu(f, g) = \rho_\mu(f - g)$, where

$$\rho_\mu(h) \quad = \quad \inf_{\alpha > 0} \arctan\left[\alpha \;+\; \mu\left\{\omega \in \Omega \;:\; |h(\omega)| > \alpha\right\}\right].$$

Some basic properties:

a. The function ρ_μ is a G-seminorm on $SM(\mathcal{S}, X)$ or a G-norm on the quotient space $SM(\mathcal{S}, X)/\mu$, making those vector spaces into topological Abelian groups.

b. In general, $SM(\mathcal{S}, X)$ is not a topological vector space. That is shown by the example below. However, in 26.12.c we shall consider a smaller subspace on which ρ_μ is indeed an F-seminorm.

Example. Let $(\Omega, \mathcal{S}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N})$, counting measure), and let $X = \mathbb{R}$. Define $f_n(j) = f(j) = j$ for all $n, j \in \mathbb{N}$, and let $c_n = \frac{1}{n}$ and $c = 0$. Then $f_n \to f$ in measure and $c_n \to c$, but $c_n f_n \not\to cf$ in measure. Thus, multiplication of scalars is not jointly continuous for this topology.

26.12. (*Optional.*) Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let $(X, |\ \ |)$ be a Banach space. A function $f : \Omega \to X$ is **totally measurable** if it is a strongly measurable function (as defined in 21.4) that also satisfies

$$\mu\left(\{\omega \in \Omega \;:\; |f(\omega)| > \varepsilon\}\right) \text{ is finite for each } \varepsilon > 0.$$

(Of course, if $\mu(\Omega) < \infty$, then *every* strongly measurable function is totally measurable.) Let $TM(\mu; X)$ denote the set of all μ-equivalence classes of totally measurable functions.

Exercises.

a. $TM(\mu, X)$ is a closed linear subspace of the G-normed space $(SM(\mu, X), \rho_\mu)$ (which is complete).

b. The finitely valued members of $TM(\mu, X)$ are dense in $TM(\mu, X)$.

c. The G-norm ρ_μ of 26.11 is an F-norm when restricted to $TM(\mu; X)$. The space $TM(\mu; X)$, equipped with this F-norm or any other equivalent F-norm, is sometimes denoted $L^0(\mu; X)$ (especially when $\mu(\Omega) < \infty$).

d. When $\mu(\Omega) < \infty$, then ρ_μ is also equivalent to this F-norm:

$$\gamma(f) \quad = \quad \int_\Omega \Gamma[\|f(\omega)\|] d\mu(\omega),$$

where Γ is any bounded remetrization function (defined as in 18.14). *Hint*: To prove γ is scalar-continuous (as in 26.3.a), use the Dominated Convergence Theorem 22.29.

e. For any $p \in (0, \infty)$, the vector space $L^p(\mu, X)$ is a linear subspace of the vector space $TM(\mu, X)$, and the F-seminorm $\|\ \ \|_p^{\min\{1,p\}}$ is stronger than ρ_μ on $L^p(\mu, X)$.

f. Dominated Convergence Theorem for TM. Let (f_j) be a sequence in $TM(\mu; X)$ that converges pointwise μ-almost everywhere to a limit f. Assume that the sequence

(f_j) is dominated by a totally measurable function — i.e., assume $g(\omega) = \sup_j |f_j(\omega)|$ is totally measurable. Then $f_j \to f$ μ-almost uniformly (hence also $f_j \to f$ in measure).

Hints: Let any $\varepsilon > 0$ be given. For each positive integer n, let $\Omega_n = \{\omega \in \Omega : g(\omega) > 1/n\}$; then $\mu(\Omega_n)$ is finite. Hence $f_j \to f$ μ-almost uniformly on Ω_n, by Egorov's Theorem. Thus there exists some set $A_n \subseteq \Omega_n$ such that $\mu(A_n) < 2^{-n}\varepsilon$, and $f_j \to f$ uniformly on $\Omega_n \setminus A_n$. Now let $A = \bigcup_{n=1}^{\infty} A_n$. Then $\mu(A) < \varepsilon$, and we shall show that $f_j \to f$ uniformly on $\Omega \setminus A$. To see this, let any $\delta > 0$ be given; we must show that for all j sufficiently large, we have $\sup_\omega |f_j(\omega) - f(\omega)| < \delta$. Choose some integer $n > 2/\delta$. Then $\Omega \setminus A = \complement A$ can be partitioned into the sets $(\complement A) \cap \Omega_n$ and $(\complement A) \cap (\complement \Omega_n)$. We have uniform convergence on $\Omega_n \setminus A_n$, hence on the smaller set $\Omega_n \setminus A = (\complement A) \cap \Omega_n$. The remaining piece, $(\complement A) \cap (\complement \Omega_n)$, is a subset of $\complement \Omega_n$, and at every ω in $\complement \Omega_n$ we have $\sup_j |f_j(\omega) - f(\omega)| \leq 2g(\omega) \leq 2/n < \delta$.

For a slightly more general treatment, see Dunford and Schwartz [1957], which permits μ to be a charge, not necessarily a measure.

26.13. (*Optional.*) (We omit the proofs of the results below; they constitute exercises that are difficult but may be within the reach of some particularly ambitious readers.)

By an **Orlicz function** we shall mean an increasing, continuous function $\varphi : [0, +\infty] \to [0, +\infty]$ that satisfies $\varphi^{-1}(0) = \{0\}$. (*Caution*: The terms "Orlicz function" and "Orlicz space" have slightly different meanings in different books and papers.) A few examples of Orlicz functions are

$$t^p \text{ (for constant } p > 0), \qquad t^p \ln(1+t), \qquad e^t - 1, \qquad \min\{1, t\}.$$

As the last example shows, we permit an Orlicz function to be bounded.

Let φ be an Orlicz function, let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let $(X, |\ \ |)$ be a Banach space. For strongly measurable functions $f : \Omega \to X$, define

$$\rho_\varphi(f) \quad = \quad \inf\left\{ r \in (0, +\infty] \ : \ \int_\Omega \varphi\left(\frac{|f(\omega)|}{r}\right) d\mu(\omega) < r \right\}.$$

Show that

a. The set $\mathcal{L}^\varphi(\mu; X) = \{f \in SM(\mathcal{S}; X) : \rho_\varphi(f) < \infty\}$ is a linear space, on which ρ_φ is a complete F-seminorm. If we take the quotient with respect to μ-equivalence classes of functions, we obtain an F-space $L^\varphi(\mu; X)$.

b. When $\varphi(t) = t^p$ for some number $p \in (0, +\infty)$, then $L^p(\mu; X)$ (defined as in 22.28) is equal to $L^\varphi(\mu; X)$ and ρ_φ is equivalent to $\|\ \ \|_p$.

c. The space $TM(\mu; X)$, defined in 26.12, is equal to the union of all the spaces $L^\varphi(\mu; X)$, as φ varies over all Orlicz functions (defined as above).

For a different treatment and references, see Rao and Ren [1991].

TAG's and TVS's

26.14. *Definitions.* Let X be an Abelian (i.e., commutative) group, with group operation $+$ and identity element 0. Let $\mathfrak{T}$ be a topology on the set X. We say $(X, \mathfrak{T})$ (or more simply, X) is a **topological Abelian group** — hereafter abbreviated **TAG** — if the group operations are continuous — i.e., if

$$\begin{array}{rcll} v & \mapsto & -v & \text{is continuous from } X \text{ into } X, \text{ and} \\ (u,v) & \mapsto & u+v & \text{is jointly continuous from } X \times X \text{ into } X. \end{array}$$

Along with the theory of TAG's, we shall also develop the slightly more specialized theory of TVS's. Let X be an vector space over the scalar field $\mathbb{F}$, and let $\mathfrak{T}$ be a topology on the set X. We say $(X, \mathfrak{T})$ is a **topological vector space** (or *topological linear space*) — hereafter abbreviated **TVS** — if the vector operations are jointly continuous; i.e., if

$$\begin{array}{rcll} (c,v) & \mapsto & cv & (\text{from } \mathbb{F} \times X \text{ into } X) \quad \text{and} \\ (u,v) & \mapsto & u+v & (\text{from } X \times X \text{ into } X) \end{array}$$

are both jointly continuous. Of course, every TVS is also a TAG.

We shall specialize further: A **locally convex space**— hereafter abbreviated **LCS** — is a topological vector space with the further property that 0 has a neighborhood basis consisting of convex sets.

Finally, a **Fréchet space** is an F-space that is also locally convex. (This should not be confused with a very different meaning given for "Fréchet space" in 16.7.)

Remarks. Clearly, any Banach space is also a Fréchet space. Some other examples are noted in 26.20.e.

It is immediate from 22.7 that any G-seminormed group (when equipped with the pseudometric topology) is a TAG. Similarly, any F-seminormed vector space is a TVS, and any seminormed vector space is an LCS. We shall see in 26.29 that TAG's, TVS's, and LCS's are not much more general than this.

In our study of TVS's in this and later chapters we shall distinguish between those theorems (such as 27.6) that require local convexity and those theorems (such as 27.26) that do not. However, this distinction is made chiefly for theoretical and pedagogical reasons — i.e., to make the basic concepts easier for the beginner. Although we do give a few examples of non-locally-convex TVS's in 26.16 and 26.17, we remark that most TVS's used in applications are in fact locally convex. Thus, it would be feasible to skip TVS's altogether and simply study LCS's, equipping some theorems with hypotheses that are slightly stronger than necessary; that approach is followed by some introductory textbooks on functional analysis.

Caution: Since most TVS's used in applications have Hausdorff topologies, some mathematicians incorporate the T_2 condition into their definition of TVS or LCS. In the present book, however, a topological space will be assumed Hausdorff only if that assumption is stated explicitly.

26.15. *Degenerate (but important) examples.*

a. The discrete topology. The topology consisting of all subsets of an Abelian group X is a TAG topology. However, if X is a vector space (other than the degenerate space $\{0\}$), then the discrete topology on X does *not* make it a TVS, because (*exercise*) multiplication of scalars times vectors is not jointly continuous. In fact, for fixed $x \neq 0$, the mapping $c \mapsto cx$ is not continuous at $c = 0$.

b. The indiscrete topology. The topology $\{\varnothing, X\}$ makes any Abelian group X into a TAG and any vector space X into an LCS. Of course, it is not Hausdorff (unless $X = \{0\}$).

26.16. *Example.* For $0 \leq p < 1$, the F-spaces $L^p[0,1]$ (defined in 26.12.c and 22.28, with μ equal to Lebesgue measure on $[0,1]$) are not locally convex. In fact, $L^p[0,1]$ has no open convex subsets other than $\varnothing$ and the entire space, and the space $L^p[0,1]^* = \{$continuous linear functionals on $L^p[0,1]\}$ is just $\{0\}$.

Proof. The F-space $L^p[0,1]$ is topologized by the F-norm $\rho(f) = \int_0^1 \Gamma\left(|f(t)|\right) dt$, where $\Gamma(s) = s^p$ in the cases of $0 < p < 1$, and Γ is any bounded remetrization function in the case of $p = 0$ (see 26.12.d). In particular, for $p = 0$, we may take $\Gamma(s) = s/(1+s)$; thus $\Gamma(s) \leq 1$ for all s in that case.

Suppose V is a nonempty open convex subset of $L^p[0,1]$. By translation, we may assume $0 \in V$. Since V is a neighborhood of 0, we have $V \supseteq \{f : \rho(f) < r\}$ for some number $r > 0$. Let g be any element of $L^p[0,1]$; we shall show that $g \in V$. Choose some integer n large enough so that $\rho(g) < rn^{1-p}$. Since the function $t \mapsto \int_0^t \Gamma(|g(s)|)ds$ is continuous, we can choose a partition $0 = t_0 < t_1 < t_2 < \cdots < t_n = 1$ such that $\int_{t_{j-1}}^{t_j} \Gamma(|g(s)|)ds = \frac{1}{n}\rho(g)$ for all j. Let $1_{(t_{j-1},t_j]}$ be the characteristic function of the interval $(t_{j-1}, t_j]$, and let $g_j = n 1_{(t_{j-1},t_j]} g$. An easy computation shows that

$$\rho(g_j) \;=\; \int_{t_{j-1}}^{t_j} \Gamma(n|g(s)|)ds \;\leq\; n^p \int_{t_{j-1}}^{t_j} \Gamma(|g(s)|)ds \;=\; n^{p-1}\rho(g) \;<\; r,$$

and thus $g_j \in V$. Since $g = \frac{1}{n}(g_1 + g_2 + \cdots + g_n)$ and V is convex, $g \in V$ also.

If Λ is a continuous linear functional on $L^p[0,1]$, then $\Lambda^{-1}\left(\{c : |c| < 1\}\right)$ is an open convex set containing 0. Hence it is all of $L^p[0,1]$; hence $\Lambda = 0$.

26.17. *Example.* If $0 < p < 1$, then the sequence space ℓ_p is not locally convex. In particular, $\{x : \|x\|_p < 1\}$ does not contain a convex neighborhood of 0. However, the set $(\ell_p)^* = \{$continuous linear functionals on $\ell_p\}$ is equal to ℓ_∞; this space is large enough to separate the points of ℓ_p.

Hints: For the first assertion, suppose $\{x : \|x\|_p < 1\}$ contains some convex neighborhood of 0, which we label V. Show that $V \supseteq \{x : \|x\|_p \leq s\}$ for some $s > 0$. Let e_j be the sequence that has a 1 in the jth place and 0s elsewhere. Then $se_j \in V$. By convexity, $v_n = \frac{1}{n}(se_1 + se_2 + \cdots + se_n) \in V$ for any positive integer n. However, show that $\|v_n\|_p > 1$ for n sufficiently large.

Any $y \in \ell_\infty$ acts as a continuous linear functional on ℓ_p, by the action $\langle x, y\rangle = \sum_{j=1}^\infty x_j y_j$; in fact, we have $\sum_j |x_j y_j| \leq \|x\|_1 \|y\|_\infty \leq \|x\|_p \|y\|_\infty$. Conversely, if $\varphi \in (\ell_p)^*$, let e_j be the sequence with 1 in the jth place and 0 elsewhere. Define a sequence $y = (y_j)$ by taking $y_j = \varphi(e_j)$. Then $|y_j| \leq |||\varphi||| \, \|e_j\|_p = |||\varphi|||$; thus y is bounded. The functionals $\langle \cdot, y \rangle$ and φ are continuous on ℓ_p, and they act the same on sequences with only finitely many terms. Such sequences are dense in ℓ_p, so $\langle \cdot, y \rangle = \varphi$ on ℓ_p.

26.18. Net characterizations of TAG's and TVS's. Let X be an Abelian group, equipped with some topology. Then X is a TAG if and only if its topology satisfies these two conditions:

(1) Whenever (x_α, y_α) is a net in $X \times X$ satisfying $x_\alpha \to x$ and $y_\alpha \to y$, then $x_\alpha + y_\alpha \to x + y$.

(2) Whenever (x_α) is a net in X satisfying $x_\alpha \to x$, then $-x_\alpha \to -x$.

More specifically, let X be a vector space, equipped with some topology. Then X is a TVS if and only if its topology satisfies conditions (1) and

(2′) Whenever (c_α, x_α) is a net in $\mathbb{F} \times X$ satisfying $c_\alpha \to c$ and $x_\alpha \to x$, then $c_\alpha x_\alpha \to cx$.

26.19. Initial object constructions of TAG's, TVS's, and LCS's. Let X be a set, let $\{(Y_\lambda, \mathcal{T}_\lambda) : \lambda \in \Lambda\}$ be a collection of topological spaces, let $\varphi_\lambda : X \to Y_\lambda$ be some mappings, and let $\mathcal{S}$ be the initial topology determined on X by the φ_λ's and $\mathcal{T}_\lambda$'s — i.e., the weakest topology on X that makes all the φ_λ's continuous (see 9.15). Show that

(i) If X is a group, the $(Y_\lambda, \mathcal{T}_\lambda)$'s are TAG's, and the φ_λ's are additive maps, then $(X, \mathcal{S})$ is a TAG.

(ii) If X is a vector space, the $(Y_\lambda, \mathcal{T}_\lambda)$'s are TVS's, and the φ_λ's are linear maps, then $(X, \mathcal{S})$ is a TVS.

(iii) If X is a vector space, the $(Y_\lambda, \mathcal{T}_\lambda)$'s are LCS's, and the φ_λ's are linear maps, then $(X, \mathcal{S})$ is an LCS.

Hints: 15.14(A), 15.24, and 26.18.

26.20. *Some important special cases of initial objects.*

a. The *product* of any collection of TAG's or TVS's or LCS's, with the product topology and product algebraic structure, is a TAG or TVS or LCS.

b. Subspace topologies are initial topologies determined by inclusion maps (see 5.15.e and 9.20). Thus, any *subgroup* of a TAG is also a TAG; and a *linear subspace* of a TVS or LCS is another TVS or LCS.

c. The *supremum*, or *least upper bound*, of a collection of topologies is the weakest topology that includes all the given topologies (see 5.23.c); it is the initial topology given by identity maps. Thus, the sup of a collection of TAG or TVS or LCS topologies is another TAG or TVS or LCS topology.

d. Let D be a collection of G-seminorms on an Abelian group X, or a collection of F-seminorms or seminorms on a vector space X. Then the gauge topology determined on

X by D is a TAG, TVS, or LCS topology, respectively. (*Hints*: As we noted in 5.23.c, the topology determined by a gauge is the supremum of the individual pseudometric topologies. It follows easily from 15.25.c and 26.18 that any sup of TAG or TVS topologies is a TAG or TVS topology.)

(A converse to this result will be given in 26.29.)

e. Let φ be a Fréchet combination of φ_j's on X (as in 26.6), and suppose that each φ_j is actually a seminorm (i.e., it is homogeneous). Then the F-normed space (X, φ) is locally convex. If it is complete, then it is a Fréchet space. These conditions are satisfied by the examples in the next few sections after 26.6.

26.21. *Change of scalar field.* Let X be a complex vector space. Then X may also be viewed as a real vector space (if we "forget" how to multiply members of X by members of $\mathbb{C} \setminus \mathbb{R}$).

Let $\mathfrak{T}$ be some topology on the set X. It is easy to show that if the complex vector space X is a TVS, then the real vector space X is also a TVS. The converse of that implication is false, however, as we now show:

A pathological example. Let $X = \{$bounded functions from $[1, +\infty)$ into $\mathbb{C}\}$. That is a complex vector space, with vector addition and scalar multiplication both defined pointwise on $[1, +\infty)$ For $f \in X$, define

$$\|f\| \quad = \quad \sup_{1 \le t < \infty} \left\{ |\operatorname{Re} f(t)| + \frac{1}{t} |\operatorname{Im} f(t)| \right\}.$$

Verify that $(X, \| \quad \|)$ is a Banach space, when we use the *real numbers* for the scalar field.

However, let f_n be the characteristic function of the interval $[n, n+1]$. Verify that $\|f_n\| = 1$ while $\|if_n\| = \frac{1}{n}$. Conclude that the topology of the Banach space $(X, \| \quad \|)$ does not make scalar multiplication jointly continuous from $\mathbb{C} \times X$ into X; hence (i) that topology does not make X into a *complex* topological vector space, and (ii) $\| \quad \|$ is not a norm on the complex vector space.

Arithmetic in TAG's and TVS's

26.22. *Arithmetic in TAG's.* Let X be a TAG, and let S, T be nonempty subsets of X. Show that

a. If S is symmetric (i.e., if $-S = S$), then $\operatorname{cl}(S)$ and $\operatorname{int}(S)$ are symmetric.

b. If S is open, then $S + T$ is open regardless of what T is.

c. If S and T are closed and S is compact, then $S + T$ is closed.

d. $\operatorname{int}(S) + \operatorname{int}(T) \subseteq \operatorname{int}(S) + T \subseteq \operatorname{int}(S + T)$.

e. $\operatorname{cl}(S) + \operatorname{cl}(T) \subseteq \operatorname{cl}(S + T)$.

Hint: If $(s_\alpha : \alpha \in A)$ is a net in S converging to x and $(t_\beta : \beta \in B)$ is a net in T converging to y, then $(s_\alpha + t_\beta : (\alpha, \beta) \in A \times B)$ is a net converging to $x + y$, where $A \times B$ has the product ordering.

f. If S is a subgroup of X, then $\mathrm{cl}(S)$ is also a subgroup, and $\mathrm{int}(S)$ is either a subgroup or empty.

26.23. *Arithmetic in TVS's.* Suppose X is a TVS and $S \subseteq X$. Then:

a. If S is balanced or convex or absolutely convex, then $\mathrm{cl}(S)$ has the same property, and so does $\mathrm{int}(S)$ if it is not empty.

b. If S is a linear subspace of X, then $\mathrm{cl}(S)$ is a linear subspace of X, and $\mathrm{int}(S)$ is either $\varnothing$ or all of X.

c. The **closed convex hull** of S, denoted here by $\mathrm{clco}(S)$, is the smallest closed convex set containing S; it is the intersection of all the closed convex sets that contain S. Show that it is also equal to the closure of the convex hull of S — that is, $\mathrm{clco}(S) = \mathrm{cl}(\mathrm{co}(S))$.

d. If S is convex, then S is connected. In particular, X itself is connected.

e. A set $S \subseteq X$ is called **midpoint convex** if $S \supseteq \frac{1}{2}S + \frac{1}{2}S$ — that is, if S contains the midpoint of any line segment whose endpoints are both in S. Clearly, any convex set is midpoint convex. Show that any *closed*, midpoint convex set is convex.

Example. The rational numbers form a subset of the reals that is midpoint convex but not convex.

f. The convex hull of a finite set is compact. More generally, if $A_1, A_2, \ldots, A_n$ are compact convex subsets of X, then $\mathrm{co}(A_1 \cup A_2 \cup \cdots \cup A_n)$ is compact. If the A_j's are also balanced, then so is $\mathrm{co}(A_1 \cup A_2 \cup \cdots \cup A_n)$.

Hints: First show that

$$C \quad = \quad \{(c_1, c_2, \ldots, c_n) \in [0,1]^n \ : \ c_1 + c_2 + \cdots + c_n = 1\}$$

is closed and bounded, hence a compact subset of $\mathbb{R}^n$. Then define $g : C \times A_1 \times A_2 \times \cdots \times A_n \to X$ by taking $g(c, a_1, a_2, \ldots, a_n) = c_1 a_1 + c_2 a_2 + \cdots + c_n a_n$. Show that $\mathrm{co}(A_1 \cup A_2 \cup \cdots \cup A_n) = \mathrm{Range}(g)$; then use the fact that the continuous image of a compact set is compact (see 17.7.h).

g. If K is a compact subset of $\mathbb{F}^n$, then the convex hull of K is also compact. *Hint*: Use Carathéodory's Theorem 12.10 and the continuity of the vector space operations.

h. (*Optional.*) The preceding result is *only* true in finite dimensions. If X is any infinite-dimensional F-space, then there exists a compact set $K \subseteq X$ whose convex hull is not closed and hence not compact.

Proof. Let ρ be a complete F-norm on X. Let $(x_n : n = 1, 2, 3, \ldots)$ be a linearly independent sequence of vectors in X. Replacing the x_n's with suitable scalar multiples, we may assume $x_n \to 0$ as $n \to \infty$. Then the set $K = \{x_0, x_1, x_2, x_3, \ldots\}$ is compact, where $x_0 = 0$. If $c_1, c_2, c_3, \ldots$ are positive numbers with sum less than 1, then the partial sums $s_n = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$ belong to the convex hull of K, hence $s = \lim_{n \to \infty} s_n$ (if it exists) belongs to the closed convex hull of K. If we choose the

c_j's small enough so that $\rho(c_j x_j) < 2^{-j}$, then s does indeed exist, by the completeness of ρ. Let "conv" denote convex hull; we shall choose the c_j's so that $s \notin \mathrm{conv}(K)$.

For each positive integer n, since c_n is positive, the vector s_n does not lie in the linear span of $x_0, x_1, x_2, \ldots, x_{n-1}$. Hence it does not lie in their convex hull, which is a compact set. Thus $r_n = \mathrm{dist}(s_n, \mathrm{conv}\{x_0, x_1, x_2, \ldots, x_{n-1}\})$ is a positive number. Note that the definition of r_n depends only on the choices of $c_1, c_2, \ldots, c_{n-1}, c_n$. Thus we may choose the c_n's and r_n's recursively: choose each c_{n+1} small enough to satisfy

$$\rho(c_{n+1}x_{n+1}) \quad < \quad 2^{-n-1} \min\{1, r_0, r_1, r_2, \ldots, r_n\}.$$

It follows that

$$\rho(s - s_n) \quad = \quad \rho(c_{n+1}x_{n+1} + c_{n+2}x_{n+2} + \cdots) \quad < \quad \sum_{j>n} 2^{-j} r_n \quad < \quad r_n.$$

Hence $\mathrm{dist}(s, \mathrm{conv}\{x_0, x_1, x_2, \ldots, x_{n-1}\}) > 0$. By 12.5.d, we have

$$\mathrm{conv}(K) \;=\; \bigcup_{n=1}^{\infty} \mathrm{conv}\{x_0, x_1, x_2, \ldots, x_{n-1}\},$$

so $s \notin \mathrm{conv}(K)$.

i. Show that in a locally convex space, the convex hull of a totally bounded set is totally bounded. Then use that to prove **Mazur's Theorem**: In a Banach space (or, more generally, in a complete LCS), the closed convex hull of a compact set is compact. (A still more general result is given in 27.3.f.)

j. Suppose X and Y are TVS's with scalar field $\mathbb{R}$, and $f : X \to Y$ is additive — that is, $f(x_1 + x_2) = f(x_1) + f(x_2)$. Also assume f is continuous. Then f is linear.

k. Let S be a convex neighborhood of 0, and let ρ be its Minkowski functional (defined in 12.27). Then ρ is continuous. If S is open, then $S = \{x \in X : \rho(x) < 1\}$.

Neighborhoods of Zero

26.24. *Discussion.* If $(X, \mathcal{T})$ is a TAG, then the topology $\mathcal{T}$ is **translation-invariant**. That is, for any $S \subseteq X$ and $x, x' \in X$, we have $S \in \mathcal{T} \iff x + S \in \mathcal{T} \iff x' + S \in \mathcal{T}$.

We saw in 5.22 that any topology $\mathcal{T}$ can be characterized in terms of its neighborhood filters $\mathcal{N}(x)$. If $\mathcal{T}$ is translation-invariant, then the neighborhood filters can be translated as well:

$$x + S \in \mathcal{N}(x) \quad \iff \quad x' + S \in \mathcal{N}(x') \quad \iff \quad S \in \mathcal{N}(0).$$

Thus, the topology is determined by its neighborhoods of 0. Some simplifications are possible at 0, so usually we just work with the neighborhood filter at 0. Likewise, we consider neighborhood bases at 0.

26.25. Neighborhood characterization of a TAG. Let X be an Abelian group, and let $\mathcal{N}$ be a proper filter on X. Then there exists a TAG topology $\mathcal{T}$ on X for which $\mathcal{N}$ is the filter of all neighborhoods of 0, if and only if $\mathcal{N}$ has these properties:

(i) For each $N \in \mathcal{N}$, there is some $A \in \mathcal{N}$ satisfying $A + A \subseteq N$.

(ii) $\mathcal{N}$ has a filterbase consisting of symmetric sets — i.e., each $N \in \mathcal{N}$ contains some $B \in \mathcal{N}$ that satisfies $-B = B$.

Moreover, in that case the topology $\mathcal{T}$ consists of the collection of all sets $T \subseteq X$ that satisfy the following condition.

$$\text{For each } x \in T, \text{ there is some } G \in \mathcal{N} \text{ such that } x + G \subseteq T.$$

The interior operator of that topology can be expressed as follows: For each $S \subseteq X$, we have

$$\operatorname{int}(S) \quad = \quad \{x \in X \ : \ x + G \subseteq S \text{ for some } G \in \mathcal{N}\}.$$

Hints: If X is a TAG with neighborhood filter $\mathcal{N}$ at 0, then (i) follows from continuity of addition; for (ii) take $B = N \cap (-N)$. Conversely, if (i) and (ii) hold, use 5.22 to show that $\mathcal{T}$, as defined above, is a topology, etc.

26.26. Neighborhood characterization of a TVS. Let X be a vector space, and let $\mathcal{N}$ be a proper filter on X. Then there exists a TVS topology $\mathcal{T}$ on X for which $\mathcal{N}$ is the filter of all neighborhoods of 0, if and only if $\mathcal{N}$ has these properties:

(i) For each $N \in \mathcal{N}$, there is some $A \in \mathcal{N}$ satisfying $A + A \subseteq N$.

(ii) $\mathcal{N}$ has a filterbase consisting of balanced sets — i.e., each $N \in \mathcal{N}$ contains some $B \in \mathcal{N}$ that is balanced (as defined in 12.3).

(iii) Every member of $\mathcal{N}$ is absorbing (as defined in 12.8).

Hints for the "only if" part: (i) X is a TAG. (ii) By continuity of scalar multiplication there is some $r > 0$ and some neighborhood G of 0 such that $|c| \leq r, x \in G \ \Rightarrow \ cx \in N$. Then $B = \bigcup_{|c| \leq r} cG$ is a balanced neighborhood of 0 that is contained in N. (iii) Continuity of multiplication.

Hints for the "if" part: First, define $\mathcal{T}$ as in 26.25, and show that $(X, \mathcal{T})$ is a TAG.

To prove continuity of $x \to cx$ for fixed $c \in \mathbb{F}$, it suffices to show continuity at $x = 0$, and we may assume $c \neq 0$. Let any balanced $B \in \mathcal{N}$ be given; it suffices to show $c^{-1}B \in \mathcal{N}$. By repeated uses of (i), there is some $H \in \mathcal{N}$ such that $mH \subseteq B$ for some integer $m > |c|$. Then $H \subseteq m^{-1}B \subseteq c^{-1}B$.

Continuity of $c \to cx$ follows from (iii); joint continuity of $(c, x) \to cx$ at (0,0) follows from (ii). Finally, joint continuity of $(c, x) \to cx$ now follows from the fact that that mapping is bilinear — if $c_\alpha \to c$ and $x_\alpha \to x$, then

$$c_\alpha x_\alpha - cx \ = \ (c_\alpha - c)x + c(x_\alpha - x) + (c_\alpha - c)(x_\alpha - x) \ \to \ 0.$$

26.27. *Further properties of neighborhood bases.* Show that

a. If X is a TAG, then $\mathrm{cl}(S) = \bigcap_{N \in \mathcal{N}}(S+N)$ for any set $S \subseteq X$, if $\mathcal{N}$ is the neighborhood filter at 0. Hence $\mathrm{cl}(S) \subseteq S + N$ for any $N \in \mathcal{N}$.

Hint: $x \in \mathrm{cl}(S)$ if and only if S meets every neighborhood of x, as we noted in 15.4.

b. If X is a TAG, then X has a neighborhood base at 0 consisting of open symmetric sets and a neighborhood base at 0 consisting of closed symmetric sets.

Hint: If N is a neighborhood of 0, choose some open neighborhood G of 0 such that $G + G \subseteq N$; then show that $G \cap (-G)$ and its closure are symmetric neighborhoods of 0 contained in N.

c. If X is a TVS, then X has a neighborhood base at 0 consisting of open balanced sets and a neighborhood base at 0 consisting of closed balanced sets. *Hint*: 26.23.a.

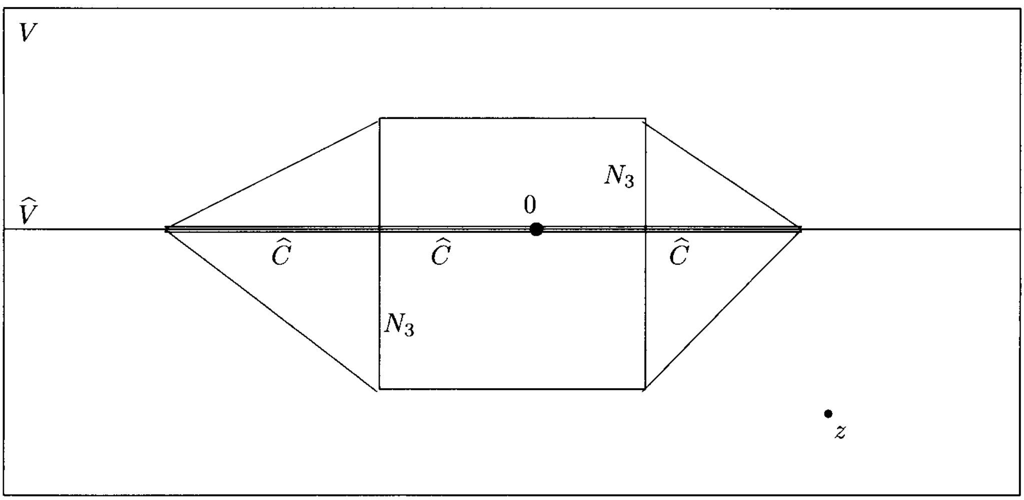
d. If X is an LCS, then X has a neighborhood base at 0 consisting of open convex balanced sets and a neighborhood base at 0 consisting of closed convex balanced sets.

26.28. *Technical lemma on subspaces* (*Dieudonné-Schwartz*). (This result will be needed in 27.41.b.)

Let V be a Hausdorff LCS. Let $\widehat{V}$ be a closed linear subspace of V, equipped with the relative topology. Let $\widehat{C}$ be a convex neighborhood of 0 in $\widehat{V}$, and let $z \in V \setminus \widehat{V}$.

Then there exists a convex neighborhood C of 0 in V such that $C \cap \widehat{V} = \widehat{C}$ and $z \notin C$.

Proof. In the following diagram, V is represented by the entire plane, and $\widehat{V}$ by a horizontal line through that plane. The set $\widehat{C}$ is represented by a line segment in that line. This line segment is drawn slightly thicker only to make it visible in the diagram; it should be interpreted as being no thicker than $\widehat{V}$.



By assumption, $V \setminus \widehat{V}$ is a neighborhood of z that does not meet $\widehat{V}$. Then $N_1 = -(V \setminus \widehat{V}) + z$ is a neighborhood of 0 such that $z - N_1$ does not meet $\widehat{V}$.

By definition of the relative topology, since $\widehat{C}$ is a neighborhood of 0 in $\widehat{V}$, we have $\widehat{C} = \widehat{V} \cap N_2$ where N_2 is some neighborhood of 0 in V.

Now, $N_1 \cap N_2$ is a neighborhood of 0 in V. Since V is locally convex, $N_1 \cap N_2$ contains some convex neighborhood N_3 of 0. Observe that $\widehat{V} \cap N_3 \subseteq \widehat{C}$ and $(z - N_3) \cap \widehat{V} = \varnothing$. In the diagram, the set N_3 is represented by a square.

Now let

$$C \quad = \quad \mathrm{co}(\widehat{C} \cup N_3) \quad = \quad \left\{\lambda u + (1-\lambda)v \ : \ \lambda \in [0,1],\ u \in \widehat{C},\ v \in N_3\right\}$$

where "co" denotes convex hull (see 12.6.c). We leave it as an exercise to verify that $C \cap \widehat{V} = \widehat{C}$ and $z \notin C$.

Characterizations in Terms of Gauges

26.29. Theorem (Birkhoff, Kakutani, and Minkowski). Let $\mathcal{T}$ be a topology on an Abelian group X. Then (i) $\mathcal{T}$ is a TAG topology if and only if $\mathcal{T}$ is the gauge topology determined by some gauge consisting of translation-invariant pseudometrics. In other words, a TAG topology is the same thing as a topology given by a collection of G-seminorms.

Suppose that X is a vector space. Then (ii) $\mathcal{T}$ is a TVS topology if and only if $\mathcal{T}$ is the gauge topology determined by some gauge consisting of F-seminorms, and (iii) $\mathcal{T}$ is a LCS topology if and only if $\mathcal{T}$ is the gauge topology determined by some gauge consisting of seminorms.

Proof. A proof of the "if" parts was sketched in 26.20.d; it remains for us to prove the "only if" parts. Suppose $\mathcal{T}$ is a TAG, TVS, or LCS topology; we shall find appropriate G-seminorms.

Case (iii) is easiest: Let $\mathcal{B}$ be a neighborhood base at 0 consisting of convex, balanced sets. Being a neighborhood of 0, each $B \in \mathcal{B}$ is also absorbing. Hence its Minkowski functional μ_B is a seminorm on X (see 12.29.g). It is easy to verify that the seminorms μ_B give the same topology as the neighborhood base $\mathcal{B}$.

For case (i), let $\mathcal{B}$ be a neighborhood base at 0 consisting of symmetric sets; for case (ii) we may assume the members of $\mathcal{B}$ are balanced sets. Temporarily fix any $B \in \mathcal{B}$, and let $B_1 = B$. Since addition is continuous, we may recursively choose $B_2, B_3, B_4, \ldots$ in $\mathcal{B}$ with $B_n \supseteq B_{n+1} + B_{n+1} + B_{n+1}$. Also let $B_0 = X$. Define relations $V_n \subseteq X \times X$ by $V_n = \{(x,y) \in X \times X : x - y \in B_n\}$. Then the V_n's are translation-invariant in this sense: $(x,y) \in V_n \iff (x+u, y+u) \in V_n$. Also, the V_n's satisfy the hypotheses of Weil's Pseudometrization Lemma 4.44. Define a pseudometric d as in that lemma; then it is also translation-invariant and thus defines a G-seminorm ρ. In fact, we have $f(x,y) = \varphi(x-y)$, where $\varphi(u) = \inf\{2^{-n} : u \in B_n\}$; hence $\rho(x) = \inf \sum_{j=1}^{m} \varphi(x_j)$ where the infimum is over all positive integers m and all decompositions $x = x_1 + x_2 + \cdots + x_m$. Since the sets B_n are symmetric or balanced, it is easy to verify that the functions φ and ρ are also symmetric or balanced. In case (ii), since the B_n's are balanced neighborhoods of 0, it is easy to verify that ρ is an F-seminorm. By Weil's Lemma, we have

$$\{x \in X \ : \ \rho(x) < 2^{-n}\} \quad \subseteq \quad B_n \quad \subseteq \quad \{x \in X \ : \ \rho(x) \leq 2^{-n}\}$$

for all positive integers n. Recall that $B_1 = B$; let us denote $\rho = \rho_B$. Then a net (x_α) in X satisfies $\rho_B(x_\alpha) \to 0$ for each $B \in \mathcal{B}$ if and only if for each $B \in \mathcal{B}$ we have eventually $x_\alpha \in B$. Thus the topology with neighborhood base at 0 given by $\mathcal{B}$ is the same as the topology given by the ρ_B's.

26.30. *Remark.* Although 16.16 and 26.29 show that every TVS is completely regular, in general a TVS need not have stronger separation properties such as normality. Indeed, an example of Stone [1948] showed that the Hausdorff locally convex space $\mathbb{R}^{\mathbb{R}}$ is not normal.

26.31. *Compatibility with a TAG or TVS.* Let $(X, \mathcal{T})$ be a TAG, TVS, or an LCS, and let $R = \{\rho_\lambda : \lambda \in \Lambda\}$ be a gauge on X consisting of G-seminorms, F-seminorms, or seminorms, respectively. We shall say that the gauge R is **compatible** with the topology $\mathcal{T}$ if $\mathcal{T}$ is the topology determined by R — that is, if

$$x_\alpha \to x \text{ in } (X, \mathcal{T}) \qquad \Longleftrightarrow \qquad \rho_\lambda(x_\alpha - x) \to 0 \text{ for each } \lambda.$$

Show that the largest gauge that is compatible with $\mathcal{T}$ is, respectively, the set of all G-seminorms, F-seminorms, or seminorms that are continuous from $(X, \mathcal{T})$ to $[0, +\infty)$. (This is a specialization of 16.20.)

26.32. Pseudometrizability criteria. Let $(X, \mathcal{T})$ be a TAG. Show that the following are equivalent:

(A) $\mathcal{T}$ is pseudometrizable.

(B) $\mathcal{T}$ can be determined by a countable collection of pseudometrics.

(C) $\mathcal{T}$ can be given by a single G-seminorm. *Hint*: 18.17.

(D) $(X, \mathcal{T})$ is first countable — i.e., it has a *countable neighborhood base at* 0 — i.e., there is a countable collection $\mathcal{N}_0$ of neighborhoods of 0, such that every neighborhood of 0 includes some element of $\mathcal{N}_0$.

Suppose, moreover, that $(X, \mathcal{T})$ is a TVS. Then the following is also equivalent:

(E) $\mathcal{T}$ can be given by a single F-seminorm.

Finally, suppose $(X, \mathcal{T})$ is an LCS. Then the following are also equivalent:

(F) $\mathcal{T}$ can be given by a countable collection of seminorms.

(G) $\mathcal{T}$ can be given by a single F-seminorm of the form

$$\rho \quad = \quad \sum_{n=1}^{\infty} 2^{-n} \arctan \rho_n,$$

where the ρ_n's are seminorms.

Remarks. We emphasize that the topology of a pseudometrizable LCS is *not* necessarily obtainable from a single seminorm. An example is given in 27.8.

Some readers will be more interested in the metrizability of a TAG, TVS, or an LCS rather than the pseudometrizability. But of course, a space is metrizable if and only if it is pseudometrizable and Hausdorff; we consider Hausdorffness in the next section.

26.33. Hausdorffness criteria. Let X be a TAG; let $\mathcal{N}$ be the neighborhood filter at 0; let R be any determining family of G-seminorms. Then the following are equivalent:

(A) $\{0\}$ is closed.

(B) $\bigcap_{N\in\mathcal{N}} N = \{0\}$.

(C) $\bigcap_{\rho\in R} \rho^{-1}(0) = \{0\}$.

(D) X is T_0 (i.e., the topology can distinguish between points of X).

(E) X is T_1 (i.e., every point is closed).

(F) X is T_2 (i.e., Hausdorff).

(G) X is T_3.

(H) X is $T_{3.5}$ (i.e., Tychonov).

26.34. *Corollary on quotients.* Suppose X is a TAG, TVS, or an LCS. Let $K \subseteq X$ be a subgroup, linear subspace, or linear subspace, respectively. Let X/K have the quotient topology. Then:

a. X/K is also a TAG, TVS, or an LCS, respectively.

Hint: The topology on X is determined by a gauge D consisting of G-seminorms, F-seminorms, or seminorms, respectively. Replacing D with an equivalent gauge, we may assume D is directed. The gauge $\hat{D}$ on X/K, defined as in 22.13.e, also consists of G-seminorms, F-seminorms, or seminorms, respectively.

b. The topology on Q is Hausdorff if and only if K is a closed subset of X (regardless of whether the topology on X is Hausdorff).

Hints: Let 0_Q denote the additive identity of Q. Refer to 26.33. Then Q is Hausdorff $\iff$ $\{0_Q\}$ is a closed set $\iff$ $K = \pi^{-1}(0_Q)$ is a closed set, by 15.31.b.

c. In particular, if $K = \mathrm{cl}(\{0\})$, then Q is the Kolmogorov quotient of X, defined in 16.5.

26.35. *A few properties of the Kolmogorov quotient.* Suppose that X is a topological Abelian group. Let $Q = X/\mathrm{cl}(\{0\})$ have the quotient topology. Then:

a. Q is a Hausdorff space.

b. Q is the Kolmogorov quotient of X.

c. For each G-seminorm ρ on X, define the corresponding G-seminorm $\widehat{\rho}$ on Q, as in 22.13. The G-seminorms that determine the topology and uniformity of X are all continuous, and so they satisfy $\rho^{-1}(0) \supseteq \mathrm{cl}(\{0\})$. By 22.13.d, the formula for $\widehat{\rho}$ simplifies to $\widehat{\rho}(\pi(x)) = \rho(x)$.

d. Let $\pi : X \to X/\mathrm{cl}(\{0\})$ be the quotient map. Let $\sigma : X/\mathrm{cl}(\{0\}) \to X$ be any selection of π^{-1} — that is, let σ be any function that satisfies $\sigma(q) \in \pi^{-1}(q)$ for all $q \in X/\mathrm{cl}(\{0\})$. Then σ is continuous.

Proof. Let (q_α) be a net converging to a limit q in $X/\mathrm{cl}(\{0\})$. Show that $\rho(\sigma(q_\alpha) - \sigma(q)) = \widehat{\rho}(q_\alpha - q) \to 0$.

e. The quotient map $\pi : X \to Q$ is open, closed, and continuous.

Uniform Structure of TAG's

26.36. *Preliminary lemmas.* Let X and Y be TAG's. By 26.29, the topologies of X and Y can be determined (not necessarily in a unique fashion) by gauges D and E consisting of G-seminorms — i.e., consisting of translation-invariant pseudometrics. Fix any such gauges D and E, and let X and Y be equipped with the uniform structure determined by those gauges. Then:

a. Let $((x_\alpha, x'_\alpha) : \alpha \in \mathbb{A})$ be a net in $X \times X$. Then $D(x_\alpha, x'_\alpha) \to 0$ in X in the sense of 18.7 if and only if $x_\alpha - x'_\alpha \to 0$ in X.

b. A function $f : X \to Y$ is uniformly continuous if and only if it has this property:

> Whenever $((x_\alpha, x'_\alpha) : \alpha \in \mathbb{A})$ is a net in $X \times X$ that satisfies $x_\alpha - x'_\alpha \to 0$ in X, then also $f(x_\alpha) - f(x'_\alpha) \to 0$ in Y;

or, equivalently, this property:

> For each neighborhood H of 0 in Y, there is some neighborhood G of 0 in X such that $x - x' \in G \Rightarrow f(x) - f(x') \in H$.

c. Any additive continuous map $f : X \to Y$ is uniformly continuous.

26.37. *Discussion: uniqueness of the uniformity.* Let $(X, \mathcal{T})$ be a topological Abelian group (TAG). As we have noted above, the topology $\mathcal{T}$ can be determined by some gauge consisting of G-seminorms. Such a collection also determines a uniformity on X. The gauge is not necessarily unique, but we can now see that the uniformity *is* unique; any two such gauges must determine the same uniformity. (*Proof.* Apply 26.36.c to the identity map.)

That unique uniformity will be called the **usual uniformity** for the topological group. It will always be understood to be in use whenever a topological Abelian group is viewed as a uniform space (unless some other arrangement is specified). It will also be in use for special kinds of TAG's — e.g., for TVS's and LCS's. Note that

> *on an Abelian group, a TAG topology and its associated usual uniformity determine each other uniquely. Consequently, we may refer to them interchangeably in discussions.*

For instance, we might say something like "the set S is a totally bounded subset of X, when X is equipped with the topology of uniform convergence on members of $\mathcal{S}$." Here we are really referring to the uniformity, not the topology, of uniform convergence on members of

$\mathcal{S}$, but in certain parts of the literature it seems to be customary to call this a "topology." No harm is done by this abuse of terminology, since the topology and uniformity determine each other uniquely.

26.38. *Remarks: nonuniqueness of the topology corresponding to the group structure.* The result developed above, on the uniqueness of the usual uniformity for a TAG, must be read carefully. It does *not* say that there is only one uniformity compatible with the given topology, nor that there is only one translation-invariant uniformity. Even if we restrict our attention to the topologies and uniformities given by translation-invariant gauges, an Abelian group X may be made into a TAG in more than one way — i.e., there may be several different pairs

$$(\mathcal{T}_1, \mathcal{U}_1),\ (\mathcal{T}_2, \mathcal{U}_2),\ (\mathcal{T}_3, \mathcal{U}_3),\ \ldots$$

consisting of a topology $\mathcal{T}_j$ that makes X into a TAG and the associated uniformity $\mathcal{U}_j$.

We illustrate this by mentioning three different uniformities on $\mathbb{R}$:

- The translation-invariant metric $d(x,y) = |x-y|$ yields a translation-invariant (and complete) uniformity on $\mathbb{R}$ and the usual topology.
- The metric $d(x,y) = |\arctan(x) - \arctan(y)|$ is not translation-invariant. It yields the usual (translation-invariant) topology on $\mathbb{R}$, but it yields a uniformity that is not translation-invariant or complete. (With this uniformity, the completion of $\mathbb{R}$ is the extended real line $[-\infty, +\infty]$.)
- The discrete metric on $\mathbb{R}$ is translation-invariant and yields a translation-invariant and complete uniformity. It yields, not the usual topology on $\mathbb{R}$, but rather the discrete topology.

It is also possible to develop a theory of topological groups that are *not* necessarily Abelian, but that theory is more complicated. A topological group that is not Abelian does not necessarily have one "preferred" uniformity, analogous to that discussed in 26.37. Examples can be found in Wilansky [1970], and in books on uniform spaces. We shall not pursue that topic here.

26.39. *Remarks: irrelevance of scalars.* A uniformity is unchanged if we replace the gauge with any uniformly equivalent gauge. In a TAG we can choose the gauge to consist of G-seminorms. In a TVS or an LCS we can do better: We can choose the gauge to consist of F-seminorms or seminorms. However, these "better" gauges do not give us more insight into the uniform structure. In the basic theory developed below, we can forget about multiplication by scalars, for it has no effect on the uniform structure; we can view our TVS's and LCS's as TAG's. (Nevertheless, the uniform structure and the operation of scalar multiplication do interact in some interesting ways; see 27.2.)

26.40. *Further properties of the usual uniformity.* Let X be a TAG, let $\mathcal{N}$ be the neighborhood filter at 0, and let $\mathcal{U}$ be the usual uniformity on X. Then:

a. The usual uniformity can be described directly in terms of the topology, as follows:

$$\mathcal{U} \quad = \quad \{S \subseteq X \times X \quad : \quad S \supseteq E_N \text{ for some } N \in \mathcal{N}\},$$

where

$$E_N \quad = \quad \{(x,y) \in X \times X \quad : \quad x - y \in N\}$$

for each neighborhood $N \in \mathcal{N}$. The sets E_N then form a filterbase for the uniformity.

b. A net $(x_\alpha : \alpha \in \mathbb{A})$ in X is Cauchy if and only if, for each neighborhood N of 0, there is some $\alpha_0 \in \mathbb{A}$ such that $\alpha, \alpha' \succcurlyeq \alpha_0 \ \Rightarrow \ x_\alpha - x_{\alpha'} \in N$.

A filter $\mathfrak{F}$ on X is Cauchy if and only if, for each neighborhood N of 0, there is some $F \in \mathfrak{F}$ satisfying $F - F \subseteq N$.

c. Let $(f_\alpha : \alpha \in \mathbb{A})$ be a net of functions from some set S into X, and let $f \in X^S$ also. Then $f_\alpha \to f$ uniformly on S if and only if for each neighborhood N of 0 there is some $\alpha_0 \in \mathbb{A}$ such that $\{f_\alpha(s) - f(s) \ : \ \alpha \succcurlyeq \alpha_0, \ s \in S\} \subseteq N$.

d. Let Ω be a topological space, and let $\{f_\lambda : \lambda \in \Lambda\}$ be a collection of functions from Ω into X. Then $\{f_\lambda\}$ is equicontinuous at a point $\omega_0 \in \Omega$ if and only if $\{f_\lambda\}$ has this property:

> For each neighborhood N of 0 in X, there is some neighborhood G of ω_0 in Ω such that $\{f_\lambda(\omega_0) - f_\lambda(\omega) : \lambda \in \Lambda, \omega \in G\} \subseteq N$.

e. Let $S \subseteq X$. Then S is totally bounded if and only if S has this property:

> For each neighborhood N of 0, there is some finite set $F \subseteq X$ (or, equivalently, some finite set $F \subseteq S$) such that $F + N \supseteq S$.

26.41. (*Optional.*) Most of the results on Riemann and Henstock integrals in Chapter 24 require norms, but the definitions and a few basic properties do not actually require norms. The definitions would make as much sense in any topological vector space X, if we replace

$$\text{for each number } \varepsilon > 0, \text{ there exists} \ldots \text{ such that} \ldots \ \|v - \Sigma[f,T]\| < \varepsilon$$

with

$$\text{for each neighborhood } N \text{ of } 0, \text{ there exists} \ldots \text{ such that} \ldots \ v - \Sigma[f,T] \in N.$$

As an exercise, readers may wish to prove the following result.

Theorem. Suppose the topological vector space X is locally convex, and assume it is complete — i.e., every Cauchy net in X converges. Then any continuous function $f : [a,b] \to X$ is Riemann integrable. (*Hint*: Any continuous function on $[a,b]$ is uniformly continuous.)

Pontryagin Duality and Haar Measure (Optional; Proofs Omitted)

26.42. *Remarks.* We now state a few further results about topological Abelian groups. We shall omit the proofs, which are not short or elementary, since these results will not be needed later in this book except in some other material marked "optional."

26.43. *Definitions.* By a **Pontryagin group** we shall mean a locally compact, Hausdorff, topological Abelian group. Some examples of Pontryagin groups are:

- $\mathbb{R}$ or $\mathbb{C}$, with the usual topology and with addition for the group operation.
- $(0, +\infty)$, with multiplication for the group operation. (This is isomorphic to $\mathbb{R}$.)
- $\mathbb{Z}$, with the usual (i.e., discrete) topology and with addition for the group operation.
- $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, with multiplication for the group operation. This group will play a special role in the theory developed below. (Of course, it is isomorphic to the group $[0, r)$ with the operation of addition modulo r, for any positive number r.)
- Any product of finitely many Pontryagin groups, with group operation defined componentwise and with the product topology. (In particular, $\mathbb{R}^n$ and $\mathbb{C}^n$.)

We may form a category by taking Pontryagin groups for the objects, with continuous group homomorphisms for the morphisms of the category. It is easy to see that this satisfies the definitions in 9.3.

For each Pontryagin group G, we now define the **dual**

$$G^* \quad = \quad \{\varphi \ : \ \varphi \text{ is a morphism from } G \text{ into } \mathbb{T}\}.$$

This is a special case of the notion of "dual" introduced in 9.55; in this context the object Δ of 9.55 is the circle group $\mathbb{T}$.

The set G^* can be made into a multiplicative group by defining products pointwise — that is, $\varphi\psi(g) = \varphi(g)\,\psi(g)$ for any $\varphi, \psi \in G^*$ and $g \in G$. The identity element of the group G^* is the constant function 1. The inverse of any element $\varphi(\cdot) \in G^*$ is the function $1/\varphi(\cdot)$. Note that $1/\varphi(g) = \overline{\varphi(g)}$, since $\varphi(g)$ takes its values in $\mathbb{T}$. Also, since φ is a group homomorphism, note that $1/\varphi(g) = \varphi(-g)$ if G is written as an additive group, or $1/\varphi(g) = \varphi(g^{-1})$ if G is written multiplicatively.

The group G^* is called the **character group** of G; the members of G^* are called the **characters** of G.

Examples. The groups $\mathbb{R}^n$ and $\mathbb{C}^n$ are isomorphic to their own character groups; the groups $\mathbb{T}$ and $\mathbb{Z}$ are isomorphic to each other's character groups.

The preceding assertions are easy to verify; the remaining ones below are not.

26.44. Pontryagin Duality Theorem. Let G be a Pontryagin group, and let G^* be its character group. Let G^* be topologized by the topology of uniform convergence on compact subsets of G. Then G^* is also a Pontryagin group; thus the mapping $G \mapsto G^*$ goes from the category of Pontryagin groups into itself. With respect to this mapping, every Pontryagin group is "reflexive" — that is, $G^{**} = G$. If $f : G_1 \to G_2$ is a morphism, then the dual map $f^* : {G_2}^* \to {G_1}^*$ (defined as in 9.3) is also a morphism. If $G^* = H$ and $H^* = G$, then G is compact if and only if H is discrete.

26.45. Theorem: existence and uniqueness of Haar measure. Let G be a Pontryagin group (as defined in 26.43). Then there exists a regular Borel measure μ on G that is translation-invariant on G — i.e., that satisfies $\mu(x+S) = \mu(S)$ for all $x \in G$ and all Borel sets $S \subseteq G$. It is unique, up to multiplication by a positive constant — i.e., if μ_1 and μ_2 are two such measures, then $\mu_1 = k\mu_2$ for some positive constant k. Any such measure is called the **Haar measure** of the group. It is bounded if and only if G is compact.

Notations. The spaces $L^p(\mu)$ may be written instead as $L^p(G)$. When the choice of the measure is clear, integration with respect to Haar measure may be written as $\int_G f(x)dx$ instead of $\int_G f(x)d\mu(x)$.

Remarks. For simplicity we have only considered commutative groups, but the notion of Haar measure generalizes to all locally compact Hausdorff groups; commutativity is not actually required. The literature contains an assortment of proofs of the existence and uniqueness of the Haar integral. They are based mainly on two proofs. One, due to Cartan, is based on an argument involving Cauchy nets and proves uniqueness while it proves existence. The other, due to Weil, is slightly shorter, uses a compactness argument, and does not prove uniqueness; it is usually supplemented by a brief proof of uniqueness due to von Neumann. Both proofs apply to noncommutative groups; both are given by Nachbin [1965]. Simpler proofs are possible if one restricts one's attention to compact groups or to commutative groups; for instance, see Izzo [1992] and references cited therein.

26.46. *Examples.* Haar measure on $\mathbb{Z}^n$ (or on any discrete group) is counting measure. Haar measure on $\mathbb{R}^n$ is n-dimensional Lebesgue measure. The circle group $[0, r)$ (with addition modulo r, as defined in 8.10.e) has Haar measure equal to the restriction of one-dimensional Lebesgue measure to subsets of $[0, r)$. Haar measure on the circle group $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ (with the operation of multiplication) can be described in terms of $[0, r)$ since those two groups are isomorphic; equivalently, Haar measure on $\mathbb{T}$ is arclength times any convenient positive constant.

Haar measure on the multiplicative group $(0, +\infty)$ can be described in terms of Lebesgue measure on the additive group $\mathbb{R}$, since those two groups are isomorphic by the mapping $(0, +\infty) \ni x \leftrightarrow \ln x \in \mathbb{R}$. That isomorphism yields this formula for Haar measure μ in $(0, +\infty)$:

$$\mu(S) \quad = \quad \int_S \frac{1}{t}\, dt \qquad \text{for Borel sets } S \subseteq (0, +\infty).$$

Here the dt is integration with respect to Lebesgue measure.

26.47. Let G and G^* be a Pontryagin group and its dual group (as defined in 26.43). Let Haar measure on both groups be denoted by dx. Of course, Haar measure is only determined up to multiplication by a positive constant; fix some particular version of Haar measure on each group.

The **Fourier transform** of a function $f : G \to \mathbb{F}$ is a corresponding function $\widehat{f} : G^* \to \mathbb{F}$. The transform is defined for $f \in L^1(G)$, for $f \in L^2(G)$, and for f in various other classes of functions by an assortment of different methods, but the different definitions agree wherever the classes of functions overlap.

The most basic of these definitions is the following:

$$\widehat{f}(\gamma) \quad = \quad \int_G f(x)\overline{\gamma(x)}dx \qquad\qquad \text{if } f \in L^1(G) \text{ and } \gamma \in G^*.$$

This makes sense for $f \in L^1(G)$, since $\overline{\gamma(x)}$ has absolute value 1 for all x.

Abstract Riemann-Lebesgue Lemma. If $f \in L^1(G)$, then $\widehat{f} \in C_0(G)$, with $\|\widehat{f}\|_\infty \le \|f\|_1$. Here $C_0(G)$ is the set of all continuous functions from G into $\mathbb{C}$ that vanish at infinity, as defined in 22.15. (This result generalizes 24.41.b; explain how.)

Plancherel Theorem. The Fourier transform, restricted to $L^1(G) \cap L^2(G)$, is a linear map from that set onto a dense subset of $L^2(G^*)$, which is distance-preserving — i.e.,

$$\|\widehat{f}\|_{L^2(G^*)} \quad = \quad c\|f\|_{L^2(G)} \tag{**}$$

(for some positive constant c that depends on the normalizations of the Haar measures; the Haar measures can be chosen so that $c = 1$). Hence that restriction extends uniquely to a linear map $f \mapsto \widehat{f}$, from $L^2(G)$ onto $L^2(G^*)$, satisfying (**). This map is sometimes called the **Plancherel transform**. It also satisfies **Parseval's Identity**:

$$c^2 \int_G f(x)\overline{g(x)}dx \quad = \quad \int_{G^*} \widehat{f}(\gamma)\overline{\widehat{g}(\gamma)}\, d\gamma \qquad\qquad \text{for} \quad f, g \in L^2(G)$$

and the **Fourier Inversion Formula**: $c^2 f(x) = \widehat{\widehat{f}}(-x)$. When $\widehat{f} \in L^1(G^*)$, then the Fourier Inversion Formula can be written in this form:

$$c^2 f(x) \quad = \quad \int_{G^*} \widehat{f}(\gamma)\gamma(x)d\gamma.$$

Examples.

a. When $G = \mathbb{R}^n$, then $G^* = \mathbb{R}^n$ also. It is convenient to define

$$\widehat{f}(\xi) \quad = \quad (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) \exp(-ix \cdot \xi) dx$$

— other constants can be used, but this constant yields $f(x) = \widehat{\widehat{f}}(-x)$. The term "**Fourier transform**" most often refers to this example.

b. The group $G = \mathbb{T}$ can be conveniently viewed as the additive group $[0, 2\pi)$, with addition modulo 2π (see 8.10.e). (Functions on $\mathbb{T}$ are also often viewed as functions on $\mathbb{R}$ that are periodic with period 2π. Intervals with length other than 2π can also be used, but the formulas are simplest for intervals of length 2π, so that is the only case we shall describe here.) The dual group is $G^* = \mathbb{Z}$, and a function on $\mathbb{Z}$ is just a sequence of numbers indexed by the integers. Thus, the transform of a function $f \in L^1(\mathbb{T})$ is the sequence of **Fourier coefficients**

$$\widehat{f}(n) \quad = \quad c_n \quad = \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx \qquad\qquad (n \in \mathbb{Z})$$

Going in the other direction, an "integral" of a function in $L^1(\mathbb{Z})$ is just a sum of real numbers. Thus we obtain

$$f(x) \quad = \quad \sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad\qquad (x \in \mathbb{T})$$

This series is to be interpreted not as a pointwise summation but as a summation in $L^2(\mathbb{T})$. That is, if $f \in L^2(\mathbb{T})$, then the partial sums $s_N(x) = \sum_{n=-N}^{N} c_n e^{inx}$ converge to f in the sense that $\lim_{N\to\infty} \|f - s_N\|_2 = 0$.

26.48. *Remarks on pointwise convergence.* Let $f \in L^2[-\pi,\pi]$, and define the Fourier coefficients $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$. Then $\sum_{n=-\infty}^{\infty} c_n e^{inx}$, the Fourier series of f, converges to f in the norm topology of $L^2[-\pi,\pi]$. Although the convergence in L^p spaces is more important for applications, it is of some historical interest to know when the Fourier series converges *pointwise* to f. For instance, a theorem of Jordan shows that if $f(-\pi) = f(\pi)$ and f has bounded variation on $[-\pi,\pi]$, then $\sum_{n=-\infty}^{\infty} c_n e^{inx} = \frac{1}{2}[f(x+) + f(x-)]$ for all x; here $f(x+)$ and $f(x-)$ are the right- and left-hand limits of f at x. If f is also continuous, then the Fourier series converges everywhere to f. Georg Cantor tried to investigate the sets of points where certain Fourier series converge; this led him to invent cardinalities and set theory.

What about functions that are not necessarily of bounded variation? It turns out that "most" continuous functions are ill-behaved at "most" points, in the following sense: Let $C_{2\pi}$ be the collection of all continuous functions $f : [-\pi,\pi] \to \mathbb{C}$ that satisfy $f(-\pi) = f(\pi)$; this is a Banach space when equipped with the sup norm. There exists a comeager set $\Phi \subseteq C_{2\pi}$ such that for each $f \in \Phi$, there exists a comeager set $E_f \subseteq [-\pi,\pi]$ such that the Fourier series of f diverges at every point of E_f.

It also turns out that "most" functions in $L^1[-\pi,\pi]$ are ill-behaved at "most" points, in a different sense: The functions whose Fourier series diverge almost everywhere in $[-\pi,\pi]$ is a comeager subset of $L^1[-\pi,\pi]$. (Kolmogorov first proved in 1926 that there exists a function in $L^1[-\pi,\pi]$ whose Fourier series diverges almost everywhere.)

However, for $p > 1$ the spaces $L^p[-\pi,\pi]$ exhibit much better behavior. If $f \in L^p[-\pi,\pi]$ for some $p > 1$, then the Fourier series for f converges almost everywhere to f. This was proved by Hunt [1968], extending methods developed earlier by Carleson for the case of $p = 2$.

For proofs or references for most of these results, see Edwards [1967]. The abstract approach to Fourier analysis is also introduced by Rudin [1960].

Ordered Topological Vector Spaces

26.49. *Definition and remarks.* A **ordered topological vector space** is a real vector space X that is equipped with both

- a topology, making X into a topological vector space, and
- an ordering, making X into an ordered vector space (as defined in 11.44).

Many different types of ordered TVS's can be defined by assuming various relations between the topology and the ordering. We shall concentrate on just a few basic types of ordered TVS's. In order of increasing specialization, these are:

$$\{\text{locally full spaces}\} \supseteq \{\text{locally solid spaces}\} \supseteq \{\text{F-lattices}\} \supseteq \{\text{Banach lattices}\}.$$

Our treatment is based largely on Fremlin [1974], Peressini [1967], and Wong and Ng [1973].

26.50. *Exercise.* Let X be an ordered TVS whose positive cone X_+ is closed. Then:

- **a.** The sets $\{x \in X : x \succcurlyeq u\}$ and $\{x \in X : x \preccurlyeq u\}$ are closed, for each $u \in X$.
- **b.** X is Hausdorff.
- **c.** X is Archimedean.

  Hint: If $\mathbb{N}y$ is bounded above by some x, then for all $n \in \mathbb{N}$ we have $x \succcurlyeq ny$, hence $\frac{1}{n}x - y \in X_+$ which is a closed set. Since X is a TVS, we have $\frac{1}{n}x - y \to -y$, and thus $y \preccurlyeq 0$.
- **d.** If $(v_\delta : \delta \in \Delta)$ is an increasing net that converges to some limit v_∞ in the topological space X, then $v_\infty = \sup_{\delta\in\Delta} v_\delta$.

  Hints: For each $\delta_0 \in \Delta$, the set $\{x \in X : x \succcurlyeq v_{\delta_0}\}$ is closed, hence contains v_∞. Also, if w is an upper bound for the set $\{v_\delta : \delta \in \Delta\}$, then v_∞ is in the closed set $\{x \in X : x \preccurlyeq w\}$.

26.51. Recall that a subset S of a preordered set $(X, \preccurlyeq)$ is **full** (or **order convex**) if $a \preccurlyeq x \preccurlyeq b$ with $a, b \in S$ implies $x \in S$ (see 4.4.a). The **full hull** of a set S is the set $\bigcup_{a,b\in S}[a,b]$; it is the smallest full set that contains S. *Exercises.*

- **a.** The full hull of any balanced subset of X is balanced.
- **b.** The full hull of any convex subset of X is convex.

26.52. *Proposition and definition.* Let X be an ordered topological vector space. Then the following conditions are equivalent; if they are satisfied we say X is **locally full** (or **ordered by a normal cone**).

(A) $(X, \mathcal{T})$ has a neighborhood base at 0 consisting of balanced, full sets.

(B) $(X, \mathcal{T})$ has a neighborhood base at 0 consisting of full sets.

(C) $(X, \mathcal{T})$ has a neighborhood base at 0 consisting of sets V with this property: If $v \in V \cap X_+$, then $[0, v] \subseteq V$.

(D) If $(x_\alpha : \alpha \in A)$ and $(y_\alpha : \alpha \in A)$ are nets in X based on the same directed set A and satisfying $0 \preccurlyeq x_\alpha \preccurlyeq y_\alpha$ and $y_\alpha \xrightarrow{\mathcal{T}} 0$, then $x_\alpha \xrightarrow{\mathcal{T}} 0$.

(E) **(The Squeeze Property.)** If $(u_\alpha), (v_\alpha), (w_\alpha)$ are nets in X based on the same directed set A and satisfying $u_\alpha \preccurlyeq v_\alpha \preccurlyeq w_\alpha$ for all $\alpha \in A$ and also satisfying $u_\alpha \xrightarrow{\mathcal{T}} p$ and $w_\alpha \xrightarrow{\mathcal{T}} p$ with the same limit p, then $v_\alpha \xrightarrow{\mathcal{T}} p$ also.

(F) If V is any neighborhood of 0, then there exists a neighborhood W of 0 with this property: If $w \in W \cap X_+$, then $[0, w] \subseteq V$.

Remark. Note the similarity between 26.52(E) and 7.40.i. Those two properties are the same in $\mathbb{R}$, since in that setting the order and topological convergences are the same.

Proof of equivalence. The implications (A) $\Rightarrow$ (B) $\Rightarrow$ (C) $\Rightarrow$ (D) are obvious. The implications (D) $\Longleftrightarrow$ (E) are an easy exercise. It suffices to show (D) $\Rightarrow$ (F) $\Rightarrow$ (A).

Let $\mathcal{N}$ be the neighborhood filter at 0.

Proof of (D) $\Rightarrow$ (F). Suppose (F) fails. Then there is some V that is a neighborhood of 0, for which there is no corresponding neighborhood W. Then for each $N \in \mathcal{N}$, there is some $w_N \in N$ such that $[0, w_N]$ is not contained in V, and hence there is some $x_N \in [0, w_N] \setminus V$. Then the net $(w_N : N \in \mathcal{N})$ converges to 0. By (D), the net $(x_N : N \in \mathcal{N})$ converges to 0 — but then eventually $x_N \in V$, a contradiction.

Proof of (F) $\Rightarrow$ (A). Let G be any neighborhood of 0; we wish to show that G contains a balanced, full neighborhood of 0. Let G' be a balanced neighborhood of 0 satisfying $G' + G' \subseteq G$; such a set is available since X is a TVS. By (F), there is some neighborhood W of 0 with this property: $\bigcap_{w \in W \cap X_+} [0, w] \subseteq G'$. Replacing W with a smaller set, we may assume $W \subseteq G'$. Now let W' be some some balanced neighborhood of 0 satisfying $W' + W' \subseteq W$. Let $K = G' \cap W'$; it is a balanced neighborhood of 0. Let F be its full hull — that is, $F = \bigcup_{a,b \in K} [a, b]$. The full hull of any balanced set is balanced; thus F is a balanced, full neighborhood of 0. It suffices to show $F \subseteq G$. Let $x \in F$. Then $a \preccurlyeq x \preccurlyeq b$ for some $a, b \in K = G' \cap W'$. Then $0 \preccurlyeq x - a \preccurlyeq b - a \in W' - W' = W' + W' \subseteq W$. Hence $b - a \in W \cap X_+$, and thus $x - a \in [0, b - a] \subseteq G'$. Finally, $x = (x - a) + a \in G' + G' \subseteq G$.

26.53. *A degenerate example.* Any topological vector space $(X, \mathcal{T})$ can be turned into a locally full space by equipping it with the degenerate ordering $x \preccurlyeq y \iff x = y$ (so that the positive cone is $\{0\}$). Indeed, with that ordering, every subset of X is full, so any neighborhood base at 0 consists of full sets.

Despite its triviality (or because of it!), this example is useful. It shows that any affine operator between topological vector spaces can be turned into a convex operator from a topological vector space into a locally full space. Thus, the results proved in this chapter for convex operators are applicable to affine operators as well.

26.54. If X is a locally convex, locally full space, then X has a neighborhood base at 0 consisting of balanced, full, convex sets.

Hint: Let N be any given neighborhood of 0 in X. Since X is locally full, we have $N \supseteq B$ where B is a balanced, full neighborhood of 0. Since X is locally convex, we have $B \supseteq C$ where C is a balanced, convex neighborhood of 0. Show that the full hull of C is a balanced, full, convex neighborhood of 0 that is contained in N.

26.55. *Definitions.* Let $(X, \preccurlyeq)$ be a Riesz space, i.e., a vector lattice. By a **Riesz F-seminorm** we shall mean an F-seminorm $\rho : X \to [0, +\infty)$ (defined as in 26.2) that also has this property:

$$/x/ \preccurlyeq /y/ \qquad \Rightarrow \qquad \rho(x) \leq \rho(y).$$

If ρ is also homogeneous (i.e., if $\rho(cx) = |c|\rho(x)$), then it is a **Riesz seminorm**. If ρ is positive-definite (i.e., if $x \neq 0 \Rightarrow \rho(x) > 0$), then it is, respectively, a **Riesz F-norm** or **Riesz norm**.

Examples. On any of the Banach spaces $L^p(\mu)$ for $1 \leq p \leq \infty$ (with $\{\text{scalars}\} = \mathbb{R}$), the norm $\| \ \|_p$ is a Riesz norm. On the F-spaces $L^p(\mu)$ for $0 < p < 1$, the F-norm $\| \ \|_p^p$ is a Riesz F-norm.

26.56. The Hahn-Banach Theorem was introduced in 12.30. Two more of its equivalents are given by the following principles:

> **(HB15) Riesz Seminorms** and **(HB16) Positive Functionals.** Let X be a Riesz space, let S be a Riesz subspace, and suppose either
>
> q is a Riesz seminorm on X, or
>
> $q : X \to \mathbb{R}$ is a positive linear functional.
>
> Let $\lambda : S \to \mathbb{R}$ be a positive linear functional, satisfying $\lambda \leq q$ on S_+. Then λ extends to a positive linear functional $\Lambda : X \to \mathbb{R}$, satisfying $\Lambda \leq q$ on X_+.

Proof that (HB2) *implies both* (HB15) *and* (HB16). In either case, the restriction of q to X_+ is a convex, isotone function $q : X_+ \to \mathbb{R}$ that satisfies $q(0) = 0$. Define $p(x) = q(x^+)$. Then p is convex; this follows from the convexity and isotonicity of q and the fact that $(ax + (1-a)y)^+ \preccurlyeq (ax)^+ + ((1-a)y)^+ = a(x^+) + (1-a)(y^+)$ if $x, y \in X$ and $a \in [0,1]$. For any $x \in S$ we have $x \preccurlyeq x^+$, hence $\lambda(x) \leq \lambda(x^+) \leq q(x^+) = p(x)$. Thus (HB2) is applicable, and λ extends to a linear functional $\Lambda : X \to \mathbb{R}$ satisfying $\Lambda \leq p$ on X; hence $\Lambda \leq p = q$ on X_+. To see that Λ is positive, note that if $x \succcurlyeq 0$, then $(-x)^+ = 0$; hence $-\Lambda(x) = \Lambda(-x) \leq p(-x) = q\left((-x)^+\right) = q(0) = 0$.

Proof that either (HB15) *or* (HB16) *implies* (HB1). In either case we take X to be the Riesz space $B(\Delta)$ and let S be the subspace consisting of those nets that are convergent in the ordinary sense. For a proof with (HB15), use the Riesz seminorm $q(x) = \|x\|_\infty = \sup\{|x(\delta)| : \delta \in \Delta\}$. For a proof with (HB16), use the positive linear functional $q(x) = \limsup_{\delta \in \Delta} x(\delta)$.

26.57. Recall from 8.42.q that, in a vector lattice, a set S is **solid** if $/x/ \preccurlyeq /y/$ and $y \in S$ imply $x \in S$. Note that any nonempty solid set is balanced.

Proposition and definition. Let $(X, \preccurlyeq)$ be a vector lattice and let $(X, \mathcal{T})$ be a topological vector space, both with the same underlying vector space X. Then the following conditions are equivalent. If one, hence all, of these conditions are satisfied, we say X is **locally solid**; some mathematicians call it a **topological Riesz space**.

(A) X has a neighborhood base at 0 consisting of solid sets.

(B) The topology $\mathcal{T}$ is the gauge topology determined by a collection of Riesz F-seminorms.

(C) The mapping $x \mapsto /x/$ is uniformly continuous from X to X (equipped with the uniform structure resulting from the topology $\mathcal{T}$).

(D) The mapping $(x, y) \mapsto x \vee y$ is uniformly continuous from $X \times X$ (with the product uniform structure) to X.

(E) The mapping $x \mapsto x^+$ is uniformly continuous from $(X, \mathcal{T})$ into $(X, \mathcal{T})$.

(F) X is locally full and the mapping $x \mapsto x^+$ is continuous at 0.

(G) For any two nets (x_α) and (y_α) in X (with the same index set), if $/x_\alpha/ \preccurlyeq /y_\alpha/$ for all α and $y_\alpha \xrightarrow{\mathcal{T}} 0$, then $x_\alpha \xrightarrow{\mathcal{T}} 0$.

Proof of (A) $\Rightarrow$ (B). The proof is similar to that of 26.29, case (ii), but we may choose the sets $B \in \mathcal{B}$ to be solid sets. Construct an F-seminorm ρ as in 26.29; we shall now show that that function is actually a Riesz F-seminorm. We know that $/x/ \preccurlyeq /y/ \Rightarrow \varphi(x) \preccurlyeq \varphi(y)$ since each B_n is solid; we are to show that $/x/ \preccurlyeq /y/ \Rightarrow \rho(x) \leq \rho(y)$. Consider any decomposition $y = y_1 + y_2 + \cdots + y_m$. Then $/x/ \preccurlyeq /y/ \preccurlyeq /y_1/ + /y_2/ + \cdots + /y_m/$, hence

$$\begin{aligned} x &\in [-/y_1/ - /y_2/ - \cdots - /y_m/,\ /y_1/ + /y_2/ + \cdots + /y_m/] \\ &= [-/y_1/, /y_1/] + [-/y_2/, /y_2/] + \cdots + [-/y_m/, /y_m/] \end{aligned}$$

by 8.36(C). Thus we can write $x = x_1 + x_2 + \cdots + x_m$ with $/x_i/ \preccurlyeq /y_i/$. Hence $\rho(x) \leq \sum_{i=1}^m \varphi(x_i) \leq \sum_{i=1}^m \varphi(y_i)$. Since $\rho(y)$ is the infimum of all such summations $\sum_{i=1}^m \varphi(y_i)$, it follows that $\rho(x) \leq \rho(y)$.

Proof of (B) $\Rightarrow$ (C). Recall from 8.42.o that $\big/ /x/ - /x'/ \big/ \preccurlyeq /x - x'/$. Hence for any Riesz F-seminorm ρ, we have $\rho(/x/ - /x'/) \leq \rho(x - x')$. If $x_\alpha - x'_\alpha \xrightarrow{\mathcal{T}} 0$, then $\rho(x_\alpha - x'_\alpha) \to 0$ for every ρ in the determining family of Riesz F-seminorms; hence $\rho(/x_\alpha/ - /x'_\alpha/) \to 0$ for each ρ; hence $/x_\alpha/ - /x'_\alpha/ \xrightarrow{\mathcal{T}} 0$.

Proof of (C) $\Rightarrow$ (D). Immediate from 8.42.l.

Proof of (D) $\Rightarrow$ (E). Obvious.

Proof of (E) $\Rightarrow$ (F). Obviously the mapping $x \mapsto x^+$ is continuous at 0. To show that X is locally full, we shall verify condition 26.52(F). Let V be any neighborhood of 0. By the uniform continuity of the mapping $x \mapsto x^+$, there is some neighborhood W of 0 such that $x - y \in W \Rightarrow x^+ - y^+ \in V$. We are to show that if $0 \preccurlyeq w \in W$, then $[0, w] \subseteq V$. Indeed, let $x \in [0, w]$. Then $x - (x - w) = w \in W$, so $x^+ - (x - w)^+ \in V$. But $x^+ = x$ since $x \succcurlyeq 0$, and $(x - w)^+ = 0$ since $w \succcurlyeq x$. Thus we have shown $x \in V$.

Proof of (F) $\Rightarrow$ (G). Since $x \mapsto x^+$ is continuous at 0, the function $x \mapsto x^- = (-x)^+$ is also continuous at 0, and therefore so is the function $x \mapsto /x/ = x^+ + x^-$. Now, suppose

$/x_\alpha/ \preccurlyeq /y_\alpha/$ and $y_\alpha \xrightarrow{\mathcal{T}} 0$. Then $/y_\alpha/ \xrightarrow{\mathcal{T}} 0$. Then

$$0 \preccurlyeq x_\alpha^+ \preccurlyeq /x_\alpha/ \preccurlyeq /y_\alpha/ \qquad \text{and} \qquad 0 \preccurlyeq x_\alpha^- \preccurlyeq /x_\alpha/ \preccurlyeq /y_\alpha/,$$

hence by 26.52(D) we have $x_\alpha^+ \xrightarrow{\mathcal{T}} 0$ and $x_\alpha^- \xrightarrow{\mathcal{T}} 0$. Hence $x_\alpha = x_\alpha^+ - x_\alpha^- \xrightarrow{\mathcal{T}} 0$.

Proof of (G) $\Rightarrow$ (A). Let W be a neighborhood of 0; we wish to show that W contains some full neighborhood of 0. Recall from 8.42.q that the *solid kernel* of W is the set

$$\mathrm{sk}(W) \quad = \quad \bigcup\{[-u,u] \ : \ [-u,u] \subseteq W\};$$

it is the largest solid subset of W. It suffices to show that $\mathrm{sk}(W)$ is a neighborhood of 0. Suppose not. Then there exists a net (y_α) that takes values outside $\mathrm{sk}(W)$ but converges to 0. Then for each α, the interval $[-/y_\alpha/, /y_\alpha/]$ is not contained in W, hence there is some $x_\alpha \in [-/y_\alpha/, /y_\alpha/] \setminus W$. Then $/x_\alpha/ \preccurlyeq /y_\alpha/$, and so $x_\alpha \xrightarrow{\mathcal{T}} 0$ by (G). But then eventually $x_\alpha \in W$, a contradiction.

26.58. *Exercise.* If X is a locally solid Riesz space that is *Hausdorff*, then its positive cone is closed; hence all the conclusions of 26.50 are applicable.

Hint: $X_+ = \varphi^{-1}(0)$, where φ is the continuous function $x \mapsto x \wedge 0$.

26.59. *Definitions.* An **F-lattice** or a **Banach lattice** will mean a lattice topologized by a Riesz F-norm, respectively a Riesz norm, which is metrically complete. Note that any F-lattice or Banach lattice satisfies condition 26.57(B) and thus is locally solid.

Theorem: Continuity of positive operators. Suppose X is an F-lattice and Y is a locally full space. Then every positive linear operator $f : X \to Y$ is continuous.

Remark. For a more general but more complicated result about convex operators, see Neumann [1985].

Proof of theorem. Suppose not. Then there exists a sequence (x_n) that converges to 0 in X, such that $f(x_n) \not\to 0$ in Y.

Then $x_n^+ \to 0$ and $x_n^- \to 0$, by 26.57(E). Since $f(x_n) = f(x_n^+) - f(x_n^-)$, at least one of the sequences $(f(x_n^+)), (f(x_n^-))$ does not converge to 0 in Y. Thus, replacing (x_n) with either (x_n^+) or (x_n^-), we may assume $x_n \succcurlyeq 0$. Replacing (x_n) with a subsequence, we may assume $f(x_n)$ stays out of some neighborhood G of 0 in Y. Choosing a smaller neighborhood, we may assume G is full. Replacing (x_n) with a subsequence, we may assume $\rho(x_n) < 4^{-n}$, where ρ is some complete Riesz F-norm that determines the topology of X. Let $u_n = 2^n x_n$; then $f(u_n) \notin 2^n G$.

By subadditivity of ρ, we have $\rho(u_n) < 2^{-n}$. Hence the series $\sum_n u_n$ converges to some limit v in X. Since the u_n's are in X_+, we have $0 \preccurlyeq u_n \preccurlyeq v$ in X, so $0 \preccurlyeq f(u_n) \preccurlyeq f(v)$ in Y. For all n sufficiently large, we have $f(v) \in 2^n G$, since G is a neighborhood of 0. But G is also full, so $f(u_n) \in 2^n G$, a contradiction.

26.60. *Corollaries.*

a. Any two complete Riesz F-norms on a vector lattice are equivalent.

Hint: The identity map is a positive operator.

b. Let X be an F-lattice, and let $f : X \to \mathbb{R}$ be a linear functional. Then f is order bounded (i.e., the image under f of any order bounded set is an order bounded set) if and only if f is continuous.

Hints: Order-bounded implies continuity, by 11.57 and 26.59. For the converse, suppose f is not order bounded. Then there is some set $B \subseteq X$ that is order bounded, such that $f(B)$ is not bounded in $\mathbb{R}$. Choose a sequence (x_n) in B with $|f(x_n)| > n^2$. Let $y_n = \frac{1}{n}x_n$; then $|f(y_n)| > n$. However, $y_n \to 0$ in X, by 27.11 and 27.2(D). Thus f is not continuous.

c. *Example of a continuous operator that is not order bounded.* Let $C[0,1] = \{$continuous functions from $[0,1]$ into $\mathbb{R}\}$ and $c_0 = \{$sequences of reals converging to $0\}$ be equipped with their sup norms; then both are Banach lattices and c_0 is Dedekind complete. Define $f : C[0,1] \to c_0$ as follows: For any $x \in C[0,1]$, let $f(x)$ be the sequence whose nth term is $\int_0^1 x(t)\sin(2\pi nt)dt$.

Hints: The sequence $f(x)$ tends to 0 by the Riemann-Lebesgue Lemma (24.41.b). It is an easy exercise to show that the operator f is continuous. The set $B = \{x \in C[0,1] : -1 \le x \le 1\}$ is order bounded. However, B contains all the functions $x_n(t) = \sin(2\pi nt)$. Observe that the nth term of the sequence $f(x_n)$ is $\frac{1}{2}$; show that $f(B)$ is not order bounded.

Chapter 27

Barrels and Other Features of TVS's

Bounded Subsets of TVS's

27.1. *Motivating exercises.*

a. Two equivalent seminorms on the same vector space yield the same collection of metrically bounded sets.

b. Show by example that equivalent F-seminorms on a vector space may yield different collections of metrically bounded sets.

27.2. *Definition.* Let X be a TVS, with scalar field $\mathbb{F}$ equal to $\mathbb{R}$ or $\mathbb{C}$. Let $S \subseteq X$. Show that the following conditions on S are equivalent. If any, hence all, of these conditions are satisfied, we say that S is **toplinearly bounded**, or **bounded in the sense of topological linear spaces**, or **bounded with respect to the TVS topology on X**.

(A) The set $\{m_s : s \in S\}$ of mappings $m_s : \mathbb{F} \to X$ defined by $m_s(c) = cs$ is equicontinuous.

(B) For each neighborhood G of 0, there is some scalar c such that $S \subseteq cG$.

(C) For each neighborhood G of 0, there is some $r > 0$ such that $S \subseteq cG$ for all scalars c with $|c| > r$.

(D) Whenever (c_n, x_n) is a sequence in $\mathbb{F} \times S$ with $c_n \to 0$, then $c_n x_n \to 0$.

(E) Whenever (c_α, x_α) is a net in $\mathbb{F} \times S$ with $c_\alpha \to 0$, then $c_\alpha x_\alpha \to 0$.

(*Hint for the proof of equivalence*: 26.27.c.)

A collection of functions $\Phi = \{\varphi_\gamma : \gamma \in \Gamma\}$, from some set Ω into X, will be called **toplinearly bounded pointwise** if for each $\omega \in \Omega$ the set $\Phi(\omega) = \{\varphi_\gamma(\omega) : \gamma \in \Gamma\}$ is toplinearly bounded in X.

Caution: Toplinear boundedness is *not* the same thing as either metric boundedness or order boundedness. In 27.5, 27.6, and 27.11 we investigate some of the relations between toplinear boundedness and the other two kinds of boundedness. It is unfortunate that the term "bounded set" has these three meanings that are sometimes quite different; the reader must strive to determine from context which meaning is intended. In the next few paragraphs, of course, "bounded" means toplinearly bounded unless some other meaning is specified.

27.3. *Basic properties of bounded sets.* Let X be a TVS, with topology $\mathcal{T}$, and let $S \subseteq X$. Show that

a. S is bounded if and only if every countable subset of S is bounded.

b. The bounded sets form an ideal: the union of finitely many bounded sets is bounded; any subset of a bounded set is bounded.

In fact, it is a proper ideal, provided that the space X does not have the indiscrete topology.

c. Any compact set is bounded.

d. The closure of a bounded set is bounded.

e. If X is locally convex, then the convex hull of any bounded subset of X is bounded.

f. A topological vector space is **quasicomplete** if each bounded, closed set is complete.

Prove this more general version of **Mazur's Theorem**: In a quasicomplete, locally convex space, the closed convex hull of a compact set is compact.

Hint: Refer to the result on totally bounded sets in 26.23.i.

g. Suppose that the topology $\mathcal{T}$ on X is the initial topology determined by a collection of linear mappings into topological vector spaces, $\varphi_\lambda : X \to (Y_\lambda, \mathcal{U}_\lambda)$. Show that $S \subseteq X$ is $\mathcal{T}$-bounded if and only if $\varphi_\lambda(S)$ is $\mathcal{U}_\lambda$-bounded for each λ.

h. Let $X = \prod_{\lambda\in\Lambda} X_\lambda$ be a product of TVS's; then (as we have noted in 26.20.a) X is also a TVS. Show that a set $B \subseteq X$ is bounded if and only if it is included in a set of the form $\prod_{\lambda\in\Lambda} B_\lambda$, where each B_λ is a bounded subset of X_λ.

i. *Change of scalar field.* Let X be a complex TVS. Then X may also be viewed as a real TVS, with the same topology, if we "forget" how to multiply vectors by nonreal scalars. However, the bounded subsets of the real TVS are the same as the bounded subsets of the complex TVS.

Proof. This may be easiest to see by considering condition 27.2(D). Any bounded subset of the complex TVS is also a bounded subset of the real TVS, since $\mathbb{R} \subseteq \mathbb{C}$. Conversely, suppose S is real-bounded, and suppose (c_n, x_n) is a sequence in $\mathbb{C} \times S$ with $c_n \to 0$. Then $c_n = a_n + ib_n$ with $a_n, b_n \to 0$ in $\mathbb{R}$. Then $a_n x_n \to 0$ and $b_n x_n \to 0$, since S is real-bounded; hence $(a_n + ib_n)x_n \to 0$.

27.4. Let X and Y be topological vector spaces. Let $f : X \to Y$ be a linear map. Suppose f is continuous (i.e., preserves convergent nets) — or, more generally, suppose f is sequentially continuous (i.e., preserves convergent sequences). Let $S \subseteq X$ be bounded. Then $f(S)$ is a bounded subset of Y.

In some contexts, a linear map is called *bounded* if it takes bounded sets to bounded sets. (This generalizes the terminology of 23.1.) With this terminology, we have just shown that

$$f \text{ is continuous} \quad\Rightarrow\quad f \text{ is sequentially continuous} \quad\Rightarrow\quad f \text{ is bounded}.$$

A partial converse is as follows:

Proposition. If X is a pseudometrizable TVS, Y is a TVS, and $f : X \to Y$ is a bounded linear map, then f is continuous.

Proof. By 26.32, the topology of X is given by an F-seminorm, ρ. By 15.34.d, it suffices to show f is sequentially continuous. Suppose not — say (x_n) is a sequence that converges to 0 in X, while $(f(x_n))$ does not converge to 0 in Y. Passing to a subsequence, we may assume $(f(x_n))$ stays out of some neighborhood G of 0 in Y. We have $\rho(x_n) \to 0$; passing to a subsequence, we may assume $\rho(x_n) < 1/n^2$. Then $\rho(nx_n) \leq 1/n$, by the subadditivity of any F-seminorm ρ. The sequence (nx_n) converges to 0, hence is bounded. Since f is bounded, the sequence $(f(nx_n)) = (nf(x_n))$ is bounded in Y. Then the sequence whose nth term is $\frac{1}{n} \cdot nf(x_n)$ must converge to 0, a contradiction.

Remark. A partial extension to nonmetrizable spaces is given in 27.41.m.

27.5. Let X be a topological vector space. Then any toplinearly bounded set B is metrically bounded, in the following sense: If ρ is any continuous F-seminorm (or, more generally, any continuous G-seminorm) on X, then $\sup_{x \in B} \rho(x) < \infty$.

Proof. Suppose not. Say there exists a sequence $(x_n) \in B$ with $\rho(x_n) > n$. Let $y_n = n^{-1}x_n$. Since B is bounded, we have $y_n \to 0$, hence $\rho(y_n) \to 0$, hence $\rho(y_n) < 1$ for n sufficiently large. But then by the subadditivity of ρ we have $\rho(x_n) = \rho(ny_n) \leq n\rho(y_n) < n$, a contradiction.

27.6. Let X be a locally convex space over scalar field $\mathbb{F}$, and let R be any family of seminorms that determines the topology of X. Let $S \subseteq X$. Then the following are equivalent:

(A) S is toplinearly bounded (as defined in 27.2).

(B) S is metrically bounded, in this sense: Each continuous seminorm on X is bounded on S.

(C) Each seminorm in the given family R is bounded on S.

(D) Each continuous linear map $f : X \to \mathbb{F}$ is bounded on S.

Proof. The implication (A) $\Rightarrow$ (B) is a special case of 27.5. The implication (B) $\Rightarrow$ (C) is trivial.

For (C) $\Rightarrow$ (A), let (x_n) be a sequence in S and let $c_n \to 0$ in $\mathbb{F}$; we wish to show that $c_n x_n \to 0$ in X. It suffices to show that $\rho(c_n x_n) \to 0$ for each $\rho \in R$. Observe that $\sup_n \rho(x_n) < \infty$, hence $\rho(c_n x_n) = |c_n|\rho(x_n) \to 0$.

For (B) $\Rightarrow$ (D), note that $\rho(x) = |f(x)|$ defines a continuous seminorm.

It remains to prove (D) $\Rightarrow$ (B). We first prove this under the additional assumption that X is a normed space. Then X^* is a complete normed space, as we noted in 23.8. The set S is a pointwise bounded set of continuous linear maps from X^* into $\mathbb{F}$. By the Uniform Boundedness Principle (23.14), S is equicontinuous. Thus S is norm bounded, when viewed as a subset of X^{**}. The canonical embedding $X \xrightarrow{\subseteq} X^{**}$ is norm-preserving (see 23.20), so S is norm bounded in X.

Now, for the general case: Let X_γ denote the vector space X with the given topology. Let ρ be any continuous seminorm on X. Let X_ρ be the seminormed space (X, ρ); this has a weaker topology than X_γ, and so the identity map $i : X_\gamma \to X_\rho$ is continuous. Form the quotient space $Z = X_\rho/\rho^{-1}(\{0\})$; its quotient topology is given by the norm $\| \quad \| = \widehat{\rho}$ defined by $\widehat{\rho}(\pi(x)) = \rho(x)$, as in 22.13.e. Let $\pi : X \to Z$ be the quotient map; it is a continuous linear map from X_ρ into $(Z, \| \quad \|)$. If λ is any continuous linear functional on Z, then the composition

$$X_\gamma \xrightarrow{i} X_\rho \xrightarrow{\pi} Z \xrightarrow{\lambda} \mathbb{F}$$

is a continuous linear functional on X_γ. Hence it is bounded on the set S. It follows that λ is bounded on $\pi(S)$. By the results of the previous paragraph, $\pi(S)$ is a norm bounded subset of the normed space $(Z, \| \quad \|)$. That is, $\widehat{\rho}$ is bounded on $\pi(S)$; hence ρ is bounded on S.

Corollary. In a normed space, a set is metrically bounded if and only if it is toplinearly bounded.

27.7. A topological vector space X is **locally bounded** if 0 has a bounded neighborhood (or, equivalently, if every point has a bounded neighborhood). Show that

a. Any locally bounded TVS is pseudometrizable.
Hint: If B is a bounded neighborhood of 0, show that $\{B, \frac{1}{2}B, \frac{1}{3}B, \frac{1}{4}B, \ldots\}$ is a neighborhood base at 0. Now use 26.32.

b. A topological vector space is seminormable (i.e., its topology can be given by a seminorm) if and only if it is both locally convex and locally bounded. Hence we have the following result.

> **Kolmogorov Normability Criterion.** A TVS is normable if and only if it is locally convex, locally bounded, and Hausdorff.

c. The product of infinitely many nondegenerate TVS's cannot be locally bounded. (Here "nondegenerate" means not having the indiscrete topology.) *Hints*: Use 15.26.a and 27.3.h.
Corollary. A product of infinitely many nondegenerate TVS's is not seminormable.

27.8. *Examples.*

a. For $0 < p \leq \infty$, $L^p(\mu; X)$ is locally bounded. (*Hint*: $\|cf\|_p = |c| \, \|f\|_p$.)
However, for $0 \leq p < 1$, the F-spaces ℓ^p and $L^p[0,1]$ are not locally convex, as we saw in 26.16; hence their topologies are not normable.

b. The space $L^0[0,1]$ is not locally bounded. (Also, as we noted in 26.16, it is not locally convex.) Hence its F-norm is not equivalent to a norm.
Proof. Let μ denote Lebesgue measure. Show that if V is a neighborhood of 0 in $L^0[0,1]$, then there is some number $\varepsilon > 0$ such that V contains all measurable functions f that satisfy $\mu(\{\omega : |f(\omega)| \geq \varepsilon\}) \leq \varepsilon$. Now, for positive integers n, define $f_n = n1_{[0,\varepsilon]}$. Then f_n lies in the set V, but the sequence $\left(\frac{1}{n} f_n\right)$ does not converge to 0 in measure, so the set V is not bounded.

c. The space $\mathbb{F}^{\mathbb{N}}$, with the product topology, is locally convex but not locally bounded; hence it is not normable. Thus the F-norm given in 26.7 is not equivalent to a norm.

d. The space $C(\mathbb{R}) = \{$continuous scalar-valued functions on $\mathbb{R}\}$, with the topology of uniform convergence on compact subsets of $\mathbb{R}$, is locally convex but not locally bounded; hence it is not normable. Thus the F-norm given in 26.8 is not equivalent to a norm.

27.9. Let Ω be a set, let $\mathcal{S}$ be a collection of subsets of Ω, and let Z be a topological Abelian group equipped with its usual uniform structure. Let Z^Ω be equipped with the product group structure (introduced in 9.18) and with the topology of uniform convergence on members of $\mathcal{S}$ (introduced in 18.26). Show that

a. Z^Ω is a topological Abelian group.

b. The same topology on Z^Ω is obtained if we replace $\mathcal{S}$ with its $\cup$-closure, defined as in 4.4.d.

For this reason, in many contexts we may freely assume that $\mathcal{S}$ is closed under finite union — or we may make this slightly weaker assumption:

> $\mathcal{S}$ is **directed by inclusion** — that is, for each $S_1, S_2 \in \mathcal{S}$ there exists some $S \in \mathcal{S}$ such that $S_1 \cup S_2 \subseteq S$.

c. Suppose that $\mathcal{S}$ is directed by inclusion and $\mathcal{B}$ is a neighborhood base at 0 in Z. Then the sets

$$G(S, B) \quad = \quad \{g \in Z^\Omega \ : \ g(S) \subseteq B\} \qquad\qquad (S \in \mathcal{S},\ B \in \mathcal{B})$$

form a neighborhood base at 0 in Z^Ω.

Now suppose that Z is a topological vector space. Then:

d. The topological Abelian group Z^Ω is not necessarily a topological vector space.

For instance, let $\mathbb{R}^\mathbb{R}$ be topologized by uniform convergence on $\mathbb{R}$ (thus $\mathcal{S} = \{\mathbb{R}\}$); then $\mathbb{R}^\mathbb{R}$ is a topological Abelian group but multiplication by scalars is not continuous.

e. Let Φ be a linear subspace of Z^Ω. Fix any $S \in \mathcal{S}$. Then the sets $f(S)$ (for $f \in \Phi$) are bounded in the topological vector space Z if and only if the sets $G(S, B) \cap \Phi$ (for $B \in \mathcal{B}$) are absorbing in the vector space Φ.

f. Let Φ be a linear subspace of Z^Ω. Then the topology of uniform convergence on members of $\mathcal{S}$ makes Φ into a topological vector space if and only if

> $(*)$ for each $f \in \Phi$ and $S \in \mathcal{S}$, the set $f(S)$ is bounded in the topological vector space Z.

Hints: The topology on Z is not affected if we replace the neighborhood base $\mathcal{B}$ with some other neighborhood base that generates the same neighborhood filter; hence we may assume members of $\mathcal{B}$ are balanced. Also, in view of 27.9.b, we may assume $\mathcal{S}$ is directed by inclusion. Then the sets $G(S, B) \cap \Phi$ are balanced sets that form a neighborhood base at 0 for the topological Abelian group Φ. Now use the preceding exercise and the characterizations of neighborhood bases in 26.25 and 26.26.

g. Assume condition $(*)$ above. Then a set $F \subseteq \Phi$ is bounded in the topological vector space Φ if and only if for each $S \in \mathcal{S}$, the set $F(S) = \{f(s) : f \in F, s \in S\}$ is bounded in Z.

h. Assume condition $(*)$ above. If Z is locally convex, then so is Φ. In particular, if Z is the scalar field ($\mathbb{R}$ or $\mathbb{C}$), then Φ is locally convex.

(*Hint*: Apply 27.9.c. If B is a convex subset of Z, then $G(S, B)$ is a convex subset of Z^{Ω}.)

27.10. *Definition.* A topological vector space is said to have the **Heine-Borel Property** if every closed, bounded subset is compact.

Examples.

a. Any finite-dimensional Hausdorff topological vector space has the Heine-Borel Property.

b. A normed vector space has the Heine-Borel Property if and only if the space is finite-dimensional; that fact will follow easily from 27.17.

c. The Fréchet space $\mathrm{Hol}(\Omega)$ described in 26.10 has the Heine-Borel Property; that fact follows easily from Montel's Theorem, stated in 25.27.

Bounded Sets in Ordered TVS's

27.11. *Exercise.* Let X be an ordered TVS that is locally full (defined in 26.52). Then every order bounded subset of X is toplinearly bounded.

Proof. Let $[a, b]$ be any order interval in X, and let G be any neighborhood of 0 in X. Then G contains some N that is a balanced full neighborhood of 0. Then $a, b \in rN$ for $r > 0$ sufficiently large. Hence $[a, b] \subseteq rN \subseteq rG$.

27.12. Theorem. Let X be a TVS, and let Y be an ordered TVS that is locally full. Let $\Omega \subseteq X$ be open and convex.

If $f : \Omega \to Y$ is convex and is continuous at some point of Ω, then f is continuous everywhere on Ω.

More generally, let Φ be a collection of convex mappings from Ω into Y. Assume Φ is pointwise toplinearly bounded — i.e., assume that for each $x \in \Omega$, the set $\Phi(x) = \{f(x) : f \in \Phi\}$ is toplinearly bounded in Y. Also assume Φ is equicontinuous at some point $x_0 \in \Omega$. Then Φ is equicontinuous at every point of Ω.

Proof (following Neumann [1985]). Let any $x_1 \in \Omega$ be given; it suffices to prove equicontinuity at x_1. We may replace the functions $f \in \Phi$ with the functions $f(\cdot + x_1) - f(x_1)$; thus we may assume that $0 = x_1 \in \Omega$ and that $f(0) = 0$ for all $f \in \Omega$. Let N be any balanced full neighborhood of 0 in Y; we are to show that there is some neighborhood G of 0 in X, contained in Ω, such that $\bigcup_{f \in \Phi} f(G) \subseteq N$.

Choose some balanced full set N' that is a neighborhood of 0 in Y and satisfies $N' + N' + N' \subseteq N$. By the assumed equicontinuity at x_0, there is some balanced neighborhood U of 0 in X, contained in Ω, such that $f(x_0 + u) - f(x_0) \in N'$ for all $f \in \Phi$ and $u \in U$. For some $\delta \in (0, 1]$ sufficiently small, we have $-\delta x_0 \in \Omega$. We know $\delta x_0 \in \Omega$ by convexity of that set. Since Φ is bounded pointwise, $B = \bigcup_{f\in\Phi}\{f(x_0), f(\delta x_0), f(-\delta x_0)\}$ is a bounded subset of Y, and hence there is some $\varepsilon \in (0,1]$ such that $\varepsilon B \subseteq N'$. Since each $f \in \Phi$ is convex and $f(0) = 0$, for all $u \in U$ we have this estimate:

$$
\begin{aligned}
-\frac{\varepsilon\delta}{1+\delta}f(x_0-u) - \frac{\varepsilon}{1+\delta}f(-\delta x_0) \;&\preccurlyeq\; -\varepsilon f\left(\frac{-\delta}{1+\delta}u\right) \;\preccurlyeq\; -f\left(\frac{-\varepsilon\delta}{1+\delta}u\right)\\
\preccurlyeq\; f\left(\frac{\varepsilon\delta}{1+\delta}u\right) \;&\preccurlyeq\; \varepsilon f\left(\frac{\delta}{1+\delta}u\right) \;\preccurlyeq\; \frac{\varepsilon\delta}{1+\delta}f(x_0+u) + \frac{\varepsilon}{1+\delta}f(-\delta x_0).
\end{aligned}
$$

The vectors on the extreme ends of this estimate belong to $N' + N' + N'$, since N' is balanced and $f(x_0 \pm u) - f(x_0) \in N'$. Since $N' + N' + N'$ is contained in N, which is full, we have $f\left(\frac{\varepsilon\delta}{1+\delta}u\right) \in N$. Let $G = \frac{\varepsilon\delta}{1+\delta}U$; this completes the proof.

27.13. *Proposition.* Let Ω be an open convex subset of a real TVS. Suppose $f : \Omega \to \mathbb{R}$ is a convex function — or, more generally, suppose that $f : \Omega \to Z$ is a convex function, where Z is some locally full ordered topological vector space.

Suppose that f is bounded above on some nonempty open set — i.e., there exists some nonempty open set $G \subseteq \Omega$ and some $z_0 \in Z$ such that $f(x) \preccurlyeq z_0$ for all $x \in G$.

Then f is continuous. (In particular, any real-valued, upper-semicontinuous, convex function on an open convex set is continuous.)

Proof. By translation we may assume $0 \in G$ and $f(0) = 0$. Replacing G with a smaller open set, we may assume G is balanced. Say $f(x) \preccurlyeq z_0$ for all $x \in G$. Then

$$
0 \;=\; f(0) \;=\; f\left(\frac{x+(-x)}{2}\right) \;\preccurlyeq\; \frac{1}{2}f(x) + \frac{1}{2}f(-x) \;\preccurlyeq\; \frac{1}{2}f(x) + \frac{1}{2}z_0,
$$

so $f(x) \succcurlyeq -z_0$ for all $x \in G$. Thus f is order bounded on G.

Let any positive integer n be given. If $x \in \frac{1}{n}G$, then $nx \in G$, so

$$
f(x) \;=\; f\left(\frac{1}{n}nx + \frac{n-1}{n}0\right) \;\preccurlyeq\; \frac{1}{n}f(nx) + \frac{n-1}{n}f(0) \;\preccurlyeq\; \frac{1}{n}z_0.
$$

Thus f is bounded above by $\frac{1}{n}z_0$ on $\frac{1}{n}G$. By the argument of the preceding paragraph, f is bounded below by $-\frac{1}{n}z_0$ on $\frac{1}{n}G$. By the Squeeze Property (26.52(E)), it follows that $\lim_{x\to 0} f(x) = 0$; thus f is continuous at 0. By 27.12, f is continuous everywhere on Ω.

Corollary. Suppose Ω is an open convex subset of $\mathbb{R}^n$. Then any convex function $f : \Omega \to \mathbb{R}$ is continuous.

Proof of corollary. For any $x \in \Omega$, let $N(x)$ be a closed n-dimensional cube centered at x, small enough to be contained in Ω. That cube has 2^n vertices $v_1, v_2, \ldots, v_{2^n}$. Each point u in $N(x)$ is a convex combination of the v_j's, and so $\sup_{u\in N(x)} f(u) \preccurlyeq \max_j f(v_j)$.

Remark. A convex function on an infinite-dimensional normed space is not necessarily continuous; see 23.6.a and 23.6.b.

27.14. *Proposition.* Let X be a topological vector space whose topology is given by an F-norm. Let Ω be an open convex subset of X. If $f : \Omega \to \mathbb{R}$ is convex and continuous, then f is locally Lipschitz.

Proof. Suppose that the topology on X is given by an F-norm ρ. Let any point $x_0 \in \Omega$ be given; we shall show f is Lipschitzian on some neighborhood of x_0. (This argument is based on Roberts and Varberg [1974].) Let B_r denote the open ball of radius r centered at x_0. Choose $r > 0$ small enough and M large enough so that $B_{2r} \subseteq \Omega$ and $f(\cdot) \leq M$ on B_{2r} and $-f(x_0) \leq M$; we shall show f is Lipschitzian on B_r with $\langle f \rangle_{\mathrm{Lip}} \leq 4M/r$. Note that for any $u \in X$ with $\rho(u) < 2r$, we have

$$\begin{array}{rcccl} -M & \leq & f(x_0) & = & f\left(\dfrac{(x_0+u)+(x_0-u)}{2}\right) \\ & & & \leq & \dfrac{f(x_0+u)}{2} + \dfrac{f(x_0-u)}{2} \quad \leq \quad \dfrac{f(x_0+u)}{2} + \dfrac{M}{2} \end{array}$$

and therefore $f(\cdot) \geq -3M$ on B_{2r}.

Let any distinct points $x_1, x_2 \in B_r$ be given; let $\alpha = \rho(x_1 - x_2)$. Let $x_3 = x_2 + \frac{r}{\alpha}(x_2 - x_1)$; note that $x_3 \in B_{2r}$. Compute

$$f(x_2) \quad = \quad f\left(\frac{r}{r+\alpha}x_1 + \frac{\alpha}{r+\alpha}x_3\right) \quad \leq \quad \frac{r}{r+\alpha}f(x_1) + \frac{\alpha}{r+\alpha}f(x_3)$$

and thus

$$f(x_2) - f(x_1) \quad \leq \quad \frac{\alpha}{r+\alpha}\left[f(x_3) - f(x_1)\right] \quad \leq \quad \frac{4M\alpha}{r+\alpha} \quad \leq \quad \frac{4M\alpha}{r} \quad = \quad \frac{4M}{r}\rho(x_1 - x_2).$$

Similarly, $f(x_1) - f(x_2) \leq \frac{4M}{r}\rho(x_1 - x_2)$.

DIMENSION IN TVS'S

27.15. Theorem (Tychonov). If X is a finite-dimensional vector space over the field $\mathbb{F}$, then there is one and only one Hausdorff TVS topology $\mathcal{T}$ on X. Moreover, it can be specified as follows: If $\{e_1, e_2, \ldots, e_n\}$ is any basis for X, then

$$f \quad : \quad (c_1,\ c_2,\ \ldots,\ c_n) \quad \mapsto \quad c_1e_1 \ + \ c_2e_2 \ + \ \cdots \ + \ c_ne_n$$

is a linear homeomorphism from $\mathbb{F}^n$ (with its product topology) onto $(X, \mathcal{T})$.

Hints: Certainly f is a linear bijection. Let $\mathcal{T}$ be any Hausdorff TVS topology on X; we wish to show f is then a homeomorphism. Certainly f is continuous since $(X, \mathcal{T})$ is a TVS and therefore the vector operations are continuous in $(X, \mathcal{T})$. The product topology on $\mathbb{F}^n$ can be given by any of the usual norms on $\mathbb{F}^n$ (see 22.11); let $\| \ \ \|$ be any of those norms. To show that f^{-1} is continuous, let $B = \{v \in \mathbb{F}^n : \|v\| < 1\}$; it suffices (why?) to show that

$f(B)$ is a neighborhood of 0 in $(X, \mathcal{T})$. Let $S = \{v \in \mathbb{F}^n : \|v\| = 1\}$. Explain: S is compact; $f(S)$ is compact; $f(S)$ is closed; $f(\mathbb{F}^n \setminus S)$ is a neighborhood of 0; $f(\mathbb{F}^n \setminus S) \supseteq V$ where V is some balanced neighborhood of 0; $\mathbb{F}^n \setminus S \supseteq f^{-1}(V)$ and $f^{-1}(V)$ is balanced; $B \supseteq f^{-1}(V)$; $f(B) \supseteq V$.

27.16. Corollaries.

a. Any finite-dimensional subspace of a Hausdorff linear topological space is complete, hence closed.

b. The only Hausdorff topological vector space that is totally bounded is the trivial space $\{0\}$.

Hint: If X contains a nonzero vector v, then it contains the one-dimensional space $\mathbb{F}v$, which is isomorphic to $\mathbb{F}$ — which is not totally bounded.

27.17. Theorem (F. Riesz). Let X be a Hausdorff TVS. Then the following conditions are equivalent:

(A) X is finite dimensional.

(B) X is locally compact.

(C) 0 has a compact neighborhood in X.

(D) 0 has a neighborhood that is totally bounded.

Proof. The implications (A) $\Rightarrow$ (B) $\Rightarrow$ (C) $\Rightarrow$ (D) are clear. It remains only to show (D) $\Rightarrow$ (A). Let U be a totally bounded neighborhood of 0 in X. Then $\frac{1}{2}U$ is also a neighborhood of 0. By 26.40.e, there is some finite set $F \subseteq U$ such that $F + \frac{1}{2}U \supseteq U$. Let M be the span of the finite set F; it suffices to show that $M = X$. We know that M is closed, by 27.16.a. Hence X/M is a Hausdorff TVS, by 26.34. Let $\pi : X \to X/M$ be the quotient map.

Note that $\pi(F) = \{0\}$. From $F + \frac{1}{2}U \supseteq U$ we deduce that $\frac{1}{2}\pi(U) \supseteq \pi(U)$; that is, $\pi(U) \supseteq 2\pi(U)$. By induction, $\pi(U) \supseteq 2^n\pi(U)$ for all n. Since π is an open mapping, $\pi(U)$ is a neighborhood of 0 in X/M; hence $\bigcup_{n=1}^{\infty} 2^n\pi(U) = X/M$. Therefore $\pi(U) = X/M$.

Since π is a uniformly continuous mapping and U is totally bounded, we deduce that $\pi(U) = X/M$ is totally bounded. By 27.16.b, $\pi(U) = \{0\}$, and thus $M = X$.

27.18. *Proposition on dimension and norms.* As usual, we assume that the scalar field $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. Assume conventional set theory (that is, ZF + AC), and the Continuum Hypothesis (CH). Then:

(i) If X is an infinite-dimensional F-space, then $\dim(X) = \mathrm{card}(X) \geq \mathrm{card}(\mathbb{R})$.

(ii) If X is an infinite-dimensional separable F-space, then $\dim(X) = \mathrm{card}(X) = \mathrm{card}(\mathbb{R})$.

(iii) If X is a vector space with $\dim(X) = \mathrm{card}(\mathbb{R})$, then there exist at least two inequivalent complete norms on X. (This conclusion should be contrasted with 27.47.b; see also 22.8.)

Proof. If $\dim(X) = \mathrm{card}(\mathbb{N})$, then X is the union of countably many finite-dimensional subspaces. However, an F-space cannot have that property, by 27.16.a and 20.16. Thus $\dim(X) > \mathrm{card}(\mathbb{N})$. By the Continuum Hypothesis, it follows that $\dim(X) \geq \mathrm{card}(\mathbb{R})$. By 11.34, it follows that $\mathrm{card}(X) = \dim(X)$. If X is separable, then $\mathrm{card}(X) \leq \mathrm{card}(\mathbb{R})$ by 15.37.a. This proves (i) and (ii).

For (iii), let X be a vector space with $\dim(X) = \mathrm{card}(\mathbb{R})$. For $1 \leq p < \infty$, define the normed space ℓ_p as in 22.25; it is complete by 22.31(i). The sequences with only finitely many nonzero terms are dense in ℓ_p; from this it follows that ℓ_p is separable. Thus $\dim(\ell_p) = \mathrm{card}(\mathbb{R})$. By 11.18.d, there exists a linear bijection f_p from X onto ℓ_p. We can define a norm $\| \;\; \|_{(p)}$ on X by taking $\|x\|_{(p)} = \|f_p(x)\|_p$; then the normed space $(X, \| \;\; \|_{(p)})$ is isomorphic to the Banach space $(\ell_p, \| \;\; \|_p)$. The norms $\| \;\; \|_{(1)}$ and $\| \;\; \|_{(2)}$ cannot be equivalent, since the Banach spaces ℓ_1 and ℓ_2 have different topological properties. (For instance, ℓ_2 is reflexive while ℓ_1 is not — see 28.41, 28.50, 28.51, and 23.10.) This argument is taken from Day [1973].

Fixed Point Theorems of Brouwer, Schauder, and Tychonov

27.19. Following are several variants of **Brouwer's Fixed Point Theorem**, in order of increasing generality.

Simplex version. Let n be a positive integer. Let Δ be the *standard n-simplex*; that is, the set

$$\Delta \quad = \quad \left\{ u \in \mathbb{R}^n \ : \ u_1, u_2, \ldots, u_n \geq 0 \text{ and } \sum_{j=1}^{n} u_j \leq 1 \right\}.$$

Then any continuous function $f : \Delta \to \Delta$ has at least one fixed point.

Convex finite-dimensional version. Let Q be a compact convex subset of $\mathbb{R}^n$. Then any continuous function $f : Q \to Q$ has at least one fixed point.

Schauder's Fixed Point Theorem. Any continuous self-mapping of a compact convex subset of a Banach space has at least one fixed point.

Tychonov's Fixed Point Theorem. Any continuous self-mapping of a compact convex subset of a Hausdorff locally convex space has at least one fixed point.

Approximate Fixed Point Theorem. Let K be a compact convex subset of a Hausdorff locally convex space X. Let $f : K \to K$ be any mapping (not necessarily continuous or measurable). Then there exists some point $\xi \in K$ that is an "approximate fixed point" of f, in the following sense:

$$\xi \quad \in \quad \bigcap_{V\in\mathcal{N}} \operatorname{clco} f(\xi+V).$$

Here $\mathcal{N}$ is the filter of neighborhoods of 0 in X, clco stands for closed convex hull, and $f(\xi+V)$ stands for $f\Big((\xi+V)\cap\operatorname{Dom}(f)\Big)$.

Remarks. Our proof is based on the Second Approximate Fixed Point Theorem in 3.37, which we proved using Maaren's Theorem in 3.36. Some analysts may prefer the proof of Rogers [1980], which is shorter but assumes familiarity with the use of Jacobians in the formula for a change of variables in integration — a formula that is well known but not at all trivial to prove. Still other mathematicians may prefer a proof by simplicial triangulations; a fairly brief, self-contained presentation of that proof is given by Border [1985].

Proof of simplex version. Let $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ be a sequence of positive numbers decreasing to 0. For each positive integer m, choose a set $S = S_m \subseteq \Delta$ satisfying the conditions of the Second Approximate Fixed Point Theorem in 3.37 with diameter less than $\varepsilon = \varepsilon_m$. Let ξ_m be any point in S_m. Since Δ is compact, the sequence (ξ_m) has a convergent subsequence. For simplicity of notation, replace (ε_m), (S_m), (ξ_m) with subsequences, so that (ξ_m) converges to a limit ξ in Δ. Since f is continuous, ξ satisfies the inequalities that were satisfied by the u's and v's in 3.37, but with ε replaced by 0. That is, $\xi_i \leq f(\xi)_i$ for $i = 1, 2, \ldots, n$, and $\sum_{i=1}^n \xi_i \geq \sum_{i=1}^n f(\xi)_i$. It follows that $f(\xi) = \xi$.

Proof of convex, finite-dimensional version. By translation and rescaling, we may assume that Q is contained in the simplex Δ. Let $\mathbb{R}^n$ be equipped with the Euclidean metric, and let $\kappa : \mathbb{R}^n \to Q$ be the closest-point projection; then κ is continuous by 22.45 or 22.51. The inclusion maps $i_1 : \Delta \to \mathbb{R}^n$ and $i_2 : Q \to \Delta$ are also continuous. Hence the composition

$$\Delta \xrightarrow{i_1} \mathbb{R}^n \xrightarrow{\kappa} Q \xrightarrow{f} Q \xrightarrow{i_2} \Delta$$

is continuous. By the simplex version of the Fixed Point Theorem, this composition has at least one fixed point ξ in Δ. Since i_1 and i_2 are inclusions and κ is idempotent with range Q, that fixed point must actually lie in Q and must be a fixed point of f.

Proof of Schauder's and Tychonov's Fixed Point Theorems. Schauder's Theorem is a special case of Tychonov's Theorem, which is in turn an easy corollary of the Approximate Fixed Point Theorem; thus it suffices to prove that result.

Proof of Approximate Fixed Point Theorem. This proof is a slight modification of an argument of Marchi and Martínez-Legaz [1991]. Most of this proof will be devoted to showing that

$$\text{for each } V \in \mathcal{N}, \text{ there exists some } x_V \in \operatorname{co} f(x_V + V). \qquad (**)$$

Let any neighborhood V of 0 be given. Replacing V with a smaller neighborhood, we may assume V is open. Define $T_V : K \to \{\text{nonempty subsets of } K\}$ by

$$T_V(x) \quad = \quad f\big((x+V)\cap K\big).$$

Since V is open, for each $y \in K$ the set $T_V^{-1}(y) = \{x \in X : y \in T_V(x)\}$ is open (easy exercise). Since K is compact, the open cover $\{T_V^{-1}(y) : y \in K\}$ has a finite subcover; say it is given by $\{T_V^{-1}(y) : y \in Y(V)\}$ for some finite set $Y(V) \subseteq K$. (The particular finite subcover is not necessarily uniquely determined, but we select some particular finite subcover. At this step and at several other steps in this proof, we make arbitrary selections, which can be justified most easily by the Axiom of Choice.)

Let $\{\beta_y : y \in Y(V)\}$ be a continuous partition of unity corresponding to this covering (see 16.29 and 17.7.g) and define $p_V : K \to \operatorname{co} Y(V)$ by

$$p_V(x) \quad = \quad \sum_{y \in Y(V)} \beta_y(x) y.$$

Then p_V is continuous. For each $y \in Y$ and $x \in K$, we have $\beta_y(x) > 0$ only if $x \in T_V^{-1}(y)$; that is, $y \in T_V(x)$. Thus we have in fact $p_V(x) \in \operatorname{co}\bigl[Y(V) \cap f\bigl((x+V) \cap K\bigr)\bigr]$ for all $x \in K$.

Since $Y(V)$ is finite, its span is a finite-dimensional subspace of X, which is isomorphic to Euclidean space by 27.15. Moreover, $\operatorname{co} Y(V)$ is compact by 26.23.g. Thus the restriction of p_V to $\operatorname{co} Y(V)$ is a continuous self-mapping of a compact convex subset of Euclidean space, which has at least one fixed point x_V by the finite-dimensional convex set version of Brouwer's Theorem. This completes the proof of (**).

Let $\mathcal{N}$ be the filter of all neighborhoods of 0 in X, ordered by reverse inclusion. Then $(x_V : V \in \mathcal{N})$ is a net in the compact set K, so it has a subnet convergent to some limit $\xi \in K$. Fix any $V \in \mathcal{N}$; it suffices to show that $\xi \in \operatorname{clco} f(\xi + V)$.

Choose a neighborhood U of 0 such that $U + U \subseteq V$. Then for all neighborhoods W sufficiently small, we have $x_W \in \xi + U$ and $W \subseteq U$, hence $x_W + W \subseteq \xi + U + U \subseteq \xi + V$, hence $x_W \in \operatorname{clco} f(x_W + W) \subseteq \operatorname{clco} f(\xi + V)$. Since $\operatorname{clco} f(\xi + V)$ is a closed set, any cluster point of the x_W's must lie in that set; in particular, ξ lies in that set.

Barrels and Ultrabarrels

27.20. *Remarks.* Ultrabarrels are a generalization of barrels. Barrels are simpler to define, but they are mainly useful in locally convex spaces; ultrabarrels can be useful in the more general setting of topological vector spaces. The theories of barrels in LCS and ultrabarrels in TVS are closely analogous; the analogy will be developed in the sections below.

The definitions of barrels and ultrabarrels involve absorbing sets (defined in 12.8). In a TVS, absorbing sets may be viewed as "generalized neighborhoods of 0" — any neighborhood of 0 is absorbing, but not every absorbing set is a neighborhood of 0. For instance, sketch a graph of $\{(x, y) \in \mathbb{R}^2 : |y| \geq x^2 \text{ or } y = 0\}$; show that this set is absorbing but is not a neighborhood of $(0, 0)$ when $\mathbb{R}^2$ is equipped with its usual topology.

27.21. *Definition.* Let X be a topological vector space. A **barrel** in X is a subset of X that is closed, convex, balanced, and absorbing. (Those terms are defined in 5.13, 12.3, and 12.8.)

27.22. *Basic properties of barrels.* Let X be a topological vector space.

a. If X is a locally convex space, then X has a neighborhood base at 0 consisting of barrels.

b. If ρ is a continuous seminorm on X and $k > 0$, then $\{x \in X : \rho(x) \leq k\}$ is a barrel.

c. If B is a barrel in X, then its Minkowski functional μ_B is a seminorm on X. Moreover, μ_B is continuous if and only if B is a neighborhood of 0.

27.23. *Definition.* Let X be a vector space. A **string** in X is a sequence of sets $(S_n : n \in \mathbb{N})$ that are balanced, absorbing, and satisfy $S_n \supseteq S_{n+1} + S_{n+1}$ for all n. The S_n's are then called the **knots** of the string.

In a topological vector space, a **closed string** is a string whose knots are closed sets; those closed sets are called **ultrabarrels**. We still have a string if we discard the first few knots of a string; hence every ultrabarrel may also be viewed as the first knot of a closed string.

In a topological vector space, a **neighborhood string** is a string, all of whose knots are neighborhoods of 0. (Some mathematicians call this a *topological string.*)

27.24. *Basic properties of strings and ultrabarrels.*

a. If B is a barrel in a TVS X, then B is also an ultrabarrel — it is a knot of the closed convex string (S_n) defined by $S_n = 2^{-n}B$.

b. If (S_n) is a string in a vector space X, then $\{S_n\}$ forms a neighborhood base at 0 for a TVS topology on X.

Conversely, if X is a TVS, then X has a neighborhood base at 0 consisting of ultrabarrels.

c. If ρ is an F-seminorm on a vector space X and k is a positive constant, then the sets $S_n = \{x \in X : \rho(x) \leq 2^{-n}\}$ form a string.

If ρ is a continuous F-seminorm on a TVS X, then the sequence (S_n) defined as above is a closed string; thus its members are ultrabarrels.

d. If (S_n) is a string in a vector space X, then there exist an F-seminorm ρ on X and positive numbers a_n, b_n decreasing to 0 that satisfy

$$\{x \in X \ : \ \rho(x) \leq a_n\} \qquad \subseteq \qquad S_n \qquad \subseteq \qquad \{x \in X \ : \ \rho(x) \leq b_n\}$$

for all n. (*Hint*: The sets $V_n = \{(x, y) \in X \times X : x - y \in S_{2n}\}$ satisfy the hypotheses of 4.44.)

Suppose, moreover, that X is a TVS. Then ρ is continuous if and only if (S_n) is a neighborhood string. (*Hint*: An F-seminorm is continuous if and only if it is continuous at 0.)

e. If (S_n) and (T_n) are strings and $S_n + T_n = U_n$, then (U_n) is a string.

27.25. *Proposition.* Suppose X is a complete metric space or, more generally, a Baire space. Then:

(i) If X is a TVS, then X is *ultrabarrelled*, as defined in 27.26(U1).

(ii) If X is an LCS, then X is *barrelled*, as defined in 27.27(B1).

Proof. It suffices to prove (i), since that result implies (ii) by 27.24.a. Let (S_n) be a closed string; we wish to show that S_1 is a neighborhood of 0. Since S_2 is absorbing, $X = \bigcup_{k=1}^{\infty} kS_2$. By 20.15(B), some kS_2 has nonempty interior. Hence S_2 has nonempty interior. Say $x_0 \in \mathrm{int}(S_2)$. Thus $x_0 + G \subseteq S_2$ where G is some neighborhood of 0. Since S_2 is symmetric, we have $0 + (G - G) = (x_0 + G) - (x_0 + G) \subseteq S_2 - S_2 = S_2 + S_2 \subseteq S_1$.

27.26. Theorem and definition. Let $(X, \mathcal{T})$ be a real or complex topological vector space. Then the following conditions on $(X, \mathcal{T})$ are equivalent. If any, hence all, of these conditions are satisfied, we say $(X, \mathcal{T})$ is **ultrabarrelled**.

(U1) Every ultrabarrel in X is a neighborhood of 0.

(U2) **(F-Seminorms Property.)** Each lower semicontinuous F-seminorm on X is continuous.

(U3) **(Banach's Closed Graph Property.)** Let Y be an F-space. Let $f : X \to Y$ be linear and have closed graph. Then f is continuous.

(U4) **(Neumann's Nonlinear Closed Graph Property.)** Let Y be a locally full F-space. Let $\Omega \subseteq X$ be an open convex set. Suppose $f : \Omega \to Y$ is a convex operator whose graph is a closed subset of $\Omega \times Y$. Then f is continuous.

(U5) **(Banach-Steinhaus Uniform Boundedness Property.)** Let Y be a topological vector space. Let Φ be a collection of continuous linear maps from X into Y that is toplinearly bounded pointwise. Then Φ is equicontinuous.

(U6) **(Neumann's Nonlinear Uniform Boundedness Property.)** Let Y be an ordered topological vector space that is locally full. Let $\Omega \subseteq X$ be an open convex set. Let Φ be a collection of continuous convex maps from Ω into Y. Suppose Φ is toplinearly bounded pointwise. Then Φ is equicontinuous.

Proof of this theorem begins in Section 27.31.

27.27. Theorem and Definition. Let $(X, \mathcal{T})$ be a real or complex locally convex space. Then the following conditions on $(X, \mathcal{T})$ are equivalent. If any, hence all, of these conditions are satisfied, we say $(X, \mathcal{T})$ is **barrelled**.

(B1) Every barrel in X is a neighborhood of 0.

(B2) **(Seminorms Property.)** Each lower semicontinuous seminorm on X is continuous.

(B3) **(Closed Graph Property.)** Let Y be a Fréchet space. Let $f : X \to Y$ be linear and have closed graph. Then f is continuous.

(B4) **(Neumann's Nonlinear Closed Graph Property.)** Let Y be a locally full Fréchet space. Let $\Omega \subseteq X$ be an open convex set. Suppose $f : \Omega \to Y$ is a convex function, whose graph is a closed subset of $\Omega \times Y$. Then f is continuous.

(B5) **(Uniform Boundedness Property.)** Let Y be a locally convex space. Let Φ be a collection of continuous linear maps from X into Y that is toplinearly bounded pointwise. Then Φ is equicontinuous.

(B6) **(Neumann's Nonlinear Uniform Boundedness Property.)** Let Y be a locally full, locally convex space. Let $\Omega \subseteq X$ be an open convex set. Let Φ be a collection of continuous convex maps from Ω into Y that is toplinearly bounded pointwise. Then Φ is equicontinuous.

Proof of this theorem begins in Section 27.31.

Remark. A seventh characterization of barrelled spaces will be given in 28.30.

27.28. Corollaries: classical versions. Let X be an F-space. By 27.25:

a. Any lower semicontinuous F-seminorm on X is continuous.

b. Uniform Boundedness Theorem. If Y is a topological vector space and Φ is a collection of continuous linear maps from X into Y such that $\Phi(x) = \{f(x) : f \in \Phi\}$ is a bounded subset of Y for each $x \in X$, then Φ is equicontinuous.

c. Closed Graph Theorem. If Y is an F-space and $f : X \to Y$ is a linear map whose graph is a closed subset of $X \times Y$, then f is continuous.

d. If Ω is an open convex subset of X and $f : \Omega \to \mathbb{R}$ is a convex function whose graph is a closed subset of $\Omega \times \mathbb{R}$, then f is continuous.

27.29. *Example application.* Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, let X be a Banach space, and let $\alpha, \beta \in (0, +\infty)$ be exponents such that $L^\alpha(\mu, X) \subseteq L^\beta(\mu, X)$. Then the inclusion $L^\alpha(\mu, X) \xrightarrow{\subseteq} L^\beta(\mu, X)$ is continuous.

Proof. It suffices to show that the inclusion map has closed graph. Suppose $f_n \to f$ in $L^\alpha(\mu, X)$ and $f_n \to g$ in $L^\beta(\mu, X)$; we are to show that $f = g$. By 22.31(ii) we may pass to subsequences such that $f_n \to f$ and $f_n \to g$ pointwise μ-almost everywhere.

Remarks. This example is taken from Villani [1985]. That paper also shows the following interesting result: Let X be a Banach space, let $(\Omega, \mathcal{S}, \mu)$ be a measure space, and let $\alpha, \beta \in (0, +\infty)$ with $\alpha < \beta$. Then

$$L^\alpha(\mu, X) \subseteq L^\beta(\mu, X) \quad \text{if and only if} \quad \inf\{\mu(S) \ : \ S \in \mathcal{S},\ \mu(S) > 0\} \ > \ 0;$$

$$L^\alpha(\mu, X) \supseteq L^\beta(\mu, X) \quad \text{if and only if} \quad \sup\{\mu(S) \ : \ S \in \mathcal{S},\ \mu(S) < \infty\} \ < \ \infty.$$

Special cases of this were given in 22.34. (Villani's paper only shows this for $X = \mathbb{R}$, but that case easily yields the general case since all the functions in $L^\alpha(\mu, X)$ or $L^\beta(\mu, X)$ are measurable, and we can separate the "regular" condition from the "not too big" condition — see the remarks in 22.28.)

27.30. *Change of scalar field.* Let X be a complex topological vector space (respectively, a complex locally convex space). Then X, with the *same topology*, may also be viewed as a real topological vector space (respectively, a real locally convex space) if we "forget" how to multiply members of X by nonreal scalars. Let us denote these two TVS's by $X_{\mathbb{C}}$ and $X_{\mathbb{R}}$. Note that the choice of scalars affects the definitions of "balanced" and "absorbing;" hence it affects the definitions of "barrel" and "ultrabarrel." Show that

a. If B is an (ultra)barrel in $X_{\mathbb{C}}$, then B is also an (ultra)barrel in $X_{\mathbb{R}}$. Likewise, any (F-)seminorm on $X_{\mathbb{C}}$ is also an (F-)seminorm on $X_{\mathbb{R}}$.

Hence, if $X_{\mathbb{R}}$ is (ultra)barrelled, then $X_{\mathbb{C}}$ is (ultra)barrelled, too.

b. It is possible for $X_{\mathbb{R}}$ to have more (ultra)barrels than $X_{\mathbb{C}}$.

For instance, show that the set $B = \{z \in \mathbb{C} : |\mathrm{Re}(z)| \leq 1 \text{ and } |\mathrm{Im}(z)| \leq 2\}$ is both a barrel and an ultrabarrel in $X_{\mathbb{R}}$, but is neither in $X_{\mathbb{C}}$ since B is not balanced in $X_{\mathbb{C}}$.

c. Nevertheless, $X_{\mathbb{R}}$ is (ultra)barrelled if and only if $X_{\mathbb{C}}$ is (ultra)barrelled.

For the moment, we shall prove this equivalence using only definitions (U2) and (B2); proofs with the other definitions in 27.26 and 27.27 will follow from the arguments given in the next subchapter.

We have already established half of this "if and only if" result. Now assume $X_{\mathbb{C}}$ is (ultra)barrelled — i.e., it satisfies condition (U2) or (B2). To show the same for $X_{\mathbb{R}}$, let ρ be any lower semicontinuous (F-)seminorm on $X_{\mathbb{R}}$; we wish to show that ρ is continuous on $X_{\mathbb{R}}$. Note that $X_{\mathbb{R}}$ and $X_{\mathbb{C}}$ differ only in their algebraic operations — they are the same set, and they have the same topology; so a function is continuous on $X_{\mathbb{R}}$ if and only if it is continuous on $X_{\mathbb{C}}$. Define $\gamma : X \to [0, +\infty)$ as in 26.5.b. As we noted in 26.5.b, this function γ is also lower semicontinuous on X, and γ is an (F-)seminorm on $X_{\mathbb{C}}$. Hence, by our assumption, γ is continuous. From the inequality $\rho \leq \gamma$, we see that ρ is continuous at 0. Since ρ is a G-seminorm, we have $|\rho(u) - \rho(v)| \leq \rho(u - v)$, and therefore ρ is continuous.

Proofs of Barrel Theorems

27.31. We now begin the somewhat lengthy proof of 27.26 and 27.27. We remark that shorter proofs of equivalence can be found in the literature (for instance, in Waelbroeck [1971]) if one omits the nonlinear conditions (U4), (U6), (B4), and (B6).

The order of proof will not be the same as the order in which the results were stated. We shall cover the barrels and ultrabarrels cases simultaneously. In the discussions below, phrases in brackets should be read or omitted for the two cases — e.g., an [F-]seminorm means a seminorm for the argument with barrels or an F-seminorm for the argument with ultrabarrels. Also, (1) will refer to either (U1) or (B1), and (2) will refer to either (U2) or (B2), etc. We shall prove the equivalence in this order:

- (1) $\Longleftrightarrow$ (2),
- (4) $\Rightarrow$ (3) $\Rightarrow$ (2),
- (6) $\Rightarrow$ (5) $\Rightarrow$ (1), and
- (1) implies both (4) and (6).

In each argument the implication will be proved with either choice of scalar field ($\mathbb{R}$ or $\mathbb{C}$); in fact, the choice of scalar field will not enter into most of the arguments. Our first few proofs are along the lines of Waelbroeck [1971]. Some researchers may also find Adasch, Ernst, and Keim [1978] helpful for further reading on this topic.

27.32. *Proof of* (1) $\Rightarrow$ (2). Let ρ be a lower semicontinuous [F-]seminorm; then the sets $S_n = \{x \in X : \rho(x) \leq 2^{-n}\}$ are closed. Each S_n is also absorbing, since ρ is scalarly continuous (see 26.3.a). It follows easily that the sequence (S_n) is a closed [convex] string. By (1), then, it is a neighborhood string; thus each S_n is a neighborhood of 0. It follows easily that ρ is continuous at 0. Since $|\rho(u) - \rho(v)| \leq \rho(u-v)$, it follows that ρ is continuous everywhere.

27.33. *Proof of* (2) $\Rightarrow$ (1). Let V_0 be an [ultra]barrel; we wish to show that V_0 is a neighborhood of 0. It suffices to produce a lower semicontinuous [F-]seminorm ρ with the property that

$$\{x \in X : \rho(x) < 1\} \quad \subseteq \quad V_0. \qquad (\natural)$$

For the locally convex case, let ρ be the Minkowski functional of V_0; that is, $\rho(x) = \inf\{k \in (0, +\infty] : k^{-1}x \in V_0\}$. Then ρ is a seminorm satisfying ($\natural$), as noted in 12.28 and 12.29.g. To show that ρ is lower semicontinuous, suppose $\rho(x_0) > r > s > 0$. Then $x_0 \notin rV_0$. Since rV_0 is closed, its complement is a neighborhood of x_0, and for x in that neighborhood we have $r^{-1}x \notin V_0$, hence $\rho(x) \geq r > s$. Thus for any s the set $\{x \in X : \rho(x) > s\}$ is open.

For the non-locally-convex case, let $(V_j : j = 0, 1, 2, 3, \ldots)$ be a closed string in X. By a **dyadic rational** in $[0, 1)$ we mean a number of the form

$$\alpha \quad = \quad \frac{t_1}{2} + \frac{t_2}{4} + \frac{t_3}{8} + \cdots + \frac{t_n}{2^n}$$

for some positive integer n, where each t_j is either 0 or 1. For each number of this type, define the set

$$W_\alpha \quad = \quad \sum_{j=1}^{n} t_j V_j \quad = \quad \sum_{\{j \in \mathbb{N} : t_j = 1\}} V_j \quad \subseteq \quad V_0.$$

Verify that the W_α's are balanced and absorbing. Also, for any dyadic rationals α, β with $\alpha + \beta < 1$ we have and $W_\alpha + W_\beta \subseteq W_{\alpha+\beta}$, hence $\mathrm{cl}(W_\alpha) + \mathrm{cl}(W_\beta) \subseteq \mathrm{cl}(W_{\alpha+\beta})$ by 26.22.e. Now define

$$\rho(x) \quad = \quad \inf\{\alpha \in [0, 1) : x \in \mathrm{cl}(W_\alpha)\},$$

with $\rho(x) = 1$ if $x \notin \bigcup_{\alpha \in [0,1)} \mathrm{cl}(W_\alpha)$. Verify (*exercise*) that ρ is an F-seminorm satisfying ($\natural$). To show that ρ is lower semicontinuous, suppose $\rho(x_0) > c$. Then $\rho(x_0) > \alpha > c$ for some dyadic rational α, and therefore $x_0 \notin \mathrm{cl}(W_\alpha)$. The complement of $\mathrm{cl}(W_\alpha)$ is an open set on which $\rho(\cdot) \geq \alpha > c$. Thus the set $\{x \in X : \rho(x) > c\}$ is open for any c.

27.34. *Proof of* (3) $\Rightarrow$ (2). Let σ be a lower semicontinuous [F-]seminorm on X. The linear subspace $K = \sigma^{-1}(0) = \{x \in X : \sigma(x) \leq 0\}$ is closed. Let $Q = X/K$ be the quotient space, and let $\pi : X \to Q$ be the canonical map. Then an [F-]norm $\widehat{\sigma}$ is defined on Q by $\widehat{\sigma}(\pi(x)) = \sigma(x)$. We topologize Q with this [F-]norm. (We do not claim that the resulting topology is the quotient topology.) Let C be the completion of the [F-]normed space $(Q, \widehat{\sigma})$. Let its [F-]norm, an extension of $\widehat{\sigma}$, again be denoted by $\widehat{\sigma}$; then $(C, \widehat{\sigma})$ is a complete [F-]normed space.

Let $i : Q \xrightarrow{\subseteq} C$ be the inclusion. We claim that the composition $i \circ \pi : X \xrightarrow{\pi} Q \xrightarrow{i} C$ has closed graph. To see this, let $((x_\alpha, q_\alpha))$ be any net in the graph of $i \circ \pi$, converging in $X \times C$ to some point (x, q); we shall show that (x, q) actually lies in the graph of $i \circ \pi$. Let any number $\varepsilon > 0$ be given; it suffices to show that $\widehat{\sigma}\Big(q - (i \circ \pi)(x)\Big) < 2\varepsilon$. Since C is the completion of Q, there is some $q' \in Q$ with $\widehat{\sigma}(q' - q) < \varepsilon$. Since $Q = \pi(X)$, we may choose some $x' \in \pi^{-1}(q')$. Now compute

$$\begin{aligned}
\widehat{\sigma}\Big(q - (i \circ \pi)(x)\Big) - \varepsilon \;&\le\; \widehat{\sigma}\,(q - q') + \widehat{\sigma}\,(q' - \pi(x)) - \varepsilon \\
&<\; \widehat{\sigma}\,(q' - \pi(x)) \;=\; \widehat{\sigma}\,(\pi(x') - \pi(x)) \;=\; \sigma(x' - x) \;\le\; \liminf_{\alpha} \sigma(x' - x_\alpha) \\
&=\; \liminf_{\alpha} \widehat{\sigma}\,(\pi(x' - x_\alpha)) \;=\; \liminf_{\alpha} \widehat{\sigma}\,(q' - q_\alpha) \;=\; \widehat{\sigma}(q' - q) \;<\; \varepsilon.
\end{aligned}$$

Hence the linear map $i \circ \pi$ does indeed have closed graph.

By (3), then, $i \circ \pi$ is continuous. Hence $\pi : X \to (Q, \widehat{\sigma})$ is continuous; hence σ is continuous on X.

27.35. *Proof of* (5) $\Rightarrow$ (1). We shall make the proof slightly longer but easier to understand by splitting it into two parts. We first prove (5) $\Rightarrow$ (1) under the additional assumption that X is a Hausdorff space. Let (V_k) be a closed [convex] string in X; we wish to show that the V_k's are neighborhoods of 0. We shall construct

(i) a [locally convex] topological vector space $\mathfrak{Z}$, and

(ii) a sequence $(\mathfrak{H}_k)$ of neighborhoods of 0 in $\mathfrak{Z}$, and

(iii) a family of continuous linear maps $\varphi_\gamma : X \to \mathfrak{Z}$ (for all γ in some index set Γ) that is toplinearly bounded pointwise and satisfies $V_k = \bigcap_{\gamma\in\Gamma} \varphi_\gamma^{-1}(\mathfrak{H}_k)$ for each k.

By our assumption of (5), it will follow that the family $\{\varphi_\gamma : \gamma \in \Gamma\}$ is equicontinuous. Then, since $\mathfrak{H}_k$ is a neighborhood of 0 in $\mathfrak{Z}$, it follows that $V_k = \bigcap_{\gamma\in\Gamma} \varphi_\gamma^{-1}(\mathfrak{H}_k)$ is a neighborhood of 0 in X. Thus, it suffices to satisfy (i), (ii), and (iii).

We shall satisfy those conditions with the index set Γ equal to the set of all [convex] neighborhood strings in X. For each [convex] string $\gamma = (U_{\gamma 1}, U_{\gamma 2}, U_{\gamma 3}, \ldots)$ belonging to Γ, let $H_{\gamma k} = U_{\gamma k} + V_k$. Then $(H_{\gamma k} : k \in \mathbb{N})$ is also a [convex] neighborhood string in X. Since X is Hausdorff, it follows from our choice of Γ that $V_k = \bigcap_{\gamma\in\Gamma} H_{\gamma k}$.

Form the external direct sums

$$\mathfrak{Z} = \bigoplus_{\gamma\in\Gamma} X \qquad \text{and} \qquad \mathfrak{H}_k = \bigoplus_{\gamma\in\Gamma} H_{\gamma k}$$

for $k \in \mathbb{N}$. Thus, $\mathfrak{Z}$ consists of all functions φ from Γ into X that vanish at all but finitely many points in Γ, and $\mathfrak{H}_k$ consists of those functions $\varphi \in \mathfrak{Z}$ satisfying the further requirement that $\varphi(\gamma) \in H_{\gamma k}$ for all γ.

Then the sequence $(\mathfrak{H}_k)$ is a [convex] string in $\mathfrak{Z}$. Hence it is a neighborhood base at 0 for a [locally convex] vector space topology on the vector space $\mathfrak{Z}$. Hereafter we shall view $\mathfrak{Z}$ as being equipped with that topology.

For each $\gamma \in \Gamma$, define a linear mapping $\varphi_\gamma : X \to \mathfrak{Z}$ as follows: $\varphi_\gamma(x)$ is the function from Γ to X defined by

$$\Big[\varphi_\gamma(x)\Big](\gamma') \quad = \quad \begin{cases} x & \text{when } \gamma' = \gamma \\ 0 & \text{when } \gamma' \neq \gamma. \end{cases}$$

Observe that $\varphi_\gamma^{-1}(\mathfrak{H}_k) = H_{\gamma k}$. Since the $\mathfrak{H}_k$'s form a neighborhood base at 0 in $\mathfrak{Z}$ and each $H_{\gamma k}$ is a neighborhood of 0, it follows that φ_γ is continuous.

Also, since the V_k's are absorbing, it follows that the functions φ_γ are pointwise bounded. Finally, we verify that $V_k = \bigcap_{\gamma\in\Gamma} H_{\gamma k} = \bigcap_{\gamma\in\Gamma} \varphi_\gamma^{-1}(\mathfrak{H}_k)$ for each k. This completes the proof, under the assumption that X is a Hausdorff space.

We now turn to the general case — i.e., where X is not necessarily Hausdorff. Let (S_n) be a closed string in X; we wish to show that S_1 is a neighborhood of 0. Let $K = \mathrm{cl}(\{0\})$. Then the quotient space X/K, equipped with the quotient topology, is Hausdorff. From 26.35 it follows easily (*exercise*) that X/K has property (5). Hence, by our preceding arguments, X/K also has property (1). The quotient mapping $\pi : X \to X/K$ is a continuous and closed mapping (see 26.35). The sequence $(\pi(S_n))$ is a closed string in X/K, hence it is a neighborhood string, hence each set $\pi^{-1}(\pi(S_n))$ is a neighborhood of 0 in X. But each S_n is a closed set, hence $S_n + K = S_n$ by 26.22.e, hence $\pi^{-1}(\pi(S_n)) = S_n$.

27.36. *Proof of* (4) $\Rightarrow$ (3) and (6) $\Rightarrow$ (5). Any vector space Y can be given a locally full ordering by taking the positive cone to be $\{0\}$. When Y is equipped with that ordering, then any linear operator from X to Y is a convex operator.

27.37. *Proof that* (1) *implies both* (4) *and* (6). The argument below, due to Neumann [1985], assumes that the scalar field is $\mathbb{R}$. The case of complex scalars can be dealt with as follows: If X satisfies (1) with scalar field $\mathbb{C}$, then X also satisfies (2) with scalar field $\mathbb{C}$, by the argument in 27.32; hence X satisfies (2) with scalar field $\mathbb{R}$, as noted in 27.30.c; hence X satisfies (1) with real scalars, by 27.33. Hence the argument in the paragraphs below is applicable; therefore (4) and (6) are valid for X with real scalars. But a glance at conditions (4) and (6) shows that those conditions do not involve the specification of the scalar field at all; therefore X also satisfies (4) and (6) with complex scalars.

We now turn to the proof of (1) implies (4) and (6), with real scalars.

The proofs of (1) $\Rightarrow$ (4) and (1) $\Rightarrow$ (6) begin the same. For both proofs we are given a pointwise bounded collection Φ of mappings $f : \Omega \to Y$, which we want to prove equicontinuous, but for (4) that collection consists of just one function. (Of course, a single function is pointwise bounded, since any single point is a toplinearly bounded set). Other differences between the proofs of (4) and (6) will be discussed when they appear, later in the argument.

Fix any $\xi \in \Omega$; it suffices to show that Φ is equicontinuous at ξ. By a translation argument given below, we may assume that

$$0 \in \Omega, \qquad \xi = 0, \qquad \text{and} \quad f(0) = 0 \quad \text{for each} \quad f \in \Phi.$$

(The translation argument is as follows: Let $\widehat{\Omega} = \Omega - \xi$. For each $f \in \Phi$ define a corresponding function $\widehat{f} : \widehat{\Omega} \to Y$ by $\widehat{f}(u) = f(u + \xi) - f(\xi)$. It suffices to show the $\widehat{f}$'s are equicontinuous at 0. By a change of notation, we may replace Ω, f with $\widehat{\Omega}, \widehat{f}$.) Next,

let Ω_0 be a closed, convex, balanced neighborhood of 0 contained in Ω.

(To prove the existence of such a set Ω_0, we do *not* need to assume X is locally convex. For instance, we could take $\Omega_0 = \frac{1}{2}\mathrm{cl}\Big(\Omega \cap (-\Omega)\Big)$; that this set has the required properties follows from 12.6.e and 26.27.a.)

Let K be any given closed neighborhood of 0 in Y. It suffices to show that $\bigcap_{f \in \Phi}\{x \in \Omega_0 : f(x) \in K\}$ is a neighborhood of 0 in X.

Let $H_0, H_1, H_2, H_3, \ldots$ be balanced, [convex,] full neighborhoods of 0 in Y, satisfying $H_0 \subseteq K$ and $H_n + H_n \subseteq H_{n-1}$ for all $n \in \mathbb{N}$; the availability of such sets follows from 27.24.b, 26.52, and 26.54. For the proof of (4), since Y is metrizable, we may also assume that the sequence (H_n) is a neighborhood base at 0.

If $u, v \in \Omega_0$ and $f \in \Phi$ and $0 \leq t \leq 1$, then (using the fact that f is convex, Ω_0 is convex and balanced, and $f(0) = 0$) we obtain

$$-tf(-u) \preccurlyeq -f(-tu) \preccurlyeq f(tu) \preccurlyeq tf(u), \tag{i}$$

$$-f(-u) - f(-v) \preccurlyeq -2f\left(-\frac{u+v}{2}\right) \preccurlyeq 2f\left(\frac{u+v}{2}\right) \preccurlyeq f(u) + f(v). \tag{ii}$$

For each integer $n \geq 0$, define

$$\widehat{U}_n \quad = \quad \bigcap_{f \in \Phi} \Big\{x \in \Omega_0 \ : \ f(x),\, f(-x) \in H_n\Big\}.$$

Note that $\mathrm{cl}(\widehat{U}_n) \subseteq \Omega_0 \subseteq \mathrm{Dom}(f)$ since Ω_0 is closed.

We now shall show that each $\widehat{U}_n$ is balanced. To see this, let any $u \in \widehat{U}_n$, $f \in \Phi$, and $t \in [0, 1]$ be given. Since H_n is balanced and contains $f(u)$ and $f(-u)$ for any $f \in \Phi$, it also contains $tf(u)$ and $-tf(-u)$. Since H_n is full, it also contains $f(tu)$ and $-f(-tu)$, by (i). Thus, $tu, -tu \in \widehat{U}_n$, so $\widehat{U}_n$ is balanced (since the scalar field is $\mathbb{R}$).

We next show that $\widehat{U}_n$ is absorbing. To see this, let any $x \in X$ be given. Since $\widehat{U}_n$ is balanced, it suffices to show $kx \in \widehat{U}_n$ for some nonzero scalar k. Since Ω_0 is a neighborhood of 0, there is some $\varepsilon > 0$ such that $\varepsilon x \in \Omega_0$. Since Φ is pointwise bounded, the set $\bigcup_{f \in \Phi}\{f(\varepsilon x), -f(-\varepsilon x)\}$ is bounded in Y. Since H_n is a neighborhood of 0 in Y, there is some $t \in (0, 1)$ such that $\bigcup_{f \in \Phi}\{tf(\varepsilon x), -tf(-\varepsilon x)\} \subseteq H_n$. As a special case of (i) we have

$$-tf(-\varepsilon x) \preccurlyeq -f(-t\varepsilon x) \preccurlyeq f(t\varepsilon x) \preccurlyeq tf(\varepsilon x).$$

Since H_n is balanced and full, we have $f(\pm t\varepsilon x) \in H_n$ for all $f \in \Phi$, and thus $t\varepsilon x \in \widehat{U}_n$. This proves $\widehat{U}_n$ is absorbing.

For the proofs of (B4) and (B6) in 27.27, we claim also that $\widehat{U}_n$ is convex. To see this, let any $x_0, x_1 \in \widehat{U}_n$ and $\lambda \in (0, 1)$ be given; let $x_\lambda = (1 - \lambda)x_0 + \lambda x_1$. Let any $f \in \Phi$ be given. Since H_n is balanced, both $-f(-x_j)$ and $f(x_j)$ lie in H_n for $j = 0, 1$. Since H_n

is convex, both $-(1-\lambda)f(-x_0)-\lambda f(-x_1)$ and $(1-\lambda)f(x_0)+\lambda f(x_1)$ lie in H_n. By the convexity of the function f and the fact that $f(0)=0$, we have

$$-(1-\lambda)f(-x_0)-\lambda f(-x_1) \quad \preccurlyeq \quad -f(-x_\lambda) \quad \preccurlyeq \quad f(x_\lambda) \quad \preccurlyeq \quad (1-\lambda)f(x_0)+\lambda f(x_1).$$

Since H_n is full, it follows that $-f(-x_\lambda)$ and $f(x_\lambda)$ both lie in H_n. Thus $x_\lambda \in \widehat{U}_n$. This proves $\widehat{U}_n$ is convex, under the hypotheses of (B4) and (B6).

Next, we shall show that $\widehat{U}_n+\widehat{U}_n \subseteq 2\widehat{U}_{n-1}$ for all $n \in \mathbb{N}$. To show this, fix any $u, v \in \widehat{U}_n$ and any $f \in \Phi$. Then $f(\pm u), f(\pm v) \in H_n$, and $H_n+H_n \subseteq H_{n-1}$, hence $f(u)+f(v)$ and $f(-u)+f(-v)$ both lie in H_{n-1}. Now apply (ii). Since H_{n-1} is balanced and full, it contains $2f\left(\pm\frac{u+v}{2}\right)$; hence it contains $f\left(\pm\frac{u+v}{2}\right)$. Thus $\frac{1}{2}(u+v) \in \widehat{U}_{n-1}$. This proves $\widehat{U}_n+\widehat{U}_n \subseteq 2\widehat{U}_{n-1}$.

Now let $U_n = 4^{-n}\widehat{U}_n$; in particular, $U_0=\widehat{U}_0$. The sets U_n $(n=0,1,2,\ldots)$ are balanced and absorbing, and satisfy $U_n+U_n \subseteq \frac{1}{2}U_{n-1} \subseteq U_{n-1}$. Then $\mathrm{cl}(U_n)+\mathrm{cl}(U_n) \subseteq \mathrm{cl}(U_{n-1})$ by 26.22.e. Hence the sequence $(\mathrm{cl}(U_n))$ is a closed [convex] string. By our assumption (1), the sets $\mathrm{cl}(U_n)$ are neighborhoods of 0 in X. Hence also the sets $\mathrm{cl}(\widehat{U}_n)$ are neighborhoods of 0 in X.

For the proof of (6), we may proceed as follows: $f(\widehat{U}_0) \subseteq H_0$, and under the hypotheses of (6) we know that each f in Φ is continuous. Hence $f(\mathrm{cl}(\widehat{U}_0)) \subseteq \mathrm{cl}(H_0) \subseteq \mathrm{cl}(K)=K$ since K is closed by assumption. Thus $\mathrm{cl}(\widehat{U}_0)$ is a neighborhood of 0 contained in $\bigcap_{f\in\Phi} f^{-1}(K)$, completing the proof of (6).

The proof of (4) will take much longer. For (4), with Φ containing just a single function f, we continue our reasoning as follows: We are to show that $f^{-1}(K)$ is a neighborhood of 0; we shall show that in fact $\mathrm{cl}(U_1) \subseteq f^{-1}(K)$. Let any $x_1 \in \mathrm{cl}(U_1)$ be given; it suffices to show $f(x_1) \in K$.

Note that $\mathrm{cl}(U_j) \subseteq U_j + \mathrm{cl}(U_{j+1})$ since $\mathrm{cl}(U_{j+1})$ is a neighborhood of 0. Hence, starting from the given vector $x_1 \in \mathrm{cl}(U_1)$ we may recursively choose vectors $u_j \in U_j$ and $x_{j+1} \in \mathrm{cl}(U_{j+1})$ so that $x_j = u_j + x_{j+1}$. Then

$$u_1+u_2+\cdots+u_j \quad \in \quad U_1+U_2+\cdots+U_j \quad \subseteq \quad U_0 \quad \subseteq \quad \mathrm{Dom}(f).$$

Now let

$$y_j \;=\; f(u_1+u_2+\cdots+u_j).$$

Note that $y_j \in f(U_0)=f(\widehat{U}_0) \subseteq H_0 \subseteq K$. If $\lim_{j\to\infty} y_j$ exists then that limit must lie in K since that set is closed. We shall show that in fact $f(x_1)=\lim_{j\to\infty} y_j$; that will complete the proof.

Let $\widehat{u}_k = 4^k u_k$. Note that $\widehat{u}_k \in \widehat{U}_k$, hence $f(\pm\widehat{u}_k) \in H_k$, hence $4^{-k}f(\pm\widehat{u}_k) \in H_k$. For $l, m \in \mathbb{N}$ with $l \le m$ it follows that $\sum_{k=l}^{m} 4^{-k}f(\pm\widehat{u}_k) \in H_{l-1}$, for all choices of the $\pm$ signs. Such sums tend to 0 as $l \to \infty$, since we chose (H_n) to be a neighborhood base at 0 in Y. Hence the sequence of partial sums $\sum_{k=1}^{m} 4^{-k}f(\pm\widehat{u}_k)$ (for $m=1,2,3,\ldots$) is Cauchy in the F-space Y. Therefore the series $\sum_{k=1}^{\infty} 4^{-k}f(\pm\widehat{u}_k)$ converges to some limit in Y, for each choice of the $\pm$ signs.

Define $t_j = 1/(2+4^{-j})$; then $\frac{1}{3}=t_0<t_1<t_2<\cdots$ and $\lim_{j\to\infty} t_j = \frac{1}{2}$. Temporarily fix any positive integers m, n with $n<m$. Then the numbers

$$\sigma = \frac{t_n}{t_m}, \qquad \tau = 2-\frac{t_m}{t_n}, \qquad \rho_m = \frac{t_{m-1}}{t_m}$$

all lie in $(0,1)$. Verify that these numbers satisfy

$$\begin{array}{lclclclcl}
\sigma & + & (1-\sigma)\sum_{k=1}^{n} 4^{-k} & + & \sum_{k=n+1}^{m} 4^{-k} & = & 1, \\
\tau & + & (1-\tau)\sum_{k=1}^{m} 4^{-k} & + & \sum_{k=n+1}^{m} 4^{-k} & = & 1, \\
\rho_m & + & (1-\rho_m)\sum_{k=1}^{m} 4^{-k} & + & 4^{-m} & = & 1.
\end{array}$$

Hence the convexity of f tells us that

$$\begin{aligned}
y_m \quad &= \quad f\left(\sigma\sum_{k=1}^{n} u_k \;+\; (1-\sigma)\sum_{k=1}^{n} 4^{-k}\widehat{u}_k \;+\; \sum_{k=n+1}^{m} 4^{-k}\widehat{u}_k\right) \\
&\preccurlyeq \quad \sigma y_n \;+\; (1-\sigma)\sum_{k=1}^{n} 4^{-k} f(\widehat{u}_k) \;+\; \sum_{k=n+1}^{m} 4^{-k} f(\widehat{u}_k) \qquad \text{(iii)}
\end{aligned}$$

and similarly

$$\begin{aligned}
y_n \quad &= \quad f\left(\tau\sum_{k=1}^{m} u_k \;+\; (1-\tau)\sum_{k=1}^{m} 4^{-k}\widehat{u}_k \;-\; \sum_{k=n+1}^{m} 4^{-k}\widehat{u}_k\right) \\
&\preccurlyeq \quad \tau y_m \;+\; (1-\tau)\sum_{k=1}^{m} 4^{-k} f(\widehat{u}_k) \;+\; \sum_{k=n+1}^{m} 4^{-k} f(-\widehat{u}_k). \qquad \text{(iv)}
\end{aligned}$$

Also, from $f(0)=0$ and the convexity of f we obtain

$$\begin{aligned}
-y_m \quad &= \quad -f(u_1+\cdots+u_m) \quad \preccurlyeq \quad f(-u_1-\cdots-u_m) \\
&= \quad f\left(4^{-1}(-\widehat{u}_1)+\cdots+4^{-m}(-\widehat{u}_m)+(1-4^{-1}-\cdots-4^{-m})(0)\right) \\
&\preccurlyeq \quad \sum_{k=1}^{m} 4^{-k} f(-\widehat{u}_k). \qquad \text{(v)}
\end{aligned}$$

Multiply line (iii) by t_m to obtain

$$t_m y_m - t_n y_n \quad \preccurlyeq \quad (t_m-t_n)\sum_{k=1}^{n} 4^{-k} f(\widehat{u}_k) \quad + \quad t_m \sum_{k=n+1}^{m} 4^{-k} f(\widehat{u}_k).$$

Also; multiply line (iv) by t_n and line (v) by $2(t_m-t_n)$ and add the results to obtain

$$t_n y_n - t_m y_m \quad \preccurlyeq \quad (t_m-t_n)\sum_{k=1}^{m} 4^{-k}\left[2f(-\widehat{u}_k)+f(\widehat{u}_k)\right] \;+\; t_n \sum_{k=n+1}^{m} 4^{-k} f(-\widehat{u}_k).$$

The right sides of the last two inequalities tend to 0 as $m, n \to \infty$, since $t_m - t_n \to 0$ and the series $\sum_{k=1}^{\infty} 4^{-k} f(\pm\widehat{u}_k)$ are convergent in Y. Since Y is locally full, it follows that $t_m y_m - t_n y_n \to 0$.

Thus $(t_n y_n : n \in \mathbb{N})$ is a Cauchy sequence in the F-space Y and therefore a convergent sequence. Since $t_n \to \frac{1}{2}$ as $n \to \infty$, the sequence (y_n) is also convergent. Let y be its limit.

To show that $y = f(x_1)$, we shall use the fact that the graph of f is closed in $\Omega \times Y$. Let G and H be any neighborhoods of 0 in X and Y, respectively; it suffices to show that $\left(x_1 + G\right) \times \left(y + H\right)$ meets the graph of f.

Note that $x_m \in \mathrm{cl}(U_m) \subseteq U_m - G$, since G is a neighborhood of 0. Therefore we can choose $v_m \in U_m$ so that $x_m - v_m \in -G$. Let $\widehat{v}_m = 4^m v_m$; then $\widehat{v}_m \in \widehat{U}_m$, hence $f(\pm\widehat{v}_m) \in H_m$, hence

$$f(\pm\widehat{v}_m) \to 0 \qquad \text{as} \quad m \to \infty.$$

Also note that $\rho_m \to 1$, and therefore the vectors

$$-y + \rho_m y_m, \qquad -y + \frac{y_m}{\rho_m}, \qquad (1 - \rho_m)\sum_{k=1}^{m} 4^{-k} f(\widehat{u}_k), \qquad (1 - \frac{1}{\rho_m})\sum_{k=1}^{m} 4^{-k} f(\widehat{u}_k)$$

all converge to 0 as $m \to \infty$. Next, observe that

$$\begin{aligned} x_1 - x_{m+1} + v_{m+1} &= u_1 + u_2 + \cdots + u_m + v_{m+1} \\ &\in U_1 + U_2 + \cdots + U_{m+1} \subseteq \Omega_0 \subseteq U_0 \subseteq \mathrm{Dom}(f). \end{aligned}$$

By convexity of f,

$$\begin{aligned} -y + f(x_1 - x_{m+1} + v_{m+1}) &= -y + f\left(\rho_m \sum_{k=1}^{m} u_k + (1 - \rho_m)\sum_{k=1}^{m} 4^{-k}\widehat{u}_k + v_{m+1}\right) \\ &\preccurlyeq -y + \rho_m y_m + (1 - \rho_m)\sum_{k=1}^{m} 4^{-k} f(\widehat{u}_k) + 4^{-m-1} f(\widehat{v}_{m+1}), \end{aligned}$$

which tends to 0 as $m \to \infty$. Also by the convexity of f,

$$\begin{aligned} y_m &= f\left(\rho_m(x_1 - x_{m+1} + v_{m+1}) + \rho_m 4^{-m-1}(-\widehat{v}_{m+1}) + (1 - \rho_m)\sum_{k=1}^{m} 4^{-k}\widehat{u}_k\right) \\ &\preccurlyeq \rho_m f(x_1 - x_{m+1} + v_{m+1}) + \rho_m 4^{-m-1} f(-\widehat{v}_{m+1}) + (1 - \rho_m)\sum_{k=1}^{m} 4^{-k} f(\widehat{u}_k) \end{aligned}$$

and consequently

$$-y + f(x_1 - x_{m+1} + v_{m+1}) \succcurlyeq -y + \frac{y_m}{\rho_m} + (1 - \frac{1}{\rho_m})\sum_{k=1}^{m} 4^{-k} f(\widehat{u}_k) - 4^{-m-1} f(\widehat{v}_{m+1}),$$

which also tends to 0 as $m \to \infty$. Since Y is locally full, $-y + f(x_1 - x_{m+1} + v_{m+1}) \to 0$ as $m \to \infty$. For m sufficiently large we have $f(x_1 - x_{m+1} + v_{m+1}) \in y + H$, so $(x_1 - x_{m+1} + v_{m+1},\ f(x_1 - x_{m+1} + v_{m+1})) \in \left(x_1 + G\right) \times \left(y + H\right)$. This completes the proof of the theorem.

Inductive Topologies and LF Spaces

27.38. *Remarks.* LF spaces are used particularly in Schwartz's distribution theory. Although we shall not develop that theory in this book, we now include a brief introduction to LF spaces, because they provide interesting examples of locally convex spaces that are barrelled but *not* metrizable. Though the definition of LF spaces is slightly complicated, we shall see in 27.46 that the LF space construction provides us with the only "natural" topology for some vector spaces.

We begin with a few results in a slightly more general setting; then we specialize to LF spaces. Examples are given in 27.42. In 27.43 we briefly sketch some of the basic ideas of distribution theory.

27.39. Theorem. Let Y be a vector space (without any topology specified yet), and let $\{(X_j, \tau_j) : j \in J\}$ be a family of locally convex topological vector spaces. For each index j, let

$$y_j \quad : \quad X_j \to Y$$

be some linear mapping. Then there exists a topology τ on Y that is locally convex and has the property that τ is the strongest locally convex topology on Y that makes all the y_j's continuous. It has these further characterizations:

(i) Let $\mathcal{B}$ be the collection of all sets $G \subseteq Y$ such that

> G is absorbing, balanced, and convex, and for each $j \in J$, the set $y_j^{-1}(G)$ is a neighborhood of 0 in X_j.

Then $\mathcal{B}$ is a neighborhood base at 0 for τ.

(ii) Let Z be another locally convex topological vector space, and let $g : Y \to Z$ be some linear map. Then g is continuous from (Y,τ) to Z if and only if each of the compositions $g \circ y_j : X_j \to Z$ is continuous. This property also uniquely determines τ.

We shall call τ the **final locally convex topology** induced by the y_j's (since it is on the *final* end of the mappings $y_j : X_j \to Y$). It is also known as the **inductive locally convex topology**.

Outline of proof.

a. Let Φ be the set of all locally convex topologies on Y for which all the y_j's are continuous. Then Φ is nonempty, since the indiscrete topology $\{\varnothing, Y\}$ is a member of Φ. Let τ be the sup of all the elements of Φ; by 26.20.c we know that τ is an LCS topology on Y. (We do not yet assert that τ is a member of Φ.)

b. Define $\mathcal{B}$ as above. Show that $\mathcal{N} = \{S \subseteq Y : S \text{ contains some element of } \mathcal{B}\}$ is the neighborhood filter at 0 for a locally convex topology σ on Y. Show that $\sigma \in \Phi$. Then $\sigma \subseteq \tau$ since $\tau = \sup \Phi$.

c. Let H be any neighborhood of 0 in (Y, τ). Using the definition of τ, show that $H \supseteq \bigcap_{\psi \in \Psi} H_\psi$, where Ψ is some finite subset of Φ (which may depend on H), and each H_ψ

is a balanced convex neighborhood of 0 in the topological space (Y, ψ). Use that fact to show that H is also a neighborhood of 0 in (Y, σ). Thus $\tau \subseteq \sigma$. This completes the proof of (i) and (ii).

d. If g is continuous from (Y, τ) to Z, then each $g \circ y_j$ is a composition of two continuous maps, and thus it is continuous. Conversely, suppose that each $g \circ y_j : X_j \to Z$ is continuous. Let H be a balanced, convex neighborhood of 0 in Z. Then $(g \circ y_j)^{-1}(H)$ is a balanced convex neighborhood of 0 in X_j, hence $g^{-1}(H)$ is a neighborhood of 0 in Y. To see that this condition uniquely determines τ, suppose that τ, τ' are two locally convex topologies on Y with this property; show that the identity map $i : Y \to Y$ is continuous in both directions between (Y, τ) and (Y, τ').

27.40. *A peculiar specialization (optional).* By taking $J = \varnothing$ in 27.39, we obtain these results:

Let Y be a vector space over the scalar field $\mathbb{F}$. Then there exist topologies on Y that make Y into a locally convex topological vector space, and among such topologies there is a strongest. It is called the **strongest** (or **finest**) **locally convex topology** on Y. It has these further properties:

a. A neighborhood base at 0 for the topology is given by the collection of all absorbing, balanced, convex sets.

b. Any linear map from Y into any other locally convex space is continuous.

27.41. *Definition.* Let X be a vector space. Let $X_1 \subsetneqq X_2 \subsetneqq X_3 \subsetneqq \cdots$ be linear subspaces with $\bigcup_{j=1}^{\infty} X_j = X$. Suppose each X_j is equipped with a topology τ_j making it a Fréchet space. Assume also that the τ_j's are *compatible*, in this sense: If $j < k$, then τ_j is the relative topology determined on X_j by the topological space (X_k, τ_k).

Let τ be the locally convex final topology on X (defined as in 27.39) determined by the inclusion maps $X_j \xrightarrow{\subseteq} X$. Then τ is called the **strict inductive limit of the τ_j's.** A locally convex space (X, τ) that can be determined in this fashion is called an **LF space.** (*Caution*: Some mathematicians use a slightly more general definition for these terms.)

Basic properties. Let (X_j, τ_j)'s and (X, τ) be as above. Then:

a. If we replace the sequence of spaces $((X_j, \tau_j))$ with any subsequence, we still obtain the same topology τ on X.

b. **Subspace lemma.** Fix any j. Suppose G_j is a convex neighborhood of 0 in X_j. Then there exists a convex neighborhood G_{j+1} of 0 in X_{j+1} such that $G_j = X_j \cap G_{j+1}$. Furthermore, if some point $y_0 \in X_{j+1} \setminus X_j$ is given, then G_{j+1} can be chosen so that $y_0 \notin G_{j+1}$. (This is immediate from 26.28.)

c. For some positive integer k, let $G_k, G_{k+1}, G_{k+2}, \ldots$ be a sequence such that G_j is a convex neighborhood of 0 in (X_j, τ_j) and $G_j = X_j \cap G_{j+1}$. Then $G = \bigcup_{j=k}^{\infty} G_j$ is a convex neighborhood of 0 in (X, τ) and $G_j = X_j \cap G$.

d. The original topology τ_j given on X_j is equal to the relative topology determined on X_j by the topological space (X, τ).

e. Each X_j is a closed subset of (X, τ).

f. (X, τ) is Hausdorff.

g. (X, τ) is not a Baire space. (*Hint*: It is the union of the X_j's, which are closed subsets with empty interiors.)

h. (X, τ) is barrelled. *Hint*: Let T be a barrel in (X, τ). Then $T \cap X_j$ is a barrel in X_j, hence a neighborhood of 0 in X_j; hence T is a neighborhood of 0 in X.

i. Let $S \subseteq X$. Then S is a bounded subset of the topological vector space (X, τ) if and only if there exists some j such that $S \subseteq X_j$ and S is a bounded subset of the topological vector space (X_j, τ_j).

Hints: Suppose S is bounded in X but is not contained in any X_j. Replacing the X_j's with a subsequence, show that there is some sequence (s_j) in S with $s_j \in X_{j+1} \setminus X_j$. Using 27.41.b, choose sets G_j so that G_j is a convex neighborhood of 0 in X_j, $G_j = X_j \cap G_{j+1}$, and $\frac{1}{j}s_j \notin G_{j+1}$. Then $G = \bigcup_{j=1}^{\infty} G_j$ is a neighborhood of 0 in X. Since S is bounded in X, we have $\frac{1}{j}s_j \to 0$ in X, hence $\frac{1}{j}s_j \in G$ for all j sufficiently large, a contradiction.

j. Let $(x_n : n \in \mathbb{N})$ be a sequence in X. Then (x_n) is convergent to some limit x_0 in X if and only if there is some j such that $\{x_n : n = 0, 1, 2, 3, \ldots\} \subseteq X_j$ and $x_n \to x_0$ in X_j.

k. X is not metrizable.

Hints: Suppose d is a metric for the topology on X. Choose a sequence (x_n) with $x_n \in X_n \setminus X_{n-1}$ (with x_1 chosen arbitrarily in X_1). Choose numbers $\varepsilon_n > 0$ small enough so that $d(\varepsilon_n x_n, 0) < \frac{1}{n}$. Then $\varepsilon_n x_n \to 0$ in X, hence $\{\varepsilon_n x_n : n \in \mathbb{N}\} \subseteq X_j$ for some j, a contradiction.

l. Let Y be another topological space. Then a map $f : X \to Y$ is sequentially continuous if and only if its restriction to each X_j is sequentially continuous.

m. Let Y be another topological vector space. Then any bounded linear map $f : X \to Y$ (defined as in 27.4) is sequentially continuous.

27.42. *Examples.* Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$).

a. $\bigsqcup_{k \in \mathbb{N}} \mathbb{F}$ is the set of all sequences of scalars that have only finitely many nonzero terms. (See 11.6.i.) It is the union of the finite dimensional subspaces $X_k = \{$sequences whose terms after the kth are zero$\}$. Thus it can be topologized as an LF space.

b. Let Ω be an open subset of $\mathbb{R}^n$ or $\mathbb{C}^n$, for some positive integer n. Let $C_c(\Omega) = \{$continuous scalar-valued functions on Ω with compact support$\}$. Then $C_c(\Omega)$ is the union of the spaces

$$C_K(\Omega) = \{f \in C_c(\Omega) \ : \ f \text{ vanishes outside } K\},$$

for compact sets $K \subseteq \Omega$. Each $C_K(\Omega)$ is a Banach space when equipped with the sup norm. We can write $\Omega = \bigcup_{j=1}^{\infty} G_j$ for some open sets G_j whose closures $K_j = \mathrm{cl}(G_j)$ are compact subsets of Ω (see 17.18.a), hence $C_c(\Omega)$ can be topologized as the strict inductive limit of the spaces $C_{K_j}(\Omega)$. (*Exercise.* The topology is not affected by the particular choice of the sequence (G_j). *Hint*: See 17.18.b.)

See also the remarks in 27.46.

27.43. *A few remarks about distribution theory.* The most important application of final locally convex spaces is in the theory of distributions, which was invented by Dirac and then formalized by L. Schwartz. This theory is particularly useful in the study of linear partial differential equations. Following is a brief sketch of how final locally convex spaces are used in that theory.

We consider a vector space consisting of "nice" functions; a typical example is

$$\mathcal{D}\left(\mathbb{R}^M\right) = \left\{\text{smooth functions from } \mathbb{R}^M \text{ into } \mathbb{C} \text{ with compact support}\right\}.$$

Ultimately, the test functions are not the real object of study, for they are fairly simple and well behaved, and well understood. The test functions are sufficiently well behaved so that they lie in the domain of many ill-behaved differential (or other) operators. Ultimately, it is these operators that are the real object of the study; we can study them by "testing" their behavior with the test functions. We equip the space of test functions with an extremely *strong* topology; then virtually any linear operator that is defined on all of the test functions — including the ill-behaved operator that we wish to study — will in fact be a *continuous* linear operator on that space of test functions. Such a continuous linear operator on the test functions is called a **distribution**. Thus, the distributions are the members of the dual space $\mathcal{D}(\mathbb{R}^M)^*$.

"Ordinary" functions f act as distributions T_f by the following rule:

$$T_f(\varphi) = \int_{\mathbb{R}^M} f(t)\varphi(t)dt \qquad \text{for} \quad \varphi \in \mathcal{D}(\mathbb{R}^M).$$

This formula makes sense for a rather wide class of f's since the φ's are so well behaved. For instance, any function $f : \mathbb{R}^M \to \mathbb{C}$ that is measurable and locally integrable (i.e., integrable on bounded subsets of $\mathbb{R}^M$) defines a distribution T_f in this fashion. The mapping $f \mapsto T_f$ is linear and injective, so it is natural to identify f and T_f; thus the ordinary functions form a subset of the distributions. Because distributions can be used like ordinary functions in some respects, distributions are often called **generalized functions**.

Many familiar operations on ordinary functions can be extended to operations on generalized functions. For instance, there is a natural way to define the derivatives of distributions. For simplicity of notation we consider only the case of $M = 1$, but the ideas below extend easily to any dimension M. If f is a continuously differentiable function, then

$$T_{(f')}(\varphi) = \int_{-\infty}^{+\infty} f'(t)\varphi(t)dt \overset{(!)}{=} -\int_{-\infty}^{+\infty} f(t)\varphi'(t)dt = -T_f(\varphi');$$

the middle equation (!) follows by integration by parts (with the boundary terms disappearing because φ has compact support). Since we identify ordinary functions with their corresponding distributions, $T_{(f')}$ is the "derivative" of T_f. We now generalize: If T is any distribution (not necessarily corresponding to some ordinary function), then the derivative of T is defined to be the distribution U given by $U(\varphi) = -T(\varphi')$. This definition makes sense because when φ is a test function, then φ' is also a test function.

It is customary to topologize the space of test functions $\mathcal{D}(\mathbb{R}^M)$ as follows: For each compact set $K \subseteq \mathbb{R}^M$, let $\mathcal{D}_K$ consist of the smooth functions that have support contained in K. We can topologize $\mathcal{D}_K$ naturally with countably many seminorms, by using the sups of the absolute values of derivatives of functions. It turns out that $\mathcal{D}_K$ is then a Fréchet space. Now $\mathcal{D}(\mathbb{R}^M)$ is the union of the $\mathcal{D}_K$'s, and in fact it is the union of countably many of the $\mathcal{D}_K$'s. Thus it can be topologized as an LF space. With that topology, $\mathcal{D}(\mathbb{R}^M)$ is not metrizable, but it inherits other, more important properties from the $\mathcal{D}_K$'s. For instance, it is barrelled. (It is also complete, but that seems to be less important.)

The various topologies on the space of distributions $\mathcal{D}(\mathbb{R}^M)^*$ are studied using duality theory, a small part of which is introduced in Chapter 28. For further reading on this classical theory, a few sources are Adams [1975], Griffel [1981], Horvath [1966], and Treves [1967].

In the classical theory (described above), distributions form a vector space but not an algebra. Although we can certainly talk about T_{fg} when f and g are ordinary functions, in general it is not possible to multiply together two distributions U and V. In recent years, however, new theories of distributions have been developed that permit multiplication of generalized functions. The theory of Colombeau [1985] is perhaps slightly simpler, but the theory of Rosinger [1990] seems to be more powerful. Both theories are based on algebraic quotients, as in 9.25. In both theories, we begin with some algebra of smooth functions, identify a suitable ideal within that algebra, and then form a quotient algebra, which then acts as a sort of completion of the "ordinary functions."

The Dream Universe of Garnir and Wright

27.44. *Remarks.* In this subchapter we consider how functional analysis is affected when we replace conventional set theory with an alternative set theory; this will help to explain certain intangibles of conventional set theory — i.e., objects that exist but lack constructible examples. This subchapter can be omitted if the reader is only interested in a conventional approach to functional analysis.

H. G. Garnir applied the term "dream space" to any normed space X with the property that every linear map from X into a normed space is continuous. (See Brunner [1987]; see also the related "good spaces" of Garnir [1974].) Any finite dimensional space is a dream space. As we noted in 23.6.b, there are no other dream spaces, under conventional set theory (ZF + AC). Garnir investigated dream spaces under some alternative set theories; later J. D. M. Wright [1975, 1977] also investigated automatic continuity under alternative set theories. Both of these mathematicians were motivated by the earlier consistency results of Solovay [1970] discussed in 14.75 of this book, but in retrospect we can say that a better motivation is given by the later consistency results of Shelah [1984] discussed in 14.74 of this book: If ZF is consistent (something we generally assume), then ZF + DC + BP is also consistent.

The theorem below improves slightly on results of Garnir and Wright by dropping unnecessary hypotheses of local convexity and Hausdorffness and generalizing to convex operators.

27.45. Garnir-Wright Closed Graph Theorem. Assume ZF + DC + BP instead of conventional set theory. Let X be an F-space. Then:

(i) If Ω is an open convex subset of X and $f : \Omega \to \mathbb{R}$ is a convex functional, then f is continuous.

(ii) If Y is any TVS and $f : X \to Y$ is any linear operator, then f is continuous.

(iii) If Ω is an open convex subset of X, Y is a locally full space, and $f : \Omega \to Y$ is a convex operator, then f is continuous.

Remarks. Perhaps it is misleading to call this a "closed graph theorem;" a more descriptive term might be "automatic continuity theorem." The fact that the graph is closed is a *conclusion*, not a hypothesis, of this theorem.

We emphasize that, in (ii), the operator f must be defined on *all* of X — not just a dense subspace — and X is *complete.* Result (ii) implies that every Banach space is a "dream space," as defined in 27.44. Contrast this with 23.6.b.

Result (ii) gives us some explanation of why every linear operator observed in applied mathematics that can be defined everywhere on a complete metric TVS is continuous. Another explanation, not requiring unconventional set theory, is given by Neumann [1980]: The operators arising in applied mathematics generally satisfy some additional condition, such as causality or positivity, which guarantees continuity. We saw an example of this in 26.59.

Our proof of the present theorem combines ideas from Neumann [1985], Wright [1975, 1977].

Proof of theorem. Obviously (i) is a special case of (iii). We can also make (ii) into a special case of (iii), as follows. Let $\Omega = X$. Any TVS Y can be equipped with the trivial ordering: $y_1 \preccurlyeq y_2$ if and only if $y_1 = y_2$. This ordering makes Y into a locally full space, and an operator $f : X \to Y$ is convex if and only if it is affine.

Thus, it suffices to prove (iii). Fix any $x_0 \in \Omega$; by 27.12 it suffices to show f is continuous at x_0. Replace f with the function $f(\cdot + x_0) - f(x_0)$; thus we may assume that $0 = x_0 \in \Omega$ and that $f(0) = 0$. Replacing Ω with the set $\Omega \cap (-\Omega)$, we may assume Ω is open, convex, and balanced. It suffices to prove f is continuous at 0.

Suppose not. Then there exists a neighborhood N_1 of 0 in Y and a sequence (x_n) that converges to 0 in X, such that $f(x_n)$ stays out of N_1. Replacing N_1 with a smaller set, we may assume N_1 is balanced and full. Let N_2 be another balanced, full neighborhood of 0 in Y, satisfying $N_2 + N_2 \subseteq N_1$. The topology on any TVS is determined by its continuous F-seminorms; thus there exists some continuous F-seminorm ρ on Y such that $\{y \in Y : \rho(y) < 1\} \subseteq N_2$. Finally, choose balanced, full neighborhood N_3 of 0 in Y, satisfying $N_3 \subseteq \{y \in Y : \rho(y) < 1\}$.

Extend f to an operator defined on all of X, still denoted f, by taking $f(x) = 0$ for all $x \in X \setminus \Omega$. (We do not assert that this new operator is convex.) Let X_0 be the closed linear span of the sequence (x_n) in X; then X_0 is a separable F-space. Hereafter we only concern ourselves with the restriction of f to X_0.

By 20.25.h and 20.30 and our assumption of BP, every subset of X_0 has the Baire property. Hence the function $\rho \circ f : X_0 \to \mathbb{R}$ is $\mathcal{BP}(X)$-measurable — i.e., measurable when X_0 is equipped with the σ-algebra of almost open sets and $\mathbb{R}$ is equipped with the σ-algebra of Borel sets. By 20.23, there exists some meager set $M \subseteq X_0$ such that the restriction of

$\rho \circ f$ to $X_0 \setminus M$ is continuous (with respect to the relative topology on $X_0 \setminus M$). Then the set

$$L \quad = \quad \bigcup_{j,k \in \mathbb{Z}} \bigcup_{n \in \mathbb{N}} (jM + kx_n),$$

being a union of countably many meager sets, is also meager in X_0. Since X_0 is a Baire space, the comeager set $X_0 \setminus L$ is dense in X_0; hence it meets the nonempty open set $X_0 \cap \Omega$. Fix any $z \in (X_0 \cap \Omega) \setminus L$. Then $\pm\frac{z}{j} \pm x_n \notin M$ for all $j, n \in \mathbb{N}$.

Since Ω is a neighborhood of 0 in X and N_3 is a neighborhood of 0 in Y, for all j sufficiently large we have $\pm\frac{z}{j} \in \Omega$ and $\pm\frac{1}{j}f(\pm z) \in N_3$. Hold any such j fixed. Since $f(0) = 0$ and f is a convex operator, we have

$$\frac{-1}{j}f(-z) \quad \preccurlyeq \quad -f\left(\frac{-z}{j}\right) \quad \preccurlyeq \quad f\left(\frac{z}{j}\right) \quad \preccurlyeq \quad \frac{1}{j}f(z).$$

Since N_3 is a full set, both $f(\frac{z}{j})$ and $-f(\frac{-z}{j})$ must belong to N_3. Hence $\rho(f(\pm\frac{z}{j})) < 1$. When $n \to \infty$, then $\pm x_n \pm \frac{z}{j} \to \pm\frac{z}{j}$ in $X_0 \setminus M$, and $\rho \circ f$ is continuous on that set (with the relative topology). Hence for all n sufficiently large, we have $\rho(f(\pm x_n \pm \frac{z}{j})) < 1$ (with all four combinations of the $\pm$ signs). Also, since Ω is open, for all n sufficiently large we have $\pm x_n \pm \frac{z}{j} \in \Omega$. Fix any such n. Then $\pm f(\pm x_n \pm \frac{z}{j}) \in N_2$ (with all eight combinations of the $\pm$ signs). Since that set is balanced, the vectors $\pm\frac{1}{2}f(\pm x_n \pm \frac{z}{j})$ also belong to N_2. By the convexity of f, we have

$$\begin{aligned} -\frac{1}{2}f\left(-x_n - \frac{z}{j}\right) - \frac{1}{2}f\left(-x_n + \frac{z}{j}\right) \quad &\preccurlyeq \quad -f\left(-x_n\right) \\ &\preccurlyeq \quad f\left(x_n\right) \quad \preccurlyeq \quad \frac{1}{2}f\left(x_n - \frac{z}{j}\right) + \frac{1}{2}f\left(x_n + \frac{z}{j}\right). \end{aligned}$$

The left and right ends of this display belong to $N_2 + N_2 \subseteq N_1$, and that set is full. Hence $f(x_n) \in N_1$, a contradiction. This completes the proof.

27.46. Corollary. Assume ZF + DC + BP in place of conventional set theory. Suppose X is a vector space, and τ is a topology on X that makes X into an LF space. Then there is no other topology besides τ that makes X into a barrelled space.

Remark. Thus, though the definition of an LF space is somewhat complicated, for some spaces (such as those in 27.42) an LF topology is in some sense the "best" one available — i.e., it is the only barrelled topology.

Proof of corollary. Say (X, τ) is the strict inductive limit of (X_j), where $X_1 \subseteq X_2 \subseteq X_3 \subseteq \cdots$ are compatible Fréchet spaces with union X. Let β be a topology on X that makes (X, β) barrelled. Each of the inclusions $X_j \xrightarrow{\subseteq} X$ is continuous from X_j (with its Fréchet topology) to (X, β), by 27.45. By 27.39(ii), then, the identity map $i : (X, \tau) \to (X, \beta)$ is continuous. Therefore its graph is closed in the product topology. Hence its inverse, the mapping $i^{-1} : (X, \beta) \to (X, \tau)$, also has closed graph. By the classical Closed Graph Theorem (which can be proved in ZF + DC), since (X, β) is barrelled, it follows that i^{-1} is continuous, too.

27.47. *Further corollaries.* Assume ZF + DC + BP in place of conventional set theory. Then:

a. If X is a reflexive Banach space — e.g., if $X = \ell_p$ for some $p \in (1, \infty)$ — then the second algebraic dual of X is equal to X. This is in contrast with 11.36.

b. Any two complete norms (or, more generally, any two complete F-norms) on a vector space are equivalent — i.e., they yield the same topology. *Proof.* The identity map $i : X \to X$ is a linear operator. (Contrast with 27.18(iii).)

Remarks. The last result explains, at least in part, the phenomenon described in 22.8. Applied mathematicians do not use the Axiom of Choice, and so they cannot prove the existence of inequivalent complete norms on a vector space. More precisely, they cannot construct two complete norms and prove that those norms are inequivalent. We emphasize that they may also be unable to prove that the two norms are equivalent. Thus, it is conceivable that an applied mathematician could equip a vector space with two different, complete, "usual" norms, which are not known (by the tools of applied mathematics) to be equivalent or inequivalent. However, that seems rather unlikely. Any two complete norms that can be constructed explicitly can probably be investigated rather extensively by constructive means as well; hence we can probably find some answer to the question of whether those norms are equivalent or not. The theorem above eliminates one of the answers — it says that, using just ZF + DC, we cannot show that the two norms are *not* equivalent.

Chapter 28

Duality and Weak Compactness

28.1. *Preview.* A topological vector space can be *retopologized* — i.e., its topology can be replaced with another, related TVS topology. Among such new topologies, one of particular interest is the so-called *weak* topology; it is weaker than the original topology, and therefore has more compact sets that can be used in existence proofs.

Although our applications later in this book are concerned with weak topologies of normed spaces, we shall introduce weak topologies in the more general setting of TVS's and LCS's because (i) that seems to be a more natural setting for the theory, (ii) the theory is not significantly harder in that setting, and (iii) some readers may be interested in other applications not covered in this book.

A highlight of this chapter is R. C. James's Sup Theorem (28.37) on weak compactness, which is quite simple to state but will take much preparation to prove.

This chapter is based partly on Floret [1980], Holmes [1975], Kelley and Namioka [1976], and Schaefer [1971].

Hahn-Banach Theorems in TVS's

28.2. *Definition.* Let X be a topological vector space, with scalar field $\mathbb{F}$. The **dual** of X is the vector space $X^* = \{$continuous linear maps from X into $\mathbb{F}\}$.

There may be several different topologies associated naturally with X^*. Unless some topology is specified for X^*, we shall view it simply as a vector space, not as a topological vector space.

28.3. Lemma. Let X be a topological vector space, and let $\Lambda \in X^*$. If $\Lambda \neq 0$, then Λ is an open mapping — that is, Λ takes any open subset of X to an open subset of $\mathbb{F}$.

Hints: Show that

- **a.** If N is a balanced subset of X, then $\Lambda(N)$ is a balanced subset of $\mathbb{F}$.
- **b.** If N is a neighborhood of 0 in X, then $\Lambda(N)$ is not just the set $\{0\}$.
- **c.** If N is a balanced neighborhood of 0 in X, then $\Lambda(N)$ is a balanced neighborhood of 0 in $\mathbb{F}$.

d. If G is open in X and $x \in G$, then $x + N \subseteq G$ for some N that is a balanced neighborhood of 0.

28.4. Following are several principles, any one of which may be referred to as "the Hahn-Banach Theorem." Considered as weak forms of the Axiom of Choice, these principles are all equivalent to each other and to the versions of HB presented in 12.31, 23.18, 23.19, 26.56, 28.14.a, and 29.32.

Many of our Hahn-Banach Theorems can be extended to complex vector spaces via the Bohnenblust-Sobczyk Correspondence (11.12): If X is a complex vector space on which λ is a linear functional, then X can also be viewed as a real vector space on which $\mathrm{Re}\,\lambda$ is a linear functional. We shall omit the details of that argument; for simplicity we shall generally only consider real vector spaces.

(HB17) Continuous Support Theorem. Let X be a real TVS. Then any continuous convex function from X into $\mathbb{R}$ is the pointwise maximum of the continuous affine functions that lie below it. That is, if $p : X \to \mathbb{R}$ is continuous and convex, then for each $x_0 \in X$ there exists some continuous affine function $f : X \to \mathbb{R}$ that satisfies $f(x) \leq p(x)$ for all $x \in X$ and $f(x_0) = p(x_0)$.

(HB18) Separation of Convex Sets in TVS's. Let A and B be disjoint nonempty convex subsets of a real topological vector space X, and suppose A is open. Then there exists $\Lambda \in X^*$ such that $\Lambda(a) < \inf_{b \in B} \Lambda(b)$ for every $a \in A$.

(HB19) Separation of Convex Sets in LCS's. Let A and B be disjoint nonempty convex subsets of a real, locally convex topological vector space X. Suppose A is compact and B is closed. Then there exists $\Lambda \in X^*$ such that $\max_{a \in A} \Lambda(a) < \inf_{b \in B} \Lambda(b)$.

(HB20) Separation of Points from Convex Sets. Let B be a nonempty closed convex subset of a real, locally convex topological vector space X. Let $x \in X \setminus B$. Then there exists $\Lambda \in X^*$ such that $\Lambda(x) < \inf_{b \in B} \Lambda(b)$.

(HB21) Intersection of Half-Spaces. Let X be a real, locally convex topological vector space. Then any closed convex subset of X is the intersection of the closed half-spaces that contain it. (By a **closed half-space** we mean a set of the form $\{x \in X : \Lambda(x) \geq r\}$, for some continuous linear functional Λ and some real number r.)

(HB22) Separation of Points. If X is a Hausdorff LCS, then X^* separates points of X. That is, if x and y are distinct points of X, then there exists some $\Lambda \in X^*$ such that $\Lambda(x) \neq \Lambda(y)$. Equivalently, if $u \in X \setminus \{0\}$, then there exists some $\Lambda \in X^*$ such that $\Lambda(u) \neq 0$.

(HB23) Separation of Subspaces. Let B be a closed linear subspace of a locally convex space X, and let $\eta \in X \setminus B$. Then there exists a member of X^* that vanishes on B but not on η.

Proof of (HB4) $\Rightarrow$ (HB17). Immediate from 27.13.

Proof of (HB17) $\Rightarrow$ (HB18). Pick any $x_0 \in B - A$. Let $C = A - B + x_0$. Show that C is a convex open neighborhood of 0, and (since A and B are disjoint) $x_0 \notin C$. Let ρ be the Minkowski functional of C. By 26.23.k we know that ρ is a continuous convex function and $C = \{x \in X : \rho(x) < 1\}$. By (HB17) there is some continuous linear functional $\Lambda : X \to \mathbb{R}$ satisfying $\Lambda \leq \rho$ everywhere on X, and satisfying $\Lambda(x_0) = \rho(x_0) \geq 1$. In particular, $\Lambda(x) < 1$ for each $x \in C$; from this it follows that $\Lambda(a) < \Lambda(b)$ for all $a \in A$ and $b \in B$. Finally, $\Lambda(A)$ is open by 28.3, so each $\Lambda(a)$ is strictly less than the supremum of the $\Lambda(a)$'s.

Proof of (HB18) $\Rightarrow$ (HB19). Each $a \in A$ has a neighborhood that is disjoint from B. (For instance, $X \setminus B$ is such a neighborhood.) That neighborhood contains a smaller neighborhood of the form $a + G_a + G_a$, where G_a is a convex open neighborhood of 0; then $(a + G_a + G_a) \cap B = \varnothing$.

The sets $(a + G_a)$ form an open cover of the compact set A. Let $\{a + G_a : a \in F\}$ be a finite subcover, where F is some finite subset of A. Show that $G = \bigcap_{a \in F} G_a$ is a convex open neighborhood of 0, satisfying $(A + G) \cap B = \varnothing$. Then apply (HB18) to the sets $A + G$ and B.

Proof of (HB19) $\Rightarrow$ (HB20) $\Rightarrow$ (HB21) $\Rightarrow$ (HB22) $\Rightarrow$ (HB9). Easy exercise.

Proof of (HB20) $\Rightarrow$ (HB23) $\Rightarrow$ (HB11). Obvious.

28.5. *Pathological example.* Because (HB17) has a topology-free analogue (HB4) in 12.31, we might be tempted to believe that (HB21) also has a topology-free analogue — i.e., that any convex set in a vector space is the intersection of the half-spaces that contain it. But that is not true; for instance, the set

$$\{(x,y) \in \mathbb{R}^2 \ : \ y > 0\} \ \cup \ \{(x,0) \in \mathbb{R}^2 \ : \ x \geq 0\}$$

is a convex set in $\mathbb{R}^2$ that is not equal to an intersection of half-spaces (easy exercise).

Bilinear Pairings

28.6. *Definitions.* A **bilinear pairing** will mean a triple $\big(X, Y, \langle\ ,\ \rangle\big)$ where X and Y are vector spaces over the scalar field $\mathbb{F}$ (without any topologies necessarily specified) and $\langle\ ,\ \rangle$ is a bilinear map from $X \times Y$ into $\mathbb{F}$ (defined as in 11.7). We may abbreviate this arrangement by $\langle X, Y\rangle$.

When $\langle X, Y\rangle$ is a bilinear pairing, then an associated bilinear pairing $[Y, X]$ can be defined by $[y, x] = \langle x, y\rangle$. However, we shall usually use the same symbol $\langle\ ,\ \rangle$ for *both* of these functions. Thus, $\langle\ ,\ \rangle$ represents two functions, one from $X \times Y$ into $\mathbb{F}$ and the other from $Y \times X$ into $\mathbb{F}$, related by $\langle x, y\rangle = \langle y, x\rangle$. This ambiguity in our notation should not cause any difficulty.

Let $\langle X, Y\rangle$ be a bilinear pairing. Then each $y \in Y$ acts as a linear map $\langle \cdot, y\rangle : X \to \mathbb{F}$; thus Y acts as a collection of functions on X. Observe that this collection of functions *separates points* of X (in the sense of 2.6) if

> for each pair of distinct points x_1, x_2 in X, there exists at least one $y \in Y$ such that $\langle x_1, y\rangle \neq \langle x_2, y\rangle$

or, equivalently (since the y's act as linear maps), if

> for each point $x \neq 0$ in X, there exists at least one $y \in Y$ such that $\langle x, y\rangle \neq 0$.

This condition may or may not be satisfied. If it is satisfied, then the elements of x act as *different* members of $\mathrm{Lin}(Y, \mathbb{F}) = \{\text{linear functionals on } Y\}$, and so we may view X as a linear subspace of $\mathrm{Lin}(Y, \mathbb{F})$. Similarly, the points of X may or may not separate the points of Y; if they do, then we may view Y as a linear subspace of $\mathrm{Lin}(X, \mathbb{F})$. We shall say that $\langle X, Y\rangle$ is a **separated pairing** if each of the sets X, Y separates the points of the other set.

Remarks. What we have called a "separated pairing" is called a "dual pairing" in many other texts, which assume the separation property throughout the entire development of duality theory. We have deviated from that conventional terminology to clarify just where the separation property is or is not needed. Admittedly, most pairings arising naturally in applications are separated, but a few are not; see 28.7.b.

28.7. *Examples.*

a. Let $X = Y = \{\text{continuous functions from } [0,1] \text{ into } \mathbb{F}\}$, and let $\langle x, y\rangle = \int_0^1 x(t)y(t)dt$. Then $\langle X, Y\rangle$ is a separated pairing.

b. Let $X = Y = \{\text{piecewise continuous functions from } [0,1] \text{ into } \mathbb{F}\}$ (defined as in 19.28), and let $\langle x, y\rangle = \int_0^1 x(t)y(t)dt$. Then $\langle X, Y\rangle$ is a bilinear pairing, but it is not separated. For instance, if x is the characteristic function of a nonempty finite set, then elements of Y do not distinguish x from 0.

c. If X is any linear space, and Y is any linear subspace of $\mathrm{Lin}(X, \mathbb{F}) = \{\text{linear functionals on } X\}$, then the evaluation map $\langle x, y\rangle = y(x)$ defines a bilinear pairing. With this pairing, X separates points of Y, but Y does not necessarily separate points of X, so $\langle\ ,\ \rangle$ is not necessarily a separated pairing.

d. The preceding case arises, in particular, if X is a topological vector space and $Y = X^*$ is its topological dual. We note several subcases:

(i) If X is not Hausdorff, then X^* does not separate points of X.

(ii) If X is a Hausdorff locally convex space, then X^* does separate points of X by (HB22) in 28.4, and so $\langle X, X^*\rangle$ is a separated pairing.

(iii) If X is a Hausdorff topological vector space that is not locally convex, then X^* may or may not separate points of X. We saw examples of those two cases in 26.17 and 26.16.

28.8. *Definition.* Let $\langle X, Y\rangle$ be a bilinear pairing. Let $\mathcal{S}$ be a collection of subsets of Y. Each element of X may be viewed as a mapping from Y into the scalar field $\mathbb{F}$, and so we may topologize X with the topology of uniform convergence on elements of $\mathcal{S}$, as defined in 18.26. We may refer to that as the **$\mathcal{S}$-topology**.

Proposition. Suppose that $\mathcal{S}$ satisfies these conditions:

(i) Each set $S \in \mathcal{S}$ is **pointwise bounded**, in the following sense: For each $x \in X$, the set of scalars $\langle x, S\rangle = \{\langle x, s\rangle : s \in S\}$ is bounded. (This condition is reformulated in 28.12.b.)

(ii) $\mathcal{S}$ is directed by inclusion — i.e., the union of any two members of $\mathcal{S}$ is contained in some member of $\mathcal{S}$.

(iii) If $S \in \mathcal{S}$ and r is a nonzero scalar, then $rS \in \mathcal{S}$.

Then the $\mathcal{S}$-topology makes X into a locally convex topological vector space. Furthermore, $\{\rho_S : S \in \mathcal{S}\}$ is a gauge that determines that topology, and $\{S^{\rhd} : S \in \mathcal{S}\}$ is a neighborhood base at 0 for that topology, where

$$\rho_S(x) \;=\; \sup_{s\in S} |\langle x, s\rangle|, \qquad\qquad S^{\rhd} \;=\; \bigcap_{s\in S}\{x \in X : |\langle x, s\rangle| \leq 1\}.$$

(The "polar" sets $S^{\rhd}$ will be studied further in 28.25 and thereafter.)

Hints: See 27.9.f, 27.9.h, and 27.9.c.

Remarks. Condition (i) is essential for a TVS topology, as we saw in 27.9.f. Conditions (ii) and (iii) are not so essential, but they are quite convenient; they yield the characterization of neighborhoods in terms of polars. Moreover, conditions (ii) and (iii) are not really restrictive: If (i) is satisfied, then we can replace $\mathcal{S}$ with a larger collection yielding the same $\mathcal{S}$-topology and also satisfying (ii) and (iii). We saw this for (ii) in 27.9.b; for (iii), replace $\mathcal{S}$ with the collection $\mathcal{T} = \{rS : r > 0,\ S \in \mathcal{S}\}$.

28.9. *Preview and definitions.* Following are four important cases of collections $\mathcal{S}$ satisfying the conditions of 28.8.

a. If $\mathcal{S} = \{\text{finite subsets of } Y\}$, then the $\mathcal{S}$-topology is denoted by $\sigma(X, Y)$ or, more briefly, σ or w. (The σ and w stand for "simple" and "weak.") It has many names — it is called the **weak topology**, the **Y-topology**, the **Y-weak topology**, the **topology of simple convergence**, or the **topology of pointwise convergence**.

In an analogous fashion, we define the $\sigma(Y, X)$ topology on Y — that is, the topology on Y given by convergence on points of X. It makes Y into a topological vector space, so it can be used to specify certain kinds of subsets of Y — for instance, the $\sigma(Y, X)$-compact sets. These are used to define some other topologies on X, described below:

b. If $\mathcal{S} = \{\text{pointwise bounded subsets of } Y\}$, where "pointwise bounded" is defined as in 28.8(i), then the resulting $\mathcal{S}$-topology is called the **strong topology** on X; it is denoted by $\beta(X, Y)$. (The β stands for "bounded.") Clearly, this collection $\mathcal{S}$ is the largest collection of subsets of Y that satisfies the conditions of 28.8, so $\beta(X,Y)$ is the the strongest topology that can be constructed as in 28.8.

c. Let $\mathcal{S}_1 = \{\sigma(Y,X)\text{-compact, convex subsets of } Y\}$ and $\mathcal{S}_2 = \{\sigma(Y,X)\text{-compact, convex, balanced subsets of } Y\}$. It can be shown that these two collections satisfy the requirements of 28.8 and that furthermore they yield the same $\mathcal{S}$-topology. (We shall not prove those assertions since they are not needed later in this book.) This topology is called the **Mackey topology** and is denoted by $\tau(X,Y)$.

d. Actually, *every* locally convex topology can be viewed as an $\mathcal{S}$-topology, by taking $\mathcal{S}$ to be the equicontinuous subsets of X^*; see 28.28.

The set X equipped with the weak, strong, or Mackey topology may be denoted X_σ, X_β, or X_τ, respectively. The topological vector space X_σ may also be denoted X_w in some contexts. It is easy to see that

$$\sigma(X,Y) \quad \subseteq \quad \tau(X,Y) \quad \subseteq \quad \beta(X,Y),$$

since a larger collection $\mathcal{S}$ yields a stronger (i.e., larger) $\mathcal{S}$-topology.

This chapter is concerned primarily with the weak topology. The Mackey and strong topologies are important in distribution theory, but they will not be considered in great depth in this book; we introduce them mainly for the sake of some information they yield about the weak topology. (See also 28.17.a.)

28.10. Retopologizations. We now describe one of the most important ways to form $\mathcal{S}$-topologies.

Let X be a vector space. Let γ be a topology that makes X into a topological vector space; let X_γ denote the vector space X equipped with that given topology. (The γ stands for "given," if you like.) Let $(X_\gamma)^*$ be its dual — i.e., the set of all continuous linear maps from X_γ into $\mathbb{F}$.

Then $\langle X, (X_\gamma)^*\rangle$ is a bilinear pairing (not necessarily separated). It can be used to define more topologies on X, most notably

the weak topology $\sigma = \sigma(X, (X_\gamma)^*)$,

the strong topology $\beta = \beta(X, (X_\gamma)^*)$, and

the Mackey topology $\tau = \tau(X, (X_\gamma)^*)$.

In this context, we may call γ the **given** topology or the **original** topology. (Some mathematicians also call it the *initial* topology, but we prefer to reserve that term for the kind of topology introduced in 9.16.)

Caution: Because the two topologies used most often are γ and σ, the beginner who studies only these two topologies may be tempted to call γ the "strong" topology, to contrast it with the "weak" topology σ. However, the term "strong" customarily refers to the topology $\beta(X, (X_\gamma)^*)$. The strong topology $\beta(X, (X_\gamma)^*)$ is at least as strong as the given topology γ, and in some cases it is strictly stronger.

We now summarize the relations between these important topologies: If X_γ is a TVS with dual $(X_\gamma)^*$, then

$$\sigma(X,(X_\gamma)^*) \quad \subseteq \quad \gamma \quad \subseteq \quad \tau(X,(X_\gamma)^*) \quad \subseteq \quad \beta(X,(X_\gamma)^*).$$

These inclusions are justified by conclusions in 28.13.b and 28.30.

Weak Topologies

28.11. *Characterizations of the weak topology.* Let $\langle X, Y\rangle$ be a bilinear pairing. As we stated in 28.9.a, the $\boldsymbol{\sigma(X,Y)}$ **topology** on X is the topology of pointwise convergence on members of Y. Thus, a net (x_α) converges to a limit x in the topological space $(X, \sigma(X,Y))$ if and only if $\langle x_\alpha, y\rangle \to \langle x, y\rangle$ for each $y \in Y$. This topology can also be characterized in other ways:

a. $\sigma(X,Y)$ is the initial topology (in the sense of 9.15, 9.16, and 15.24) generated by elements of Y. In other words, it is the weakest topology that makes all the mappings $\langle \cdot, y\rangle : X \to \mathbb{F}$ continuous.

b. One gauge that determines the topology $\sigma(X,Y)$ is the collection of seminorms $\{\rho_y : y \in Y\}$, where $\rho_y(x) = |\langle x, y\rangle|$.

c. A neighborhood subbasis at 0 for this topology is given by the sets

$$S_y(\varepsilon) = \{x \in X : |\langle x, y\rangle| \le \varepsilon\}, \qquad \text{for} \quad y \in Y, \quad \varepsilon > 0.$$

That is, a set is a neighborhood of 0 in this topology if and only if that set contains the intersection of finitely many sets of the form $S_y(\varepsilon)$, for various y's and ε's.

28.12. *Basic properties of the weak topology.* Let $\langle X, Y\rangle$ be a bilinear pairing, and let $\sigma = \sigma(X,Y)$ be the resulting weak topology. Show that

a. X_σ is a locally convex topological vector space.

b. A set $B \subseteq X$ is weakly bounded (i.e., bounded in the topological vector space X_σ, in the sense of 27.2) if and only if each y in Y is a bounded function on B — that is, if and only if $\sup_{b\in B} |\langle b, y\rangle| < \infty$ for each $y \in Y$. (Thus, the "bounded pointwise" requirement introduced in 28.8(i) is the requirement that each $S \in \mathcal{S}$ be $\sigma(Y,X)$-bounded.)

c. Every member of Y is a continuous linear map from X_σ into $\mathbb{F}$. Thus, $y \mapsto \langle \cdot, y\rangle$ is a linear mapping from Y into $(X_\sigma)^*$.

d. That mapping $y \mapsto \langle \cdot, y\rangle$, from Y into $(X_\sigma)^*$, is surjective. That is, *every* continuous linear map $\lambda : X_\sigma \to \mathbb{F}$ is represented by at least one member of Y.

Hints: $\{x \in X : |\lambda(x)| \le 1\}$ is a σ-neighborhood of 0. Use 28.11.c to show that there exists a finite set $F \subseteq Y$ such that $\bigcap_{y\in F} \mathrm{Ker}(y) \subseteq \mathrm{Ker}(\lambda)$. By the Common Kernel Lemma 11.16, we have $\lambda \in \mathrm{span}(F)$.

e. Suppose X separates points of Y. Then the mapping $y \mapsto \langle \cdot, y \rangle$ from Y into $(X_\sigma)^*$, described in the last two exercises, is also injective. Thus it is a linear bijection. Allowing a change of notation, we therefore have $(X_\sigma)^* = Y$.

f. X_σ is Hausdorff if and only if Y separates points of X. In that case, X may be viewed as a subset of $\mathrm{Lin}(Y, \mathbb{F})$, as explained in 28.6. Then $\mathbb{F}^Y \subseteq \mathrm{Lin}(Y, \mathbb{F}) \subseteq \mathbb{F}^Y$. Show that the topology $\sigma(X, Y)$ is the relative topology on X determined by the product topology on $\mathbb{F}^Y$.

28.13. *Basic properties of weak retopologizations.* Let X be a vector space; let X_γ be a TVS formed by equipping that vector space with some given topology γ; let $(X_\gamma)^*$ be the resulting dual space; let $\sigma = \sigma(X, (X_\gamma)^*)$ be the resulting weak topology; let X_σ be the resulting TVS — that is, the weak retopologization of X_γ, as discussed in 28.10. Then:

a. The weak topology $\sigma(X, (X_\gamma)^*)$ is a locally convex topology — whether the given topology γ is locally convex or not.

b. The weak topology σ has fewer open sets and fewer closed sets than the original topology γ. (Here we use "fewer" in the extended sense of mathematics — i.e., meaning "fewer or as many.")

In brief, *every weakly open set is open*, and *every weakly closed set is closed.*

c. For any set $S \subseteq X$, we have

$$\sigma\text{-int}(S) \subseteq \gamma\text{-int}(S) \subseteq S \subseteq \gamma\text{-cl}(S) \subseteq \sigma\text{-cl}(S).$$

d. The weak topology has *more* (i.e., at least as many) compact sets, bounded sets, and convergent nets than the original topology. Thus, *every compact set is weakly compact*, and *every bounded set is weakly bounded*, and

$$x_\alpha \xrightarrow{\gamma} x \qquad \Rightarrow \qquad x_\alpha \xrightarrow{\sigma} x.$$

Compact sets are often used in existence proofs; that is one of our main reasons for studying weak topologies.

e. By 28.12.d, the original topology γ and the weak topology σ have the *same* set X^* of continuous linear functionals. That is, $(X_\gamma)^* = (X_\sigma)^*$. Therefore, *in discussions of the original and weak topologies, we may refer to the dual simply as* X^*.

Hence $\sigma(X, (X_\sigma)^*) = \sigma(X, (X_\gamma)^*)$. That is, *the weak topology of the weak topology is the weak topology.* Thus, repeating this weak retopologization procedure cannot get us another, different, still weaker topology.

We may say that a topological vector space X_γ already *has the weak topology* if $X_\gamma = X_\sigma$ — that is, if $\gamma = \sigma(X, (X_\gamma)^*)$.

28.14. *Weak retopologization of locally convex spaces.* Additional conclusions can be drawn if the original topology is locally convex. Let X_γ be a locally convex topological vector space and let X_σ be its weak retopologization. Then:

a. We have this principle, which is another equivalent of the Hahn-Banach Theorem:

(HB24) Weak closures. In a locally convex space, every γ-closed, convex set is σ-closed. In brief, *every closed convex set is weakly closed.*

For a proof, refer to 28.4; it is easy to see that (HB20) $\Rightarrow$ (HB24) $\Rightarrow$ (HB21).

Remark. (HB24) will be used to prove the next few results below. The Hahn-Banach Theorem and its consequences are needed frequently in duality theory, and will be used heavily throughout the remainder of this chapter. *Hereafter we shall use the Hahn-Banach Theorem freely*; we shall discontinue our practice of keeping track of its uses and its equivalents.

b. Let S be a convex subset of X. Then $\sigma\text{-cl}(S) = \gamma\text{-cl}(S)$.

c. A convex subset of X is closed if and only if it is weakly closed; a linear subspace of X is closed if and only if it is weakly closed; a linear subspace of X is dense in X if and only if it is weakly dense; X is separable if and only if it is weakly separable.

d. Every weakly bounded set is bounded. Thus, X_γ and X_σ have the same bounded sets. (*Hint*: 27.6.)

e. Any weakly convergent sequence is bounded.

28.15. Let X be a TVS with scalar field $\mathbb{F}$, and let X^* be its dual — i.e., the vector space of continuous linear maps from X into $\mathbb{F}$. One of the most important topologies on X^* is the $\sigma(X^*, X)$ topology — i.e., the topology of pointwise convergence on members of X. It is called the **weak-star topology** (or **weak-* topology**); it is often abbreviated as $X^*_{w^*}$.

We caution the reader against referring to "the weak-star topology on Y" since we may have $Y = X^*$ for more than one choice of X. Different choices of X may yield the same *set* $Y = X^*$ but may nevertheless yield different weak-star topologies on that set.

Here are some basic properties of the weak-star topology:

a. $X^*_{w^*}$ is a locally convex space.

b. $X^*_{w^*}$ is Hausdorff, whether X is Hausdorff or not. (Indeed, X separates the points of X^*, by definition of the set X^*.) The topology on $X^*_{w^*}$ is the relative topology determined by viewing X^* as a subset of $\mathbb{F}^X$, when that product is equipped with the product topology.

c. Each $x \in X$ determines a continuous linear map $\Lambda_x : X^*_{w^*} \to \mathbb{F}$, by the rule $\Lambda_x(f) = f(x)$. (This is the **evaluation map** at x.) Furthermore, *every* continuous linear functional on $X^*_{w^*}$ can be written in this form. (*Hint*: 28.12.d.) Thus $x \mapsto \Lambda_x$ is a surjective linear map from X onto $(X^*_{w^*})^*$.

If X^* separates the points of X, then the mapping $x \mapsto \Lambda_x$ is also injective, so (allowing a change of notation) we have $(X^*_{w^*})^* = X$. That is, X is equal to the dual of its own weak dual.

Remarks. What about other retopologizations besides the weak one? A locally convex space X_γ is called *semireflexive* if $(X^*_{\beta(X^*,X)})^*$ and X are equal as sets; here

β indicates the strong topology. The space is *reflexive* if in addition $\gamma = \beta(X, X^*)$ — that is, if the strong dual of the strong dual of (X, γ) is equal to (X, γ). We shall not study reflexivity in such a general setting; in 28.41 we shall study reflexivity in the more specialized setting of normed spaces.

Weak Topologies of Normed Spaces

28.16. Let $(X, \|\ \ \|)$ be a normed space over the scalar field $\mathbb{F}$. The *usual* topology on X is the norm topology — i.e., the metric topology given by the metric $d(x, y) = \|x - y\|$. We use it to define continuous linear functionals $\lambda : X \to \mathbb{F}$; they make up the dual X^*. We use that dual to define the weak topology $\sigma(X, X^*)$ as in 28.10.

The norm and weak topologies are the two topologies used most often on a normed vector space. The weak topology is weaker than the norm topology. Some of the exercises in 28.18 below show that if X is infinite-dimensional, then the weak topology is *strictly* weaker than the norm topology.

28.17. *Exercises.* Let $(X, \|\ \ \|)$ be a normed space. Show that

- **a.** The norm topology on X is the same as the strong topology $\beta(X, X^*)$ and the Mackey topology $\tau(X, X^*)$, defined in 28.10.
- **b.** The norm topology on X is a locally convex topology. Hence X and X_σ have the same dual, the same bounded sets, and the same closed convex sets.
- **c.** The weak topology $\sigma(X, X^*)$ is a Hausdorff, locally convex topology on X.
- **d.** If (x_α) is a net converging weakly to some limit x, then $\|x\| \leq \liminf_{\alpha \to \infty} \|x_\alpha\|$. (This result should not be confused with 28.14.e.)
- **e.** If X is finite dimensional, then the weak topology and the norm topology on X are identical. *Hint*: 27.15.

28.18. *Proposition.* Suppose $(X, \|\ \ \|)$ is an infinite dimensional normed space. Then there exists a directed set $\mathcal{F}$ and a net $(x_F : F \in \mathcal{F})$ in X such that $x_F \to 0$ in the weak topology but $\|x_F\| \to \infty$.

Hints: We know X^* is infinite-dimensional, by Kottman's Theorem (23.22). Let H be a vector basis for X^*; then H is an infinite set. Let $\mathcal{F} = \{$finite subsets of $H\}$, directed by inclusion. For each $F \in \mathcal{F}$, choose some $v \in H \setminus F$; then use the Common Kernel Lemma (11.16)(Q_k) to find some vector in X that vanishes on F but not on v. Let x_F be a suitable scalar multiple of that vector, chosen so that $\|x_F\| \geq \mathrm{card}(F)$.

Corollaries. Suppose $(X, \|\ \ \|)$ is an infinite dimensional normed space. Then:

- **a.** The weak closure of the unit sphere $S = \{x \in X : \|x\| = 1\}$ is the unit ball $B = \{x \in X : \|x\| \leq 1\}$.

Hints: Suppose $\|z\| < 1$. Choose a net (x_F) as in the preceding proposition. Let $z_F = z + r_F x_F$ for some real number r_F; show that a suitable choice of r_F yields $z_F \in S$ and $|r_F| \le \dfrac{(1+\|z\|)}{\|x_F\|} \to 0$. Hence (z_F) is a net in S converging weakly to z.

b. If $S \subseteq X$ is bounded, then $\sigma\text{-int}(S) = \varnothing$. That is, in the weak topology, any bounded set has empty interior.

Hints: Suppose p is in the weak interior of S. Replacing S with $S - p$, we may assume $p = 0$. There is a net that converges weakly to 0 but stays out of the bounded set S.

c. The topology of X_σ is not metrizable. (Of course, certain small subsets of X_σ, equipped with their relative topologies, may be metrizable.)

Hint: If d is a metric for that topology, use the preceding proposition to find a sequence (x_n) satisfying $d(0, x_n) < 1/n$ and $\|x_n\| > n$. This contradicts 28.14.e.

28.19. *Proposition.* Let X be a locally uniformly convex Banach space — see 22.38. Let (x_α) be a net in X, and also let $x_\infty \in X$. Then the following are equivalent:

(A) $\|x_\alpha - x_\infty\| \to 0$.

(B) $x_\alpha \to x_\infty$ weakly and $\limsup_\alpha \|x_\alpha\| \le \|x_\infty\|$.

Proof. The proof of (A) $\Rightarrow$ (B) is trivial. Assume (B); we shall prove (A). Note that $x_\alpha \to x_\infty$ weakly implies $\|x_\infty\| \le \liminf_\alpha \|x_\alpha\|$, and thus $\|x_\alpha\| \to \|x_\infty\|$. We may assume $x_\infty \ne 0$ (why?), and that all the x_α's are also nonzero (why?). Replacing x_α with $x_\alpha/\|x_\alpha\|$, we may assume $\|x_\alpha\| = 1$ for all α and $\|x_\infty\| = 1$ (explain). By the Hahn-Banach Theorem (HB8) in 23.18, there exists some $\lambda \in X^*$ such that $\|\lambda\| = \lambda(x_\infty) = 1$. Then $2 \ge \|x_\alpha + x_\infty\| \ge |\lambda(x_\alpha + x_\infty)| \to 2$. Thus $\|x_\alpha + x_\infty\| \to 2$. By local uniform convexity, $x_\alpha \to x_\infty$.

28.20. Recall from 23.10 that the dual of the Banach space ℓ_1 is ℓ_∞. Hence a net (x_α) converges weakly in ℓ_1 to a limit x_∞ if and only if $\sum_{j=1}^\infty x_{\alpha,j} z_j \to \sum_{j=1}^\infty x_{\infty,j} z_j$ for each $z = (z_1, z_2, z_3, \ldots)$ in ℓ_∞.

The space ℓ_1 has an unusual property, not shared by most Banach spaces.

Schur's Theorem. Let (x_n) be a *sequence* converging to a limit x_∞ in the weak topology of ℓ_1. Then also $x_n \to x_\infty$ in the norm topology.

Remarks. We emphasize that Schur's Theorem applies to only to sequences, not to nets. That is clear from 28.18, for instance.

The proof below is direct. Some mathematicians may prefer a proof using Baire category, such as that given by Conway [1969].

Outline of proof of theorem. Assume that $x_n \to x_\infty$ weakly but not in norm; we shall obtain a contradiction. Say $x_n = (x_{n,1}, x_{n,2}, x_{n,3}, \ldots)$. Then:

a. We may assume $x_\infty = 0$. (Replace each x_n with $x_n - x_\infty$.)

b. We may assume $\|x_n\| = 1$ for all n. *Hints*: The sequence $\big(\|x_n\| : n \in \mathbb{N}\big)$ is bounded and does not converge to 0. Replacing (x_n) with a subsequence, we may assume that the numbers $\|x_n\|$ are all positive and converge to some positive number c. Since the weak topology makes ℓ_1 a topological vector space, we may replace each x_n with $x_n/\|x_n\|$ (explain).

c. For each finite set $S \subseteq \mathbb{N}$, we have $\sum_{j\in S} |x_{n,j}| \to 0$ as $n \to \infty$.

d. Replacing (x_n) with a subsequence, show that there exist disjoint finite sets $S(1)$, $S(2)$, $S(3)$, ... contained in $\mathbb{N}$ such that $\sum_{j\in S(n)} |x_{n,j}| > 2/3$.

e. There exists some $z \in \ell_\infty$ such that $\|z\|_\infty = 1$ and $|z_j x_{n,j}| = |x_{n,j}|$ whenever $j \in S(n)$.

f. $|\sum_{j=1}^{\infty} x_{n,j} z_j| > 1/3$ for all n, contradicting the fact that $x_n \to x_\infty$ weakly.

28.21. Let $[a,b]$ be a compact interval in $\mathbb{R}$, and let $C[a,b] = \{$continuous scalar-valued functions on $[a,b]\}$; this is a Banach space with the sup norm. We emphasize that the following result is for *sequences*, not for nets.

Proposition. In $C[a,b]$, a sequence (f_n) converges weakly to a limit f if and only if the sequence (f_n) is uniformly bounded and $f_n \to f$ pointwise on $[a,b]$.

Proof. If $f_n \to f$ weakly, then (f_n) is bounded by 28.14.e and each pointwise evaluation mapping $g \mapsto g(t)$ is a continuous linear functional on $C[a,b]$, so $f_n \to f$ pointwise. Conversely, suppose $f_n \to f$ pointwise and boundedly. By 29.34, each continuous linear functional on $C[a,b]$ is represented by a scalar-valued measure μ on the Borel sets; it suffices to show that $\int f_n\,d\mu \to \int f\,d\mu$. If the scalar field is $\mathbb{C}$, we may consider the real and complex parts of μ; thus it suffices to consider real-valued μ. By the Jordan Decomposition, it suffices to consider finite positive measures μ. Then $\int f_n\,d\mu \to \int f\,d\mu$ by the Dominated Convergence Theorem (22.29).

28.22. When no topology is specified for X^*, then X^* is generally understood to be equipped with its norm topology, using the operator norm as in 23.7. That topology on X^* is usually used to define the second dual — i.e., the vector space X^{**}.

In addition to the norm topology, two other topologies on X^* that are occasionally useful are the *weak topology* $\sigma(X^*, X^{**})$ and the *weak-star topology* $\sigma(X^*, X)$.

Exercises.

a. The weak topology $\sigma(X^*, X^{**})$ and the weak-star topology $\sigma(X^*, X)$ are Hausdorff, locally convex topologies on X^*. The weak-star topology is weaker than (or equal to) the weak topology. (In 28.41(B), we shall consider the conditions under which these two topologies are equal.)

b. The norm-closed unit ball, $\{v \in X^* : \|v\| \le 1\}$, is closed in both the weak and weak-star topologies.

28.23. *Example.* Let $X = c_0 = \{$sequences of scalars converging to $0\}$; we have seen in 23.10 that $X^* = \ell_1$ and $X^{**} = \ell_\infty$.

Recall from 21.11.b that a *probability measure on* $\mathbb{N}$ is a sequence (p_n) with $p_n \ge 0$ for all n and $\sum_{n=1}^{\infty} p_n = 1$. Let P be the set of all such probability measures. Show that P is a closed convex subset of ℓ_1, when that space is given its norm topology. Hence P is also weakly closed — i.e., closed in the $\sigma(\ell_1, \ell_\infty)$ topology — by 28.14.a.

However, P is not weak-star closed — that is, P is not closed in the $\sigma(\ell_1, c_0)$ topology. Indeed, we have $0 \notin P$, but the sequence

$$(1,0,0,0,0,\ldots), \quad (0,1,0,0,0,\ldots), \quad (0,0,1,0,0,\ldots), \quad (0,0,0,1,0,\ldots), \quad \ldots$$

is easily shown to be weak-star convergent to 0.

28.24. As we have shown above, a normed space, equipped with a weak or weak-star topology, generally is not metrizable. Nevertheless, certain subsets of that nonmetrizable TVS may be metrizable, when equipped with the relative topology. Here are two particularly important special cases. Let V be a normed space.

a. If V is separable and Φ is a norm-bounded subset of V^*, then the relative topology on Φ determined by the weak-star topology is metrizable.

b. If V^* is separable and Φ is a norm-bounded subset of V, then the relative topology on Φ determined by the weak topology is metrizable.

Hints: Show that convergence in Φ, pointwise on the separable space mentioned, is equivalent to convergence pointwise on some dense subset of that separable space, since Φ is bounded. Convergence on a countable set can be determined by a countable collection of seminorms.

Polar Arithmetic and Equicontinuous Sets

28.25. *Definition.* Let $\langle X, Y\rangle$ be a bilinear pairing. For each set $R \subseteq X$, we define the **polar** of R to be the set

$$R^{\triangleleft} \quad = \quad \{y \in Y : |\langle x, y\rangle| \leq 1 \text{ for all } x \in R\}.$$

Similarly, we may define the polar of any set $S \subseteq Y$ to be the set

$$S^{\triangleright} \quad = \quad \{x \in X : |\langle x, y\rangle| \leq 1 \text{ for all } y \in S\}.$$

These operations are a special case of 4.10(D).

Caution: Notations differ. For instance, some mathematicians call the objects above the *absolute polars* of R and S, and use $\mathrm{Re}\langle x, y\rangle$ instead of $|\langle x, y\rangle|$ to define "polar." Moreover, among many mathematicians, $R^{\triangleleft}$ and $S^{\triangleright}$ are denoted by R^o and S^o; we have introduced separate notations to reduce confusion among beginners.

28.26. *Elementary properties.* We state results mainly for $\triangleleft$; analogous results obviously hold for $\triangleright$.

a. $\varnothing^{\triangleleft} = Y$, $X^{\triangleleft} = \{0\}$, $R^{\triangleleft\triangleright} \supseteq R$, and $R \subseteq S \Rightarrow R^{\triangleleft} \supseteq S^{\triangleright}$.

b. $(rR)^{\triangleleft} = r^{-1}(R^{\triangleleft})$, for any real number $r > 0$.

c. If $(X, \| \quad \|)$ is a normed space and $(Y, \| \quad \|)$ is its dual, then the polar of the closed ball $\{x \in X : \|x\| \leq r\}$ is the closed ball $\{y \in Y : \|y\| \leq r^{-1}\}$, for any real number $r > 0$. In particular, the polar of the closed unit ball is the closed unit ball.

d. $R^{\lhd}$ is a $\sigma(Y, X)$-closed, convex, balanced subset of Y, for any set $R \subseteq X$.

e. $R^{\lhd}$ is absorbing in Y — hence a barrel in $\sigma(Y, X)$ — if and only if R is $\sigma(X, Y)$-bounded in X.

f. Let $R \subseteq X$, and let R_1 be its $\sigma(X, Y)$-closed, convex, balanced hull. Then $R^{\lhd} = R_1^{\lhd}$.

g. Let X be a topological vector space with topology γ; use the bilinear pairing $\langle X, X^* \rangle$. Then a set $V \subseteq X^*$ is equicontinuous (from X_γ to $\mathbb{F}$) if and only if $V^{\rhd}$ is a γ-neighborhood of 0 in X.

h. Let X be a topological vector space with topology γ. Let $V \subseteq X^*$ be equicontinuous (from X_γ to $\mathbb{F}$). Then the $\sigma(X^*, X)$-closed convex balanced hull of V is also equicontinuous. (*Proof.* Immediate from the preceding two exercises.)

28.27. The Bipolar Theorem. $S^{\lhd\rhd}$ is the $\sigma(X, Y)$-closed, convex, balanced hull of S — that is, the smallest $\sigma(X, Y)$-closed, convex, balanced set containing S. In particular, $S^{\lhd\rhd} = S$ if and only if S is $\sigma(X, Y)$-closed, convex, and balanced.

Proof. Let C be the $\sigma(X, Y)$-closed, convex, balanced hull of S. Then $S \subseteq C \subseteq S^{\lhd\rhd}$. Suppose $x_0 \in S^{\lhd\rhd} \setminus C$; we shall obtain a contradiction. Consider X as a locally convex space equipped with the $\sigma(X, Y)$ topology. By the Hahn-Banach Separation Theorem (HB20) in 28.4 there is some $y_0 \in Y$ that satisfies $\sup_{x \in C} \mathrm{Re}\langle x, y_0 \rangle < \mathrm{Re}\langle x_0, y_0 \rangle$. (We may omit the "Re" if the scalar field is $\mathbb{R}$.) Since C is balanced, the left side of this inequality equals $\sup_{x \in C} |\langle x, y_0 \rangle|$. Replacing y_0 with cy_0 for some suitable scalar c, we may assume that

$$\sup_{x \in C} |\langle x, y_0 \rangle| \quad < \quad 1 \quad < \quad |\langle x_0, y_0 \rangle| .$$

Then $y_0 \in C^{\lhd} \subseteq S^{\lhd}$, hence $x_0 \in S^{\lhd\rhd} \subseteq \{y_0\}^{\rhd}$, a contradiction (explain).

28.28. *Proposition: the original topology is an $\mathcal{S}$-topology.* Let X be a locally convex space. Then the topology of X is equal to the topology of uniform convergence on equicontinuous subsets of X^*; it is also equal to the topology of uniform convergence on $\sigma(X^*, X)$-closed, convex, balanced equicontinuous subsets of X^*. (Here a subset of X^* is considered equicontinuous if it is equicontinuous as a collection of maps from X with the given topology to the scalar field.)

Proof. Let $\mathbb{F}$ be the scalar field. Let $\mathcal{S} = \{$equicontinuous subsets of $X^*\}$. Refer to the characterization of neighborhoods of 0 given in 28.8. If $S \in \mathcal{S}$, then $S^{\rhd}$ is a γ-neighborhood of 0, by 28.26.g. Conversely, if N is a γ-neighborhood of 0, then N contains a set B that is a γ-closed, convex, balanced neighborhood of 0, by 26.27.d. Then B is also weakly closed, by 28.14.a. By the Bipolar Theorem (28.27), then, $B = B^{\lhd\rhd}$. By 28.26.g, the set $S = B^{\lhd}$ is equicontinuous, hence belongs to $\mathcal{S}$, and $N \supseteq S^{\rhd}$.

Finally, note that the $\sigma(X^*, X)$-closed, convex, balanced hull of any member of $\mathcal{S}$ is also a member of $\mathcal{S}$, by 28.26.h.

28.29. Banach-Alaoglu-Bourbaki Theorems. The Ultrafilter Principle, introduced in 6.32 and studied further especially in 17.22, is equivalent to the following principles:

> **(UF26)** Let X be a topological vector space with scalar field $\mathbb{F}$. Let V be an equicontinuous set of maps from X into $\mathbb{F}$. Then V is relatively compact in $\sigma(X^*, X)$ — that is, V is contained in a $\sigma(X^*, X)$-compact set. (Therefore every equicontinuous, $\sigma(X^*, X)$-closed set is $\sigma(X^*, X)$-compact.)

> **(UF27)** Let X be a topological vector space with topology γ and with dual X^*. If S is a neighborhood of 0 in X, then $S^{\lhd}$ is $\sigma(X^*, X)$-compact, where polars are with respect to the bilinear pairing $\langle X, X^* \rangle$.

> **(UF28)** Let $(X, \| \quad \|)$ be a normed space with scalar field $\mathbb{F}$. Let X^* be its dual, and let V be the closed unit ball of that dual — that is, $\{v \in X^* : \|v\| \leq 1\}$. Then V is $\sigma(X^*, X)$-compact.

For brevity in the proofs, let the $\sigma(X^*, X)$ topology be denoted by w^*. The reader may find it helpful to review 17.15. The equivalence of these principles with other forms of UF was first announced, without proof, by Rubin and Scott [1954]; our proof is based on Luxemburg [1969].

Proof of (UF19) $\Rightarrow$ (UF26). The w^* topology is the relative topology induced on X^* by considering that set as a subset of $\mathbb{F}^X$ with the product topology. The closure of an equicontinuous set is equicontinuous, as we noted in 18.33.a. For each $x \in X$, the set $V(x) = \{(x, v) : v \in V\}$ is a bounded subset of $\mathbb{F}$, since V is equicontinuous. Hence $\mathrm{cl}[V(x)]$ is a compact subset of $\mathbb{F}$. Then V is contained in the set $\prod_{x \in X} \mathrm{cl}[V(x)]$, which is compact by (UF19). Therefore $\mathrm{cl}(V)$ is itself a compact subset of $\mathbb{F}^X$. It remains to show that $\mathrm{cl}(V)$ is actually a subset of X^*. Any pointwise limit of linear functions is linear, so each member of $\mathrm{cl}(V)$ is linear. Also, since V is equicontinuous, any pointwise limit of members of V is continuous. This completes the proof.

Proof of (UF26) $\Rightarrow$ (UF27). By 28.26.g we know that $S^{\lhd}$ is an equicontinuous set of maps from X into $\mathbb{F}$. By 28.26.d we know $S^{\lhd}$ is $\sigma(X^*, X)$-closed.

Proof of (UF27) $\Rightarrow$ (UF28). $V = S^{\lhd}$ where S is the closed unit ball of X.

Proof of (UF28) $\Rightarrow$ (UF1). Let Ω be a nonempty set, and let $\mathcal{F}$ be a proper filter of subsets of Ω; we wish to show that $\mathcal{F}$ is contained in an ultrafilter.

Let $X = B(\Omega) = \{$bounded functions from Ω into $\mathbb{R}\}$; this is a real Banach space when equipped with the sup norm. Let V be the closed unit ball of the dual of X. There is a natural injective mapping $\varphi : \Omega \to V$, as follows: $\varphi_\omega(x) = x(\omega)$ for $\omega \in \Omega$ and $x \in X$. Thus we may view Ω as a subset of V; then members of $\mathcal{F}$ are subsets of V.

Let $\mathcal{K} = \{w^*\text{-cl}(F) \ : \ F \in \mathcal{F}\}$. Members of $\mathcal{K}$ are subsets of V, since V is w^*-closed. Hence members of $\mathcal{K}$ are w^*-compact. The collection $\mathcal{K}$ has the finite intersection property since $\mathcal{F}$ does. Therefore $\mathcal{K}$ has nonempty intersection. Choose some v_0 in the intersection of $\mathcal{K}$. Then $v_0 \in V$, so v_0 is a linear map from X into $\mathbb{R}$ with $\|v_0\| \leq 1$. Define $\mu : \mathcal{P}(\Omega) \to \mathbb{R}$ by taking $\mu(S) = v_0(1_S)$.

Fix any $F \in \mathcal{F}$. Since $v \in w^*\text{-cl}(F)$, there is some net $(\omega(\alpha) : \alpha \in A)$ in F such that $\varphi_{\omega(\alpha)} \xrightarrow{w^*} v_0$. Thus, for each $x \in X$, we have $x(\omega(\alpha)) \to v_0(x)$.

In particular, for each $S \in \mathcal{P}(\Omega)$ we have $\mu(S) = v_0(1_S) = \lim_{\alpha \in A} 1_S(\omega(\alpha))$. Then μ is a charge on the measurable space $(\Omega, \mathcal{P}(\Omega))$, and its range is contained in $\{0, 1\}$ since a net of 0s and 1s can only have limit 0 or 1. Moreover, $\mu(F) = \lim_{\alpha \in A} 1_F(\omega(\alpha)) = 1$.

The preceding conclusions about μ are valid for each $F \in \mathcal{F}$; in particular $\mu(\Omega) = 1$. Thus μ is a two-valued probability charge that takes the value 1 on $\mathcal{F}$, so μ is the characteristic function of an ultrafilter that contains $\mathcal{F}$. (This proof is modified from Luxemburg [1969].)

Remark. The Ultrafilter Principle and its consequences are needed frequently in duality theory and will be used heavily throughout the remainder of this chapter. *Hereafter we shall use the Ultrafilter Principle freely*; we shall discontinue our past practice of keeping track of its uses and its equivalents.

28.30. Let X be a topological vector space with topology γ and with dual X^*. Consider polars with respect to the bilinear pairing $\langle X, X^* \rangle$. Let $S \subseteq X^*$. We consider some conditions that might be satisfied by S:

(A) S is equicontinuous from X_γ to the scalar field $\mathbb{F}$ — that is, $S^\triangleright$ is a γ-neighborhood of 0.

(B) S is contained in some $\sigma(X^*, X)$-compact, convex, balanced subset of X^*.

(C) S is contained in some $\sigma(X^*, X)$-compact subset of X^*.

(D) S is $\sigma(X^*, X)$-bounded; that is, $S^\triangleright$ is absorbing.

Proposition. In any TVS X_γ we have (A) $\Rightarrow$ (B) $\Rightarrow$ (C) $\Rightarrow$ (D).

Moreover, suppose X_γ is a locally convex space. Then X_γ is barrelled if and only if (D) $\Rightarrow$ (A). In other words, a locally convex space X is barrelled if and only if it satisfies this condition (compare with 27.27(B5)):

> **(B5′) Another Uniform Boundedness Property.** Let Φ be a collection of continuous linear maps from X into the scalar field that is bounded pointwise. Then Φ is equicontinuous.

Proof. For (A) $\Rightarrow$ (B), note that $S \subseteq S^{\triangleright\triangleleft}$, and use 28.26.d and (UF27) in 28.29. The implication (B) $\Rightarrow$ (C) is trivial. The implication (C) $\Rightarrow$ (D) is just 27.3.c.

To show that (D) $\Rightarrow$ (A) in any barrelled space, note that $S^\triangleright$ is a $\sigma(X, X^*)$-closed, convex, balanced subset of Y; hence it is also X_γ-closed (see 28.13.b). If $S^\triangleright$ is absorbing, then it is a γ-barrel, hence a γ-neighborhood of 0.

On the other hand, suppose that (D) $\Rightarrow$ (A); let us show X_γ is barrelled. Let $R \subseteq X$ be a barrel — i.e., a γ-closed, convex, balanced, absorbing set; we wish to show R is a γ-neighborhood of 0. Since R is γ-closed and convex, it is also weakly closed — i.e., it is $\sigma(X, X^*)$-closed; see 28.14.a. Hence $R = R^{\triangleleft\triangleright}$. Let $S = R^\triangleleft$; then $R = S^\triangleright$. Now apply (D) $\Rightarrow$ (A).

28.31. Orlicz-Pettis Theorem. Let $\sum_{j=1}^{\infty} x_j$ be a series in a Banach space $(X, \| \quad \|)$. Then $\sum_{j=1}^{\infty} x_j$ is unconditionally convergent (as in conditions (A) through (F) of 23.26) if and only if

(G) Each subseries $\sum_{k=1}^{\infty} x_{j_k}$ is weakly convergent. That is, if $S \subseteq \mathbb{N}$, then there exists some $y_S \in X$ such that the series $\sum_{j \in S} x_j$ converges weakly to y_S.

Proof. Obviously (F) $\Rightarrow$ (G). We shall prove (G) $\Rightarrow$ (C). We may assume that the scalar field is $\mathbb{R}$. For any $f \in X^*$, the subseries

$$\sum_{f(x_j)>0} f(x_j) \qquad \text{and} \qquad \sum_{f(x_j)<0} f(x_j)$$

are both convergent; hence $\sum_{j=1}^{\infty} |f(x_j)| < \infty$. Thus $V : f \mapsto \{f(x_j) : j \in \mathbb{N}\}$ defines a linear map $V : X^* \to \ell_1$.

To show that V is continuous, we shall use the Closed Graph Theorem (27.28.c): Suppose that $f^{(k)} \to f$ in X^* and $V\left(f^{(k)}\right) \to (c_j : j \in \mathbb{N})$ in ℓ_1, as $k \to \infty$. For fixed j, the former hypothesis yields $f^{(k)}(x_j) \to f(x_j)$, while the latter hypothesis yields $f^{(k)}(x_j) \to c_j$. Thus $c_j = f(x_j)$, V has closed graph, and therefore V is continuous.

Let U be the closed unit ball of X^*; then $V(U)$ is a bounded subset of ℓ_1.

We wish to show that $V(U)$ is relatively compact. Let (u_k) be any sequence in U; we wish to show that $(V(u_k) : k = 1, 2, 3, \ldots)$ has a subsequence that is convergent in ℓ_1. By Schur's Theorem (28.20), it suffices to show that $(V(u_k))$ has a subsequence that is weakly convergent in ℓ_1.

Let X_0 be the closed span of the sequence (x_j); then X_0 is separable. Note that X_0 is also weakly closed; hence $y_S \in X_0$ for each set $S \subseteq \mathbb{N}$.

For each k, let $\widehat{u}_k$ be the restriction of u_k to X_0; thus $\widehat{u}_k$ is a member of the closed unit ball of $X_0{}^*$. By 28.24.a and (UF28) in 28.29, that closed unit ball is a compact metrizable space, when equipped with the $\sigma(X_0{}^*, X_0)$ topology. Therefore, the sequence $(\widehat{u}_k : k = 1, 2, 3, \ldots)$ has a subsequence $(\widehat{u}_{k(p)} : p = 1, 2, 3, \ldots)$ that is $\sigma(X_0{}^*, X)$-convergent to some limit $\widehat{u}_0$ in that closed unit ball. That is, $\widehat{u}_{k(p)}(y) \to \widehat{u}_0(y)$ for each y in X. In particular, $\widehat{u}_{k(p)}(y_S) \to \widehat{u}_0(y_S)$ for each $S \subseteq \mathbb{N}$.

By the Hahn-Banach Theorem (HB7) in 23.18, we can extend the functional $\widehat{u}_0 : X_0 \to \mathbb{R}$ to a continuous linear functional $u_0 : X \to \mathbb{R}$ with the same norm; then $u_0 \in U$.

It suffices to show that the corresponding subsequence $(V(u_{k(p)}) : p = 1, 2, 3, \ldots)$ converges weakly in ℓ_1 to $V(u_0)$. That is, we shall show that for each $\varphi \in \ell_\infty$, we have

$$\left\langle \varphi,\ V(u_{k(p)}) \right\rangle \quad \to \quad \langle \varphi,\ V(u_0) \rangle \qquad \text{as } p \to \infty. \tag{$*$}$$

It suffices to show $(*)$ for all φ in a dense subset of ℓ_∞, since the $V(u_{k(p)})$'s are bounded. By linearity, it suffices to show $(*)$ for all φ in a set whose span is dense in ℓ_∞. One such set is the set of all characteristic functions of subsets of $\mathbb{N}$; thus it suffices to show $(*)$ whenever $\varphi = 1_S$ for some $S \subseteq \mathbb{N}$. Unwinding the notation, for any $u \in U$ we find that $\langle 1_S,\ V(u) \rangle = \sum_{j \in S} u(x_j) = u(y_S) = \widehat{u}(y_S)$, where $\widehat{u}$ is the restriction of u to X_0. Thus

$$\left\langle 1_S,\ V(u_{k(p)}) \right\rangle \ = \ \widehat{u}_{k(p)}(y_S) \quad \to \quad \widehat{u}_0(y_S) \ = \ \langle 1_S,\ V(u_0) \rangle$$

when $p \to \infty$. This completes the proof.

Duals of Product Spaces

28.32. Lemma. The dual of a product of TVS's is (algebraically) equal to the external direct sum of their duals; that is, $\left(\prod_{\lambda\in\Lambda} X_\lambda\right)^* = \coprod_{\lambda\in\Lambda}(X_\lambda{}^*)$.

In more detail, let $\mathbb{F}$ be the scalar field. For each $\lambda \in \Lambda$, let X_λ be a topological vector space, with dual $X_\lambda{}^* = \{$continuous linear maps from X_λ into $\mathbb{F}\}$. Let $P = \prod_{\lambda\in\Lambda} X_\lambda$ be the product topological vector space. Let $Q = \coprod_{\lambda\in\Lambda}(X_\lambda{}^*)$ be the external direct sum of the duals — i.e., Q consists of those functions $q \in \prod_{\lambda\in\Lambda}(X_\lambda{}^*)$ such that $q_\lambda = 0$ for all but finitely many λ's. For each $q \in Q$, define a corresponding mapping $\widehat{q} : P \to \mathbb{F}$ by $\widehat{q}(p) = \sum_{\lambda\in\Lambda} q_\lambda(p_\lambda) = \sum_{\lambda\in\Lambda} q_\lambda(\pi_\lambda(p))$; here $\pi_\lambda : P \to X_\lambda$ is the λth coordinate projection. Then the mapping $q \mapsto \widehat{q}$ is an algebraic isomorphism (i.e., a linear bijection) from Q onto P^*. (We do not consider any topologies on Q or P^* here.)

Hints: Each mapping $P \xrightarrow{\pi_\lambda} X_\lambda \xrightarrow{q_\lambda} \mathbb{F}$ is a composition of two continuous maps and therefore continuous; thus each $\widehat{q}$ is continuous. Purely algebraic considerations (i.e., without regard to topology) show that the mapping $\widehat{q}$ is linear and that the mapping $q \mapsto \widehat{q}$ is linear and injective. It suffices to show that this mapping is also surjective. Let any $\psi \in P^*$ be given.

For each $\lambda \in \Lambda$ there is a continuous linear injection $\iota_\lambda : X_\lambda \to P$, defined by taking $\iota_\lambda(x)$ to be the vector whose λth coordinate is x and whose other coordinates are all 0. The composition $q_\lambda : X_\lambda \xrightarrow{\iota_\lambda} P \xrightarrow{\psi} \mathbb{F}$ is a continuous linear functional on X_λ, and thus a member of X^*.

Since $\psi : P \to \mathbb{F}$ is continuous, the set $N = \{p \in P : |\psi(p)| < 1\}$ is a neighborhood of 0 in P. By the basic properties of the product topology (see 15.24.a), $N \supseteq \prod_{\lambda\in\Lambda} N_\lambda$ where each N_λ is a neighborhood of 0 in X_λ and $N_\lambda = X_\lambda$ for all but finitely many λ's. Say we have $N_{\lambda_j} \subsetneqq X_{\lambda_j}$ for $j = 1, 2, \ldots, m$. If $p \in P$ is a point that vanishes in all the coordinates $\lambda_1, \lambda_2, \ldots, \lambda_m$, then $rp \in N$ for every scalar r, and therefore $\psi(p) = 0$. Thus, for any $p \in P$, the value of $\psi(p)$ only depends on the coordinates $p_{\lambda_1}, p_{\lambda_2}, \ldots, p_{\lambda_m}$. It follows that $\psi(p) = q_{\lambda_1}(p_{\lambda_1}) + q_{\lambda_2}(p_{\lambda_2}) + \cdots + q_{\lambda_m}(p_{\lambda_m})$. This completes the proof.

28.33. *A nonmetrizable example.* Let $\mathbb{F}$ be the scalar field. Then $\mathbb{F}^{\mathbb{N}}$ is the space of all sequences of scalars, and $\coprod_{\mathbb{N}} \mathbb{F} = \{y \in \mathbb{F}^{\mathbb{N}} : y_j = 0$ for all but finitely many $j\}$. These two spaces have a natural separated pairing: $\langle x, y\rangle = \sum_{j=1}^{\infty} x_j y_j$.

When equipped with the product topology, $\mathbb{F}^{\mathbb{N}}$ is a Fréchet space, as discussed in 26.7, 26.20.a, and thereafter. By 28.32, its dual is $\coprod_{\mathbb{N}} \mathbb{F}$. Conversely, a natural topology for $\coprod_{\mathbb{N}} \mathbb{F}$ is the inductive limit topology, making it an LF space, as discussed in 27.42; then it is barrelled but not metrizable. We shall now show that its dual is $\mathbb{F}^{\mathbb{N}}$.

First, consider any $x \in \mathbb{F}^{\mathbb{N}}$. It defines a linear functional $\langle x, \cdot\rangle : \coprod_{\mathbb{N}} \mathbb{F} \to \mathbb{F}$ by the bilinear pairing defined above. This linear functional is continuous on each of the finite-dimensional subspaces $S_n = \{y \in \mathbb{F}^{\mathbb{N}} : y_j = 0$ for all $j > n\}$; hence it is continuous on $\coprod_{\mathbb{N}} \mathbb{F}$.

Conversely, let ψ be a continuous linear functional on $\coprod_{\mathbb{N}} \mathbb{F}$. Define $x_j = \psi(e_j)$, where $e_j = (0, 0, \ldots, 0, 1, 0, 0, \ldots)$ is the sequence with 1 in the jth place and 0s elsewhere. Then $\langle x, \cdot\rangle = \psi(\cdot)$ on $\coprod_{\mathbb{N}} \mathbb{F}$.

28.34. *Proposition.* The weak topology on a product is the product of the weak topologies.

In more detail, let $P = \prod_{S\in\mathcal{S}} Y_S$ be a product of topological vector spaces, and let P have the resulting product topology, thus making it a TVS also. Let $(Y_S)_\sigma$ and P_σ be the spaces Y_S and P equipped with their weak topologies, respectively. Then the topology of P_σ is equal to the product topology on $\prod_{S\in\mathcal{S}}((Y_S)_\sigma)$.

Proof. Let $\pi_S : P \to Y_S$ be the Sth coordinate projection. Recall from 28.32 that $P^* = \coprod_{S\in\mathcal{S}}(Y_S{}^*)$. Let (p_α) be a net in P, and let $p \in P$. Then

$$\begin{aligned}
p_\alpha \to p \text{ in } P_\omega \quad &\iff \quad \psi(p_\alpha - p) \to 0 \text{ in } \mathbb{F} \text{ for each } \psi \in P^* \\
&\iff \quad \sum_{j=1}^{n} \varphi_j\left(\pi_{S_j}(p_\alpha - p)\right) \to 0 \text{ for each finite set} \\
&\qquad\qquad \{S_1, S_2, \ldots, S_n\} \subseteq \mathcal{S} \text{ and each collection of} \\
&\qquad\qquad \text{functionals } \varphi_1 \in (Y_{S_1})^*, \ldots, \varphi_n \in (Y_{S_n})^* \\
&\iff \quad \varphi\left(\pi_S(p_\alpha - p)\right) \to 0 \text{ for each } S \in \mathcal{S} \text{ and} \\
&\qquad\qquad \text{each } \varphi \in (Y_S)^* \\
&\iff \quad \pi_S(p_\alpha) \to \pi_S(p) \text{ in } (Y_S)_\sigma \text{ for each } S \in \mathcal{S} \\
&\iff \quad p_\alpha \to p \text{ in } \prod_{S\in\mathcal{S}}((Y_S)_w).
\end{aligned}$$

28.35. Theorem on embedding an LCS in a product of Banach spaces. Let X be a Hausdorff locally convex space. It is sometimes convenient to represent X as a linear subspace of a product of Banach spaces B_S, as follows:

Let $\mathbb{F}$ be the scalar field, and let X^* be the dual of X. Let $\mathcal{S}$ be the collection of all equicontinuous subsets of X^*; here equicontinuous refers to mappings from X with its given topology to $\mathbb{F}$. Then the seminorms ρ_S, defined as in 28.8, determine the topology of X, and they separate the points of X.

For each $S \in \mathcal{S}$, the seminorm ρ_S determines a norm $\|\cdot\|_S = \widehat{\rho}_S(\cdot)$ on the quotient space $X/\rho_S^{-1}(0)$, as in 22.13.e. Let $\pi_S : X \to X/\rho_S^{-1}(0)$ be the quotient map. Let B_S be the completion of the normed space $(X/\rho_S^{-1}(0), \|\cdot\|_S)$; the norm of the Banach space B_S will also be denoted by $\|\cdot\|_S$. Let $P = \prod_{S\in\mathcal{S}} B_S$ be the product, equipped with the product topology. Define a map $\iota : X \to P$ by taking the Sth coordinate of $\iota(x)$ to be $\pi_S(x)$. Show that

(i) The mapping ι is injective. Thus we may view ι as an inclusion; then X is a subset of the product P. The quotient maps $\pi_S : X \to X/\rho_S^{-1}(0)$ are just restrictions of the coordinate projections $\pi_S : P \to B_S$.

(ii) The given topology on X is the relative topology determined by P.

(iii) Let "σ" denote weak topology; then X_σ has the relative topology determined by $P_\sigma = \prod_{S\in\mathcal{S}}((B_S)_\sigma)$. (The last equation was proved in 28.34.)

Proofs. X is Hausdorff, so the seminorms ρ_S separate points of X. It follows that ι is injective. The rest of (i) is obvious.

For (ii) and (iii), let any net (x_α) be given in X, and let $x_\infty \in X$; we must show that $x_\alpha \to x_\infty$ in the given topology (respectively, in the weak topology) if and only if for each $S \in \mathcal{S}$ we have $\pi_S(x_\alpha) \to \pi_S(x_\infty)$ in the norm topology of B_S (respectively, in the weak topology of B_S). By linearity, we may assume $x_\infty = 0$.

For (ii) the argument is quite simple: $x_\alpha \to 0$ in the given topology $\iff$ $\rho_S(x_\alpha) \to 0$ for each S $\iff$ $\|\pi_S(x_\alpha)\|_S \to 0$ for each S $\iff$ $\pi_S(x_\alpha) \to 0$ in B_S for each S.

For (iii), first suppose that $x_\alpha \to 0$ weakly in X. Let any $S \in \mathcal{S}$ be given, and let $f \in (B_S)^*$. Then $X \xrightarrow{\iota} P \xrightarrow{\pi_S} B_S \xrightarrow{f} \mathbb{F}$ is a composition of continuous linear maps, if X is equipped with its given topology and X_S with the norm topology and P with the product topology. Thus $f \circ \pi_S \circ \iota$ is a member of X^*. Since $x_\alpha \to 0$ weakly in X, we have $(f \circ \pi_S \circ \iota)(x_\alpha) \to 0$ in $\mathbb{F}$. Thus $\pi_S(x_\alpha) \to 0$ weakly in B_S.

Conversely, suppose that $\pi_S(x_\alpha) \to 0$ weakly in B_S for each S. Fix any $f \in X^*$; we must show $f(x_\alpha) \to 0$. We may assume that f is not the constant function 0. Note that the singleton $\{f\}$ is itself an equicontinuous subset of X^*; let us denote it by S. Then $\rho_S(x) = |f(x)|$. Then $X/\rho_S^{-1}(0) = X/f^{-1}(0) \approx \text{Range}(f) = \mathbb{F}$, where $\approx$ denotes an algebraic isomorphism. Thus B_S is one-dimensional, so weak convergence in B_S is the same as norm convergence. Hence $|f(x_\alpha)| = \rho(x_\alpha) = \|\pi_S(x_\alpha)\|_S \to 0$.

Characterizations of Weak Compactness

28.36. Eberlein-Smulian-Grothendieck Theorem. Let X be a Hausdorff locally convex TVS, and let X^* be its dual. In the conditions below, an equicontinuous subset of X^* means a collection of linear maps that is equicontinuous from the given topology on X to the usual topology on the scalar field.

Let Φ be a bounded convex subset of X that is complete (in the given topology on X). Then the following are equivalent:

(A) **(Iterated limit condition.)** For every net $(\varphi_\alpha : \alpha \in \mathbb{A})$ in Φ and every equicontinuous net $(s_\beta : \beta \in \mathbb{B})$ in X^*, we have

$$\lim_{\alpha \in \mathbb{A}} \lim_{\beta \in \mathbb{B}} \langle \varphi_\alpha, s_\beta \rangle \quad = \quad \lim_{\beta \in \mathbb{B}} \lim_{\alpha \in \mathbb{A}} \langle \varphi_\alpha, s_\beta \rangle$$

whenever both sides of the equation exist — i.e., whenever all the indicated limits exist.

(B) **(Sequential iterated limit condition.)** For every sequence (φ_m) in Φ and every equicontinuous sequence (s_n) in X^*, we have

$$\lim_{m \to \infty} \lim_{n \to \infty} \langle \varphi_m, s_n \rangle \quad = \quad \lim_{n \to \infty} \lim_{m \to \infty} \langle \varphi_m, s_n \rangle$$

whenever both sides of the equation exist.

(C) Φ is weakly compact. That is, each net in Φ has a subnet that converges weakly to some member of Φ.

(D) Φ is weakly countably compact. That is, each sequence in Φ has a subnet that converges weakly to some member of Φ.

Moreover, if X is a Banach space, then those conditions are equivalent to this one:

(E) Φ is weakly sequentially compact. That is, each sequence in Φ has a subsequence that converges weakly to some member of Φ.

Proof. Note that Φ is complete, hence closed. Also Φ is convex, hence Φ is weakly closed.

The implications (C) $\Rightarrow$ (D) and (E) $\Rightarrow$ (D) and (A) $\Rightarrow$ (B) are obvious.

We may prove (D) $\Rightarrow$ (A) as follows: By (UF26) in 28.29, any equicontinuous subset of X^* is relatively compact in $\sigma(X^*, X)$. Hence the sequence (s_n) has a subnet that converges in $\sigma(X^*, X)$ to some $s \in X^*$. By assumption (D), (φ_m) has a subnet that converges weakly to some $\varphi \in \Phi$. If both $\lim_{m\to\infty} \lim_{n\to\infty} \langle \varphi_m, s_n \rangle$ and $\lim_{n\to\infty} \lim_{m\to\infty} \langle \varphi_m, s_n \rangle$ exist, it follows easily that those limits must both be equal to $\langle \varphi, s \rangle$.

Next we prove that (B) implies (C) and (E) when X is a normed space. Let S be the closed unit ball of the dual normed space $(X^*, \| \quad \|)$. Since the span of S is all of X^*, the weak topology on X is precisely the same as the topology of convergence pointwise on S. Let S be equipped with the $\sigma(X^*, X)$ topology; then S is $\sigma(X^*, X)$ compact, by (UF26) in 28.29. Since Φ is bounded, the set of scalars $\{\langle \varphi, s \rangle : \varphi \in \Phi, s \in S\}$ is bounded, and thus its closure in the scalar field $\mathbb{F}$ is a compact metric space M. Members of Φ may be viewed as distinct maps from S into M; it is easy to verify that those maps are continuous when S has the $\sigma(X^*, X)$ topology. Then $\Phi \subseteq C(S, M) \subseteq M^S$ as in 17.50. Then condition 17.50(B) is satisfied, hence conditions 17.50(C) and 17.50(E) are also satisfied. Moreover, Φ is weakly closed, by assumption, so the weak limits guaranteed by 17.50 all lie in Φ.

Finally we prove (B) $\Rightarrow$ (C), in general. For that purpose we employ the notation and conclusions of 28.35. Note that Φ is complete in X, hence Φ is complete when considered as a subset of P, hence Φ is closed in P. Also Φ is convex, hence Φ is closed in P_σ. Since the topology of X_σ is the relative topology determined by P_σ, we merely need to show that Φ is contained in some compact subset of P_σ. Recall from 28.34 that $P_\sigma = \prod_{S\in\mathcal{S}}((B_S)_w)$. Certainly Φ is contained in $\prod_{S\in\mathcal{S}} \pi_S(\Phi)$. By Tychonov's Theorem (AC21) in 17.16 (or by the weaker version (UF19) in 17.22) it suffices to show that for each S, the set $\pi_S(\Phi)$ is compact in $(B_S)_\sigma$. Applying the results of the previous paragraph, it suffices to show $\pi_S(\Phi)$ satisfies condition (B) in B_S; we hold S fixed throughout the remainder of this argument. Let (y_m) be a sequence in $\pi_S(\Phi)$, and let (t_n) be an equicontinuous sequence in $B_S{}^*$. Since B_S is a normed space, the equicontinuity of (t_n) simply means that (t_n) is bounded in norm; thus for any $\varepsilon > 0$ there is some δ such that

$$y \in B_S, \qquad \|y\|_S \leq \delta, \qquad n \in \mathbb{N} \qquad \Rightarrow \qquad t_n(y) \leq \varepsilon.$$

Define the compositions $s_n : X \xrightarrow{\iota} P \xrightarrow{\pi_S} B_S \xrightarrow{t_n} \mathbb{F}$. Then they satisfy

$$x \in X, \qquad \rho_S(x) \leq \delta, \qquad n \in \mathbb{N} \qquad \Rightarrow \qquad s_n(x) \leq \varepsilon.$$

Thus the sequence (s_n) is an equicontinuous subset of X^*. For each y_m in $\pi_S(\Phi)$, choose any $\varphi_m \in \Phi \cap \pi_S^{-1}(y_n)$. Then the scalars $\langle \varphi_m, s_n \rangle$ are the same as the scalars $\langle y_m, t_n \rangle$, so hypothesis (B) for Φ implies condition (B) for $\pi_S(\Phi)$. This completes the proof.

28.37. James's Sup Theorem. Let X be a Hausdorff, locally convex, real TVS, and let X^* be its dual. Let B be a bounded, weakly closed subset of X. Assume that the closed convex hull of B is complete. Then the following are equivalent:

(A) B is weakly compact.

(B) Each member of X^* attains a maximum on B.

(C) Each continuous (not necessarily linear) map from X_σ to $\mathbb{R}$ attains a maximum on B.

Of course, (A) $\Rightarrow$ (C) is just 17.7.i, and (C) $\Rightarrow$ (B) is trivial. The new result here is (B) $\Rightarrow$ (A).

Remarks. Note that the set B is not required to be convex.

James's Theorem is something of a "supertheorem." Its proof is quite long, but it makes into easy corollaries several substantial theorems that were proven in the years before James's Theorem. For instance, in 28.38 we shall use James's Theorem to prove the Kreĭn-Šmulian Theorem.

James originally proved this theorem for normed spaces; the proof was simplified slightly and extended to locally convex spaces by Pryce [1966]. For separable Banach spaces X, a substantially shorter proof was later given by Simons [1972]; that proof can also be found in Deville, Godefroy, and Zizler [1993]. However, for the more general setting considered here, Pryce's proof (given below) still seems to be the shortest.

Proof of theorem. We may replace B with its closed convex hull, since (i) each $f \in X^*$ attains the same maximum on that set, and (ii) if that set is weakly compact, then so is B since B is weakly closed. Thus, we may assume B itself is closed, convex, and complete. Replacing X with its completion, we may assume X is complete.

Assume B is not weakly compact; we shall eventually reach a contradiction. By 28.36, there is a sequence (z_n) in B and an equicontinuous sequence (e_j) in X^* such that the iterated limits $\lim_j \lim_n e_j(z_n)$ and $\lim_n \lim_j e_j(z_n)$ exist and are unequal. We may assume that

$$\lim_j \lim_n e_j(z_n) \quad > \quad a \quad > \quad b \quad > \quad \lim_n \lim_j e_j(z_n)$$

for some real numbers a, b. Let Γ be the linear span of the sequence (e_j), and let

$$\Phi \quad = \quad \{\text{positively homogeneous, continuous functions from } X \text{ to } \mathbb{R}\}.$$

Then Γ, X^*, Φ are linear spaces, with $\Gamma \subseteq X^* \subseteq \Phi$.

Each $f \in \Phi$ is continuous, and therefore is bounded on some neighborhood of 0. It follows that f is bounded on bounded sets, and in particular f is bounded on B. Hence the number

$$\rho(f) \quad = \quad \sup\{f(x) : x \in B\}$$

is finite. In this fashion we obtain a mapping $\rho : \Phi \to \mathbb{R}$. We easily verify that ρ is a sublinear functional on the linear space Φ. Also, we verify that if $A \subseteq \Phi$ is equicontinuous, then A is uniformly bounded on some neighborhood of 0, and hence uniformly bounded on B. Therefore ρ is bounded on any equicontinuous subset of Φ.

Let (f_i) be a subsequence of (e_j) — in the next paragraph we shall be more specific about our choice of this subsequence, but let us first consider properties of *any* subsequence. Define functions $\underline{\varphi}, \overline{\varphi} : X \to [-\infty, +\infty]$ by

$$\underline{\varphi}(x) \;=\; \liminf_{i\to\infty} f_i(x), \qquad\qquad \overline{\varphi}(x) \;=\; \limsup_{i\to\infty} f_i(x).$$

It is easy to verify that

$$\max\left\{|\underline{\varphi}(x) - \underline{\varphi}(y)|,\ |\overline{\varphi}(x) - \overline{\varphi}(y)|\right\} \;\leq\; \sup_i |f_i(x) - f_i(y)|.$$

Since the f_i's are equicontinuous and vanish at 0, it follows that $\underline{\varphi}$ and $\overline{\varphi}$ are real-valued and continuous. Furthermore, it is easy to verify that both $\underline{\varphi}$ and $\overline{\varphi}$ are positively homogeneous. Thus, they belong to the linear space Φ.

We shall now show that the subsequence (f_i) can be chosen so that

$$\rho(h - \underline{\varphi}) \;=\; \rho(h - \overline{\varphi}) \qquad\qquad \text{for all } h \in \Gamma. \tag{1}$$

To see this, let $(h_k : k \in \mathbb{N})$ be a sequence that is dense in Γ, when Γ is equipped with the topology of uniform convergence on B. (For instance, the linear combinations of e_j's with rational coefficients form a countable dense set; arrange it into a sequence.) For any fixed k, by 10.38 the given sequence (e_j) has a subsequence (f_i) that satisfies $\rho(h_k - \underline{\varphi}) = \rho(h_k - \overline{\varphi})$. By a diagonal subsequence argument (similar to that in 17.27), the given sequence has a subsequence (f_i) that satisfies $\rho(h_k - \underline{\varphi}) = \rho(h_k - \overline{\varphi})$ simultaneously for *every* positive integer k. Since the sequence (h_k) is dense in Γ, we obtain (1).

Since (f_i) is a subsequence of (e_j), we have

$$\lim_i \lim_n f_i(z_n) \;>\; a \;>\; b \;>\; \lim_n \lim_i f_i(z_n).$$

Deleting the first few terms of the sequences (f_i) and (z_n), we may assume that

$$\lim_{n\to\infty} f_i(z_n) > a \text{ for each } i, \qquad\qquad \lim_{i\to\infty} f_i(z_n) < b \text{ for each } n.$$

Then for each fixed k, we have

$$\lim_n \left[f_k(z_n) - \lim_i f_i(z_n)\right] \;>\; a - b \;>\; 0. \tag{2}$$

For $n = 0, 1, 2, 3, \ldots$, let K_n be the convex hull of the set $\{f_{n+1}, f_{n+2}, f_{n+3}, \ldots\}$. Then $\Phi \supseteq X^* \supseteq \Gamma \supseteq K_0 \supseteq K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots$. Next we claim that

$$\rho(f - \underline{\varphi}) \;>\; a - b \qquad\qquad \text{for each } f \in K_0. \tag{3}$$

To see this, fix any $f \in K_0$. Then $f = \sum_{k=1}^N \lambda_k f_k$ for some positive integer N and some $\lambda_1, \lambda_2, \ldots, \lambda_N \in [0,1]$ with $\sum_{k=1}^N \lambda_k = 1$. By (2), choose an integer n large enough so that $f_k(z_n) - \lim_{m\to\infty} f_m(z_n) > a - b$ for all $k = 1, 2, \ldots, N$. Then

$$\begin{aligned}
\rho(f - \underline{\varphi}) \;&=\; \sup_{x\in B}\left[f(x) - \underline{\varphi}(x)\right] \;\geq\; f(z_n) - \underline{\varphi}(z_n) \\
&=\; f(z_n) - \liminf_{k\to\infty} f_k(z_n) \;=\; f(z_n) - \lim_{m\to\infty} f_m(z_n) \\
&=\; \sum_{k=1}^N \lambda_k \left[f_k(z_n) - \lim_{m\to\infty} f_m(z_n)\right] \;>\; \sum_{k=1}^N \lambda_k(a-b) \;=\; a - b,
\end{aligned}$$

which proves (3).

Next we shall recursively choose $g_1 \in K_1$, $g_2 \in K_2$, $g_3 \in K_3$, ... satisfying

$$\rho\left[\sum_{n=1}^{N} \frac{g_n - \underline{\varphi}}{n!}\right] + \frac{1}{2} \cdot \frac{a-b}{(N+1)!} \quad < \quad \inf_{g \in K_N} \rho\left[\sum_{n=1}^{N} \frac{g_n - \underline{\varphi}}{n!} + \frac{g - \underline{\varphi}}{(N+1)!}\right] \tag{4}$$

for all integers $N \geq 0$ (with the convention that any summation $\sum_{n=1}^{0}$ is 0). For $N = 0$, inequality (4) is immediate from (3). Now assume that

$$\rho\left[\sum_{n=1}^{N-1} \frac{g_n - \underline{\varphi}}{n!}\right] + \frac{1}{2} \cdot \frac{a-b}{N!} \quad < \quad \inf_{g \in K_{N-1}} \rho\left[\sum_{n=1}^{N-1} \frac{g_n - \underline{\varphi}}{n!} + \frac{g - \underline{\varphi}}{N!}\right]$$

for some integer $N \geq 1$. Since $K_{N-1} \supseteq K_N$, we have also

$$\rho\left[\sum_{n=1}^{N-1} \frac{g_n - \underline{\varphi}}{n!}\right] + \frac{1}{2} \cdot \frac{a-b}{N!} \quad < \quad \inf_{g \in K_N} \rho\left[\sum_{n=1}^{N-1} \frac{g_n - \underline{\varphi}}{n!} + \frac{g - \underline{\varphi}}{N!}\right].$$

Now apply the lemma in 12.26, with

$$\alpha = \frac{1}{N!}, \qquad \beta = \frac{1}{(N+1)!}, \qquad \gamma = \frac{a-b}{2}, \qquad \xi = \sum_{n=1}^{N-1} \frac{g_n - \underline{\varphi}}{n!},$$

and with the convex set equal to $K_N - \underline{\varphi}$. Take $g_N = \eta + \underline{\varphi}$; this proves (4) and completes the recursive choice of the g_n's. Substituting $g = g_{N+1}$ in (4), we obtain also

$$\rho\left[\sum_{n=1}^{N} \frac{g_n - \underline{\varphi}}{n!}\right] + \frac{1}{2} \cdot \frac{a-b}{(N+1)!} \quad < \quad \rho\left[\sum_{n=1}^{N+1} \frac{g_n - \underline{\varphi}}{n!}\right]. \tag{5}$$

Next we shall show that this sequence (g_n) satisfies

$$\underline{\varphi}(x) \leq \liminf_{n\to\infty} g_n(x) \leq \limsup_{n\to\infty} g_n(x) \leq \overline{\varphi}(x) \qquad \text{for each } x \in X. \tag{6}$$

Fix any $x \in X$; temporarily fix any $n \in \mathbb{N}$. Since g_n is a convex combination of finitely many of the functions $f_{n+1}, f_{n+2}, f_{n+3}, \ldots$, there exists at least one $i \in \{n+1, n+2, n+3, \ldots\}$ with $f_i(x) \leq g_n(x)$. This proves $\liminf_{i\to\infty} f_i(x) \leq \liminf_{n\to\infty} g_n(x)$, which is half of (6). The other half is proved similarly.

Next, note that K_0 is the convex hull of the equicontinuous set $\{f_1, f_2, f_3, \ldots\}$. By 28.30, the $\sigma(X^*, X)$-closure of K_0 in X^* is $\sigma(X^*, X)$-compact. The sequence (g_n) lies in K_0 and therefore has a $\sigma(X^*, X)$-cluster point φ_0 in X^*. It follows that

$$\liminf_{n\to\infty} g_n(x) \quad \leq \quad \varphi_0(x) \quad \leq \quad \limsup_{n\to\infty} g_n(x) \qquad \text{for each } x \in X. \tag{7}$$

Combine this with (6) to obtain $\underline{\varphi}(x) \leq \varphi_0(x) \leq \overline{\varphi}(x)$ for all $x \in X$, and therefore

$$h(x) - \overline{\varphi}(x) \leq h(x) - \varphi_0(x) \leq h(x) - \underline{\varphi}(x) \qquad \text{for all } x \in X \text{ and } h \in \Gamma.$$

Take the supremum over all $x \in B$ to obtain $\rho(h - \overline{\varphi}) \le \rho(h - \varphi_0) \le \rho(h - \underline{\varphi})$ for all $h \in \Gamma$. Combine that with (1) to obtain $\rho(h - \varphi_0) = \rho(h - \underline{\varphi})$ for each $h \in \Gamma$. Since ρ is positively homogeneous, $\rho(\alpha h - \alpha\varphi_0) = \rho(\alpha h - \alpha\underline{\varphi})$ for each $\alpha > 0$ and $h \in \Gamma$. In particular,

$$\rho\left[\sum_{n=1}^{N} \frac{g_n - \varphi_0}{n!}\right] \quad = \quad \rho\left[\sum_{n=1}^{N} \frac{g_n - \underline{\varphi}}{n!}\right]$$

for any integer $N \ge 0$. Let $\gamma_n = g_n - \varphi_0$; from (5) we now obtain

$$\rho\left[\sum_{n=1}^{N} \frac{\gamma_n}{n!}\right] + \frac{1}{2} \cdot \frac{a-b}{(N+1)!} \quad < \quad \rho\left[\sum_{n=1}^{N+1} \frac{\gamma_n}{n!}\right].$$

Since the γ_n's lie in the equicontinuous set $K_0 - \varphi_0$, the summation $\psi = \sum_{n=1}^{\infty} \gamma_n/n!$ defines a functional $\psi \in X^*$. By the hypothesis of the theorem, ψ attains a maximum at some point $u \in B$. Also, since $K_0 - \varphi_0$ is equicontinuous and B is bounded,

$$\text{the number} \qquad \beta \quad = \quad \sup_{f \in K_0 - \varphi_0} \ \sup_{x \in B} \ |f(x)| \qquad \text{is finite.}$$

For each $N \in \mathbb{N}$ we have

$$\begin{aligned}
\sum_{n=1}^{N} \frac{\gamma_n(u)}{n!} \quad &= \quad \psi(u) - \sum_{n=N+1}^{\infty} \frac{\gamma_n(u)}{n!} \quad = \quad \rho(\psi) - \sum_{n=N+1}^{\infty} \frac{\gamma_n(u)}{n!} \\
&\ge \quad \rho\left[\sum_{n=1}^{N} \frac{\gamma_n}{n!}\right] - \rho\left[-\sum_{n=N+1}^{\infty} \frac{\gamma_n}{n!}\right] - \sum_{n=N+1}^{\infty} \frac{\gamma_n(u)}{n!} \\
&\ge \quad \rho\left[\sum_{n=1}^{N} \frac{\gamma_n}{n!}\right] - \sum_{n=N+1}^{\infty} \frac{2\beta}{n!} \quad > \quad \rho\left[\sum_{n=1}^{N-1} \frac{\gamma_n}{n!}\right] + \frac{1}{2} \cdot \frac{a-b}{N!} - \sum_{n=N+1}^{\infty} \frac{2\beta}{n!} \\
&\ge \quad \sum_{n=1}^{N-1} \frac{\gamma_n(u)}{n!} + \frac{1}{2} \cdot \frac{a-b}{N!} - \sum_{n=N+1}^{\infty} \frac{2\beta}{n!}.
\end{aligned}$$

Subtract $\sum_{n=1}^{N-1} \gamma_n(u)/n!$ from both ends of this computation and then multiply the resulting inequality by $N!$; that leaves

$$\gamma_N(u) \quad > \quad \frac{a-b}{2} \; - \; N! \sum_{n=N+1}^{\infty} \frac{2\beta}{n!}.$$

Since $0 = \lim_{N\to\infty} N! \sum_{n=N+1}^{\infty} 2\beta/n!$ (*exercise*), we have $\liminf_{N\to\infty} \gamma_N(u) \ge \frac{a-b}{2} > 0$, contradicting (7). This completes the proof of the theorem.

28.38. Kreĭn-Smulian Theorem. The closed convex hull of a weakly compact subset of a Banach space is weakly compact.

More generally, let X be a Hausdorff, locally convex space. Let $B \subseteq X$, let C be the closed convex hull of B, and suppose C is complete (in the given topology). Also suppose B is weakly compact. Then C is weakly compact.

Proof. Without loss of generality we may assume the scalar field is $\mathbb{R}$. We know B is bounded, by 27.3.c and 28.14.d; hence C is bounded, by 27.3.d and 27.3.e. Let $\lambda \in X^*$. Since B is weakly compact, we know that λ attains a maximum M on B. It follows easily that M is also the maximum of λ on C. Now apply James's Theorem.

Some Consequences in Banach Spaces

28.39. *Exercise.* A subset of ℓ_1 is compact if and only if it is weakly compact.

Proof. Immediate from 17.33, 28.20, and 28.36.

28.40. Goldstine-Weston Theorem. Let $(X, \| \ \|)$ be a normed space, with dual X^* and bidual X^{**}. Let B and B^{**} be the closed unit balls of X and X^{**}. Then B is $\sigma(X^{**}, X^*)$-dense in B^{**}.

Proof. Note that B^{**} is closed and convex, hence $\sigma(X^{**}, X^*)$-closed. Also, $B \subseteq B^{**}$. Let C be the $\sigma(X^{**}, X^*)$-closure of B in X^{**}; then $C \subseteq B^{**}$. It suffices to show that $C = B^{**}$. Suppose that $\xi \in B^{**} \setminus C$. We know that C is convex by 26.23.a, and $\sigma(X^{**}, X^*)$-compact by (UF28) in 28.29. By the Hahn-Banach Theorem (HB20) applied to the locally convex space X^{**} with the $\sigma(X^{**}, X^*)$ topology, there is some $f \in X^*$ with $c = \max_{\eta \in C} \operatorname{Re} f(\eta) < \operatorname{Re} f(\xi)$. Then $\sup_{x \in B} |f(x)| \leq c$, hence $\|f\| \leq c$, and hence $|f(\xi)| \leq \|f\| \, \|\xi\| \leq c$, a contradiction.

28.41. Theorem (Banach, Smulian, James, et al.) Let X be a Banach space, and let B be its closed unit ball. Then the following are equivalent:

(A) X is **reflexive**. That is, the canonical embedding $X \xrightarrow{\subseteq} X^{**}$ is surjective — i.e., each continuous linear functional on X^* acts the same as the evaluation map $f \mapsto f(x)$ for some $x \in X$. This condition is generally abbreviated as: $X^{**} = X$.

(B) On X^*, the weak topology $\sigma(X^*, X^{**})$ and the weak-star topology $\sigma(X^*, X)$ are equal.

(C) B is weakly compact.

(D) Every closed, convex, bounded subset of X is weakly compact.

(E) Whenever Q is a nonempty, closed, convex subset of X and $x \in X$, then there is at least one point $q \in Q$ that is closest to x — i.e., that satisfies $\|x - q\| = \operatorname{dist}(x, Q)$.

(F) Any nonempty, closed, convex subset of X contains at least one point of minimum norm. (Compare with 22.39(E).)

(G) For each $f \in X^*$, we have $\|f\| = \max\{|f(x)| : x \in B\}$.

(H) B is weakly complete — i.e., complete when equipped with the uniformity of the weak topology.

(I) The weak topology on X is quasicomplete (as defined in 27.3.f).

Remarks. In condition (G), we emphasize that a maximum is given, not just a supremum. Contrast this with 23.7 and (HB8) in 23.18.

Also, we note that some of the conditions are purely topological — i.e., they are unaffected if we replace the given norm on X with some equivalent norm. Therefore *all* the conditions above are purely topological. Thus, reflexivity should not be viewed as a property of certain normed vector spaces; rather, it is a property of certain topological vector spaces whose topologies are normable.

In condition (A), we emphasize that the isomorphism between X and X^{**} cannot be just any isomorphism; it must be given by the **canonical embedding** of $X \xrightarrow{\subseteq} X^{**}$, which was described in 9.57 and 23.20. It can be shown that the space J introduced in 22.26 is isomorphic to J^{**} — i.e., there exists a linear homeomorphism between J and J^{**} — but that isomorphism is not given by the canonical embedding, and in fact J does not satisfy any of the equivalent conditions listed above. This was proved by James [1951]; additional discussion on this subject can be found in James [1982].

Proof of (A) $\Rightarrow$ (B). Obvious.

Proof of (B) $\Rightarrow$ (C). By the Alaoglu Theorem (UF28) in 28.29.

Proof of (C) $\Rightarrow$ (D). Any closed, convex set is weakly closed and contained in the compact set rB for some $r > 0$.

Proof of (D) $\Rightarrow$ (E). Let $r = \operatorname{dist}(x, Q)$. Then the sets $S_n = \{q \in Q : \|x - q\| \le r + \frac{1}{n}\}$ are closed, convex, and bounded, hence weakly compact. Since $S_1 \supseteq S_2 \supseteq S_3 \supseteq \cdots$ and each S_n is nonempty, the intersection of the S_n's is nonempty. Any point q in that intersection satisfies $\|x - q\| = r$.

Proof of (E) $\Rightarrow$ (F). Take $x = 0$.

Proof of (F) $\Rightarrow$ (G). We know that $\|f\| = \sup\{\mathrm{Re} f(x) : x \in B\}$, and we wish to show that that supremum is actually a maximum. We may assume $f \ne 0$. Let $Q = \{x \in X : \mathrm{Re} f(x) \ge \|f\|\}$. Let q be a member of Q with smallest norm. It suffices to show that $\|q\| \le 1$. (Some readers may wish to take that assertion as an exercise before reading further.) Choose a sequence (b_n) with $\|b_n\| = 1$ and $\mathrm{Re} f(b_n) \to \|f\|$. Let $q_n = (\mathrm{Re} f(b_n))^{-1} \|f\| b_n$; then $q_n \in Q$. Since q is the member of Q with smallest norm, $\|q_n\| \ge \|q\|$ for all n. We have $\|q_n\| \to 1$, hence $1 \ge \|q\|$.

Proof of (G) $\Rightarrow$ (C). Immediate from James's Theorem 28.37.

Proof of (C) $\Rightarrow$ (H). Obvious.

Proof of (H) $\Rightarrow$ (I). The weak and norm topologies have the same bounded sets. Any TVS topology is invariant under multiplication by a positive scalar. Hence it suffices to consider subsets of B. Any weakly closed subset of B is (for the weak topology) a closed subset of a complete set, hence complete.

Proof of (I) $\Rightarrow$ (A). Let $\xi \in X^{**}$ be given. By the Goldstine-Weston Theorem, there is some bounded net $(x_\lambda : \lambda \in \Lambda)$ that is $\sigma(X^{**}, X^*)$-convergent to ξ. Then (x_λ) is $\sigma(X^{**}, X^*)$-Cauchy, hence $\sigma(X, X^*)$-Cauchy, hence $\sigma(X, X^*)$-convergent to some $x \in X$. It follows that $\xi(f) = f(x)$ for all $f \in X^*$.

28.42. *Exercise.* Let X be a Banach space, with dual X^*. Then X is reflexive if and only if X^* is reflexive.

Hints: Let X^{**} and X^{***} be the second and third dual spaces. Let $T : X \to X^{**}$ and $U : X^* \to X^{***}$ be the canonical embeddings; these maps are linear and distance-preserving. We are to show that (i) T is surjective if and only if (ii) U is surjective.

The proof of (i) $\Rightarrow$ (ii) involves little more than unwinding the notation and "chasing some arrows around a diagram." Let ξ be any member of X^{***}. The composition $\lambda : X \xrightarrow{T} X^{**} \xrightarrow{\xi} \mathbb{F}$ is a member of X^*. Using the definitions of T and U, verify that $U(\lambda) = \xi$.

The proof of (ii) $\Rightarrow$ (i) is a bit more substantial — it uses the Hahn-Banach Theorem. Suppose T is not surjective. Then $T(X)$ is a proper closed subspace of X^{**}; say $\varphi \in X^{**} \setminus T(X)$. By (HB11) in 23.18, there exists some continuous linear functional $\xi \in X^{***}$ that vanishes on $T(X)$ but not on φ. Unwind the notation to arrive at a contradiction.

28.43. *Exercise.* Show that the weak topology on ℓ_1 is sequentially complete but it is not quasicomplete.

28.44. Let $(X, \| \quad \|)$ be a real Banach space, with dual space X^*. (For simplicity we consider only real scalars.) The **normalized duality map** of X is the map $J : X \to \{\text{subsets of } X^*\}$ defined by

$$J(x) \quad = \quad \left\{\lambda \in X^* \ : \ \|\lambda\| = \|x\| \text{ and } \lambda(x) = \|x\|^2\right\}.$$

Such a map will be used in 30.20 and thereafter. Here are some of its basic properties:

a. The set $J(x)$ is nonempty, by (HB8) in 23.18.

b. The set $J(x)$ is convex and weak-star compact (hence also norm-closed).
Hint: Show that it is the intersection of the two sets $\{\lambda \in X^* : \|\lambda\| \leq \|x\|\}$ and $\{\lambda \in X^* : \lambda(x) \geq \|x\|^2\}$, both of which are convex and weak-star closed. Refer to (UF28) in 28.29.

c. $J(cx) = cJ(x)$ for any real number c.

d. If X^* is strictly convex, then J is single-valued — i.e., $J(x)$ is a singleton for each $x \in X$.
Hint: $J(x)$ is a convex subset of the sphere $\{\lambda \in X^* : \|\lambda\| = \|x\|\}$; see 22.39(C).

e. *Example.* If X is a Hilbert space with inner product $\langle\ ,\ \rangle$, then $X^* = X$ (see 28.50), and $J(x)$ is the singleton $\{\lambda_x\}$, where $\lambda_x(u) = \langle x, u\rangle$.

f. *Example.* Let $X = L^p(\mu)$ for some measure space $(\Omega, \mathcal{S}, \mu)$ and $1 < p < \infty$. Then $X^* = L^q(\mu)$ (see 28.50), and $J(x)$ is a singleton $\{y\}$, where $y \in L^q(\mu)$ is given by

$$y(\omega) \quad = \quad \begin{cases} x(\omega)|x(\omega)|^{p-2}\|x\|_p^{2-p} & \text{wherever } x(\omega) \neq 0 \\ 0 & \text{wherever } x(\omega) = 0. \end{cases}$$

g. *Example.* Let μ be σ-finite, and let $X = L^1(\mu)$; then $X^* = L^\infty(\mu)$ by 28.51. For any $x \in X$, the set $J(x)$ consists of all measurable real-valued functions y that satisfy these two conditions:

$$\begin{array}{rcll} y(\omega) & = & \|x\|_1 \operatorname{sign}(x(\omega)) & \quad \text{whenever } x(\omega) \neq 0, \\ |y(\omega)| & \leq & \|x\|_1 & \quad \text{whenever } x(\omega) = 0. \end{array}$$

More about Uniform Convexity

28.45. *Remark.* The results below, and more on this subject, can be found in Goebel and Reich [1984]. Some of the results below and in Chapter 29 can be proved more "constructively," in one sense or another of that word — e.g., without relying on the Banach-Alaoglu Theorem, the Hahn-Banach Theorem, and other weak forms of the Axiom of Choice. See Ishihara [1988], for instance. However, the nonconstructive arguments used below are quicker and (in this author's opinion) prettier.

28.46. Milman-Pettis Theorem. Any uniformly convex Banach space is reflexive.

Actually, reflexivity is a topological property, but uniform convexity is not. Thus we might reword the theorem in this slightly more precise fashion:

If X is a topological vector space whose topology can be given by a norm, and at least one such norm is uniformly convex, then the topological vector space X is reflexive.

Proofs of theorem. This is immediate from 28.41(E) and 22.45. However, we proved 28.41 using James's Theorem 28.37, which had a very long proof. For readers who wish to skip James's Theorem, we shall present another, more elementary proof, from Ringrose [1959]:

Let $(X, \| \ \|)$ be a uniformly convex Banach space. We have $X \subseteq X^{**}$ with the canonical embedding, as in 23.20. Let B be the closed unit ball of X. Let $\xi \in X^{**}$; we want to show $\xi \in X$. By rescaling, we may assume $\|\xi\| = 1$. By the Goldstine-Weston Theorem 28.40, there is some net $(x_\lambda : \lambda \in \Lambda)$ in B that converges in $\sigma(X^{**}, X^*)$ to ξ. Then $\|\xi\| \leq \liminf_\lambda \|x_\lambda\|$, as in 28.17.d, so we have $\|x_\lambda\| \to 1$. We may replace the vectors x_λ with the vectors $x_\lambda / \|x_\lambda\|$; thus we may assume $\|x_\lambda\| = 1$ for all λ.

We shall show that this net is Cauchy in the norm topology of X. Let any $\varepsilon > 0$ be given; let $\delta = \delta(\varepsilon)$ be the modulus of convexity of the space (as in 22.40). By definition of the norm of X^{**}, there is some $f \in X^*$ with $\|f\| = 1$ and $\operatorname{Re}\langle \xi, f \rangle > 1 - \delta$. Then for all $\lambda \in \Lambda$ sufficiently large, $\operatorname{Re}\langle x_\lambda, f \rangle > 1 - \delta$. Hence for all λ, μ sufficiently large, we have $\frac{1}{2}\|x_\lambda + x_\mu\| \geq \operatorname{Re}\left\langle \frac{1}{2}(x_\lambda + x_\mu), f \right\rangle > 1 - \delta$. Therefore $\|x_\lambda - x_\mu\| \leq \varepsilon$.

Thus the net (x_λ) is Cauchy in the normed space $(X, \| \ \|)$, which is complete. Therefore (x_λ) is norm convergent to some $x_0 \in X$. Since (x_λ) is also $\sigma(X^{**}, X^*)$-convergent to ξ, it follows that $x_0 = \xi$, and thus $\xi \in X$.

28.47. Lemma on asymptotic centers. Let C be a nonempty, closed, convex subset of a uniformly convex Banach space $(X, | \ |)$. Let (x_n) be a bounded sequence in X. Define a mapping $f : C \to [0, +\infty)$ by $f(u) = \limsup_{n\to\infty} |u - x_n|$. Then there is a unique point $u_0 \in C$ satisfying $f(u_0) = \min\{f(u) : u \in C\}$.

(The point u_0 is called the **asymptotic center** of the sequence (x_n) with respect to the set C.)

Proof. The function $f = \lim_{m\to\infty} \sup_{n\geq m} |x_n - x|$ is a limit of convex functions; hence it is convex. Also, it is continuous — in fact, it is nonexpansive, for if we take limsup on both sides of the inequality $|x_n - x| \leq |x_n - x'| + |x' - x|$ we obtain $f(x) \leq f(x') + |x' - x|$. Also, it is easy to see that if C is not bounded then $\lim_{|u|\to\infty} f(u) = \infty$. Let $I = \inf\{f(u) : u \in C\}$. Then the sets $C_n = \{u \in C : f(u) \leq I + \frac{1}{n}\}$ are nonempty, closed, convex, and bounded; hence they are nonempty and weakly compact. Moreover, $C_1 \supseteq C_2 \supseteq C_3 \supseteq \cdots$. Hence the C_n's have nonempty intersection, so f does assume a minimum — i.e., $f(u_0) = I$ for at least one point $u_0 \in C$.

To show that u_0 is unique, suppose that $f(u_0) = f(u_1) = I$ where $u_0 \neq u_1$. Let $v = \frac{1}{2}(u_0 + u_1)$. Let any $\varepsilon > 0$ be given. By the definition of f, there is some $N(\varepsilon)$ such that

$$|x_n - u_0| \leq I + \varepsilon, \qquad |x_n - u_1| \leq I + \varepsilon \qquad \text{whenever } n \geq N(\varepsilon).$$

Therefore

$$|x_n - v| \quad \leq \quad (I + \varepsilon)\left(1 - \delta\left(\frac{|u_0 - u_1|}{I + \varepsilon}\right)\right) \qquad \text{whenever } n \geq N(\varepsilon),$$

where δ is the modulus of convexity. Taking the limsup, we obtain

$$f(v) \quad \leq \quad I\left(1 - \delta\left(\frac{|u_0 - u_1|}{I}\right)\right) \quad < \quad I,$$

contradicting the fact that I is the minimum value of f. This proves the uniqueness of the asymptotic center.

28.48. Browder-Göhde-Kirk Fixed Point Theorem. Let C be a closed, convex, bounded subset of a uniformly convex Banach space. Let $g : C \to C$ be nonexpansive. Then g has at least one fixed point.

In fact, if x_0 is any point in C, and a sequence (x_n) is defined by $x_{n+1} = g(x_n)$, then the asymptotic center of the sequence (x_n) with respect to C is a fixed point of g.

Proof. Let u be that asymptotic center. Since g is nonexpansive, $|x_{n+1} - g(u)| \leq |x_n - u|$ for all n, and thus $\limsup_n |x_n - g(u)| \leq \limsup_n |x_n - u|$. Since $\limsup_n |x_n - (\cdot)|$ achieves its unique minimum at u, we have $g(u) = u$.

28.49. *Optional example.* We cannot replace "uniformly convex" with "strictly convex" in the preceding theorem.

Let $C[0,1] = \{$continuous scalar-valued functions on $[0,1]\}$; this is a Banach space when equipped with the sup norm. Show that $\|f\|_s = \|f\|_\infty + \|f\|_2$ is a strictly convex norm on $C[0,1]$ that is equivalent to the usual sup norm $\| \quad \|_\infty$. Also show that

$$F \quad = \quad \{f \in C[0,1] \ : \ f(0) = 0, \ f(1) = 1, \text{ and } \mathrm{Range}(f) \subseteq [0,1]\}$$

is a closed, convex, bounded subset of $C[0,1]$. Show that $(\varphi f)(t) = tf(t)$ defines a mapping $\varphi : F \to F$ that is nonexpansive when the norm $\| \quad \|_s$ is used, but φ has no fixed point.

Duals of the Lebesgue Spaces

28.50. Theorem. Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space (not necessarily finite or σ-finite). Then the Banach spaces $L^p(\mu)$ and $L^q(\mu)$ are the norm duals of each other, with elements of one space representing elements of the other space's dual by the bilinear pairing

$$[T(y)](x) \quad = \quad \langle x, y\rangle \quad = \quad \int_\Omega x(\omega)y(\omega)d\mu(\omega).$$

In particular, $L^2(\mu)$ is its own dual. In view of 22.56, any Hilbert space is its own dual.

Remark. Compare this result with the remark at the end of 29.21.

Proof of theorem. We shall show that $(L^p(\mu))^* = L^q(\mu)$. More precisely, for each $y \in L^q(\mu)$, let $T(y)$ be the mapping $\langle \cdot, y\rangle : L^p(\mu) \to \{\text{scalars}\}$ defined above; we shall show that T is an isomorphism from $L^q(\mu)$ onto $(L^p(\mu))^*$.

It follows from Hölder's Inequality that if $y \in L^q(\mu)$, then $T(y) = \langle \cdot, y\rangle$ (defined as above) is a continuous linear functional on $L^p(\mu)$, with operator norm $|||T(y)||| \leq \|y\|_q$. To show that we have equality here, define a function $x \in L^p(\mu)$ by choosing $x(\omega)$ to satisfy $|x(\omega)|^p = |y(\omega)|^q$ and $x(\omega)y(\omega) \geq 0$ for all ω. Then we can apply Hölder's Equality (see 22.33); it is easy to verify that $\langle x, y\rangle = \|x\|_p\|y\|_q$; from this it follows that $|||T(y)||| \geq \|y\|_q$. Thus the mapping $y \mapsto T(y)$ is a norm-preserving linear map from $L^q(\mu)$ into $(L^p(\mu))^*$; it suffices to prove that this map is surjective.

Suppose not. Then $T(L^q(\mu))$ is a proper, closed, linear subspace of $(L^p(\mu))^*$. By the Hahn-Banach Theorem (HB23) in 28.4, there is some ξ_0 in $(L^p(\mu))^{**}$ that vanishes on $T(L^q(\mu))$ but does not vanish on some $v_0 \in (L^p(\mu))^*$.

By 22.41.a and 28.46, $L^p(\mu)$ is reflexive. Thus there is some $x_0 \in L^p(\mu)$ that represents ξ_0, in this sense: We have $u(x_0) = \xi_0(u)$ for every $u \in (L^p(\mu))^*$. Define a function $y_0 \in L^q(\mu)$ by taking $|y_0(\omega)|^q = |x_0(\omega)|^p$ for all ω, with $x_0(\omega)y_0(\omega) \geq 0$. Since ξ_0 vanishes on $T(L^q(\mu))$, we have $0 = \xi_0(T(y_0)) = [T(y_0)](x_0) = \langle x_0, y_0\rangle = \|x_0\|_p^p = \|\xi_0\|^p$. Thus $\xi_0 = 0$, contradicting our assertion that ξ_0 does not vanish on some v_0.

28.51. Theorem. Let $(\Omega, \mathcal{S}, \mu)$ be a σ-finite measure space. Then the dual of $L^1(\mu)$ is $L^\infty(\mu)$.

Proof. Let $\mathbb{F}$ be the scalar field. For any $y \in L^\infty(\mu)$, we may define a mapping $T_y : L^1(\mu) \to \mathbb{F}$ by the rule

$$T_y(x) \quad = \quad \int_\Omega x(\omega)y(\omega)d\mu(\omega) \qquad\qquad \text{for } x \in L^1(\mu),$$

and obviously $|||T_y||| \leq \|y\|_\infty$. Now let any $T \in L^1(\mu)^*$ be given. It suffices to show that (i) $T = T_y$ for some $y \in L^\infty(\mu)$, and (ii) $\|y\|_\infty \leq |||T|||$.

Let us first fix our attention on any measurable set $\Omega_0 \subseteq \Omega$ that has finite measure. Let μ_0 denote the restriction of μ to Ω_0 and its measurable subsets. If $f \in L^1(\mu_0)$, then we may extend f to a member of $L^1(\mu)$ by defining $f = 0$ on $\Omega \setminus \Omega_0$. We have $L^2(\mu_0) \subseteq L^1(\mu_0)$ with continuous inclusion, by Hölder's inequality: $\|g\|_1 = \|1g\|_1 \leq \|g\|_2\|1\|_2 = \sqrt{\mu(\Omega_0)}\|g\|_2$. Hence the composition $\varphi : L^2(\mu_0) \xrightarrow{\subseteq} L^1(\mu_0) \xrightarrow{\subseteq} L^1(\mu) \xrightarrow{T} \mathbb{F}$ is continuous, and thus φ is a member of $L^2(\mu_0)$. However, $L^2(\mu_0)^* = L^2(\mu_0)$ by 28.50. Unwinding the notation, we see that there is a uniquely determined function $y_0 \in L^2(\mu_0)$ that satisfies $T(x) = \int_{\Omega_0} xy_0 d\mu_0$ for all $x \in L^2(\mu_0)$.

We claim that $|y_0(\cdot)| \leq |||T|||$ almost everywhere on Ω_0. Indeed, suppose on the contrary that $\{\omega \in \Omega_0 : |y_0(\omega)| > |||T|||\}$ has positive measure. Then the set

$$S \quad = \quad \{\omega \in \Omega_0 \ : \ r < |y_0(\omega)|\}$$

has positive measure, for some number $r > |||T|||$. Define

$$x(\omega) \quad = \quad \begin{cases} |y_0(\omega)|/y_0(\omega) & \text{when } \omega \in S \\ 0 & \text{when } \omega \in \Omega_0 \setminus S. \end{cases}$$

Then the function x belongs to $L^\infty(\mu_0) \subseteq L^2(\mu_0)$. Then

$$r\mu(S) \ \leq \ \int_S |y_0(\cdot)|d\mu_0 \ = \ \int_M xy_0\, d\mu_0 \ = \ T(x) \ \leq \ |||T|||\,\|x\|_1 \ = \ |||T|||\,\mu(S),$$

a contradiction. This proves our claim; we have $\|y_0\|_\infty \leq |||T|||$.

By our choice of y_0, we have $\int_{\Omega_0} xy_0\, d\mu = T(x)$ for all $x \in L^2(\mu_0)$. However, both sides of that equation are continuous functions of $x \in L^1(\mu_0)$, and $L^2(\mu_0)$ is dense in $L^1(\mu_0)$. Thus, that equation is valid for all $x \in L^1(\mu_0)$.

The function y_0 is uniquely determined on each set Ω_0 of finite measure. By covering Ω with an increasing sequence of sets of finite measure, we see that there is a measurable function $y : \Omega \to \mathbb{F}$, with $\|y\|_\infty \leq |||T|||$, such that $\int xy\, d\mu = T(x)$ whenever x is a member of $L^1(\mu)$ that vanishes outside some set of finite measure. Such functions are dense in $L^1(\mu)$, and both sides of the equation $\int xy\, d\mu = T(x)$ are continuous in x. Hence the equation holds for all $x \in L^1(\mu)$.

28.52. If $(\Omega, \mathcal{S}, \mu)$ is not σ-finite, then $(L^1(\mu))^*$ is *not* necessarily equal to $L^\infty(\mu)$.

Example (from Holmes [1975]). Let $(\Omega, \mathcal{S}, \mu)$ be the interval $[0, 1]$ equipped with counting measure. Here $\mathcal{S} = \mathcal{P}(\Omega)$, so every function $f : [0, 1] \to \mathbb{F}$ is measurable (where $\mathbb{F}$ is the scalar field $\mathbb{R}$ or $\mathbb{C}$). The integral of a function $f : [0,1] \to \mathbb{F}$ is the sum $\sum_{t\in[0,1]} f(t) \in [0,+\infty]$, provided that $\|f\|_1 = \sum_{t\in[0,1]} |f(t)|$ is finite.

Now let $\mathcal{S}_0$ be the σ-algebra of countable or cocountable subsets of $[0,1]$, and let μ_0 be the restriction of μ to $\mathcal{S}_0$. Thus a function $f : [0,1] \to \mathbb{F}$ is measurable with respect to $\mathcal{S}_0$ if and only if, for each Borel set $B \subseteq \mathbb{F}$, the set $f^{-1}(B)$ is countable or cocountable. The integral of such functions is the same as in the previous paragraph. If $f \in L^1(\mu_0)$, then the function $g(t) = tf(t)$ is not necessarily measurable with respect to $\mathcal{S}_0$, but we still have $|g(t)| \leq |f(t)|$, and so $g \in L^1(\mu)$. Thus we can define a bounded linear functional $\Lambda : L^1(\mu_0) \to \mathbb{F}$ by $\Lambda(f) = \sum_{t\in[0,1]} tf(t)$. It is an easy exercise that there does not exist a function $h \in L^\infty(\mu_0)$ satisfying $\Lambda(f) = \sum_{t\in[0,1]} f(t)h(t)$ for all $f \in L^1(\mu_0)$.

28.53. *Remark.* The dual of $L^\infty(\mu)$ will be characterized in 29.31.c.

Chapter 29

Vector Measures

Basic Properties

29.1. *Definition.* Let $(X, |\ \ |)$ be a Banach space. By an X-valued **charge** we mean a finitely additive mapping $\mu : \mathcal{A} \to X$, where $\mathcal{A}$ is an algebra of sets. By an X-valued **measure** we mean a countably additive mapping $\mu : \mathcal{S} \to X$, where $\mathcal{S}$ is a σ-algebra of sets. See 11.37. Such objects will be investigated in this chapter. Vector measures are used in spectral theory, but that subject will not be pursued in this book. This chapter covers only a very small portion of the subject of vector measures; the interested reader should refer to the encyclopedic work of Diestel and Uhl [1977] for a broader treatment.

Caution: The terminology varies. For instance, what we have called "charge" and "measure" are what Diestel and Uhl [1977] call, respectively, "vector measure" and "countably additive vector measure on a σ-algebra."

29.2. *Observation.* If $\mu : \mathcal{S} \to X$ is a Banach-space-valued measure and $S_1, S_2, S_3, \ldots$ are disjoint measurable sets, then the series $\mu(\bigcup_{j=1}^{\infty} S_j) = \sum_{j=1}^{\infty} \mu(S_j)$ is unconditionally convergent (defined in 23.26).

Proof: $\bigcup_{j=1}^{\infty} S_j$ is not affected if we change the order of the S_j's.

29.3. *Theorem.* Any Banach-space-valued measure is bounded.

Proof of proposition. Let $\lambda : \mathcal{S} \to X$ be a Banach-space-valued measure, where $\mathcal{S}$ is a σ-algebra of subsets of Ω. Then $|\lambda(S)|$ is a finite number for each $S \in \mathcal{S}$; we must show that $\sup_{S \in \mathcal{S}} |\lambda(S)|$ is finite as well. Suppose not. Call a set $S \in \mathcal{S}$ "fat" if it has the property that $\sup\{|\lambda(A)| : A \in \mathcal{S}, A \subseteq S\} = \infty$; then our assumption is that Ω is fat. We now recursively choose fat sets $S_0 \supseteq S_1 \supseteq S_2 \supseteq \cdots$ with $|\lambda(S_{n+1})| > |\lambda(S_n)| + 1$, by the following procedure: Let $S_0 = \Omega$; this is fat. Given a fat set S_n, choose some measurable set $B \subseteq S_n$ such that $|\lambda(B)| > 2|\lambda(S_n)| + 1$. Since S_n is fat, at least one of the sets $B, S_n \setminus B$ must be fat. Call that set S_{n+1}; we easily verify that $|\lambda(S_{n+1})| > |\lambda(S_n)| + 1$. This completes the recursion. Now, for $n = 1, 2, 3, \ldots$, let $T_n = S_{n-1} \setminus S_n$. Then the T_n's are disjoint and have union equal to Ω, hence $\sum_{n=1}^{\infty} \lambda(T_n) = \lambda(\Omega)$. However,

$$|\lambda(T_n)| \quad = \quad |\lambda(S_{n-1}) - \lambda(S_n)| \quad \geq \quad |\lambda(S_n)| - |\lambda(S_{n-1})| \quad \geq \quad 1,$$

so the series $\sum_{n=1}^{\infty} \lambda(T_n)$ diverges. This contradiction proves $\sup_{S\in\mathcal{S}} |\lambda(S)| < \infty$.

29.4. Theorem and definition. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space; assume μ is finite. Let $(X, |\ \ |)$ be a Banach space, and let $\lambda : \mathcal{S} \to X$ be a vector charge. Then the following two conditions are equivalent.

(A) $\lim_{\mu(S)\to 0} \lambda(S) = 0$. That is, for each number $\varepsilon > 0$, there exists some number $\delta > 0$ such that if $S \in \mathcal{S}$ with $\mu(S) < \delta$, then $|\lambda(S)| < \varepsilon$.

(B) λ is a measure, and λ vanishes on sets of μ-measure 0 — that is, $\mu(S) = 0 \Rightarrow \lambda(S) = 0$.

If either (hence both) of these conditions is satisfied, we say that λ is **absolutely continuous** with respect to μ or that it is **μ-continuous**; this is abbreviated $\boldsymbol{\lambda << \mu}$. Some examples will be given in 29.7 and 29.10.

Proof of equivalence. First assume (A). Obviously $\mu(S) = 0 \Rightarrow \lambda(S) = 0$. Let the sets $E_1, E_2, E_3, \ldots$ be disjoint with union E; we want to show that $\sum_{j=1}^{\infty} \lambda(E_j) = \lambda(E)$. Let $F_j = E \setminus (E_1 \cup E_2 \cup \cdots \cup E_j)$; we want to show that $\lambda(F_j) \to 0$ as $j \to \infty$. We know that $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$ and the F_j's have empty intersection. Hence $\mu(F_j) \downarrow 0$. By μ-continuity, $|\lambda(F_j)| \to 0$. Thus (A) $\Rightarrow$ (B).

We shall prove (B) $\Rightarrow$ (A) first in the case where λ is real-valued. Note that if $\mu(S) = 0$, then μ vanishes on every measurable subset of S; hence so does λ; hence so does $/\lambda/$; hence so do λ^+ and λ^-. By the Jordan Decomposition, it suffices to consider λ^+ and λ^-; thus we may assume $\lambda \geq 0$. Suppose that λ does not satisfy the condition in (A). Then there exist a number $\varepsilon > 0$ and measurable sets S_n such that $\lambda(S_n) > \varepsilon$ and $\mu(S_n) \to 0$. Passing to a subsequence, we may assume $\mu(S_n) < 2^{-n}$. Let $T = \limsup_{n\to\infty} S_n = \bigcap_{j=1}^{\infty} \bigcup_{n=j}^{\infty} S_n$. Then for each j we have $T \subseteq \bigcap_{n=j}^{\infty} S_n$, hence $\mu(T) \leq \sum_{n=j}^{\infty} \mu(S_n) < 2^{-j+1}$, hence $\mu(T) = 0$. On the other hand, by 21.25.c we have $\varepsilon \leq \limsup_{n\to\infty} \lambda(S_n) \leq \lambda(\limsup_{n\to\infty} S_n) = \lambda(T)$. This is a contradiction, proving (B) $\Rightarrow$ (A) in the case of real-valued λ.

Now we prove (B) $\Rightarrow$ (A) for an arbitrary Banach space X. We may assume the scalar field is $\mathbb{R}$. Suppose (B) holds but not (A); thus there exist measurable sets S_k with $|\lambda(S_k)| > \varepsilon$ and $\mu(S_k) < 2^{-k}$. We have $\lim_{m\to\infty} \mu\left(\bigcup_{k\geq m} S_k\right) = 0$. For each $u \in X^*$, the real-valued measure $u\lambda$ is μ-continuous and therefore satisfies

$$\lim_{m\to\infty} \sup \left\{ |u\lambda(S)| \ :\ S \in \mathcal{S},\ S \subseteq \bigcup_{k\geq m} S_k \right\} \quad = \quad 0.$$

Recursively choose positive integers $m(p)$ as follows: Let $m(0) = 1$. Given $m(p-1)$, use the Hahn-Banach Theorem (HB8) to choose some $u_p \in X^*$ with $|u_p| = 1$ and $u_p\lambda(S_{m(p-1)}) > \varepsilon$. Then choose $m(p) > m(p-1)$ large enough so that

$$\sup \left\{ |u_p\lambda(S)| \quad : \quad S \in \mathcal{S}\,,\ S \subseteq \bigcup_{k\geq m(p)} S_k \right\} \quad < \quad \frac{1}{2}\varepsilon\,.$$

This completes the recursion. Now define the sets $F_p = S_{m(p-1)} \setminus \bigcup_{k\geq m(p)} S_k$. The sets F_p are disjoint, and $|\lambda(F_p)| \geq |u_p\lambda(F_p)| > \frac{1}{2}\varepsilon$. Since λ is a measure, we have

$\lambda\left(\bigcup_{p=1}^{\infty} F_p\right) = \sum_{p=1}^{\infty} \lambda(F_p)$, and therefore the series $\sum_{p=1}^{\infty} \lambda(F_p)$ is convergent, and therefore $\lim_{p\to\infty} |\lambda(F_p)| = 0$, a contradiction.

THE VARIATION OF A CHARGE

29.5. *Definition.* Let $\mathcal{A}$ be an algebra of subsets of Ω, let $(X, |\ \ |)$ be a Banach space, and let $\lambda : \mathcal{A} \to X$ be a charge. The **variation** of λ is the function $/\lambda/ : \mathcal{A} \to [0, +\infty]$ defined by

$$/\lambda/(A) \quad = \quad \sup\left\{|\lambda(S_1)| + |\lambda(S_2)| + \cdots + |\lambda(S_n)|\right\}$$

where the supremum is over all positive integers n and partitions of A into disjoint subsets $S_1, S_2, \ldots, S_n \in \mathcal{A}$. (Much of the literature denotes the variation by $|\lambda|$, but we prefer $/\lambda/$ for reasons indicated in 8.39.)

For further clarification, we may refer to $/\lambda/$ as the **variation in the sense of charges or measures** to distinguish it from another type of "variation" introduced in 19.21. The relation between the two notions of "variation" will be considered in 29.33 and 29.34.

If $/\lambda/(\Omega) < \infty$, we say λ has **bounded variation**. The number $/\lambda/(\Omega)$ may also be written as $\mathrm{Var}(\lambda)$ or as $\mathrm{Var}(\lambda, \mathcal{A})$ if several different algebras of sets are being considered.

29.6. *Basic properties of the variation of a charge.*

a. $/\lambda/$ is a positive charge.

b. $\sup\{|\lambda(S)| : S \in \mathcal{A}\} \leq /\lambda/(\Omega)$. Thus, any charge with bounded variation is a bounded charge.

c. The **space of all X-valued charges on $\mathcal{A}$ with bounded variation** is a Banach space, with the variation for a norm. We may denote this space by $\boldsymbol{BV(\mathcal{A}, X)}$.

Proof of completeness. Apply 22.17 with $\Gamma = \mathcal{A}$, with Φ consisting of all functions of the form $\varphi(\lambda) = |\lambda(S_1)| + |\lambda(S_2)| + \cdots + |\lambda(S_n)|$ for disjoint sets $S_j \in \mathcal{A}$.

Remark. For a still larger space, see the space of bounded charges in 29.29.f.

d. If $X = \mathbb{R}$, then $/\lambda/(A) \;=\; \sup\{|\lambda(S)| + |\lambda(A \setminus S)| \;:\; S \in \mathcal{S},\ S \subseteq A\}$. Thus the definition of "variation" given in this chapter agrees with the definition of "variation" given in 11.47. As we noted in 11.47, any bounded real charge has bounded variation; hence any real-valued measure has bounded variation.

e. If λ is countably additive, then $/\lambda/(A)$ is also equal to $\sup \sum_{j=1}^{\infty} |\lambda(S_j)|$, where the supremum is over all partitions of A into countably many disjoint sets $S_j \in \mathcal{S}$.

Hints: Any finite partition can be written as a countable partition, by taking all but finitely many of the S_j's to be empty; thus $/\lambda/(A) \leq \sup \sum_{j=1}^{\infty} |\lambda(S_j)|$. On the other hand, if $/\lambda/(A) < r < \sum_{j=1}^{\infty} |\lambda(S_j)|$ for some countable partition (S_j) and some real number r, then choose N large enough to satisfy $r < \sum_{j=1}^{N} |\lambda(S_j)|$; then we have $/\lambda/(A) < |\lambda(A \setminus \bigcup_{j=1}^{N} S_j)| + \sum_{j=1}^{N} |\lambda(S_j)|$, a contradiction.

f. If λ is countably additive, then $/\lambda/$ is too. Thus, if λ is a measure, then $/\lambda/$ is too.

Hints: Let $A_1, A_2, A_3, \ldots$ be disjoint members of $\mathcal{A}$ with union $A \in \mathcal{A}$. Since $/\lambda/$ is a positive charge, for any positive integer N we have

$$\sum_{j=1}^{N} /\lambda/(A_j) \quad = \quad /\lambda/\left(\bigcup_{j=1}^{N} A_j\right) \quad \leq \quad /\lambda/(A),$$

and taking limits we obtain $\sum_{j=1}^{\infty} /\lambda/(A_j) \leq /\lambda/(A)$. For the reverse inequality, let $(B_k : k \in \mathbb{N})$ be any partition of A into countably many disjoint members of $\mathcal{A}$. Then $(A_j \cap B_k : k \in \mathbb{N})$ is a partition of A_j, and $(A_j \cap B_k : j \in \mathbb{N})$ is a partition of B_k. Hence $\sum_k |\lambda(B_k)| = \sum_k \left|\sum_j \lambda(A_j \cap B_k)\right| \leq \sum_{j,k} |\lambda(A_j \cap B_k)| \leq \sum_j /\lambda/(A_j)$. Taking the supremum over all choices of the sequence (B_k) yields $/\lambda/(A) \leq \sum_{j=1}^{\infty} /\lambda/(A_j)$.

g. If λ is a vector measure with bounded variation, then $\lambda << /\lambda/$.

h. Any real-valued measure — or, more generally, any measure taking values in a finite-dimensional Banach space — has bounded variation. (*Proof.* 29.3 and 29.6.d.)

29.7. *Example: a pathological vector measure.* We exhibit a bounded vector measure that has infinite variation on every nontrivial set. (This example is taken from Diestel and Uhl [1977].)

Let $(\Omega, \mathcal{S}, \mu)$ be the measure space $[0,1]$ with Lebesgue measure on the Lebesgue-measurable sets. Let $X = L^2[0,1]$. Define $\lambda : \mathcal{S} \to X$ by $\lambda(S) = 1_S$ (i.e., the characteristic function of S). Then λ is μ-continuous — i.e., λ vanishes on sets that have μ-measure 0.

To show that λ is countably additive, verify that if (E_n) is a sequence of disjoint measurable subsets of $[0,1]$, then

$$\left\|\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) - \sum_{n=1}^{N} \lambda(E_n)\right\|_X^2 \quad = \quad \mu\left(\bigcup_{n=N+1}^{\infty} E_n\right),$$

which tends to 0 as $N \to \infty$.

On the other hand, if E is a measurable set with $\mu(E) > 0$, we shall show that $/\lambda/(E) = \infty$. Indeed, let N be any positive integer. By 24.25, $\int_0^u 1_E(t)dt$ is a continuous function of u, so it must pass through each number between 0 and $\mu(E)$. Hence we can partition E into disjoint measurable sets $E_1, E_2, \ldots, E_N$ that have equal Lebesgue measure — i.e., they all have $\mu(E_j) = \frac{1}{N}\mu(E)$. Then

$$/\lambda/(E) \quad \geq \quad \sum_{j=1}^{N} |\lambda(E_j)|_X \quad = \quad \sum_{j=1}^{N} \sqrt{\int_0^1 \left[1_{E_j}(t)\right]^2 dt} \quad = \quad \sqrt{N\mu(E)},$$

which can be made arbitrarily large.

29.8. Nikodym Convergence Theorem (optional). Let $(X, |\ \ |)$ be a Banach space, and let $\{\lambda_n : n \in \mathbb{N}\}$ be a sequence of X-valued measures on a measurable space $(\Omega, \mathcal{S})$. Assume $\lambda(S) = \lim_{n\to\infty} \lambda_n(S)$ exists in X for each $S \in \mathcal{S}$. Then λ is a measure. Furthermore:

The λ_n's are uniformly countably additive. That is, if $S_k \downarrow \varnothing$ in $\mathcal{S}$, then $\lambda_n(S_k) \to 0$ (as $k \to \infty$) uniformly in n. In other words, the sequence-valued function

$$\Lambda(S) \quad = \quad \Big(\lambda_1(S),\ \lambda_2(S),\ \lambda_3(S),\ \ldots\Big)$$

is a $c(X)$-valued measure.

Explanation of notations. The expression $S_k \downarrow \varnothing$ means that $S_1 \supseteq S_2 \supseteq S_3 \supseteq \cdots$ and $\bigcap_{k=1}^{\infty} S_k = \varnothing$. Also, $c(X)$ means the Banach space of all convergent sequences in X; it is normed by $\|(x_1, x_2, x_3, \ldots)\| = \sup_{n\in\mathbb{N}} |x_n|$.

Proof of theorem. We first prove this theorem in the classical case, where X is the scalar field $\mathbb{F}$. In that case, each variation $/\lambda_n/$ is a positive, finite measure. Hence

$$\mu(S) \quad = \quad \sum_{n=1}^{\infty} \frac{/\lambda_n/(S)}{2^n\ /\lambda_n/(\Omega)}$$

defines a probability μ on $(\Omega, \mathcal{S})$ with the further property that $\lambda_n << \mu$ for all n.

Define a pseudometric d on $\mathcal{S}$, by

$$d(S,T) \ = \ \mu(S \triangle T) \ = \ \int_\Omega |1_S(\omega) - 1_T(\omega)|\ d\mu(\omega) \ = \ \|1_S - 1_T\|_1\,;$$

the last expression is a norm in the Lebesgue space $L^1(\mu)$. The space $L^1(\mu)$ is complete, by 22.31(i); and $\{1_S : S \in \mathcal{S}\} = \Big\{f \in L^1(\mu) \ : \ \text{Range}(f) \subseteq \{0,1\}\Big\}$ is a closed subset of $L^1(\mu)$, by 22.31(ii); hence the pseudometric space $(\mathcal{S}, d)$ is complete. By the Baire Category Theorem (20.16), $(\mathcal{S}, d)$ is a Baire space, and so any comeager subset of $\mathcal{S}$ is dense and thus nonempty.

From $\lambda_n << \mu$ it follows easily that each λ_n is a continuous map from the pseudometric space $(\mathcal{S}, \mu)$ into $\mathbb{R}$, and each λ_n is a continuous map from $(\mathcal{S}, \mu)$ into X. By the Baire-Osgood Theorem (20.8), (λ_n) is equicontinuous on a subset of $\mathcal{S}$ that is comeager, hence nonempty.

Say (λ_n) is equicontinuous at some particular $T \in \mathcal{S}$. Let $S_k \downarrow \varnothing$. We verify that the sequences $(T \cup S_k)$ and $(T \setminus S_k)$ both converge to T in the metric space $(\mathcal{S}, \mu)$. Since the sequence (λ_n) is equicontinuous at T, the sequences $\lambda_n(T \cup S_k)$ and $\lambda_n(T \setminus S_k)$ both converge to $\lambda_n(T)$ uniformly in n as $k \to \infty$. Then $\lambda_n(S_k) = \lambda_n(T \cup S_k) - \lambda_n(T \setminus S_k)$ converges to 0 uniformly in n as $k \to \infty$. Thus (λ_n) is uniformly countably additive. It follows easily that the limit λ is a measure. This completes the proof in the case where X is the scalar field.

We now turn to the general case. For each $u \in X^*$, we know that $u \circ \lambda$ is a measure, by the scalar case. Let us next show that λ itself is a measure: If (T_n) is any sequence of disjoint measurable sets, we know that $\sum_{n=1}^{\infty} \lambda(T_n)$ converges weakly to $\lambda\left(\bigcup_{n=1}^{\infty} T_n\right)$. Since the same type of conclusion holds when we replace (T_n) with a subsequence, we know that any subseries of $\sum_{n=1}^{\infty} \lambda(T_n)$ converges weakly. By the Orlicz-Pettis Theorem (28.31), it follows that $\sum_{n=1}^{\infty} \lambda(T_n)$ converges in X to a limit. That limit can only be $\lambda\left(\bigcup_{n=1}^{\infty} T_n\right)$, since X^* separates points of X. Thus λ is a measure.

Replacing (λ_n) with the sequence $(\lambda_n - \lambda)$, we may assume $\lambda = 0$. That is, $\lambda_n(S) \to 0$ for each $S \in \mathcal{S}$.

Suppose the sequence (λ_n) is not uniformly countably additive. Thus there exists a sequence $S_k \downarrow \varnothing$ in $\mathcal{S}$, which does not satisfy $\sup_n |\lambda_n(S_k)| \to 0$ as $k \to \infty$. That is, some constant $\varepsilon > 0$ satisfies $\limsup_{k\to\infty} \sup_{n\in\mathbb{N}} |\lambda_n(S_k)| > \varepsilon$. Replacing (S_k) with a subsequence, we may assume $\sup_n |\lambda_n(S_k)| > \varepsilon$ for all k. Thus, for each k there is some $n(k)$ satisfying $|\lambda_{n(k)}(S_k)| > \varepsilon$.

Since $S_k \downarrow \varnothing$, we also have $\max_{1\le n\le N} |\lambda_n(S_k)| \to 0$ as $k \to \infty$ for each fixed N. Thus, the sequence $(n(k))$ cannot take all its values in some finite set $\{1, 2, \ldots, N\}$. That is, the sequence $(n(k))$ is unbounded. Replacing (S_k) and $(n(k))$ with subsequences, we may assume $n(1) < n(2) < n(3) < \cdots$. Replacing (λ_n) with the subsequence $(\lambda_{n(k)})$, we may assume $|\lambda_k(S_k)| > \varepsilon$ for all k.

Choose some $u_k \in X^*$ satisfying $|u_k\lambda_k(S_k)| > \varepsilon$ and $|u_k| = 1$. For each fixed $S \in \mathcal{S}$, we have $|u_k\lambda_k(S)| \le |\lambda_k(S)| \to 0$. By the scalar case that we have already proved,

$$\Lambda(S) \quad = \quad (u_1\lambda_1(S),\ u_2\lambda_2(S),\ u_3\lambda_3(S),\ \ldots)$$

is a c-valued measure. Since $S_k \downarrow \varnothing$, it follows that $\Lambda(S_k) \to 0$. However, $|\Lambda(S_k)| \ge |u_k\lambda_k(S_k)| > \varepsilon$, a contradiction.

29.9. Corollary: Nikodym Boundedness Theorem. Let $(X, |\ \ |)$ be a Banach space, and let Λ be a collection of X-valued measures on a measurable space $(\Omega, \mathcal{S})$. If $\sup_{\lambda\in\Lambda} |\lambda(S)| < \infty$ for each $S \in \mathcal{S}$, then in fact $\sup_{\lambda\in\Lambda} \sup_{S\in\mathcal{S}} |\lambda(S)| < \infty$.

Proof. Suppose not. Then we may choose sequences (λ_n) in Λ and (S_n) in $\mathcal{S}$, with $|\lambda_n(S_n)| > n^2$. The measures $\gamma_n = \frac{1}{n}\lambda_n$ satisfy $|\gamma_n(S_n)| > n$ and $\lim_{n\to\infty} \gamma_n(S) = 0$ for each S. By the Nikodym Convergence Theorem (29.8), $\Gamma(S) = (\gamma_1(S), \gamma_2(S), \gamma_3(S), \ldots)$ defines a $c(X)$-valued measure. By 29.3, any Banach-space-valued measure is bounded; but that contradicts $|\Gamma(S_n)| > n$.

Remark. With a longer proof, a slightly weaker hypothesis suffices; see Diestel and Uhl [1977].

Indefinite Bochner Integrals and Radon-Nikodym Derivatives

29.10. Example: the Bochner integral as a vector measure. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space, let $(X, |\ |)$ be a Banach space, and let $h \in L^1(\mu, X)$. We shall show that the function $\lambda : \mathcal{S} \to X$ defined by the Bochner integral

$$\lambda(S) \quad = \quad \int_S h\, d\mu \quad = \quad \int_\Omega 1_S(\omega)h(\omega)d\mu(\omega)$$

is an X-valued measure on $\mathcal{S}$. Obviously it is μ-continuous (as defined in 29.4). It is sometimes called the **indefinite integral** of h. Also, we say that h is the **Radon-Nikodym**

derivative of λ with respect to μ. We say "the derivative" rather than "a derivative," because — as we shall show below — there is at most one $h \in L^1(\mu, X)$ satisfying this relation with measures λ and μ. Thus we may write $h = \dfrac{d\lambda}{d\mu}$.

Proofs. Let $h \in L^1(\mu, X)$. Obviously λ is a vector charge. To show that it is a vector measure (i.e., countably additive) and is absolutely continuous, we may reason as follows: The function $|h(\cdot)|$ is measurable and is a member of $L^1(\mu, \mathbb{R})$ (see 22.28). By 21.38(i), the function γ defined by $\gamma(S) = \int_S |h(\cdot)|\, d\mu$ is a finite positive measure. Since $|\lambda(S)| \leq \gamma(S)$ for every measurable set S, it is easy to see that the vector charge λ satisfies $\lambda << \gamma$ by criterion 29.4(A). Therefore $\lambda << \gamma$ by criterion 29.4(B), and so λ is a vector measure. We have $\lambda << \mu$ by criterion 29.4(B).

To show h is uniquely determined, suppose h_1, h_2 are both Radon-Nikodym derivatives of λ with respect to μ. Then $g = h_1 - h_2$ is a member of $L^1(\mu, X)$ that satisfies $\int_S g\, d\mu = 0$ for every $S \in \mathcal{S}$; we are to show that $g = 0$. Suppose that g is not the zero function. Altering g on a set of measure 0, we may assume g has separable range. Let X_0 be the closed span of the range of g; then X_0 is a separable closed subspace of X. We can cover $X_0 \setminus \{0\}$ with countably many closed balls B_j that do not contain 0. The set $S = g^{-1}(B_j)$ has positive measure for at least one j; fix that j. Since B_j is a closed convex set that does not contain 0, by the Hahn-Banach Theorem (HB20) in 28.4 there exists some functional $\varphi \in X^*$ that is positive everywhere on B_j. Then

$$0 \quad = \quad \varphi(0) \quad = \quad \varphi\left(\int_S g\, d\mu\right) \quad = \quad \int_S (\varphi \circ g) d\mu \quad > \quad 0,$$

a contradiction.

29.11. *Proposition.* Suppose $h = \dfrac{d\lambda}{d\mu}$, as above. Then λ has bounded variation. In fact,

$$/\lambda/(S) \; = \; \int_S |h(\cdot)| d\mu, \qquad \text{and in particular} \qquad /\lambda/(\Omega) \; = \; \|h\|_1.$$

Proof. It will be helpful to denote λ instead by λ_h, to display its dependence on h. The inequality $/\lambda_h/(S) \leq \int_S |h(\cdot)|\, d\mu$ follows easily from the definition of the variation. Therefore the mapping $h \mapsto \lambda_h$ is a nonexpansive linear map from $L^1(\mu, X)$ into the Banach space of X-valued functions of bounded variation, normed by the variation as described in 29.6.c. Since the map $h \mapsto \lambda_h$ is continuous, it suffices to prove the equation $/\lambda_h/(S) = \int_S |h(\cdot)|\, d\mu$ for all h in some dense subset of $L^1(\mu, X)$. The integrable simple functions are dense in $L^1(\mu, X)$, and the proof is easy for such functions.

Remarks. A converse question is this: If $\lambda << \mu$ and λ has bounded variation, when does $d\lambda/d\mu$ exist? That question is addressed in 29.20 through 29.26.

29.12. *Some further properties of the Radon-Nikodym derivative.*

a. **Change of Variables Formula.** If $g = d\mu/d\nu$ where μ, ν are positive, finite measures, then $\int_S h\, d\mu = \int_S hg\, d\nu$ for any function $h \in L^1(\mu, X)$ and any measurable set S. (Compare with 25.17.)

Hints: Show this first when h is a simple function. Then prove this for any measurable function $h : \Omega \to [0, +\infty)$, by taking limits of simple functions and using the Monotone Convergence Theorem. Finally, prove it for arbitrary $f \in L^1(\mu, X)$ by taking limits of simple functions and using the Dominated Convergence Theorem; the dominating functions take their values in $[0, +\infty)$.

b. Chain Rule for Vector Measures. From the Change of Variables Formula it follows that

$$\frac{d\lambda}{d\nu} = \frac{d\lambda}{d\mu}\frac{d\mu}{d\nu}$$

whenever the right side exists. More precisely, if $h = d\lambda/d\mu$ and $g = d\mu/d\nu$, where μ, ν are positive finite measures and λ is a vector measure, then the Radon-Nikodym derivative $d\lambda/d\nu$ also exists; it is equal to hg. (Compare with 25.6.)

Conditional Expectations and Martingales

29.13. *Notations/assumptions.* Throughout this subchapter, we assume $(X, |\ \ |)$ is a Banach space and $(\Omega, \mathcal{S}, \mu)$ is a probability space — i.e., a set equipped with a σ-algebra of subsets and a probability measure. We consider various sub-σ-algebras $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots \subseteq \mathcal{S}$. The restriction of μ to those sub-σ-algebras will be denoted $\mu_{\mathcal{A}}, \mu_{\mathcal{B}}, \mu_{\mathcal{C}}$, etc. Thus $L^p(\mu_{\mathcal{A}}, X)$ consists of those members of $L^p(\mu, X)$ that are equivalence classes of functions measurable from $(\Omega, \mathcal{A})$ to the Borel subsets of X. Note that $L^p(\mu_{\mathcal{A}}, X)$ is a closed linear subspace of $L^p(\mu, X)$; this follows from 21.3 and 22.31(ii).

Conditional expectations and martingales are used extensively in probability theory; this book will use them to prove Theorem 29.26.

29.14. *Proposition and definition.* Let $f \in L^1(\mu, X)$, and let $\mathcal{A} \subseteq \mathcal{S}$ be a sub-σ-algebra. Then there exists a unique (up to μ-equivalence) function $g \in L^1(\mu_{\mathcal{A}}, X)$ with this property: $\int_A g\, d\mu = \int_A f\, d\mu$ for every $A \in \mathcal{A}$. Such a function will be denoted by $E(f|\mathcal{A})$; it is called the **conditional expectation** of f with respect to $\mathcal{A}$. In this fashion we define the conditional expectation operator

$$E(\cdot|\mathcal{A}) \quad : \quad L^1(\mu, X) \to L^1(\mu_{\mathcal{A}}, X).$$

It has these further properties:

(i) It is linear.

(ii) It is nonexpansive from $L^1(\mu, X)$ to $L^1(\mu_{\mathcal{A}}, X)$ — that is, $\|E(f|\mathcal{A})\|_1 \leq \|f\|_1$.

(iii) It is idempotent — that is, $E(\cdot|\mathcal{A}) \circ E(\cdot|\mathcal{A}) = E(\cdot|\mathcal{A})$.

(iv) Let $\mathbb{F}$ be the scalar field. If $f \in L^2(\mu, \mathbb{F})$, then $E(f|\mathcal{A})$ is the closest vector to f in the closed linear subspace $L^2(\mu_{\mathcal{A}}, \mathbb{F})$.

Proof. We first prove uniqueness. Suppose that $g_1, g_2 \in L^1(\mu_{\mathcal{A}}, X)$ satisfy $\int_A g_1\, d\mu = \int_A g_2\, d\mu$ for every $A \in \mathcal{A}$. Then $h = g_1 - g_2$ is a member of $L^1(\mu_{\mathcal{A}}, X)$ that satisfies $\int_A h\, d\mu = 0$ for every $A \in \mathcal{A}$. It suffices to show that $h = 0$ almost everywhere. Suppose, on the contrary, that $\{\omega \in \Omega : h(\omega) \neq 0\}$ has positive measure. Since the range of h is separable, its points are separated by the functions $\operatorname{Re}\psi_n$ for some sequence (ψ_n) in X^* — see 23.24. Then $\{\omega \in \Omega : \operatorname{Re}\psi_n h(\omega) \neq 0\}$ has positive measure for at least one n. Replacing ψ_n with $-\psi_n$ if necessary, we may assume that the set $A = \{\omega \in \Omega : \operatorname{Re}\psi_n h(\omega) > 0\}$ has positive measure. But $\int_A \operatorname{Re}\psi_n h\, d\mu = \operatorname{Re}\psi_n \int_A h\, d\mu = 0$, a contradiction. This proves uniqueness.

Let us define a linear map $E(\cdot|\mathcal{A})$ from *some* linear subspace of $L^1(\mu, X)$ into $L^1(\mu_{\mathcal{A}}, X)$, by writing $g = E(f|\mathcal{A})$ whenever $g \in L^1(\mu_{\mathcal{A}}, X)$ satisfies $\int_A g\, d\mu = \int_A f\, d\mu$ for all $A \in \mathcal{A}$. (The fact that the operator's domain is all of $L^1(\mu, X)$ will not be established until the end of this proof.) Clearly, $E(f|\mathcal{A})$ exists and equals f whenever $f \in L^1(\mu_{\mathcal{A}}, X)$; thus the domain of the conditional expectation operator contains $L^1(\mu_{\mathcal{A}}, X)$ and the operator is idempotent.

We next show that $E(\cdot|\mathcal{A})$ is nonexpansive. Say $g = E(f|\mathcal{A})$. Define measures $\lambda_{\mathcal{S}}, \lambda_{\mathcal{A}}$ on $\mathcal{S}, \mathcal{A}$, respectively, by

$$\lambda_{\mathcal{S}}(S) = \int_S f\, d\mu \text{ for } S \in \mathcal{S}, \qquad \lambda_{\mathcal{A}}(A) = \int_A g\, d\mu \text{ for } A \in \mathcal{A}.$$

Then $\|f\|_1 = \operatorname{Var}(\lambda_{\mathcal{S}}) = /\lambda_{\mathcal{S}}/(\Omega)$ and $\|g\|_1 = \operatorname{Var}(\lambda_{\mathcal{A}}) = /\lambda_{\mathcal{A}}/(\Omega)$, by 29.11. However, $\lambda_{\mathcal{A}}$ is just the restriction of $\lambda_{\mathcal{S}}$ to the smaller σ-algebra $\mathcal{A}$, so $\operatorname{Var}(\lambda_{\mathcal{A}}) \leq \operatorname{Var}(\lambda_{\mathcal{S}})$; thus $\|E(f|\mathcal{A})\|_1 \leq \|f\|_1$.

Next we consider $f \in L^2(\mu, \mathbb{F})$. Let g be the closest point to f in the closed linear subspace $L^2(\mu_{\mathcal{A}}, \mathbb{F})$. Then $g - f$ is orthogonal to $L^2(\mu_{\mathcal{A}}, \mathbb{F})$, as we noted in 22.51. That is, $\int_\Omega (g - f)h\, d\mu = 0$ for every $h \in L^2(\mu_{\mathcal{A}}, \mathbb{F})$. In particular, taking $h = 1_A$ shows that $\int_A g\, d\mu = \int_A f\, d\mu$ for every $A \in \mathcal{A}$. (We can take $h = 1_A$ because μ is a probability measure, hence $1_A \in L^2(\mu, \mathbb{F}) \subseteq L^1(\mu, \mathbb{F})$.) Thus, for scalar-valued functions, the domain of $E(\cdot|\mathcal{A})$ contains $L^2(\mu, \mathbb{F})$; note that that linear space is dense in $L^1(\mu, \mathbb{F})$. Since $E(\cdot|\mathcal{A}) : L^2(\mu, \mathbb{F}) \to L^1(\mu_{\mathcal{A}}, \mathbb{F})$ is nonexpansive for the $\|\ \ \|_1$ norms (as noted in the preceding paragraph), the operator extends uniquely to a nonexpansive linear operator $\Gamma : L^1(\mu, \mathbb{F}) \to L^1(\mu_{\mathcal{A}}, \mathbb{F})$ (which, however, we have not yet established to be a conditional expectation operator). For fixed $A \in \mathcal{A}$, the mappings $f \mapsto \int_A f\, d\mu$ and $f \mapsto \int_A \Gamma f\, d\mu$ are continuous linear maps from $L^1(\mu, \mathbb{F})$ into $\mathbb{F}$, and they agree on the dense set $L^2(\mu, \mathbb{F})$, hence they agree on $L^1(\mu, \mathbb{F})$. This proves Γ is indeed a conditional expectation operator, defined everywhere on $L^1(\mu, \mathbb{F})$.

For $f \in L^1(\mu, X)$ with a general Banach space X, we construct $E(f|\mathcal{A})$ first in the case where f is a simple function. Say $f = \sum_{j=1}^n 1_{S_j}(\cdot)x_j$ where the $1_{S_j}(\cdot)$'s are characteristic functions of disjoint sets $S_j \in \mathcal{S}$, and the x_j's are members of X. Define $E(f|\mathcal{A}) = \sum_{j=1}^n E\left(1_{S_j}(\cdot)|\mathcal{A}\right) x_j$. It is easy to verify that this defines a linear, nonexpansive mapping $E(\cdot|\mathcal{A})$ from the simple functions in $L^1(\mu, X)$ into $L^1(\mu_{\mathcal{A}}, X)$, satisfying $\int_A E(f|\mathcal{A})\, d\mu = \int_A f\, d\mu$ for all $A \in \mathcal{A}$. Those properties are preserved when we take limits, and the simple functions are dense in $L^1(\mu, X)$; thus the conditional expectation operator extends to a mapping with those properties on all of $L^1(\mu, X)$.

29.15. *Corollaries.*

a. $\int_K |E(f|\mathcal{A})(\omega)|d\mu(\omega) \leq \int_K |f(\omega)|d\mu(\omega)$ for any $K \in \mathcal{A}$. That is, the conditional expectation operator is nonexpansive from $L^1(\mu_{\mathcal{S}\cap K}, X)$ to $L^1(\mu_{\mathcal{A}\cap K}, X)$, where $\mathcal{S}\cap K$ and $\mathcal{A} \cap K$ denote the traces of the σ-algebras $\mathcal{S}$ and $\mathcal{A}$ on the set K.

Proof. Replace μ with its restriction to K, and apply the preceding results.

b. If $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{S}$, then $E\big(E(f|\mathcal{B})|\mathcal{A}\big) = E(f|\mathcal{A})$.

c. *Example.* Let $S_1, S_2, S_3, \ldots$ be disjoint members of $\mathcal{S}$ with union equal to Ω. Let $\mathcal{A}$ be the σ-algebra generated by the S_j's; thus $\mathcal{A} = \{$unions of S_j's$\}$. Verify that $g = E(f|\mathcal{A})$ can be represented as follows:

$$g(\omega) \quad = \quad \begin{cases} \dfrac{\int_{S_j} f\,d\mu}{\mu(S_j)} & \text{if } \omega \in S_j \text{ and } \mu(S_j) > 0 \\[2ex] 0 & \text{if } \omega \in S_j \text{ and } \mu(S_j) = 0. \end{cases}$$

29.16. *Definition.* Let Γ be a collection of sub-σ-algebras of $\mathcal{S}$ that is directed by inclusion — i.e., assume that for any $\mathcal{A}, \mathcal{B} \in \Gamma$ there exists some $\mathcal{C} \in \Gamma$ with $\mathcal{C} \supseteq \mathcal{A} \cup \mathcal{B}$. An X-valued **martingale** indexed by Γ will mean a net $(g_{\mathcal{A}} : \mathcal{A} \in \Gamma)$ in $L^1(\mu, X)$ satisfying

$$g_{\mathcal{A}} \;=\; E\left(g_{\mathcal{B}}|\mathcal{A}\right) \qquad \text{whenever} \quad \mathcal{A} \subseteq \mathcal{B}$$

— that is, satisfying

$$\int_A g_{\mathcal{A}}\,d\mu \;=\; \int_A g_{\mathcal{B}}\,d\mu \qquad \text{whenever} \quad \mathcal{A} \subseteq \mathcal{B} \text{ and } A \in \mathcal{A}.$$

An important special case is that in which Γ consists of an increasing sequence of sub-σ-algebras $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_3 \subseteq \cdots \subseteq \mathcal{S}$, with conditional expectation operators $E_j = E(\cdot|\mathcal{A}_j)$. Then the corresponding functions form a sequence $g_1, g_2, g_3, \ldots$ that satisfies $g_j = E_j(g_k)$ whenever $j \leq k$ — that is, $\int_A g_j\,d\mu = \int_A g_k\,d\mu$ whenever $j \leq k$ and $A \in \mathcal{A}_j$. We may refer to such martingales as **sequential martingales**.

Two examples of methods for constructing martingales are given in 29.17 and 29.24.

29.17. Mean Convergence Theorem for martingales. Let $(g_{\mathcal{A}} : \mathcal{A} \in \Gamma)$ be a net in $L^1(\mu, X)$ where Γ is a collection of sub-σ-algebras directed by inclusion. Then these two conditions are equivalent:

(A) $(g_{\mathcal{A}} : \mathcal{A} \in \Gamma)$ is a martingale that converges in $L^1(\mu, X)$ to some limit g_∞.

(B) There is some function $g \in L^1(\mu, X)$ such that $g_{\mathcal{A}} = E(g|\mathcal{A})$ for each $\mathcal{A} \in \Gamma$.

Moreover, if those two conditions are satisfied, then

$$g_\infty = E(g|\mathcal{S}_\infty) \tag{$*$}$$

where $\mathcal{S}_\infty$ is the σ-algebra generated by $\bigcup_{\mathcal{A}\in\Gamma} \mathcal{A}$. (*Remark.* The function g_∞ is determined uniquely almost everywhere by these conditions, but the function g might not be.)

Proof of (A) $\Rightarrow$ (B). Fix any $\mathcal{A} \in \Gamma$. Since $E(\cdot|\mathcal{A}) : L^1(\mu, X) \to L^1(\mu_{\mathcal{A}}, X)$ is a nonexpansive mapping and $0 = \lim_{\mathcal{B}} \|g_{\mathcal{B}} - g_\infty\|_1$, it follows that $0 = \lim_{\mathcal{B}} \|E(g_{\mathcal{B}}|\mathcal{A}) - E(g_\infty|\mathcal{A})\|_1$.

For $\mathcal{B}$ sufficiently large, we have $\mathcal{B} \supseteq \mathcal{A}$, hence $E(g_{\mathcal{B}}|\mathcal{A}) = g_{\mathcal{A}}$; thus $0 = \lim_{\mathcal{B}} \|g_{\mathcal{A}} - E(g_\infty|\mathcal{A})\|_1$. But $\|g_{\mathcal{A}} - E(g|\mathcal{A})\|_1$ does not depend on $\mathcal{B}$, so $g_{\mathcal{A}} = E(g_\infty|\mathcal{A})$.

Proof of (B) $\Rightarrow$ (A) and (*). Let $g_{\mathcal{A}} = E(g|\mathcal{A})$. That $(g_{\mathcal{A}} : \mathcal{A} \in \Gamma)$ is a martingale follows immediately from 29.15.b.

Let $g_\infty = E(g|\mathcal{S}_\infty)$; it remains to show that $0 = \lim_{\mathcal{A}} \|g_{\mathcal{A}} - g_\infty\|_1$. Note that if $B \in \mathcal{B}$ then $B \in \mathcal{S}_\infty$, hence

$$\int_B g_\infty \, d\mu \quad = \quad \int_B g \, d\mu \quad = \quad \int_B g_{\mathcal{B}} \, d\mu;$$

this proves $E(g_\infty|\mathcal{B}) = g_{\mathcal{B}}$. The function g_∞ is $\mathcal{S}_\infty$-measurable, hence by 22.30 we can approximate g_∞ arbitrarily closely in $L^1(\mu, X)$ by an $\mathcal{S}_\infty$-measurable simple function. The algebra of sets $\mathcal{U} = \bigcup_{\mathcal{A}\in\Gamma} \mathcal{A}$ generates the σ-algebra $\mathcal{S}_\infty$, so by 21.26 each member of $\mathcal{S}_\infty$ can be approximated arbitrarily closely in measure by some member of $\mathcal{U}$. Combining these two results, we can approximate g_∞ arbitrarily closely in $L^1(\mu, X)$ by a simple function of the form $h = \sum_{j=1}^n 1_{U_j}(\cdot)x_j$ where the U_j's belong to $\mathcal{U}$; here 1_{U_j} is the characteristic function of U_j. Thus we can satisfy $\|g_\infty - h\|_1 < \varepsilon$ for any given ε. Temporarily hold ε fixed; then we may fix n, h, and some particular $\mathcal{A}$ that contains all of $U_1, U_2, \ldots, U_n$. For all $\mathcal{B} \in \Gamma$ sufficiently large, we have $\mathcal{B} \supseteq \mathcal{A}$, and therefore $E(h|\mathcal{B}) = h$. Since $E(\cdot|\mathcal{B})$ is nonexpansive, we obtain

$$\begin{aligned} \|g_{\mathcal{B}} - g_\infty\|_1 \quad &\le \quad \|g_{\mathcal{B}} - h\|_1 + \|g_\infty - h\|_1 \\ &= \quad \|E(g_\infty|\mathcal{B}) - E(h|\mathcal{B})\|_1 + \|g_\infty - h\|_1 \quad \le \quad 2\|g_\infty - h\|_1 \quad < \quad 2\varepsilon. \end{aligned}$$

Thus $\limsup_{\mathcal{B}} \|g_{\mathcal{B}} - g_\infty\|_1 \le 2\varepsilon$. Now let $\varepsilon \downarrow 0$; this proves $\lim_{\mathcal{B}} \|g_{\mathcal{B}} - g_\infty\|_1 = 0$.

29.18. Maximal lemma for martingales. Let (g_n) be a sequential martingale, with σ-algebras $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_3 \subseteq \cdots \subseteq \mathcal{S}$. Let some number $\varepsilon > 0$ be given. Then

$$\mu\left(\left\{\omega \in \Omega \; : \; \sup_{n\in\mathbb{N}} |g_n(\omega)| > \varepsilon\right\}\right) \quad \le \quad \frac{1}{\varepsilon} \sup_{n\in\mathbb{N}} \int_\Omega |g_n(\cdot)| d\mu.$$

(Compare with 24.43.)

Proof. Let $B = \{\omega \in \Omega \; : \; \sup_n |g_n(\omega)| > \varepsilon\}$. A point ω belongs to B if and only if $|g_n(\omega)| > \varepsilon$ for some $n \in \mathbb{N}$. We shall classify the ω's in B by considering which is the *first* value of n satisfying this condition. In other words, let $B_1 = \{\omega \in \Omega : \varepsilon < |g_1(\omega)|\}$, and for $n > 1$ let

$$B_n \quad = \quad \left\{\omega \in \Omega \; : \; \max_{1\le j<n} |g_j(\omega)| \le \varepsilon < |g_n(\omega)|\right\}.$$

Then the B_n's form a partition of B, so $\mu(B) = \sum_{n=1}^\infty \mu(B_n)$. Since g_n is $\mathcal{A}_n$-measurable and the $\mathcal{A}_n$'s are increasing, it follows that $B_n \in \mathcal{A}_n$. Observe that

$$\mu(B_n) \; = \; \int_{B_n} 1 \, d\mu \; \le \; \int_{B_n} \frac{|g_n(\omega)|}{\varepsilon} \, d\mu(\omega) \; \le \; \int_{B_n} \frac{|g_m(\omega)|}{\varepsilon} \, d\mu(\omega)$$

for any $m \geq n$; the last inequality follows from 29.15.a since $g_n = E(g_m|\mathcal{A}_n)$. Therefore, for any integer $m \geq 1$ we have

$$\sum_{n=1}^{m} \mu(B_n) \;\leq\; \sum_{n=1}^{m} \int_{B_n} \frac{|g_m(\omega)|}{\varepsilon}\, d\mu(\omega) \;\leq\; \frac{1}{\varepsilon} \sup_{n\in\mathbb{N}} \int_\Omega |g_n(\cdot)|\, d\mu.$$

Finally, take limits as $m \to \infty$.

29.19. Pointwise Convergence Theorem. Let $((f_n, \mathcal{A}_n) : n = 1, 2, 3, \ldots)$ be a sequential X-valued martingale, which converges in $L^1(\mu, X)$ to a limit f. Then also $f_n \to f$ pointwise μ-almost everywhere.

Proof. Say the conditional expectation operators are $E_n = E(\cdot|\mathcal{A}_n)$; thus the functions are $f_n = E_n(f)$. Let any numbers $\delta, \varepsilon > 0$ be given. By 22.30 and 21.26, we have $\|f-g\|_1 < \delta\varepsilon/2$ for some simple function g that is $\mathcal{A}_k$-measurable for some positive integer k. Then $E_n g = g$ for all $n \geq k$. For all $m, n \geq k$, we have

$$\begin{aligned} |f_n - f_m| \;&=\; |E_n f - E_m g| \\ &\leq\; |E_n g - E_m g| + |E_n(f-g) - E_m(f-g)| \;\leq\; 2\sup_{j\geq 1} |E_j(f-g)| \;\stackrel{\text{def}}{=}\; h. \end{aligned}$$

Taking limits, we have $\limsup_{m,n\to\infty} |f_n(\omega) - f_m(\omega)| \leq h(\omega)$ for each ω. Therefore

$$\begin{aligned} \mu\left(\left\{\omega \in \Omega \,:\, \limsup_{m,n\to\infty} |f_n(\omega) - f_m(\omega)| > \varepsilon\right\}\right) \;&\leq\; \mu(\{\omega \in \Omega \,:\, h(\omega) > \varepsilon\}) \\ =\; \mu\left(\left\{\omega \in \Omega \,:\, \sup_{j\geq 1} |E_j(f(\omega) - g(\omega))| > \frac{\varepsilon}{2}\right\}\right) \;&\leq\; \frac{2}{\varepsilon}\|f-g\|_1 \;<\; \delta, \end{aligned}$$

where the next-to-last inequality follows from Lemma 29.18 with $p = f - g$. Letting $\delta \downarrow 0$ shows that $\mu\left(\left\{\omega \in \Omega \,:\, \limsup_{m,n\to\infty} |f_n(\omega) - f_m(\omega)| > \varepsilon\right\}\right) = 0$. Since ε is arbitrary, this shows that $\limsup_{m,n\to\infty} |f_n(\omega) - f_m(\omega)| = 0$ almost everywhere, and thus $\lim_{n\to\infty} f_n(\omega)$ exists almost everywhere. We established earlier in this proof that f_n converges to f in $L^1(\mu, X)$; hence $f_n \to f$ pointwise almost everywhere.

EXISTENCE OF RADON-NIKODYM DERIVATIVES

29.20. Classical Radon-Nikodym Theorem. Let $(\Omega, \mathcal{S}, \mu)$ be a finite measure space. Let λ be a scalar-valued measure on $\mathcal{S}$. Assume that λ is μ-continuous (as defined in 29.4). Then there exists a Radon-Nikodym derivative $h = d\lambda/d\mu$, as defined in 29.10.

Remark. An analogous result can be proved for σ-finite measures μ, but to keep our definitions simple we shall only consider finite measures μ.

Proof of theorem. Any complex measure can be decomposed into its real and imaginary components; for any real measure λ we have the Jordan Decomposition $\lambda = \lambda^+ - \lambda^-$. Thus, it suffices to consider the case of $\lambda \geq 0$. By assumption, λ is scalar-valued, so $\lambda(\Omega) < \infty$.

We prove existence of h first under the additional assumption that $\lambda(S) \leq \mu(S)$ for all $S \in \mathcal{S}$. In that case, it is clear that $\int_\Omega |f(\cdot)| d\lambda \leq \int_\Omega |f(\cdot)| d\mu$ for every measurable function f. Thus we have the inclusion $\iota : \mathcal{L}^1(\mu) \xrightarrow{\subseteq} \mathcal{L}^1(\lambda)$, and in fact that inclusion is a continuous linear map from one seminormed space into the other, with operator norm ≤ 1.

Note that if two measurable functions f, g are μ-equivalent, they are also λ-equivalent, since λ is μ-continuous. Therefore the inclusion $\iota : \mathcal{L}^1(\mu) \xrightarrow{\subseteq} \mathcal{L}^1(\lambda)$ determines a continuous linear mapping $\widehat{\iota} : L^1(\mu) \to L^1(\lambda)$ with norm ≤ 1. (We do not assert that this is an injective map — after all, two functions which are not μ-equivalent may possibly be λ-equivalent.)

The integral mapping $I : f \mapsto \int_\Omega f\, d\lambda$ is a continuous linear map from $L^1(\lambda)$ into the scalar field $\mathbb{F}$. The composition $I \circ \widehat{\iota} : L^1(\mu) \to L^1(\lambda) \to \mathbb{F}$ is a continuous linear map. By 28.51, we know that that linear map is given by some $h \in L^1(\mu)^* = L^\infty(\mu)$. Thus, there is some function $h \in L^\infty(\mu)$, determined uniquely up to μ-equivalence, such that

$$\int_\Omega f\, d\lambda \quad = \quad \int_\Omega fh\, d\mu \tag{1}$$

for every $f \in L^1(\mu)$. Since μ and λ are nonnegative, it follows easily that h is nonnegative. By the Monotone Convergence Theorem, (1) holds for any positive measurable function f, whether it is integrable or not.

In particular, when f is the characteristic function of a set $S \in \mathcal{S}$, we find $\lambda(S) = \int_S h\, d\mu$. In particular, $\int_\Omega h\, d\mu = \lambda(\Omega) < \infty$, so $h \in L^1(\mu)$. This completes the proof of existence of the Radon-Nikodym derivative $d\lambda/d\mu$ in the case where $\lambda \leq \mu$.

We now remove the assumption that $\lambda(S) \leq \mu(S)$ for all S. (We continue to assume that $\lambda \geq 0$. The reduction used here is taken from Bradley [1989].) Let $\pi = \lambda + \mu$. Then λ and μ are scalar-valued, π-continuous measures, and $\lambda, \mu \leq \pi$. By the preceding arguments, there exist Radon-Nikodym derivatives $d\lambda/d\pi$ and $d\mu/d\pi$; these are members of $L^1(\pi)$. Note that $d\mu/d\pi$ satisfies the condition analogous to (1); thus,

$$\int_\Omega f\, d\mu \quad = \quad \int_\Omega f\, \frac{d\mu}{d\pi}\, d\pi \tag{2}$$

for every nonnegative measurable f. Now define the set $\Omega_0 = \left\{\omega \in \Omega : \dfrac{d\mu}{d\pi}(\omega) = 0\right\}$ and the nonnegative measurable function

$$h(\omega) \quad = \quad \begin{cases} \dfrac{\dfrac{d\lambda}{d\pi}(\omega)}{\dfrac{d\mu}{d\pi}(\omega)} & \text{when } \omega \in \Omega \setminus \Omega_0 \\[2ex] 0 & \text{when } \omega \in \Omega_0. \end{cases}$$

Then $\mu(\Omega_0) = \int_{\Omega_0} \frac{d\mu}{d\pi}\, d\pi = 0$. Since λ is μ-continuous, we have also $\lambda(\Omega_0) = 0$. Apply (2) with $f(\cdot) = 1_S(\cdot) h(\cdot)$. By the definitions of $d\lambda/d\pi$, $d\mu/d\pi$, and h, for any measurable set $S \subseteq \Omega \setminus \Omega_0$ we have

$$\lambda(S) \;=\; \int_S \frac{d\lambda}{d\pi}\, d\pi \;=\; \int_S h\, \frac{d\mu}{d\pi}\, d\pi \;=\; \int_S h\, d\mu.$$

Since λ and $\int_{(\cdot)} h\,d\mu$ both vanish on subsets of Ω_0, we have $\lambda(S) = \int_S h\,d\mu$ for all $S \in \mathcal{S}$. This completes the proof in the case of $\mu(\Omega) < \infty$.

29.21. *Definition.* We say that a Banach space $(X, |\ \ |)$ has the **Radon-Nikodym Property**, or **RNP**, if this condition is satisfied:

(A) (RNP with respect to arbitrary measure.) Suppose $(\Omega, \mathcal{S}, \mu)$ is a finite measure space and $\lambda : \mathcal{S} \to X$ is a μ-continuous measure with bounded variation. Then there exists some $g \in L^1(\mu; X)$ such that $\lambda(S) = \int_S g\,d\mu$ for all $S \in \mathcal{S}$.

Summary of results. We saw in 29.20 that the scalar field has the RNP. It follows easily that any finite-dimensional Banach space has the RNP. More generally, we shall prove in 29.26 that every reflexive Banach space has the RNP. We shall see by simple examples that some nonreflexive Banach spaces have the RNP (see 29.22) and some do not (see 29.23).

Other characterizations (proofs omitted). Many other conditions are equivalent to the Radon-Nikodym Property. Here are a few of them:

(B) (RNP with respect to Lebesgue measure.) Let $(\Omega, \mathcal{S}, \mu)$ be the unit interval equipped with Lebesgue measurable sets and Lebesgue measure. Suppose that $\lambda : \mathcal{S} \to X$ is a μ-continuous measure with bounded variation. Then there exists some $g \in L^1(\mu; X)$ such that $\lambda(S) = \int_S g\,d\mu$ for all $S \in \mathcal{S}$.

(C) (Riesz Representation Property.) Suppose $(\Omega, \mathcal{S}, \mu)$ is a finite measure space and $T : L^1(\mu) \to X$ is a continuous linear operator. Then there exists some $g \in L^\infty(\mu; X)$ such that $T(f) = \int_\Omega fg\,d\mu$ for all $f \in L^1(\mu)$.

(D) (Huff and Morris Property.) If D is a nonempty closed bounded subset of X, then some continuous real-linear functional on X assumes a maximum value on D.

Many more formulations and the proofs of equivalence can be found in Diestel and Uhl [1977]. That book also includes this interesting result:

> Let X be a Banach space, let $(\Omega, \mathcal{S}, \mu)$ be a finite measure space, and let $p \in [1, +\infty)$ and $q \in (1, +\infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then the dual of $L^p(\mu, X)$ is $L^q(\mu, X^*)$ if and only if the dual space X^* has the Radon-Nikodym Property.

Thus, $(L^p(\mu, X))^* = L^q(\mu, X^*)$ is true for "nice" Banach spaces X, but not more generally.

29.22. *Example.* The space ℓ_1 has the RNP.

Proof. Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$). Let $(\Omega, \mathcal{S}, \mu)$ be a finite measure space and let $\lambda : \mathcal{S} \to \ell_1$ be a μ-continuous vector measure with bounded variation. We are to exhibit a function $g \in L^1(\mu, \ell_1)$ satisfying $\lambda(S) = \int_S g\,d\mu$ for all $S \in \mathcal{S}$.

We may write $\lambda(S) = \Big(\lambda_1(S), \lambda_2(S), \lambda_3(S), \ldots\Big)$; then each $\lambda_j : \mathcal{S} \to \mathbb{F}$ is a μ-continuous, scalar-valued measure with bounded variation. It follows easily from the definition of the norm of ℓ_1 and the definition of the variation of a charge, that

$$/\lambda/(S) \quad = \quad /\lambda_1/(S) \; + \; /\lambda_2/(S) \; + \; /\lambda_3/(S) \; + \; \cdots .$$

(*Hint*: To prove that $\sum_{j=1}^{\infty} /\lambda_j/(S) \leq /\lambda/(S)$, it suffices to show that $\sum_{j=1}^{N} /\lambda_j/(S) \leq /\lambda/(S)$ for any positive integer N.)

Since the scalar field has the RNP, for each j we have $\lambda_j(S) = \int_S g_j \, d\mu$ for some $g_j \in L^1(\mu; \mathbb{F})$. Define $g(\omega) = \Big(g_1(\omega),\, g_2(\omega),\, g_3(\omega),\, \ldots\Big)$. Observe that

$$\int_\Omega \left(\sum_{j=1}^{\infty} |g_j(\omega)|\right) d\mu(\omega) \quad = \quad \sum_{j=1}^{\infty} \int_\Omega |g_j(\omega)| d\mu(\omega) \quad = \quad \sum_{j=1}^{\infty} /\lambda_j/(\Omega) \quad = \quad /\lambda/(\Omega),$$

which is finite; this proves that $g \in L^1(\mu, \ell_1)$.

Define a "truncation map" $T_N : \ell_1 \to \ell_1$ by

$$T_N(x_1, x_2, x_3, \ldots) \quad = \quad (x_1, x_2, \ldots, x_{N-1}, x_N, 0, 0, 0, \ldots).$$

Then for any $x \in \ell_1$, we have $\lim_{N\to\infty} T_N(x) = x$ — or, more precisely,

$$\lim_{N\to\infty} \|x - T_N(x)\|_1 \quad = \quad 0\,;$$

this follows from the definition of ℓ_1 and its norm. Now, for any $S \in \mathcal{S}$ and $N \in \mathbb{N}$ it is easy to verify that $T_N\Big(\lambda(S)\Big) = T_N\left(\int_S g \, d\mu\right)$. Taking limits yields $\lambda(S) = \int_S g \, d\mu$.

29.23. *Example.* The space c_0 lacks the RNP.

Let $(\Omega, \mathcal{S}, \mu)$ be the interval $[0, 2\pi]$ equipped with Lebesgue measure. For $S \in \mathcal{S}$, let $\lambda(S)$ be the sequence whose nth term is

$$\lambda_n(S) \quad = \quad \int_0^{2\pi} 1_S(t) \sin(nt) d\mu(t) \qquad\qquad (n = 1, 2, 3, \ldots).$$

By the Riemann-Lebesgue Lemma (24.41.b), $\lambda_n(S) \to 0$ as $n \to \infty$. Thus λ is a map from $\mathcal{S}$ into the Banach space $c_0 = \{$sequences of real numbers converging to $0\}$, which we equip with the sup norm as usual. Obviously λ is finitely additive — i.e., a vector charge. Also, $\|\lambda(S)\| \leq \mu(S)$, so λ is a μ-continuous vector measure with bounded variation. However, we shall show that there does not exist an integrable function $g : [0, 2\pi] \to c_0$ with the property that

$$\lambda(S) \quad = \quad \int_S g(t) d\mu(t) \qquad\qquad \text{for every } S \in \mathcal{S}.$$

Indeed, suppose there were such a function. Then $g(t)$ is a member of c_0 — i.e., a sequence $(g_1(t), g_2(t), g_3(t), \ldots)$. Applying the nth coordinate projection to the equation above, we obtain

$$\int_S \sin(nt) d\mu(t) \quad = \quad \int_S g_n(t) d\mu(t) \qquad\qquad \text{for every } S \in \mathcal{S}$$

and therefore $g_n(t) = \sin(nt)$ for almost every t. However, we shall show that the function $g(t) = (\sin(t), \sin(2t), \sin(3t), \ldots)$ defined in this fashion generally does not take values in c_0 — in fact, we shall show that $g(t) \in c_0$ only for t in a set of measure 0. Fix any small number $\varepsilon > 0$, and let $E_n = \{t \in [0, 2\pi] : |g_n(t)| \geq \varepsilon\}$. It is easy to show that $\mu(E_n) = 2\pi - 4\arcsin(\varepsilon)$ for each n. Define limsups as in 7.48 and 21.25.c; then

$$\{t : g(t) \notin c_0\} \quad \supseteq \quad \{t : |g_n(t)| \geq \varepsilon \text{ for infinitely many } n\} \quad = \quad \limsup_{n\to\infty} E_n.$$

Hence

$$\begin{aligned}\mu(\{t \in [0, 2\pi] : g(t) \notin c_0\}) \quad &\geq \quad \mu(\limsup_{n\to\infty} E_n) \\ &\geq \quad \limsup_{n\to\infty} \mu(E_n) \quad = \quad 2\pi - 4\arcsin(\varepsilon).\end{aligned}$$

Now take limits as $\varepsilon \downarrow 0$.

29.24. *Definition/example.* The rather complicated definitions in this section and the related lemma in the next section are in preparation for the major theorem given in 29.26; some readers may find it helpful to glance ahead to that result.

The collection $\mathfrak{F} = \{$finite subalgebras of $\mathcal{S}\}$ is directed by inclusion. Let μ be a probability measure on $(\Omega, \mathcal{S})$, let λ be an X-valued measure on $(\Omega, \mathcal{S})$, and suppose that λ is μ-continuous — that is, λ vanishes on μ-null sets. We use λ to define a martingale $(g_{\mathcal{A}} : \mathcal{A} \in \mathfrak{F})$ as follows.

For each finite subalgebra $\mathcal{A} \subseteq \mathcal{S}$, observe that $\pi(\mathcal{A}) = \{$minimal nonempty members of $\mathcal{A}\}$ is a finite partition $\pi(\mathcal{A})$ of Ω, and $\mathcal{A} = \{$unions of members of $\pi(\mathcal{A})\}$. Define a simple (i.e., finitely valued) integrable function $g_{\mathcal{A}} : \Omega \to X$ by

$$g_{\mathcal{A}}(\omega) \quad = \quad \begin{cases} \dfrac{\lambda(T)}{\mu(T)} & \text{if } \omega \in T \in \pi(\mathcal{A}) \text{ and } \mu(T) > 0 \\[2ex] 0 & \text{if } \omega \in T \in \pi(\mathcal{A}) \text{ and } \mu(T) = 0. \end{cases}$$

Some observations:

- **a.** The function $g_{\mathcal{A}}$ is defined uniquely everywhere on Ω, not just μ-almost everywhere.
- **b.** The restriction of λ to $\mathcal{A}$ has bounded variation and has Radon-Nikodym derivative equal to $g_{\mathcal{A}}$; thus we obtain $\mathrm{Var}(\lambda, \mathcal{A}) = \|g_{\mathcal{A}}\|_1$.
- **c.** $\int_A g_{\mathcal{A}}\, d\mu \;=\; \lambda(A)$ when $A \in \mathcal{A}$.
- **d.** $(g_{\mathcal{A}} : \mathcal{A} \in \mathfrak{F})$ is an X-valued martingale. For purposes of the discussion in the next few sections, we shall call this **the full sieve martingale** associated with λ.
- **e.** If $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_3 \subseteq \cdots$ is an increasing sequence of finite subalgebras of $\mathcal{S}$, then the sequence $g_n = g_{\mathcal{A}_n}$ $(n = 1, 2, 3, \ldots)$, with σ-algebras $\mathcal{A}_n$, is a martingale. We shall call it **a sequential sieve martingale** associated with λ. Different sequential sieve martingales are obtained from different sequences $(\mathcal{A}_n)$.

29.25. Sieve Convergence Lemma. Let $(\Omega, \mathcal{S}, \mu)$ be a probability space. Let λ be an X-valued measure on $(\Omega, \mathcal{S})$ that is μ-continuous. Define sieve martingales as in 29.24. Then the following conditions are equivalent:

(A) There exists a Radon-Nikodym derivative $h = d\lambda/d\mu$. That is, there exists some $h \in L^1(\mu, X)$ that satisfies $\lambda(S) = \int_S h\, d\mu$ for every $S \in \mathcal{S}$. Hence $g_{\mathcal{A}} = E(h|\mathcal{A})$ for each $\mathcal{A} \in \mathfrak{F}$.

(B) The full sieve martingale $(g_{\mathcal{A}} : \mathcal{A} \in \mathfrak{F})$ associated with λ converges in $L^1(\mu, X)$ to some limit h.

(C) For each increasing sequence of algebras $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_3 \subseteq \cdots$ contained in $\mathcal{S}$, the sequential sieve martingale (g_n) associated with λ converges in $L^1(\mu, X)$ to some limit g_∞.

(D) For each increasing sequence of algebras $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_3 \subseteq \cdots$ contained in $\mathcal{S}$, there exists some function $g_\infty \in L^1(\mu, X)$ with the following property: For each $\varphi \in X^*$, the scalar-valued sequential sieve martingale $(\varphi \circ g_n : n = 1, 2, 3, \ldots)$ converges in $L^1(\mu, \mathbb{F})$ to $\varphi \circ g_\infty$. (Here $\mathbb{F}$ is the scalar field.)

Furthermore, when these conditions are satisfied, then the functions h in (A), (B) are the same, and the limits g_∞ in (C),(D) are equal to $E(h|\mathcal{A}_\infty)$, where $\mathcal{A}_\infty$ is the σ-algebra generated by $\bigcup_{n\in\mathbb{N}} \mathcal{A}_n$.

Proof of theorem. The implication (A) $\Rightarrow$ (C) follows from 29.17; the implication (C) $\Rightarrow$ (B) follows from 19.7. To prove (B) $\Rightarrow$ (A), fix any set $S \in \mathcal{S}$. For all $\mathcal{A} \in \mathfrak{F}$ sufficiently large, we have $S \in \mathcal{A}$, hence $\int_S g_{\mathcal{A}}\, d\mu = \lambda(S)$. Take limits as $g_{\mathcal{A}} \to h$ in $L^1(\mu, X)$ to obtain $\int_S h\, d\mu = \lambda(S)$.

The implication (C) $\Rightarrow$ (D) is obvious. For (D) $\Rightarrow$ (C), let any sequence $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_3 \subseteq \cdots \subseteq \mathcal{S}$ be given, and suppose the conclusion of (D) holds. Temporarily fix any $\varphi \in X^*$. It is easy (*exercise*) to verify that the sequence $(\varphi \circ g_n)$ is a martingale in $L^1(\mu, \mathbb{F})$. Since that martingale converges in $L^1(\mu, \mathbb{F})$ to $\varphi \circ g_\infty$, we know from the implication (A) $\Rightarrow$ (B) in 29.17 that $\varphi \circ g_n = E(\varphi \circ g_\infty|\mathcal{A}_n)$ for each n. That is, $\varphi \circ g_n$ is $\mathcal{A}_n$-measurable, and $\int_S \varphi \circ g_n\, d\mu = \int_S \varphi \circ g_\infty\, d\mu$ for each set $S \in \mathcal{A}_n$. Since φ is continuous and linear, it commutes with the Bochner integral; thus we obtain $\varphi\left(\int g_n\, d\mu\right) = \varphi\left(\int g_\infty\, d\mu\right)$ for each $\varphi \in X^*$. Since X^* separates the points of X, it follows that $\int g_n\, d\mu = \int g_\infty\, d\mu$ for each $S \in \mathcal{A}_n$. That is, $g_n = E(g_\infty|\mathcal{A}_n)$. By the implication (B) $\Rightarrow$ (A) in 29.17, it follows that $g_n \to g_\infty$ in $L^1(\mu, X)$.

29.26. Theorem (Phillips). Every reflexive Banach space has the RNP.

Proof (following Rønnow [1967] and Chatterji [1968]). Let $(X, |\ \ |)$ be a reflexive Banach space. Let $(\Omega, \mathcal{S}, \mu)$ be a positive measure space, and let λ be an X-valued μ-continuous vector measure that has bounded variation. We are to show that the Radon-Nikodym derivative $d\lambda/d\mu$ exists.

If $\mathcal{S}$ is not complete, we may extend λ and μ to the completion of $\mathcal{S}$, by taking them both to be 0 on any μ-null set. Thus we may replace $\mathcal{S}$ with its completion — we may assume $\mathcal{S}$ is complete (i.e., every null set is measurable).

Since $\lambda << \mu$, we know that $/\lambda/ << \mu$. By the classical Radon-Nikodym Theorem, we know that the Radon-Nikodym derivative $\dfrac{d/\lambda/}{d\mu}$ exists. It suffices to show that the derivative

$\dfrac{d\lambda}{d/\lambda/}$ exists, for then we may apply the Chain Rule 29.12.b to obtain $\dfrac{d\lambda}{d\mu} = \dfrac{d\lambda}{d/\lambda/}\dfrac{d/\lambda/}{d\mu}$. Hence we may replace μ with $/\lambda/$; thus we may assume hereafter that $\mu = /\lambda/$. By rescaling, we may also assume that μ is a probability measure.

Let $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{A}_3 \subseteq \cdots$ be an increasing sequence of finite algebras contained in $\mathcal{S}$. Define the X-valued martingale (g_n) as in 29.24.e It suffices to verify condition 29.25(D).

Let X_0 be the closed linear span of the union of the ranges of the g_n's; then X_0 is separable and weakly closed. Since $|\lambda(S)| \leq /\lambda/(S) = \mu(S)$, from our definition of the g_n's in 29.24 we see that $|g_n(\omega)| \leq 1$ for all n, ω. Let B be the closed unit ball of X; the g_n's take their values in $B \cap X_0$. The set B is weakly compact by 28.41(C); hence $B \cap X_0$ is weakly compact. Therefore $B \cap X_0$ is weakly sequentially compact, by 28.36(E).

Temporarily fix any $\omega \in \Omega$. The sequence $(g_n(\omega))$ has a subsequence $(g_{n_k}(\omega))$ that converges weakly to some limit $g_\infty(\omega) \in B \cap X_0$. Many choices of $g_\infty(\omega)$ may be possible, using different subsequences; we use the Axiom of Choice to select some particular $g_\infty(\omega)$. Making a choice for each ω, we define a function $g_\infty : \Omega \to B$. (Different ω's may use different subsequences; we do not yet assert anything about measurability of g_∞.)

Let $\mathbb{F}$ be the scalar field. Temporarily fix any $\varphi \in X^*$. By the classical Radon-Nikodym Theorem, the scalar-valued measure $\varphi \circ \lambda$ has a Radon-Nikodym derivative $\dfrac{d(\varphi \circ \lambda)}{d\mu}$. By the implication (A) $\Rightarrow$ (C) in 29.25 (with the scalar field used as a Banach space), the sequence $(\varphi \circ g_n : n \in \mathbb{N})$ converges in $L^1(\mu, \mathbb{F})$ to some limit h_φ. By 29.19, we also have $\varphi \circ g_n \to h_\varphi$ almost everywhere. Thus there is some set N_φ with measure 0, such that for any $\omega \in \Omega \setminus N_\varphi$ we have $\langle \varphi, g_n(\omega)\rangle \to h_\varphi(\omega)$. On the other hand, holding ω fixed, some subsequence $(g_{n_k}(\omega))$ converges weakly to $g_\infty(\omega)$. It follows that

$$h_\varphi(\omega) \quad = \quad \langle \varphi, g_\infty(\omega)\rangle \qquad \text{for all } \omega \in \Omega \setminus N_\varphi.$$

Since h_φ is measurable and N_φ is a null set, the function $\omega \mapsto \langle \varphi, g_\infty(\omega)\rangle$ is measurable.

Thus g_∞ is weakly measurable. It is also separably valued and bounded, since it takes its values in $B \cap X_0$. Therefore it is strongly measurable (by 23.25) and belongs to $L^\infty(\mu, X)$.

As we noted above, for each $\varphi \in X^*$ we have $\varphi \circ g_n \to h_\varphi = \varphi \circ g_\infty$ as $n \to \infty$. This completes the verification of 29.25(D), and thus the proof of the theorem.

Semivariation and Bartle Integrals

29.27. *Notations.* Throughout the next few sections, we shall assume $\mathcal{A}$ is an algebra of subsets of some set Ω. Also, we assume $(X, |\ \ |)$ is a Banach space, and U is the closed unit ball of the dual of X — that is,

$$U \quad = \quad \left\{u \in X^* \ :\ |u|_{X^*} \leq 1\right\}.$$

29.28. *Definition.* With notations as in 29.27, let $\lambda : \mathcal{A} \to X$ be a vector charge. Then for each $u \in U$, the function $u\lambda = u \circ \lambda$ is a scalar-valued charge; its variation is the positive charge $/u\lambda/$. The **semivariation** of λ is the function $\iddots\lambda\iddots : \mathcal{A} \to [0, +\infty]$ defined by

$$\iddots\lambda\iddots(A) \quad = \quad \sup_{u \in U} /u\lambda/(A).$$

Caution: A more commonly used notation is $\|\lambda\|$; see for instance Diestel and Uhl [1977]. Our notation $\iddots\lambda\iddots$ is unconventional, but perhaps more suggestive.

29.29. *Basic properties of semivariations.*

a. $\iddots\lambda\iddots$ is monotone — that is, $A \subseteq B \Rightarrow \iddots\lambda\iddots(A) \leq \iddots\lambda\iddots(B)$. Hence the largest value taken by $\iddots\lambda\iddots$ is the (not necessarily finite) number $\iddots\lambda\iddots(\Omega)$.

b. $\iddots\lambda\iddots$ is **finitely subadditive** — that is, $\iddots\lambda\iddots\left(\bigcup_{j=1}^n C_j\right) \leq \sum_{j=1}^n \iddots\lambda\iddots(C_j)$ whenever n is a positive integer and $C_1, C_2, \ldots, C_n$ are members of $\mathcal{A}$.

c. If the scalar field is $\mathbb{R}$, our formula for the semivariation can be rewritten

$$\iddots\lambda\iddots(A) \quad = \quad \sup\Big\{|u\lambda(S)| + |u\lambda(A \setminus S)| \ : \ u \in U,\ S \in \mathcal{S},\ S \subseteq A\Big\}.$$

This follows from 29.6.d.

d. For any set $A \in \mathcal{S}$, we have

$$\frac{1}{4}\iddots\lambda\iddots(A) \quad \leq \quad \sup\Big\{|\lambda(S)| \ : \ S \in \mathcal{A},\ S \subseteq A\Big\} \quad \leq \quad \iddots\lambda\iddots(A).$$

If the scalar field is $\mathbb{R}$, we can use $\frac{1}{2}$ in place of $\frac{1}{4}$.

Hints: This follows from the previous observation; for the second inequality use also (HB8) in 23.18.

e. In particular, $\frac{1}{4}\iddots\lambda\iddots(\Omega) \leq \sup_{A \in \mathcal{A}} |\lambda(A)| \leq \iddots\lambda\iddots(\Omega)$. Thus, λ is bounded (i.e., has bounded range) if and only if its semivariation over Ω is finite. Instead of saying that a charge has "bounded semivariation," we may just say that the charge is bounded.

f. The **space of all bounded X-valued charges on $\mathcal{A}$** is a Banach space, when normed by either $\|\lambda\| = \sup\{|\lambda(A)| : A \in \mathcal{A}\}$ or $\|\lambda\| = \iddots\lambda\iddots(\Omega)$; these two norms are equivalent. The norm $\iddots\lambda\iddots(\Omega)$ is more convenient for some purposes, particularly 29.30 below. The space of bounded charges will be denoted $\boldsymbol{ba}(\mathcal{A}, X)$. (For proof of completeness, apply 22.17 with $\Gamma = \mathcal{A}$.)

In general, $ba(\mathcal{A}, X)$ is larger than the space $BV(\mathcal{A}, X)$ of charges with bounded variation, which was introduced in 29.6.c. However, when X is finite-dimensional, then the two spaces $ba(\mathcal{A}, X)$ and $BV(\mathcal{A}, X)$ are the same, and the variation is another norm equivalent to the semivariation. When X is the scalar field, then the variation is *equal* to the semivariation.

g. Let $\mathcal{S}$ be a σ-algebra of subsets of Ω. The **space of X-valued measures** (i.e., countably additive charges) on $\mathcal{S}$ will be denoted $\boldsymbol{ca}(\mathcal{A}, X)$. It is a closed subspace of $ba(\mathcal{A}, X)$ and is normed by the semivariation or the sup norm, as in 29.29.f.

(Again, we note that if X is finite-dimensional, then another equivalent norm is given by the variation; if X is the scalar field, then the semivariation is equal to the variation.)

29.30. *Notations.* For the discussions below, let Ω be a nonempty set, and let $\mathcal{A}$ be an algebra (but not necessarily a σ-algebra) of subsets of Ω. Let $\mathbb{F}$ be the scalar field (equal to $\mathbb{R}$ or $\mathbb{C}$). Let

$$B(\Omega) \quad = \quad \{\text{bounded functions from } \Omega \text{ into } \mathbb{F}\};$$

this is a Banach space when equipped with the sup norm.

In this context, a *simple function* will mean a function $f : \Omega \to \mathbb{F}$ whose range is a finite set and that satisfies $f^{-1}(c) \in \mathcal{A}$ for each $c \in \mathbb{F}$. Equivalently, f is a linear combination of finitely many characteristic functions of members of $\mathcal{A}$.) Observe that the simple functions form a linear subspace of the Banach space $B(\Omega)$. Now define

$$\mathrm{Unif}(\mathcal{A}) \quad = \quad \{\text{uniform limits of simple functions}\};$$

thus $\mathrm{Unif}(\mathcal{A})$ is the closure of the simple functions in $B(\Omega)$. It is a closed linear subspace of $B(\Omega)$, and thus it is a Banach space when equipped with the sup norm.

Elementary exercise. If the algebra $\mathcal{A}$ is a σ-algebra, then $\mathrm{Unif}(\mathcal{A})$ is equal to the set of all bounded measurable real-valued functions on Ω.

Definition. Let $(X, |\ \ |)$ be a Banach space, with scalar field $\mathbb{F}$. Let Ω, $\mathcal{A}$, $\mathrm{Unif}(\mathcal{A})$, etc., be as above. Define the Banach space $ba(\mathcal{A}, X)$ of bounded charges as in 29.29.f, with the semivariation for the norm: $\|\lambda\| = \vdots\lambda\vdots(\Omega)$. By the **Bartle integral** we shall mean a continuous bilinear map

$$\int_\Omega (\cdot)\, d(\cdot) \qquad : \qquad \mathrm{Unif}(\mathcal{A}) \times ba(\mathcal{A}, X) \quad \to \quad X$$

defined as follows. When f is a simple function (i.e., a finitely valued measurable function), then define the integral in the obvious fashion, as in 11.42 — that is, $\int_\Omega f\, d\lambda = \sum_c \lambda\left(f^{-1}(c)\right) c$. It is easy to verify that

$$\left|\int_\Omega f\, d\lambda\right| \quad \leq \quad \|f\|_\infty \, \|\lambda\|. \tag{!}$$

Thus the mapping $f \mapsto \int_\Omega f\, d\lambda$ is continuous, and so it extends uniquely (see 23.2.e) to a linear map on all of $\mathrm{Unif}(\mathcal{A})$, also satisfying (!).

We emphasize that the charge λ need not be scalar-valued or countably additive, but the integrand f must be scalar-valued and bounded. (Contrast this with the Bochner integral — defined in 23.16 — for which the measure must be positive and countably additive, but whose integrand may be vector-valued and unbounded.)

Remarks. The definition given above is convenient for our purposes, but the literature sometimes uses the term "Bartle integral" with wider choices of f and λ. In particular, Bartle himself permitted all of f, λ, and $\int f\, d\lambda$ to take values in vector spaces X, Y, Z; where we have used multiplication of a scalar times a vector, he used a bilinear map $\langle\ ,\ \rangle : X \times Y \to Z$. What we have called the "Bartle integral" is what some mathematicians would call the **Radon integral**, but that term has other meanings, too.

29.31. *Further properties of the Bartle integral.* Let Ω, $\mathcal{A}$, $\mathrm{Unif}(\mathcal{A})$, etc., be as in 29.30.

a. In the fashion indicated above, each $\lambda \in ba(\mathcal{A}, X)$ defines a continuous linear map $\beta_\lambda : \mathrm{Unif}(\mathcal{A}) \to X$, defined by $\beta_\lambda(f) = \int_\Omega f\,d\lambda$. The preceding argument shows that this linear map has operator norm $|||\beta_\lambda||| \leq \|\lambda\|$.

Actually, from the definition of semivariation it follows easily (*exercise*) that $|||\beta_\lambda|||$ is *equal* to $\|\lambda\|$. That is,

$$\sup\left\{\left|\int_\Omega f\,d\lambda\right| \;:\; f \in \mathrm{Unif}(\mathcal{A}),\ \|f\|_\infty \leq 1\right\} \quad = \quad \vdots\lambda\vdots(\Omega).$$

(For this and related purposes, the semivariation works much better than the sup of the charge, even though the semivariation and sup are equivalent norms.)

b. Furthermore, *every* continuous linear map from $\mathrm{Unif}(\mathcal{A})$ to X is of this form. Thus, the mapping $\lambda \mapsto \beta_\lambda$ is an isomorphism (i.e., norm-preserving linear bijection) from $ba(\mathcal{A}, X)$ to the operator-normed space

$$BL\Big(\mathrm{Unif}(\mathcal{A}), X\Big) \quad = \quad \{\text{bounded linear maps from } \mathrm{Unif}(\mathcal{A}) \text{ to } X\}$$

(defined as in 23.1).

In particular, $ba(\mathcal{A}, \mathbb{F}) = \big(\mathrm{Unif}(\mathcal{A})\big)^*$.

c. The dual of $L^\infty(\mu)$. Let $(\Omega, \mathcal{S}, \mu)$ be a measure space (i.e., assume $\mathcal{S}$ is a σ-algebra and μ is countably additive). Show that the mapping $\lambda \mapsto \beta_\lambda$ gives an isomorphism from $ba(\mu, X)$ onto $BL\,(L^\infty(\mu), X))$, where $ba(\mu, X)$ is the subspace of $ba(\mathcal{S}, X)$ consisting of those charges that vanish on μ-null sets.

(*Hints*: Since $\mathcal{S}$ is a σ-algebra, $\mathrm{Unif}(\mathcal{S})$ is the space of bounded measurable functions. Then $L^\infty(\mu)$ is a quotient space of $\mathrm{Unif}(\mathcal{S})$, obtained by identifying those functions that are μ-equivalent.)

In particular, $ba(\mu, \mathbb{F}) = \big(L^\infty(\mu)\big)^*$.

29.32. The following principles are equivalent to the Hahn-Banach Theorems, which were presented in 12.31, 23.18, 23.19, 26.56, 28.4, and 28.14.a. Notation is as in 29.30.

> **(HB25) Banach's Generalized Integral.** Let $\Omega, \mathcal{A}, B(\Omega), \mathbb{F}$, etc., be as in 29.30, with $X = \mathbb{F} = \mathbb{R}$. Let λ be a bounded charge on $\mathcal{A}$. Then the Bartle integral $f \mapsto \int f\,d\lambda$, already defined on $\mathrm{Unif}(\mathcal{A})$ in 29.30, can be extended (not necessarily uniquely) to a continuous linear map $\S : B(\Omega) \to \mathbb{R}$, satisfying $|\S f| \leq \|f\|_\infty \vdots\lambda\vdots$. If λ is a positive charge, then $\S$ can be chosen so that it is also a positive linear functional.

> **(HB26) Banach's Charge.** Let $\mathcal{A}$ be an algebra of subsets of a set Ω, and let λ be a bounded real-valued charge on $\mathcal{A}$. Then λ can be extended to a real-valued charge Λ on all of $\mathcal{P}(\Omega)$. If λ is a positive charge, then we can also choose Λ to be a positive charge.

(See also the related remarks in 21.23.)

Proof that (HB2) and (HB7) imply (HB25). The first statement is obvious from (HB7); the result about positive charges will take a little more work. Observe that the mapping

$f \mapsto \|f^+\|_\infty$ is convex, since $\|f^+\|_\infty$ is the supremum of the linear functionals 0 and $f(\omega)$ (for $\omega \in \Omega$). For $f \in \mathrm{Unif}(\mathcal{A})$, we have $\int f\,d\lambda \leq \int f^+\,d\lambda \leq \|f^+\|_\infty \vdots\lambda\vdots$, hence by (HB2) we can extend the Bartle integral to a linear map $\S$ that satisfies $\S f \leq \|f^+\|_\infty \vdots\lambda\vdots$. When $f \leq 0$, then $\|f^+\|_\infty \leq 0$; this proves $\S$ is a positive linear functional. However, we still need to show that a functional chosen in this fashion will also satisfy the inequality $|\S f| \leq \|f\|_\infty \vdots\lambda\vdots$. Any $f \in B(\Omega)$ can be expressed in terms of its Jordan Decomposition: $f = f^+ - f^-$ with

$$f^+(\omega) = f(\omega) \text{ and } f^-(\omega) = 0 \text{ wherever } f(\omega) \geq 0,$$

$$f^-(\omega) = -f(\omega) \text{ and } f^+(\omega) = 0 \text{ wherever } f(\omega) \leq 0.$$

Then $\S(f^+)$ and $\S(f^-)$ are both nonnegative, so

$$\begin{array}{rclcl} |\S f| & = & |\S(f^+) - \S(f^-)| & \leq & \max\{\S(f^+), \S(f^-)\} \\ & & & \leq & \max\{\|f^+\|_\infty, \|f^-\|_\infty\} \vdots\lambda\vdots \quad = \quad \|f\|_\infty \vdots\lambda\vdots \end{array}$$

Proof of (HB25) $\Rightarrow$ (HB26). Define $\Lambda(S) = \S(1_S)$.

Proof of (HB26) $\Rightarrow$ (HB12). Let $\mathfrak{I} = \{\Omega \setminus F : F \in \mathfrak{F}\}$; this is the proper ideal that is dual to $\mathfrak{F}$. Let $\mathcal{A} = \mathfrak{F} \cup \mathfrak{I} = \{S \subseteq \Omega \ : \ S \in \mathfrak{F} \text{ or } S \in \mathfrak{I}\}$; verify that $\mathcal{A}$ is an algebra of sets. Define $\lambda : \mathcal{A} \to \{0, 1\}$ by taking $\lambda(F) = 1$ for each $F \in \mathfrak{F}$ and $\lambda(I) = 0$ for each $I \in \mathfrak{I}$; verify that this is a positive charge on $\mathcal{A}$. Now extend to $\Lambda = \mu$.

Measures on Intervals

29.33. Theorem: scalar-valued measures on an interval. Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$). Let $\varphi : [a, b] \to \mathbb{F}$ be a function that has bounded variation (in the sense of intervals — i.e., as in 19.21). Then:

- **a.** The Henstock-Stieltjes integral $\mu_\varphi(S) = \int_a^b 1_S(t)d\varphi(t)$ exists for every Borel set $S \subseteq [a, b]$, and thus defines a scalar-valued measure μ_φ on those sets. (That measure has bounded variation in the sense of measures, as in 29.5 — see 29.6.h.)
- **b.** Let $f : [a, b] \to \mathbb{F}$ be bounded and measurable (from the Borel sets to the Borel sets). Then the Henstock-Stieltjes integral $\int_a^b f\,d\varphi$ exists and is equal to the Bartle integral $\int_{[a,b]} f\,d\mu_\varphi$.

Proof. If φ is complex-valued, we may write it as $\operatorname{Re}\varphi + i \operatorname{Im}\varphi$; thus it suffices to consider real-valued φ. Any real-valued function of bounded variation can be written as the difference of two increasing functions. Thus we can apply 24.35; μ_φ is a linear combination of positive finite measures on the Borel sets, and thus it is a scalar-valued measure.

By the definition of μ_φ, the equation $\int_a^b f\,d\varphi = \int_{[a,b]} f\,d\mu_\varphi$ is clear when f is the characteristic function of a measurable set — hence also when f is a simple function, by linearity. The simple functions are dense in $\mathcal{L}^\infty(\mathcal{S})$, and the Henstock-Stieltjes and Bartle integrals are continuous on $\mathcal{L}^\infty(\mathcal{S})$ (see 24.17 and 29.30). Take limits to prove $\int_a^b f\,d\varphi = \int_{[a,b]} f\,d\mu_\varphi$ for all $f \in \mathcal{L}^\infty(\mathcal{S})$.

29.34. Riesz Representation Theorem for intervals. Let $\Omega = [a,b]$ be a compact interval in $\mathbb{R}$, and let $\mathcal{B}$ be its σ-algebra of Borel subsets. Let $\mathbb{F}$ be the scalar field ($\mathbb{R}$ or $\mathbb{C}$), and let

$$C[a,b] \quad = \quad \{\text{continuous functions from } [a,b] \text{ into } \mathbb{F}\}.$$

Then these three Banach spaces are isomorphic:

- the space $C[a,b]^*$ of continuous linear functionals $\Lambda : C[a,b] \to \mathbb{F}$, equipped with the operator norm;
- the space $ca(\mathcal{B},\mathbb{F})$ of scalar-valued measures μ, normed by the variation $/\mu/$ as in 29.29.g;
- the space $NBV([a,b],\mathbb{F})$ of normalized functions of bounded variation, equipped with the norm $\|\varphi\| = \mathrm{Var}(\varphi,[a,b])$ as in 22.19.d.

In fact, the following maps are linear norm-preserving bijections:

- the mapping $M : \varphi \mapsto \mu_\varphi$ from $NBV([a,b],\mathbb{F})$ to $ca(\mathcal{B},\mathbb{F})$ defined by the Henstock-Stieltjes integral $\mu_\varphi(S) = \int_a^b 1_S\,d\varphi$ as in 29.33;
- the mapping $B : \mu \mapsto \beta_\mu$ from $ca(\mathcal{B},\mathbb{F})$ to $C[a,b]^*$ defined by the Bartle integral $\beta_\mu(f) = \int_\Omega f\,d\mu$, as in 29.30;
- the mapping $\Lambda : \varphi \mapsto \lambda_\varphi$ from $NBV([a,b],\mathbb{F})$ onto $C[a,b]^*$ given by the Riemann-Stieltjes integral $\lambda_\varphi(f) = \int_a^b f\,d\varphi$, as in 24.26(ii).

(In fact, the mapping Λ is actually equal to $B \circ M$.) Thus, the two kinds of "variations" defined in 19.21 and 29.5 are equal, for φ and μ corresponding as above.

Proof (based on Limaye [1981]). The equation $\Lambda = B \circ M$ follows from 29.33. It follows from 24.28 that the mapping Λ is injective when considered only from $NBV([a,b],\mathbb{F})$ to $C[a,b]^*$. We saw in 29.30 and 24.16.c that $|||B||| \le 1$ and $|||M||| \le 1$; hence $|||\Lambda||| \le 1$. It suffices to show that the mapping $\Lambda : NBV([a,b],\mathbb{F}) \to C[a,b]^*$ is surjective and that $|||\Lambda^{-1}||| \le 1$. Thus, let any $\lambda \in C[a,b]^*$ be given; it suffices to show that there is some $\varphi \in NBV([a,b],\mathbb{F})$ satisfying $\lambda = \lambda_\varphi$ and $\|\varphi\| \le \|\lambda\|$.

Let $\widehat{\lambda}$ be any Hahn-Banach extension of λ to $\mathcal{L}^\infty(\mathcal{B})$ — that is, let $\widehat{\lambda} : \mathcal{L}^\infty(\mathcal{B})$ be any continuous linear map from $\mathcal{L}^\infty(\mathcal{B})$ to $\mathbb{F}$ that extends λ and satisfies $\|\widehat{\lambda}\| = \|\lambda\|$. Define $\psi(t) = \widehat{\lambda}(1_{(a,t]})$, where $1_{(a,t]}$ is the characteristic function of the interval $(a,t]$. In particular, $\psi(a) = \widehat{\lambda}(1_\varnothing) = \widehat{\lambda}(0) = 0$.

Note that if $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ is any partition and $k_1, k_2, \ldots, k_n$ are any constants, then the function $u : [a,b] \to \mathbb{F}$ defined by

$$u(\cdot) \quad = \quad \sum_{j=1}^n k_j\,1_{(t_{j-1},t_j]}(\cdot) \quad = \quad \sum_{j=1}^n k_j\left(1_{(a,t_j]}(\cdot) - 1_{(a,t_{j-1}]}(\cdot)\right) \qquad (**)$$

satisfies $\widehat{\lambda}(u) = \sum_{j=1}^{n} k_j\,[\psi(t_j) - \psi(t_{j-1})]$.

We claim next that $\psi \in BV([a,b],\mathbb{F})$, with $\mathrm{Var}(\psi) \le \|\widehat{\lambda}\|$. To see this, let any partition of $[a,b]$ be given; define u as in $(**)$ with constants

$$k_j \quad = \quad \begin{cases} \dfrac{|\psi(t_j) - \psi(t_{j-1})|}{\psi(t_j) - \psi(t_{j-1})} & \text{if } \psi(t_j) \ne \psi(t_{j-1}) \\[2ex] 0 & \text{if } \psi(t_j) = \psi(t_{j-1}). \end{cases}$$

We may assume that the k_j's are not all zero; hence $\|u\|_\infty = 1$ and

$$\sum_{j=1}^{n} |\psi(t_j) - \psi(t_{j-1})| \quad = \quad \sum_{j=1}^{n} k_j\,[\psi(t_j) - \psi(t_{j-1})] \quad = \quad \widehat{\lambda}(u) \quad \le \quad \|\widehat{\lambda}\|.$$

This proves our claim.

Next we claim that $\int_a^b f\,d\psi = \lambda(f)$ for every $u \in C[a,b]$. Indeed, since f is continuous and ψ has bounded variation, the integral $\int_a^b f\,d\psi$ is a Riemann-Stieltjes integral, not just a Henstock-Stieltjes integral. Thus, in the approximating Riemann sums $\Sigma[f,T,\psi]$, we may take the tags τ_j to be any points in the subintervals $[t_{j-1}, t_j]$. In particular, we may take $\tau_j = t_j$. Let $k_j = f(t_j)$, and define u as in $(**)$; then $\Sigma[f,T,\psi] = \sum_{j=1}^{n} k_j[\psi(t_j) - \psi(t_{j-1})] = \widehat{\lambda}(u)$. The Riemann sums of this type converge to the Riemann-Stieltjes integral $\int_a^b f\,d\psi$. Meanwhile, the approximating functions u defined in $(**)$ converge uniformly to the continuous function f, and so $\widehat{\lambda}(u)$ converges to $\widehat{\lambda}(f) = \lambda(f)$. This proves our claim.

By 24.28 we may write $\psi = \psi_1 + \psi_2$ where $\psi_1 \in C[a,b]^\perp$, $\psi_2 \in NBV([a,b],X)$, and $\mathrm{Var}(\psi_2) \le \mathrm{Var}(\psi)$. Take $\varphi = \psi_2$; this completes the proof of the present theorem.

29.35. We now state a more general theorem. We omit the proof, which is quite long; it can be found in books on measure and integration.

Riesz Representation Theorem (general version; proof omitted). Let Ω be a locally compact, Hausdorff topological space. Let $C_c(\Omega)$ be the ordered vector space of all real-valued, continuous functions on Ω that have compact support. Then each positive linear functional Λ on $C_c(\Omega)$ is of the form

$$\Lambda(f) \quad = \quad \int_\Omega f\,d\mu$$

where μ is a positive measure on the Borel subsets of Ω. There may be more than one measure satisfying this requirement, but there is only one satisfying the following further conditions: Each compact subset of Ω has finite measure; μ is *outer regular*, in the sense that $\mu(B) = \inf\{\mu(G) : G \text{ is an open superset of } B\}$ for each Borel set B; and μ is *inner regular*, in the sense that $\mu(G) = \sup\{\mu(K) : K \text{ is a compact subset of } G\}$ for each open set G.

29.36. Theorem. Let φ be some mapping from an interval $[a,b]$ into the scalar field ($\mathbb{R}$ or $\mathbb{C}$). Then the following conditions are equivalent:

(A) $\varphi(x) = \varphi(a) + \int_a^x g(t)dt$ for some function $g \in L^1[a,b]$.

(B) φ is **absolutely continuous in the classical sense**; that is, for each number $\varepsilon > 0$ there exists some number $\delta > 0$ such that

> whenever $a \le s_1 \le t_1 \le s_2 \le t_2 \le \cdots \le s_n \le t_n \le b$ with $\sum_{j=1}^n |s_j - t_j| < \delta$, then $\sum_{j=1}^n |\varphi(s_j) - \varphi(t_j)| < \varepsilon$.

In this case we might also describe *phi* as **absolutely continuous in the sense of intervals.**

(C) φ is continuous, and φ has bounded variation in the sense of intervals (see 19.21). Moreover, if we define a measure μ_φ on the Lebesgue measurable sets by the Henstock-Stieltjes integral $\mu_\varphi(S) = \int_a^b 1_S(t)d\varphi(t)$, then μ_φ is absolutely continuous (in the sense of measures, as in 29.4) with respect to Lebesgue measure.

Proof. Throughout the proof, let Λ denote Lebesgue measure on $[a,b]$.

Proof of (A) $\Rightarrow$ (B). Define a positive finite measure γ by $\gamma(A) = \int_A |g(\cdot)|\, d\Lambda$, as in the proof of 29.10. Compute

$$\sum_{j=1}^n |\varphi(s_j) - \varphi(t_j)| \quad = \quad \sum_{j=1}^n \left| \int_{s_j}^{t_j} g(r)dr \right| \quad \le \quad \sum_{j=1}^n \int_{s_j}^{t_j} |g(r)|dr \quad = \quad \gamma(A)$$

where $A = \bigcup_{j=1}^n (s_j, t_j)$. Since $\gamma << \Lambda$ (as noted in 29.10), for each $\varepsilon > 0$ there is some $\delta > 0$ such that $\Lambda(A) < \delta \Rightarrow \gamma(A) < \varepsilon$.

Proof of (B) $\Rightarrow$ (C). That φ is continuous and has bounded variation in the sense of intervals is an easy exercise. To prove $\mu_\varphi << \Lambda$, let $\varepsilon > 0$ be given; choose $\delta > 0$ as in (B). Let S be any Lebesgue-measurable subset of $[a,b]$ with $\Lambda(S) < \delta$; we shall show that $|\mu_\varphi(S)| \le \varepsilon$.

By 24.40, there is some open set $G \supseteq S$ with $\Lambda(G) < \delta$. (Here "open" refers to the relative topology on $[a,b]$; thus $[a,b]$ itself is open.) By 15.37.d we know that $G = \bigcup_{j=1}^\infty H_j$ for some disjoint open intervals H_j (not necessarily arranged from left to right across the the interval $[a,b]$ as j increases); again "open" refers to the relative topology.

Form a gauge $\gamma : [a,b] \to (0, +\infty)$ with the following property: Whenever τ is a point in H_j for some $j \in \mathbb{N}$, then $\gamma(\tau)$ is a positive number small enough so that $[a,b] \cap [\tau - \gamma(\tau), \tau + \gamma(\tau)] \subseteq H_j$. Let $T = (m, t_i, \tau_i)$ be any tagged division of $[a,b]$ that is γ-fine; let $I = \{i \in \{1,2,\ldots,m\} : \tau_i \in S\}$. Then $\mu_\varphi(S) = \int_a^b 1_S(t)d\varphi(t)$ is approximated by the Riemann-Stieltjes sum

$$\Sigma[1_S, T, \varphi] \quad = \quad \sum_{i=1}^m 1_S(\tau_i)\,[\varphi(t_i) - \varphi(t_{i-1})] \quad = \quad \sum_{i \in I} [\varphi(t_i) - \varphi(t_{i-1})]\,.$$

Whenever $i \in I$, then $\tau_i \in G$, hence $\tau_i \in H_j$ for some j, and therefore $[t_{i-1}, t_i] \subseteq H_j$ by our choice of the gauge γ. The intervals (t_{i-1}, t_i) are disjoint, and therefore the sum of the lengths of the (t_{i-1}, t_i)'s (for $i \in I$) is less than or equal to the sum of the lengths of the H_j's. That sum is less than or equal to $\Lambda(G)$, and thus is less than δ. Hence, by our hypothesis (B), we have $\sum_{i \in I} |\varphi(t_i) - \varphi(t_{i-1})| < \varepsilon$. Thus $|\Sigma[1_S, T, \varphi]| < \varepsilon$. Taking limits as the tagged division T becomes finer, we see that $|\mu_\varphi(S)| \le \varepsilon$.

Proof of (C) $\Rightarrow$ (A). The measure μ_φ has bounded variation by 29.6.h, and $\mu_\varphi << \Lambda$. By the Radon-Nikodym Theorem (29.20), there is some $g \in L^1[a,b]$ such that $\mu_\varphi(S) = \int_S g\, d\Lambda$ for all Lebesgue-measurable sets $S \subseteq [a,b]$. In particular, when S is the interval $[a,x]$, then by 24.22.b and the continuity of φ we have $\varphi(x)-\varphi(a) = \int_a^b 1_S(t)d\varphi(t) = \mu_\varphi(S) = \int_a^x g(t)dt$.

Pincus's Pathology (Optional)

29.37. The following three principles are equivalent to one another, in the sense of weak forms of the Axiom of Choice. All of these principles assert the existence of intangibles — i.e., we can use the Axiom of Choice to prove that these objects exist, but we cannot find explicitly constructible examples of any of these objects. Condition (A), below, is a consequence of the Hahn-Banach Theorem, as we noted in 23.10.

Recall that a *measurable space* is a set equipped with a σ-algebra of subsets.

(A) $(\ell_\infty)^* \supsetneqq \ell_1$ — that is, there exists a bounded linear functional on ℓ_∞ that cannot be represented by a member of ℓ_1.

(B) There exists a measurable space $(\Omega, \mathcal{S})$ and a bounded scalar-valued charge on $\mathcal{S}$ that is not a measure.

(C) There exists a probability charge on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ that vanishes on finite sets.

Proof (based on Wagon [1985]). For (C) $\Rightarrow$ (A), it is an easy exercise to show that if $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ is a bounded charge that is not countably additive, then the Bartle integral $\Lambda(f) = \int_\mathbb{N} f\, d\mu$ (defined as in 29.30) satisfies the requirements of (A). For (A) $\Rightarrow$ (B), it is an easy exercise to show that if $\Lambda : \ell_\infty \to \mathbb{F}$ satisfies (A), then $\mu(S) = \Lambda(1_S)$ satisfies (B), where 1_S is the characteristic function of any set $S \subseteq \mathbb{N}$.

It remains to show (B) $\Rightarrow$ (C). Let κ be the given scalar-valued charge. If the scalar field is $\mathbb{C}$, then we may write $\kappa = \mathrm{Re}\,\kappa + i\mathrm{Im}\,\kappa$; thus at least one of $\mathrm{Re}\,\kappa$, $\mathrm{Im}\,\kappa$ is a real-valued bounded charge μ that is not countably additive. Next, use the Jordan Decomposition $\mu = \mu^+ - \mu^-$ (see 8.42.f and 11.47); thus at least one of μ^+, μ^- is a positive, real valued, bounded charge ν that is not countably additive. Since ν is positive and finitely additive, for any disjoint measurable sets $S_1, S_2, S_3, \ldots \subseteq \Omega$ we have

$$\sum_{j=1}^{N} \nu(S_j) \quad = \quad \nu\left(\bigcup_{j=1}^{N} S_j\right) \quad \le \quad \nu\left(\bigcup_{j=1}^{\infty} S_j\right),$$

hence $\sum_{j=1}^\infty \nu(S_j) \le \nu\left(\bigcup_{j=1}^\infty S_j\right)$, and similarly $\sum_{j\in J}^\infty \nu(S_j) \le \nu\left(\bigcup_{j\in J}^\infty S_j\right)$ for any set $J \subseteq \mathbb{N}$. On the other hand, since ν is not countably additive, there is some sequence (S_j)

that satisfies $\sum_{j=1}^{\infty}\nu(S_j) \neq \nu\left(\bigcup_{j=1}^{\infty} S_j\right)$, and therefore $\sum_{j=1}^{\infty}\nu(S_j) < \nu\left(\bigcup_{j=1}^{\infty} S_j\right) < \infty$. Define

$$\lambda(J) \quad = \quad \nu\left(\bigcup_{j\in J} S_j\right) - \sum_{j\in J}\nu(S_j) \qquad\qquad \text{for all } J\subseteq\mathbb{N}.$$

Then λ takes values in $[0,+\infty)$ and, in particular, $\lambda(\mathbb{N})>0$. Also, λ is finitely additive but vanishes on finite subsets of $\mathbb{N}$, since the mappings $J\mapsto\nu\left(\bigcup_{j\in J}S_j\right)$ and $J\mapsto\sum_{j\in J}\nu(S_j)$ are both finitely additive. Finally, let $\pi(J)=\nu(J)/\nu(\mathbb{N})$; then π is the required probability.

29.38. We shall now show that any of the equivalent principles listed in 29.37 implies the following principle:

> **(NBP)** There exists a subset of $\{0,1\}^{\mathbb{N}}$ that lacks the Baire property.

Here $\{0,1\}$ has the discrete topology and $\{0,1\}^{\mathbb{N}}$ has the product topology, as usual.

This implication was first stated without proof in Solovay [1970]; the first published proof apparently is that of Pincus [1974]. The slightly shorter proof below is essentially that of Taylor; it was published in Wagon [1985].

Proof. We assume 29.37(C). As usual, we shall identify $\mathcal{P}(\mathbb{N})=\{\text{subsets of }\mathbb{N}\}$ with $\{0,1\}^{\mathbb{N}}=\{\text{sequences of 0s and 1s}\}$, by identifying each subset of $\mathbb{N}$ with its characteristic function. However, for clarity we shall use different notations in these two settings. For any sequence $a=(a_1,a_2,a_3,\ldots)\in\{0,1\}^{\mathbb{N}}$, the corresponding member of $\mathcal{P}(\mathbb{N})$ is the set $N(a)=\{j\in\mathbb{N} \,:\, a_j=1\}$. Similarly, for any finite sequence $a=(a_1,a_2,\ldots,a_m)$ of 0s and 1s, the corresponding set is $N_m(a)=\{j\in\{1,2,\ldots,m\} \,:\, a_j=1\}$.

The given probability measure $\mu \;:\; \mathcal{P}(\mathbb{N})\to[0,1]$ yields a corresponding function $\nu \,:\, \{0,1\}^{\mathbb{N}}\to[0,1]$. Let $T=\{a\in\{0,1\}^{\mathbb{N}} \,:\, \nu(a)=0\}$. We shall show that T lacks the Baire property. Assume, on the contrary, that T has the Baire property; we shall obtain a contradiction.

Since the measure μ vanishes on finite subsets of $\mathbb{N}$, it follows that ν takes the same value on any two sequences that differ in only finitely many terms; thus T is a tail set in $\{0,1\}^{\mathbb{N}}$. By 20.33, T is either meager or comeager.

Note that if a' is the sequence obtained from a by switching all the 0s to 1s and all the 1s to 0s, then $\nu(a)+\nu(a')=1$, since μ is finitely additive. Also, the mapping $a\mapsto a'$ is a homeomorphism from $\{0,1\}^{\mathbb{N}}$ onto itself, preserving all topological properties. If T is comeager, then the set

$$U \quad = \quad \{0,1\}^{\mathbb{N}}\setminus T \quad = \quad \{a\in\{0,1\}^{\mathbb{N}} \,:\, \nu(a)>0\}$$

is meager, and so the set

$$U' \quad = \quad \{a' \,:\, a\in U\} \quad = \quad \{b\in\{0,1\}^{\mathbb{N}} \,:\, \nu(b)<1\}$$

is also meager. But then $\{0,1\}^{\mathbb{N}}=U\cup U'$ is also meager, contradicting 20.17. Thus T cannot be comeager.

Hence T is meager. Say $T=\bigcup_{p\in\mathbb{N}}Q_p$, where each Q_p is nowhere-dense in $\{0,1\}^{\mathbb{N}}$.

When a and b are finite sequences, let $a \oplus b$ denote their concatenation — i.e., the sequence a followed by the sequence b. For instance, $(1,2,3) \oplus (4,5) = (1,2,3,4,5)$. An *extension* of a finite sequence $(a_1, a_2, \ldots, a_m)$ is a longer sequence, either finite or infinite, whose first m terms are $a_1, a_2, \ldots, a_m$ in that order. For any finite or infinite sequence $a = (a_1, a_2, a_3, \ldots)$, let $\pi_j(a) = a_j$.

We shall now recursively construct integers $0 = \lambda(0) < \lambda(1) < \lambda(2) < \cdots$ and functions $F_p : \{0,1\}^p \to \{0,1\}^{\lambda(p)}$ with certain properties described below.

(i) The function F_p maps sequences of 0s and 1s of length p to sequences of 0s and 1s of length $\lambda(p)$. (The function F_0 maps the empty sequence to the empty sequence.)

(ii) If $q > p$ and $b \in \{0,1\}^q$ is an extension of $a \in \{0,1\}^p$, then $F_q(b)$ is an extension of $F_p(a)$.

(iii) When p is a positive integer, no infinite sequence contained in Q_p will be an extension of any of the finite sequences in $\mathrm{Range}(F_p)$.

(iv) For each positive integer $j \in \{1, 2, \ldots, \lambda(p)\}$, there is at most one sequence a in $\{0,1\}^p$ with the property that $\pi_j(F_p(a)) = 1$. In other words, the sets $N_{\lambda(p)}(F_p(a))$ $(a \in \{0,1\}^p)$ are disjoint.

The definition of $\lambda(0)$ and F_0 is clear. Now suppose that some $\lambda(p)$ and F_p have already been specified, satisfying the conditions above; we wish to determine $\lambda(p+1)$ and F_{p+1}. They will be constructed in several steps.

As a first step, define a function G_p on $\{0,1\}^{p+1}$ by taking

$$G_p(a \oplus b) \;=\; F_p(a) \oplus a \oplus b \qquad \text{for } a \in \{0,1\}^p \text{ and } b \in \{0,1\}.$$

In other words,

$G_p(000\cdots00)$	consists of the sequence	$F_p(000\cdots0)$	followed by	$000\cdots00$
$G_p(000\cdots01)$	consists of the sequence	$F_p(000\cdots0)$	followed by	$000\cdots01$
$G_p(000\cdots10)$	consists of the sequence	$F_p(000\cdots1)$	followed by	$000\cdots10$
$G_p(000\cdots11)$	consists of the sequence	$F_p(000\cdots1)$	followed by	$000\cdots11$
	$\cdots$		$\cdots$	
$G_p(111\cdots10)$	consists of the sequence	$F_p(111\cdots1)$	followed by	$111\cdots10$
$G_p(111\cdots11)$	consists of the sequence	$F_p(111\cdots1)$	followed by	$111\cdots11$

Thus, all the sequences $G_p(a)$ (for $a \in \{0,1\}^{p+1}$) have length $\lambda(p) + 2^{p+1}$.

For our next step, we shall extend all these sequences $G_p(a)$ to new sequences $H_p(a)$, all of which have different lengths. Let $a_1, a_2, \ldots, a_{2^{p+1}}$ be the 2^{p+1} elements of $\{0,1\}^{p+1}$, listed in any convenient order. Say we have already formed extensions

$$H_p(a_1),\; H_p(a_2),\; \ldots,\; H_p(a_{n-1})$$

for some n (or take $n = 1$ if we haven't formed any of these extensions yet). We now wish to extend $G_p(a_n)$ to a longer sequence $H_p(a_n)$ having certain properties. First, add 0s to the end of the sequence $G_p(a_n)$, to make a sequence that is as long as any of the extensions $H_p(a_1)$, $H_p(a_2)$, ..., $H_p(a_{n-1})$ already formed (or skip this operation if $n = 1$). Now, by 20.5.c, we can extend the sequence further, to get a new sequence $H_p(a_n)$ having the property that no infinite extension of this new sequence $H_p(a)$ is a member of Q_{p+1}.

With this construction, we now have

$$[\text{length of } H_p(a_1)] \quad < \quad [\text{length of } H_p(a_2)] \quad < \quad \cdots \quad < \quad [\text{length of } H_p(a_{2^{p+1}})].$$

Thus, the longest of these sequences is $H_p(a_{2^{p+1}})$. Take the length of that sequence as our definition of $\lambda(p+1)$. Now extend all the other $H_p(a_n)$'s to that length by adding 0s to the ends of those sequences. The resulting sequences are our definition of the $F_{p+1}(a_n)$'s. This completes our recursive construction of the F_p's and $\lambda(p)$'s.

Now define a mapping $F_\infty : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ as follows: For each $a = (a_1, a_2, a_3, \ldots)$ in $\{0,1\}^{\mathbb{N}}$, let $F_\infty(a)$ be the sequence that is an extension of all the finite sequences $F_p(a_1, a_2, \ldots, a_p)$ for $p = 1, 2, 3, \ldots$. The set $B = \text{Range}(F_\infty) \subseteq \{0,1\}^{\mathbb{N}}$ now has these properties:

(a) The set B is disjoint from all the Q_p's and hence from T. (This follows from property (iii) of the F_p's.)

(b) If a, b are two distinct members of B, then $N(a) \cap N(b) = \{j \in \mathbb{N} : a_j = b_j = 1\}$ is a finite set. (This follows from property (iv) of the F_p's.)

(c) B is uncountable. (Indeed, F_∞ is injective, because of our method of constructing the G_p's from the F_p's.)

For each $b \in B$, we have $b \notin T$, and therefore $\nu(b) > 0$. Hence $\nu(b) > \frac{1}{k}$ for some positive integer b. Since B is uncountable, there is some positive integer k such that $B_k = \{b \in A : \nu(b) > \frac{1}{k}\}$ is uncountable. Hence that B_k has at least k distinct elements $b_1, b_2, \ldots, b_k \in \{0,1\}^{\mathbb{N}}$. If we change finitely many 1s to 0s in any fashion whatsoever, then the resulting new sequences $c_1, c_2, \ldots, c_k \in \{0,1\}^{\mathbb{N}}$ will also satisfy $\nu(c_j) > \frac{1}{k}$. However, by property (b) it is possible to choose the c_j's so that their corresponding sets of integers $S_j = N(c_j)$ are pairwise disjoint. Thus $\mu(S_j) > \frac{1}{k}$ while

$$\mu(S_1) + \mu(S_2) + \cdots + \mu(S_k) \quad = \quad \mu(S_1 \cup S_2 \cup \cdots \cup S_k) \quad \leq \quad 1,$$

a contradiction.

Chapter 30

Initial Value Problems

30.1. *Overview.* This chapter is concerned with initial value problems in a Banach space X; these take the form

$$\left.\begin{array}{ll} u'(t) = f(t, u(t)) & (0 \le t \le T), \\ u(0) = x_0. & \end{array}\right\} \qquad \text{(IVP)}$$

Here x_0 is a given element of X, called the *initial value*, and f is a given mapping (not necessarily linear or continuous in either variable) from some subset of $\mathbb{R} \times X$ into X. The function u is the *solution* of the initial value problem — it is not necessarily given; it may be viewed as the "answer" of the initial value problem. The problem (IVP) is called *nonautonomous* because the function $f(t, x)$ may depend on both t and x. We shall devote some extra attention to the *autonomous* problem

$$\left.\begin{array}{ll} u'(t) = f(u(t)) & (0 \le t \le T), \\ u(0) = x_0 & \end{array}\right\} \qquad \text{(AIVP)}$$

wherein f is a mapping from some subset of X into X.

The vector $u(t)$ may represent the state of some system at time t. The differential equation $u'(t) = f(t, u(t))$ is also called an *evolution equation*, because it describes how the system evolves as time passes. In the real world, all things change as time passes, so the initial value problem described above is very general; it represents many different phenomena.

Various questions can be asked about the solution u. For instance: What additional assumptions about X, f, x_0 are sufficient to ensure that a solution exists? That the solution is unique? That the solution depends continuously on the data — i.e., that a small change in f or x_0 results in only a small change in the solution? Can the solution be found exactly, or approximated by some effective algorithm? How quickly do the approximations converge? Some of these questions can be addressed and partially answered in a very general and abstract setting. This chapter will concentrate mainly on the existence of solutions.

Here is a preview of some basic results:

- If f is continuous and X is finite-dimensional, then a solution exists — see 30.12.
- Continuity of f does not imply existence of solutions in infinite-dimensional spaces — see 30.4.

- However, in any Banach space, *most* differential equations with continuous right-hand sides have solutions — see 30.11.
- Existence of solutions is guaranteed by certain additional assumptions — e.g., Lipschitz conditions (see 30.9), compactness (see 30.12), isotonicity (see 30.14), or dissipativeness (see 30.28). Some of these conditions do not require continuity of f.

A more extensive survey can be found in Schechter [1989].

Elementary Pathological Examples

30.2. Nonuniqueness of solutions. Even if an initial value problem has a solution, the solution might not be unique. For instance, in the Banach space $X = \mathbb{R}$, consider the autonomous initial value problem

$$\begin{cases} u'(t) = \sqrt{u(t)} & (t \geq 0), \\ u(0) = 0. \end{cases}$$

This problem has infinitely many solutions. For each number $c \geq 0$, a solution $u : [0, +\infty) \to \mathbb{R}$ is given by the function

$$u(t) \quad = \quad \begin{cases} 0 & (0 \leq t \leq c) \\ \frac{1}{4}(t-c)^2 & (c \leq t). \end{cases}$$

30.3. Existence only locally, not globally. Even if a differential equation has a solution for every initial value, the solution might not exist for all time. For instance, in the Banach space $X = \mathbb{R}$, consider the autonomous initial value problem

$$\begin{cases} u'(t) = \big(u(t)\big)^2 & (t \geq 0), \\ u(0) = x_0. \end{cases}$$

It is easy to verify that the solution is given by

$$\begin{array}{lll} u(t) = (x_0^{-1} - t)^{-1} & \text{for all } t \geq 0 & \text{if } x_0 < 0, \\ u(t) = 0 & \text{for all } t \geq 0 & \text{if } x_0 = 0, \\ u(t) = (x_0^{-1} - t)^{-1} & \text{for all } t \in [0, x_0^{-1}) & \text{if } x_0 > 0. \end{array}$$

Moreover, the solution is *uniquely* determined by x_0; that fact will follow from 30.9. Note that if $x_0 > 0$, then a solution does not exist for all positive time; rather, the solution $u(t)$ blows up when t increases to the finite time x_0^{-1}. Thus, we say a solution exists *locally* in time, but perhaps not *globally*.

30.4. Existence not even locally. Let X be a Banach space, and let $f : X \to X$ be continuous. The autonomous initial value problem (AIVP) in 30.1 has a solution under certain additional assumptions — e.g., if f is locally Lipschitzian (see 30.9) or X is finite-dimensional (see 30.12) — but without additional assumptions, a solution might not exist even locally. The following example was given by Dieudonné [1950].

Let $X = c_0 = \{\text{sequences that converge to } 0\}$; this is a Banach space when equipped with the sup norm. Define

$$f(x_1, x_2, x_3, \ldots) \quad = \quad \left(\sqrt{|x_1|}, \quad \sqrt{|x_2|}, \quad \sqrt{|x_3|}, \quad \ldots\right).$$

It is easy to verify that f is a continuous map from X into X. However, we claim that the autonomous initial value problem (AIVP) with initial value $x_0 = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right)$ does not have any solution for any $T > 0$.

Indeed, suppose that $u(t) = (u_1(t), u_2(t), u_3(t), \ldots)$ were such a solution. Then each component u_j would be the solution of this one-dimensional initial value problem:

$$\left.\begin{array}{ll} u_j'(t) = \sqrt{|u_j(t)|} & (0 \leq t \leq T), \\ u_j(0) = 1/j. & \end{array}\right\}$$

Since the function $x \mapsto \sqrt{|x|}$ is positive, any solution u_j of this problem must be increasing. Since $x \mapsto \sqrt{|x|}$ is Lipschitzian on $[\frac{1}{j}, +\infty)$, there is a unique solution for $t \geq 0$, by 30.9. It is easy to verify that that solution is

$$u_j(t) \quad = \quad \left(\frac{t}{2} + \frac{1}{\sqrt{j}}\right)^2.$$

But then $u_j(t) > t^2/4$, so $u(t) \notin c_0$ for any $t > 0$.

Further remarks. After Dieudonné published this example for c_0, other mathematicians gave similar examples in other spaces. Finally Godunov [1975] proved that

> if X is any infinite-dimensional Banach space, then there exists $x_0 \in X$ and a continuous function $f : \mathbb{R} \times X \to X$ such that the initial value problem (IVP) does not have any solution for any $T > 0$.

The proof of that result is long and complicated; it will not be given in this book.

Carathéodory Solutions

30.5. *A precise notion of "solution."* The term "solution" has many different meanings in the literature. Precision was not needed for the preceding elementary examples — they would make sense with any reasonable notion of "solution" — but for the theorems developed later in this chapter we will need precise definitions.

The most obvious kind of "solution" for a differential equation is a continuously differentiable function that satisfies the equation. However, we will find it advantageous to study a slightly weaker and more general notion of "solution:"

Let X be a Banach space, and let f be a function (not necessarily continuous) from some subset of $\mathbb{R} \times X$ into X. By a **Carathéodory solution** of the differential equation $u'(t) = f(t, u(t))$ on an interval $J \subseteq \mathbb{R}$, we mean a function $u : J \to X$ such that $\mathrm{Graph}(u) \subseteq \mathrm{Dom}(f)$ and

> (*) whenever $a < b$ in J, then the Bochner integral $\int_a^b f(t, u(t))dt$ exists and equals $u(b) - u(a)$.

This notion of "solution" will suffice for most of this chapter. (More general notions of "solution" will be introduced in 30.16.)

Any Carathéodory solution is continuous, by 24.41.a. Furthermore, if u is a Carathéodory solution of $u'(t) = f(t, u(t))$ on J, then

> (!) for almost every $t \in J$, the Fréchet derivative $u'(t)$ exists and equals $f(t, u(t))$,

by 25.16. We remark that condition (!) is slightly weaker than (*). For many purposes, integrals are easier to work with than derivatives; thus they yield a simpler and more satisfactory theory. We retain the differential equation $u'(t) = f(t, u(t))$ as a shorthand notation and as a source of intuition, but the theory developed below is really concerned with integral equations.

30.6. The question of whether a solution exists globally in time (as in 30.2) often can be separated into two component questions: (i) Do solutions exist at least locally in time (as in 30.3)? and (ii) Can solutions be continued further in time (as in the next theorem below)?

Definition. Let Ω be a subset of a Banach space X, and let $f : \mathbb{R} \times \Omega \to X$ be some mapping. We shall say that f is **locally generative** on Ω if:

> For each $(t_0, x_0) \in \mathbb{R} \times \Omega$, there exist some $\varepsilon > 0$ and some Carathéodory solution of $u'(t) = f(t, u(t))$ on the interval $[t_0, t_0 + \varepsilon]$ with $u(t_0) = x_0$.

Remarks. (1) We emphasize that no assertion is made about *uniqueness* of the solution. (2) Necessary and sufficient conditions for a mapping f to be locally generative are not known. Later in this chapter we shall give several different *sufficient* conditions for f to be locally generative, using Lipschitzness, compactness, or isotonicity conditions.

Continuability Theorem. Let X, Ω, f be as above; assume f is locally generative on Ω. Then for each $(t_0, x_0) \in \mathbb{R} \times \Omega$, there exists some Carathéodory solution of $u'(t) = f(t, u(t))$ on an interval $[t_0, t_1)$, satisfying the initial condition $u(t_0) = x_0$, which is "noncontinuable" (i.e., it cannot be continued further) because it satisfies at least one of the following three conditions:

(i) $t_1 = +\infty$.

(ii) $u(t_1) = \lim_{t \uparrow t_1} u(t)$ exists and lies outside Ω.

(iii) $\int_{t_0}^{t_1} \|f(t, u(t))\| dt = +\infty$, and thus the Bochner integral $\int_{t_0}^{t_1} f(t, u(t))dt$ cannot exist.

Remarks. In many applications, additional information about u or f or Ω allows us to eliminate possibilities (ii) or (iii). For instance, (ii) cannot occur if Ω is closed, and (iii) cannot occur if f is bounded and t_1 is finite.

Proof of theorem. Let any $(t_0, x_0) \in \mathbb{R} \times \Omega$ be given. Consider all Carathéodory solutions u of the differential equation $u'(t) = f(t, u(t))$ with initial condition $u(t_0) = x_0$, on intervals of the form $[t_0, t_1]$ or $[t_0, t_1)$. Partially order these solutions by inclusion of their graphs — i.e., $u_1 \preccurlyeq u_2$ if $\mathrm{Graph}(u_1) \subseteq \mathrm{Graph}(u_2)$. It is easy to see that Zorn's Lemma is applicable, and thus there exists a maximal solution, which we shall denote by u.

First, we show that the domain of u is not an interval of the form $[t_0, t_1]$ for some $t_1 < \infty$. Indeed, if it were, we could use the local generativeness of f to find a solution on $[t_1, t_1 + \varepsilon]$ with initial value $u(t_1)$. Combining the two solutions yields a solution on $[t_0, t_1 + \varepsilon]$, contradicting the maximality of $[t_0, t_1]$.

Thus, a maximal solution u exists, with domain of the form $[t_0, t_1)$. Now assume that none of conditions (i), (ii), or (iii) are satisfied; we shall obtain a contradiction.

Since $\int_{t_0}^{t_1} \|f(t, u(t))\| dt < \infty$, the quantity $\int_a^b \|f(t, u(t))\| dt$ must become small as a and b increase to t_1. Therefore $\|u(b) - u(a)\| = \|\int_a^b f(t, u(t)) dt\|$ becomes small, and $u(t)$ is Cauchy as $t \uparrow t_1$. Thus $u(t_1) = \lim_{t \uparrow t_1} u(t)$ exists. By our assumption, $u(t_1) \in \Omega$. But then we have a Carathéodory solution on $[t_0, t_1]$, contradicting the maximality of $[t_0, t_1)$. This completes the proof.

30.7. Carathéodory solutions as fixed points. A Carathéodory solution of the initial value problem (IVP) in 30.1 is a function $u : [0, T] \to X$ that satisfies the integral equation

$$u(t) = x_0 + \int_0^t f(s, u(s)) ds \qquad \text{for all} \quad t \in [0, T]. \tag{IE}$$

This integral equation often can be transformed to a fixed point problem in the following fashion: Let

$$C([0, T], X) \quad = \quad \{\text{continuous functions from } [0, T] \text{ into } X\}.$$

Define a operator Φ from some appropriate subset of $C([0, T], X)$ into $C([0, T], X)$ by the formula

$$[\Phi(u)](t) \quad = \quad x_0 + \int_0^t f(s, u(s)) ds \qquad \text{for all } t \in [0, T].$$

Then a solution of the integral equation (IE) is the same thing as a function u that satisfies $\Phi(u) = u$; that is, a fixed point of Φ.

Thus, to prove the existence and other basic properties of a function $u : [0, T] \to X$ in an abstract setting, we shall apply theorems about fixed points for a function $\Phi : \mathrm{Dom}(\Phi) \to C([0, T], X)$ in an even more abstract setting. In the next few pages, we shall apply fixed point theorems of Banach, Vidossich, Schauder, and Tarski, to obtain several different easy results about initial value problems. Deeper theorems about differential equations can be proved by longer, more specialized methods, which do not involve fixed point theorems. One such result is the Crandall-Liggett Theorem, presented in 30.28.

Lipschitz Conditions

30.8. Gronwall's Lemma. Let δ, λ be two mappings from an interval $[0, S]$ into $[0, +\infty)$. Assume δ is continuous, λ is integrable, $C \in [0, +\infty)$, and

$$\delta(t) \quad \leq \quad C + \int_0^t \lambda(s)\delta(s)ds \qquad \text{for all } t \in [0, S].$$

Then

$$\delta(t) \quad \leq \quad C \exp\left[\int_0^t \lambda(s)ds\right] \qquad \text{for all } t \in [0, S].$$

Proof. Let $R(t) = C + \int_0^t \lambda(s)\delta(s)ds$ and $Q(t) = R(t) \exp\left[-\int_0^t \lambda(s)ds\right]$. Then $\delta(t) \leq R(t)$, and $R'(t) = \lambda(t)\delta(t) \leq \lambda(t)R(t)$ for almost all t. Also,

$$Q'(t) \;=\; [R'(t) - \lambda(t)R(t)] \exp\left[-\int_0^t \lambda(s)ds\right] \;\leq\; 0.$$

Integrating yields $Q(t) \leq Q(0) = C$; the conclusion of the lemma follows immediately.

30.9. Cauchy-Lipschitz Existence Theorem. Let G be an open subset of a Banach space X, and let $f : G \to X$ be locally Lipschitzian. Then for any $x_0 \in G$, there exists a solution of the autonomous initial value problem (AIVP) for some $T > 0$. The solution is unique and depends continuously on the initial value. In fact, if u_1, u_2 are two solutions of $u'(t) = f(u(t))$ on some interval $[0, T]$, and f has Lipschitz constant λ on the compact set $\text{Range}(u_1) \cup \text{Range}(u_2)$, then

$$\|u_1(t) - u_2(t)\| \quad \leq \quad e^{\lambda t} \, \|u_1(0) - u_2(0)\|$$

for all $t \in [0, T]$. If $G = X$ and f is Lipschitzian, then the solution can be continued for all positive time.

More generally: Let G be an open subset of a Banach space X. Let $f : [0, T] \times G \to X$ be integrably locally Lipschitz, as defined in 22.36. Then for each $x_0 \in G$, there exists a Carathéodory solution of the initial value problem

$$\begin{cases} u'(t) = f(t, u(t)) & (0 \leq t \leq S), \\ u(0) = x_0 \end{cases}$$

for some number $S = S(x_0)$ with $0 < S \leq T$. The solution is unique and depends continuously on the initial data: If u_1, u_2 are two solutions of $u'(t) = f(t, u(t))$ on any interval $[0, S]$, then

$$\|u_1(t) - u_2(t)\| \;\leq\; \|u_1(r) - u_2(r)\| \exp\left(\int_r^t \lambda_K(s)ds\right) \tag{CD}$$

for all $[r, t] \subseteq [0, S]$, where $K = \text{Range}(u_1) \cup \text{Range}(u_2)$ and λ_K is as in 22.36. If $G = X$ and f is integrably Lipschitz (as defined in 22.36), then we can take $S = T$ — that is, the solution is continuable across the entire time interval where f is defined.

Proof. First we shall prove continuous dependence. Suppose u_1 and u_2 are two Carathéodory solutions of $u'(t) = f(t, u(t))$ on an interval $[0, S]$, and let $K = \text{Range}(u_1) \cup \text{Range}(u_2)$. Let $\delta(t) = \|u_1(t) - u_2(t)\|$. Then

$$\delta(t) \;=\; \left\| u_1(r) - u_2(r) + \int_r^t [f(u_1(s)) - f(u_2(s))]ds \right\| \;\leq\; \delta(r) + \int_r^t \lambda_K(s)\delta(s)ds$$

for any $[r, t] \subseteq [0, S]$. Hence Gronwall's Inequality 30.8 applies; it proves the continuous dependence condition (CD).

Next we prove local existence. Let $x_0 \in G$ be given. Let B be the closed ball centered at x_0 with radius R; for R sufficiently small we have $B \subseteq G$. By the proposition in 22.36, with R sufficiently small there is some integrable function φ such that whenever $u, v : [0, T] \to B$ are continuous functions, then

$$\|f(t, u(t)) - f(t, v(t))\| \leq \varphi(t)\|u(t) - v(t)\| \qquad \text{for almost all } t.$$

In particular, taking $v = x_0$, we see that u also satisfies

$$\|f(t, u(t))\| \quad \leq \quad \|f(t, x_0)\| + (R + \|x_0\|)\, \varphi(t) \quad \stackrel{\text{def}}{=} \quad \gamma(t);$$

the function γ thus defined is integrable too. Now choose some $S > 0$ small enough so that $S \leq T$ and $\int_0^S \varphi(t)dt < 1$ and $\int_0^S \gamma(t)dt \leq R$. Let $C([0, S], X)$ and $C([0, S], B)$ be the sets of all continuous functions from $[0, S]$ into X and into B, respectively; then $C([0, S], X)$ is a Banach space (with the sup norm) and $C([0, S], B)$ is a closed subset of that Banach space. Since $B \subseteq G$, for each $u \in C([0, S], B)$ we can define

$$(\Phi u)(t) \quad = \quad x_0 + \int_0^t f(s, u(s))ds \qquad (0 \leq t \leq S).$$

Since $\int_0^S \gamma \leq R$, it follows easily that Φ maps $C([0, S], B)$ into itself. Also, for any $u_1, u_2 \in C([0, S], B)$ we have

$$\begin{aligned} \|(\Phi u_1)(t) - (\Phi u_2)(t)\| \quad &= \quad \left\| \int_0^t [f(s, u_1(s)) - f(s, u_2(s))]ds \right\| \\ &\leq \quad \int_0^t \|f(s, u_1(s)) - f(s, u_2(s))\|ds \quad \leq \quad \|u_1 - u_2\|_\infty \int_0^t \varphi(s)ds \end{aligned}$$

and therefore $\langle \Phi \rangle_{\text{Lip}} \leq \int_0^S \varphi \cdot < 1$. By Banach's Theorem on strict contractions (19.39), Φ has at least one fixed point $u \in C([0, S], B)$; thus the initial value problem has at least one solution.

If f is integrably Lipschitz and $X = G$, we shall modify the local existence argument of the preceding paragraph to prove the continuability result. Take $R = \infty$ and $B = X$. Since f is integrably Lipschitz, there is a function $\varphi \in L^1[a, b]$ (which does not depend on the choice of x_0) such that

$$\|f(t, x_1) - f(t, x_2)\| \;\leq\; \varphi(t)\, \|x_1 - x_2\| \qquad \text{for all } t \in [a, b],\; x_1, x_2 \in X.$$

The condition $\int_0^S \gamma(t)dt \leq R$ may be omitted, since $R = \infty$. Form a partition $0 = t_0 < t_1 < t_2 < \cdots < t_m = T$ fine enough so that $\int_{t_{j-1}}^{t_j} \varphi(t)dt < 1$ for each j. The argument of the preceding paragraph establishes the existence of a Carathéodory solution to each of the initial value problems

$$\begin{cases} u'(t) = f(t, u(t)) & (t_{j-1} \leq t \leq t_j), \\ u(t_{j-1}) = x_{j-1} \end{cases}$$

for $j = 1, 2, \ldots, m$. Use the final value of one initial value problem as the initial value of the next problem. Put the m solutions together to obtain a solution on $[0, T]$.

30.10. Theorem on continuous dependence. Let G be an open subset of a Banach space $(X, \|\ \ \|)$. Let $g : [0, T] \times G \to X$ be a function that is integrably locally Lipschitz (as defined in 22.36). Let $f_1, f_2, f_3, \ldots$ be functions from $[0, T] \times G$ into X (not necessarily satisfying any Lipschitz conditions or other regularity conditions). Let $u_1, u_2, u_3, \ldots$ and v be Carathéodory solutions of

$$u_n'(t) = f_n(t, u_n(t)), \qquad\qquad v'(t) = g(t, v(t))$$

on $[0, T]$. Suppose that $u_n(0) \to v(0)$ as $n \to \infty$, and $f_n \to g$ uniformly on $[0, T] \times G$ as $n \to \infty$. Then $u_n \to v$ uniformly on $[0, T]$ as $n \to \infty$.

Proof. Since Range(v) is a compact set, by 22.36 there exist some number $r > 0$ and some function $\varphi \in L^1[0, T]$ with this property: Whenever p_1 and p_2 are continuous functions from $[0, T]$ into G such that

$$\max_{j=1,2} \max_{0 \leq t \leq T} \operatorname{dist}(p_j(t), \operatorname{Range}(v)) \quad \leq \quad r,$$

then $\|g(t, p_1(t)) - g(t, p_2(t))\| \leq \varphi(t)\|p_1(t) - p_2(t)\|$ for almost all t.

Choose some partition $0 = t_0 < t_1 < t_2 < \cdots < t_m = T$ that is fine enough so that $\int_{t_{j-1}}^{t_j} [1 + 2\varphi(s)]ds < 1/2$ for all j. Fix any j, and assume that $u_n(t_{j-1}) \to v(t_{j-1})$ as $n \to \infty$; it suffices to show that $u_n \to v$ uniformly on $[t_{j-1}, t_j]$.

Let any ε in $(0, r/2)$ be given; choose n large enough so that $\|u_n(t_{j-1}) - v(t_{j-1})\| \leq \varepsilon$ and $\|f_n - g\|_\infty \leq \varepsilon$; it suffices to show that $\|u_n(t) - v(t)\| \leq 2\varepsilon$ for all $t \in [t_{j-1}, t_j]$. Suppose the contrary; let τ be the first point in $[t_{j-1}, t_j]$ satisfying $\|u_n(\tau) - v(\tau)\| \geq 2\varepsilon$. Then for all s in $[t_{j-1}, \tau)$, we have $\|u_n(s) - v(s)\| < 2\varepsilon < r$. For such s we have

$$\begin{aligned} &\|f_n(s, u_n(s)) - g(s, v(s))\| \\ &\quad \leq \quad \|f_n(s, u_n(s)) - g(s, u_n(s))\| + \|g(s, u_n(s)) - g(s, v(s))\| \\ &\quad \leq \quad \|f_n - g\|_\infty + \varphi(s)\|u_n(s) - v(s)\| \quad \leq \quad \varepsilon + 2\varepsilon\varphi(s). \end{aligned}$$

Therefore

$$\begin{aligned} 2\varepsilon \quad &= \quad \|u_n(\tau) - v(\tau)\| \\ &= \quad \left\| u_n(t_{j-1}) - v(t_{j-1}) + \int_{t_{j-1}}^{\tau} [f_n(s, u_n(s)) - g(s, v(s))]ds \right\| \\ &\leq \quad \|u_n(t_{j-1}) - v(t_{j-1})\| + \int_{t_{j-1}}^{\tau} [\varepsilon + 2\varepsilon\varphi(s)]ds \quad \leq \quad \varepsilon + \frac{1}{2}\varepsilon, \end{aligned}$$

which is a contradiction. This completes the proof.

Generic Solvability

30.11. *Notation.* In the following discussion, whenever Ω is a topological space and X is a Banach space, let

$$BC(\Omega, X) \quad = \quad \{\text{bounded, continuous functions from } \Omega \text{ into } X\};$$

this is a Banach space when equipped with the sup norm.

Theorem on generic solvability. Let X be a Banach space, let $x_0 \in X$, and let T be a positive number. Then there exists a comeager set $F \subseteq BC([0,T] \times X,\ X)$ with the following properties: For each $f \in F$, the initial value problem (IVP) in 30.1 has a unique solution $u_f \in BC([0,T],\ X)$, and the solution map $f \mapsto u_f$ is continuous from F into $BC([0,T],\ X)$. (Thus, "most" differential equations with continuous right-hand sides have unique solutions, and the solutions depend continuously on the right-hand sides.)

Remark. A similar result was first proved by Lasota and Yorke [1973]; our own proof follows that of Vidossich [1974]. For related results and additional references, see Myjak [1983].

Proof of theorem. Each f in $B([0,T] \times X,\ X)$ can be used to define a continuous mapping $\Phi_f\ :\ BC([0,T],\ X) \to BC([0,T],\ X)$ by the rule

$$\Big[\Phi_f(u)\Big](t) \quad = \quad x_0 + \int_0^t f(s, u(s))ds.$$

A solution of (IVP) is the same thing as a fixed point of Φ_f. It suffices to verify the hypotheses of 20.10. Note that different f's yield different Φ_f's. Indeed, if $f_1(\tau,\xi) \neq f_2(\tau,\xi)$, then any continuous function u with $u(\tau) = \xi$ will yield $\Phi_{f_1}(u) \neq \Phi_{f_2}(u)$; verifying this is an easy exercise.

Let Ψ be the set of all such mappings Φ_f. Then Ψ is a bijective copy of $B([0,T]\times X,\ X)$, and so we may topologize Ψ by copying the topology of $BC([0,T] \times X,\ X)$. Then Ψ is a complete metric space. It is easy to verify that the topology on Ψ is stronger than the topology of uniform convergence on $BC([0,T],\ X)$.

When f is locally Lipschitz, then (IVP) has a unique solution — i.e., Φ_f has a unique fixed point. Define Ψ_0 as in 20.10; then Ψ_0 contains the locally Lipschitz functions, by 30.10. The locally Lipschitz maps from $[0,T] \times X$ into X are dense in $BC([0,T] \times X,\ X)$, by 18.6.c. This completes the proof.

Compactness Conditions

30.12. Peano's Existence Theorem. Let Ω be an open subset of a Banach space X.

Let $f : [0, +\infty) \times \Omega \to X$ be jointly continuous (or, more generally, assume f is jointly measurable and $f(t, \cdot) : \Omega \to X$ is continuous for each fixed t). Suppose that

> (CM) f is a **compact mapping** — i.e., f maps bounded sets to relatively compact sets.

Let $x_0 \in \Omega$. Then the initial value problem (IVP) in 30.1 has at least one solution for some $T > 0$.

Remarks. We can weaken (CM) slightly: it suffices to assume that f maps *closed*, bounded sets to relatively compact sets. This condition can be omitted altogether if X is finite-dimensional and f is jointly continuous, since this condition follows from those assumptions — see 17.7.h and 17.17. The finite-dimensional case can be found in most books on ordinary differential equations. Hypothesis (CM) can *not* be removed when X is infinite-dimensional; see 30.4.

Proof of theorem. Let R be some positive number small enough so that the closed ball B centered at x_0 with radius R is contained in Ω. Then $f([0,1] \times B)$ is relatively compact, hence bounded; say $\|u\| \leq M$ for all $u \in f([0,1] \times B)$. Choose some positive number $T \leq \min\{1, R/M\}$. Let $C([0,T], X)$ and $C([0,T], B)$ be the sets of continuous functions from $[0,T]$ into X and into B, respectively; then $C([0,T], X)$ is a Banach space (with the sup norm) and $C([0,T], B)$ is a closed convex subset of that Banach space. Define a mapping $\Phi : C([0,T], B) \to C([0,T], X)$ by

$$(\Phi v)(t) \quad = \quad x_0 + \int_0^t f(s, v(s))ds \qquad\qquad (0 \leq t \leq T).$$

It is easy to show that this mapping is continuous. By our choice of T, it follows easily (*exercise*) that Φ maps $C([0,T], B)$ into itself. Furthermore, $\|(\Phi v)(t) - (\Phi v)(r)\| = \| \int_r^t f(s, v(s))ds\| \leq |t - r|M$, and thus the range of Φ is equicontinuous.

Recall from 26.23.i that, in a Banach space, the closed convex hull of a compact set is compact. The set $f([0,T] \times B)$ is relatively compact; hence its closed convex hull is a compact set $K_1 \subseteq X$. Any function Φv has range contained in $x_0 + TK_1$, which is also a compact subset of X. By the Arzela-Ascoli Theorem (18.35), the range of Φ is contained in a compact set $\mathcal{K}_1 \subseteq C([0,T], X)$. Let $\mathcal{K}_2$ be the closed convex hull of $\mathcal{K}_1$; then Φ has range contained in $\mathcal{K}_3 = \mathcal{K}_2 \cap C([0,T], B)$, which is a compact convex subset of $C([0,T], X)$.

The restriction of Φ to $\mathcal{K}_3$ is a continuous self-mapping of the compact convex set $\mathcal{K}_3$. By Schauder's Fixed Point Theorem (27.19), Φ has at least one fixed point in $\mathcal{K}_3$; that fixed point is a solution of (IVP).

30.13. *Remarks on generalizations.* Instead of assuming that f maps bounded sets to relatively compact sets, we could make the weaker assumption that $\gamma\big(f(t,S)\big) \leq \omega\big(t, \gamma(S)\big)$ for all t and all bounded sets S; here γ is one of the measures of noncompactness α or β (defined in 19.19) and ω is some suitable function. Some results in this direction are given by Mönch and von Harten [1982], Heinz [1983], Banaś [1985], Song [1987], and other papers cited by those.

Isotonicity Conditions

30.14. Biles-Schechter Theorem. Let $(X, \|\ \ \|, \preccurlyeq)$ be a Dedekind complete Banach lattice. Let $[0,T]$ and X be equipped with their σ-algebra of Lebesgue-measurable sets and Borel sets, respectively. Let $f : [0,T] \times X \to X$ be a mapping with the following properties:

(i) f is jointly measurable and maps separable sets to separable sets (or, more generally, f has the property that whenever $x : [0,T] \to X$ is continuous, then the mapping $t \mapsto f(t, x(t))$ is measurable and separably valued).

(ii) For each fixed t, the function $f(t, \cdot) : X \to X$ is increasing.

(iii) There exist functions $b, c \in L^1([0,T], X)$ such that $b(t) \preccurlyeq f(t,x) \preccurlyeq c(t)$ for all $(t,x) \in [0,T] \times X$.

Then for each $x_0 \in X$, there exists a Carathéodory solution to (IVP). Among the solutions there is a largest; it is the pointwise supremum of all the solutions. We may refer to it as the *maximal solution.* In fact, it is also equal to the pointwise supremum of all the solutions of the integral inequality

$$u(t) \quad \preccurlyeq \quad x_0 + \int_0^t f(s, u(s))ds \qquad\qquad (0 \le t \le T).$$

Remarks. If we also assume

$(\natural)$ there exists a function $m \in L^1[0,T]$ such that $\|f(t,x)\| \le m(t)$ for all $(t,x) \in [0,T] \times X$,

then we do not need X to be a lattice; it suffices to assume that X is a Dedekind complete ordered Banach space whose positive cone is closed and whose topology and ordering make X a locally full space (defined in 26.52). Condition $(\natural)$ can be replaced by still other, weaker, more complicated conditions, but we shall not pursue those here.

Isotonicity conditions have not yet been used extensively in applications in the literature. We include this theorem not so much for its usefulness, but for its theoretical interest. The present argument was first given in finite dimensions by Biles [1995]; it was subsequently extended to Banach lattices by Schechter [1996].

Proof of theorem. We shall first use the fact that X is a Banach lattice to prove condition $(\natural)$; in fact we shall prove it with $m(t) = \|b(t)\| + \|c(t)\|$. The proof is just an application of ordinary lattice arithmetic. For any t, x we have

$$\begin{array}{rcccccl} f(t,x) & \preccurlyeq & c(t) & \preccurlyeq & /c(t)/ & \preccurlyeq & /b(t)/ + /c(t)/ \qquad \text{and} \\ -f(t,x) & \preccurlyeq & -b(t) & \preccurlyeq & /b(t)/ & \preccurlyeq & /b(t)/ + /c(t)/, \end{array}$$

hence

$$/f(t,x)/ \;=\; f(t,x) \vee (-f(t,x)) \;\preccurlyeq\; /b(t)/ + /c(t)/ \;=\; \Big/ /b(t)/ + /c(t)/ \Big/.$$

Then

$$\|f(t,x)\| \le \Big\| /b(t)/ + /c(t)/ \Big\| \le \Big\| /b(t)/ \Big\| + \Big\| /c(t)/ \Big\| = \|b(t)\| + \|c(t)\|.$$

This proves (♮).

Let $C([0,T],X) = \{$continuous functions from $[0,T]$ into $X\}$. Let $\sqsubseteq$ denote the pointwise ordering on $C([0,T],X)$ — that is, $u \sqsubseteq v$ if $u(t) \preccurlyeq v(t)$ for all $t \in [0,T]$. Observe that if $v \in C([0,T],X)$, then the mapping $s \mapsto f(s,v(s))$ is measurable and $\|\ \ \|$-dominated by the integrable function m; hence it is integrable. Therefore we can define the function

$$(\Phi v)(t) = x_0 + \int_0^t f(s,v(s))ds.$$

Then Φ maps $C([0,T],X)$ into itself. It is clear that a solution of the initial value problem is the same as a fixed point of Φ, and a solution of the integral inequality is the same as a solution of $u \sqsubseteq \Phi u$. Since each mapping $f(s,\cdot) : X \to X$ is $\preccurlyeq$-increasing, it follows that $\Phi : C([0,T],X) \to C([0,T],X)$ is $\sqsubseteq$-increasing. Let

$$\beta(t) = x_0 + \int_0^t b(s)ds, \qquad \gamma(t) = x_0 + \int_0^t c(s)ds;$$

those are continuous functions of t. Define the set

$$\mathcal{V} = \Big\{ v \in C\Big([0,T],X\Big) : \beta(t) \preccurlyeq v(t) \preccurlyeq \gamma(t) \quad \text{and} \\ \textstyle\int_r^t b(s)ds \preccurlyeq v(t) - v(r) \preccurlyeq \int_r^t c(s)ds \quad \text{for all} \quad [r,t] \subseteq [0,T] \Big\}.$$

Clearly, Φ maps $C([0,T],X)$ into $\mathcal{V}$; hence $\mathcal{V}$ is nonempty and Φ maps $\mathcal{V}$ into $\mathcal{V}$.

We next show $(\mathcal{V},\sqsubseteq)$ is a complete lattice. (We note that $C([0,T],X)$ is *not* Dedekind complete, in general; thus β and γ are essential for the following argument.) Let V be any nonempty subset of $\mathcal{V}$, and define $\sigma(t) = \sup\{v(t) : v \in V\}$ and $\iota(t) = \inf\{v(t) : v \in V\}$; these functions are well defined since X is Dedekind complete. It suffices to show that $\sigma \in \mathcal{V}$ (for then $\iota \in \mathcal{V}$ by similar reasoning). Clearly, $\beta(t) \preccurlyeq \sigma(t) \preccurlyeq \gamma(t)$. Fix any $[r,t] \subseteq [0,T]$. For each $v \in V$ we have

$$v(r) + \int_r^t b(s)ds \preccurlyeq v(t) \qquad \text{and} \qquad v(t) \preccurlyeq v(r) + \int_r^t c(s)ds,$$

hence

$$v(r) + \int_r^t b(s)ds \preccurlyeq \sigma(t) \qquad \text{and} \qquad v(t) \preccurlyeq \sigma(r) + \int_r^t c(s)ds,$$

hence (taking the supremum on the left side)

$$\sigma(r) + \int_r^t b(s)ds \preccurlyeq \sigma(t) \qquad \text{and} \qquad \sigma(t) \preccurlyeq \sigma(r) + \int_r^t c(s)ds,$$

and therefore $\int_r^t b(s)ds \preccurlyeq \sigma(t) - \sigma(r) \preccurlyeq \int_r^t c(s)ds$. We shall use that inequality, finally, to prove that σ is continuous. To show that σ is continuous from the right, let $t_n \downarrow r$; then $\int_r^{t_n} b(s)ds \to 0$ and $\int_r^{t_n} c(s)ds \to 0$ (see 26.52(E)). Since X is a Banach lattice (or, more generally, since X is locally full), it follows that $\sigma(t_n) - \sigma(r) \to 0$. Similarly, σ is left continuous. Thus $\mathcal{V}$ is order complete.

We can now apply Tarski's Fixed Point Theorem in the form of 4.30; this completes the proof of the present theorem.

30.15. Corollary on comparison of solutions. Let f_1, f_2 be two functions satisfying the conditions of the preceding theorem, let $x_1, x_2 \in X$, and let u_1, u_2 be the maximal solutions of the initial value problems

$$\left.\begin{array}{ll} u_j'(t) = f_j(t, u_j(t)) & (0 \le t \le T), \\ u_j(0) = x_j & \end{array}\right\} \qquad (\mathrm{IVP}_j)$$

for $j = 1, 2$. Suppose that $x_1 \preccurlyeq x_2$, and $f_1(t,x) \preccurlyeq f_2(t,x)$ for all $(t,x) \in [0,T] \times X$. Then $u_1(t) \preccurlyeq u_2(t)$ for all $t \in [0,T]$.

Proof. We have $u_1(t) = x_1 + \int_0^t f_1(s, u_1(s))ds \preccurlyeq x_2 + \int_0^t f_2(s, u_1(s))ds$. Thus u_1 is a solution of the integral inequality given by f_2 and x_2. However, u_2 is the largest solution of that integral inequality.

Generalized Solutions

30.16. The preceding subchapters were concerned mainly with Carathéodory solutions of differential equations. Such solutions are differentiable almost everywhere, as we noted in 30.5. In the remainder of this chapter we consider "generalized solutions" — i.e., functions that are are not necessarily differentiable, but nevertheless "solve" the differential equation in some natural sense. We shall briefly discuss why generalized solutions are sometimes needed; then we discuss some of the main types of generalized solutions.

Let us begin with the world's simplest partial differential equation:

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \qquad \text{or, more briefly,} \qquad u_t = u_x .$$

We seek real-valued solutions $u = u(t,x)$, defined for real t, x. It is easy to verify that a solution is given by $u(t,x) = p(t+x)$, if p is any real-valued differentiable function. We could refer to $u_t = u_x$ as a very simple **wave equation**, because the function $u(t,x) = p(t+x)$ behaves much like a wave at the seashore: it retains its shape while moving horizontally. (*Caution*: The term "wave equation" is commonly applied to several other, more complicated equations that model water waves more accurately.)

Classically, a solution $u = u(t,x)$ is viewed as a real-valued function — i.e., a mapping $u : \mathbb{R}^2 \to \mathbb{R}$. A different viewpoint, closer to the ideas at the beginning of this chapter, views $u(t,x)$ as a continuous function of x for each fixed t. Thus, for each t, $u(t,\cdot)$ is a member of some space of continuous functions — e.g., the Banach space $BC(\mathbb{R})$ of bounded, continuous functions from $\mathbb{R}$ into $\mathbb{R}$, equipped with the sup norm. Then we may suppress the x variable from our notation and write $u(t,\cdot)$ instead as $u(t) \in BC(\mathbb{R})$. We may view u as a Banach-space-valued function $u : [a,b] \to BC(\mathbb{R})$. With this viewpoint, we may attempt to apply theorems like the ones developed earlier in this chapter.

However, it is important to understand that the Fréchet derivative $v = u'(t_0)$ of the Banach-space-valued function $u : [a,b] \to BC(\mathbb{R})$ is a much stronger derivative than the classical, pointwise derivative $v(x) = u_t(t_0, x)$ of the real-valued function $u : \mathbb{R}^2 \to \mathbb{R}$. In both cases we have

$$\frac{u(t_0+h,x) - u(t_0,x)}{h} \to v(x) \qquad \text{as } h \to 0.$$

For the pointwise derivative the convergence is pointwise in x; for the Fréchet derivative the convergence is uniform in x. If p is a differentiable function but $p' \notin BC(\mathbb{R})$, then $u(t,x) = p(t+x)$ only satisfies $u_t = u_x$ in the classical (pointwise) sense, not in the sense of Fréchet derivatives. Thus theorems of the type developed earlier in this chapter are not directly applicable.

It is natural to view $u(t,x) = p(t+x)$ as the "solution" of the initial value problem

$$\left.\begin{array}{ll} u_t(t,x) = u_x(t,x) & (t \geq 0), \\ u(0,x) = p(x) & \end{array}\right\} \qquad \text{(WIVP)}$$

for any differentiable function p. By taking limits, we may extend this definition; it is natural to view $u(t,x) = p(t+x)$ as the "solution" of the wave initial value problem (WIVP) for *any* function p — even one that is not differentiable. Thus, some differential equations have natural "solutions" that are not differentiable in any sense. Such nondifferentiable solutions turn out to be the correct answers to many physical problems.

30.17. The need for nondifferentiable solutions becomes even more evident when we turn to nonlinear problems, such as **Burgers's Equation**:

$$u_t \quad = \quad u u_x \, .$$

Even if the initial data $u(0,\cdot)$ is continuously differentiable, the solution $u(t,\cdot)$ may develop discontinuities at some later time t. We shall demonstrate this with some simple examples. Let $q : \mathbb{R} \to \mathbb{R}$ be some continuously differentiable function that satisfies $q(z) = z$ for all z in $[-1,1]$, and $q'(z) \geq 1$ for all $z \in \mathbb{R}$. The particular choice of q will not affect our main reasoning below, but we mention a couple of examples for concreteness. A trivial example is given by $q(z) = z$; more complicated examples can be devised by the reader.

For any fixed t in $[0,1)$, the function $\psi(t,z) = q(z) - tz$ satisfies $\frac{\partial}{\partial z}\psi(t,z) \geq 1 - t > 0$. Hence $\psi(t,\cdot)$ is strictly increasing and is a bijection from $\mathbb{R}$ onto $\mathbb{R}$. Let $u(t,\cdot)$ be its inverse; thus

$$u(t,\ q(z) - tz) \quad = \quad z\,. \qquad (\natural\natural)$$

(In the example of $q(z) = z$ we obtain $\psi(t,z) = (1-t)z$ and $u(t,x) = (1-t)^{-1}x$.)

Now differentiate both sides of $(\natural\natural)$ with respect to t to obtain

$$0 \quad = \quad \frac{d}{dt}\left[u(t, q(z) - tz) - z\right] \quad = \quad u_t(t, q(z) - tz) - u_x(t, q(z) - tz)z.$$

Then substitute $q(z) - tz = x$ to obtain $0 = u_t(t,x) - u_x(t,x)u(t,x)$. Thus $u(t,x)$ is a solution of Burgers's Equation $u_t = uu_x$, at least for $0 \leq t < 1$. The initial data is $u(0,x) = q^{-1}(x)$, which is continuously differentiable since q is.

Observe that $u(t,x) \leq -1$ for all $x \leq -1+t$, and $u(t,x) \geq 1$ for all $x \geq 1-t$. If u extends to a continuous function on $0 \leq t \leq 1$, it must satisfy $u(1,x) \leq -1$ for all $x \leq 0$ and $u(1,x) \geq 1$ for all $x \geq 0$, a contradiction.

Thus the solution $u(t,x)$, which is continuously differentiable for all (t,x) in $[0,1) \times \mathbb{R}$, becomes discontinuous at time $t = 1$; we say that it develops **shocks**. After time $t = 1$, the solution may still be physically meaningful, but its mathematical theory becomes more complicated. We shall not pursue that theory here, other than to mention that one must deal with "generalized solutions" — i.e., discontinuous functions $u(t,x)$ that correspond somehow to the equation $u_t = uu_x$ but do not satisfy it in a classical sense. The development of shocks is quite typical of nonlinear partial differential equations and is explained further in books on that subject — see Lax [1973], for instance.

Semigroups and Dissipative Operators

30.18. *Motivation for the Crandall-Liggett Theorem.* Let A be an operator for which the differential equation $u'(t) = A(u(t))$ has "solutions" of some sort. More precisely, suppose that M is a subset of a Banach space, and for each $x_0 \in M$ there is a *unique* solution $u : [0,+\infty) \to M$ of the initial value problem

$$\begin{cases} u'(t) = A(u(t)) & (t \geq 0), \\ u(0) = x_0. \end{cases}$$

We may denote that solution by $u(t) = S(t)x_0$ to display its dependence on both the time t and the initial value x_0. In this fashion we define a family of mappings $S(t) : M \to M$ for $t \geq 0$, with $S(0)x = x$.

For most reasonable notions of "solution," the solutions of the two initial value problems

$$\left\{ \begin{array}{ll} u'(t) = A(u(t)) & (0 \leq t), \\ u(0) = x_0 & \end{array} \right\} \qquad \text{and} \qquad \left\{ \begin{array}{ll} v'(t) = A(v(t)) & (0 \leq t), \\ v(0) = u(t_1) & \end{array} \right\}$$

are related by $u(t_1 + s) = v(s)$. From this it follows that the mappings $S(t)$ satisfy the **semigroup property**: $S(t)S(s)z = S(t+s)z$.

In some of the most interesting cases, the semigroup of mappings satisfies an **exponential growth condition**:

$$\langle S(t) \rangle_{\mathrm{Lip}} \leq \exp(\omega t)$$

for some constant ω. For example, if $A : X \to X$ is a Lipschitzian mapping, then A generates a semigroup satisfying an exponential growth condition; that follows from 30.9.

However, a semigroup arising from differential equations may satisfy an exponential growth condition even if the operator A is not Lipschitzian — in fact, even if the operator A is not continuous. In 30.24 we shall show that if the semigroup $S(t)$ is differentiable at

$t = 0$, then the operator A must satisfy a *dissipativeness* condition; this is a generalization of Lipschitzness. Conversely, even for semigroups that are not differentiable, an operator A that satisfies a dissipativeness condition plus a mild "range condition" must generate a semigroup that satisfies an exponential growth condition; this is established in 30.28.

We might denote the semigroup $S(t)$ instead by $S_A(t)$, to display its dependence also on the choice of the operator A. A still more suggestive notation is $S(t)x = e^{tA}x$, where A is the operator appearing in the differential equation. If A is a continuous linear operator, then e^{tA} can be defined in several different equivalent fashions:

$$e^{tA} \quad = \quad \sum_{n=0}^{\infty} \frac{(tA)^n}{n!} \quad = \quad \lim_{n\to\infty} \left(I + \frac{tA}{n}\right)^n \quad = \quad \lim_{n\to\infty} \left(I - \frac{tA}{n}\right)^{-n}.$$

If the operator A is discontinuous and/or nonlinear, then most of these formulas become meaningless or incorrect, but the limit of $\left(I - \frac{t}{n}A\right)^{-n}$ may still be meaningful and useful. Even if A is a badly behaved operator — e.g., a differential operator, which is discontinuous in most of the usual Banach spaces of functions — the operator $(I - \lambda A)^{-1}$ may be quite well behaved when λ is a small positive number — e.g., it may be an integral operator, which is continuous or even compact on many of the usual Banach spaces.

30.19. Although the abstract theory applies to both linear and nonlinear operators, for illustrative purposes we shall give just one very elementary linear example. (For more advanced examples, the reader should consult books devoted specifically to partial differential equations and evolution equations.) Let us use the Banach space $C_0(\mathbb{R})$ of continuous functions from $\mathbb{R}$ to $\mathbb{R}$ that vanish at infinity (as explained in 22.15), with the sup norm. Let A be the operator $\frac{d}{dx}$, with domain $D(A)$ equal to the set of all functions $f \in C_0(\mathbb{R})$ such that f is differentiable and $f' \in C_0(\mathbb{R})$. Then the differential equation $u'(t) = A(u(t))$ is a reformulation of "the world's simplest partial differential equation," discussed in 30.16. We shall now show that $(I - \lambda A)^{-1}$ is very well behaved for any positive number λ.

Let $g \in C_0(\mathbb{R})$ be given; then $(I - \lambda A)^{-1}g$ has what value? Assuming that it exists, it is a function $f \in D(A)$ that satisfies $f - \lambda Af = g$. Let us find that function f. Rewrite its equation as $f(x) - \lambda f'(x) = g(x)$. Multiply both sides of this equation by $-\lambda^{-1}\exp(-x/\lambda)$, to obtain

$$\frac{d}{dx}\left[f(x)\exp\!\left(\frac{-x}{\lambda}\right)\right] \quad = \quad \left[f'(x) - \frac{1}{\lambda}f(x)\right]\exp\!\left(\frac{-x}{\lambda}\right) \quad = \quad -\frac{1}{\lambda}g(x)\exp\!\left(\frac{-x}{\lambda}\right).$$

Integrate both sides — starting from $x = 0$, say — to obtain

$$f(x)\exp\!\left(\frac{-x}{\lambda}\right) \quad = \quad C - \frac{1}{\lambda}\int_0^x g(t)\exp\!\left(\frac{-t}{\lambda}\right)dt$$

for some constant C. To find the value of C, take limits on both sides of this equation as $x \to \infty$. We have $f(x) \to 0$ since f vanishes at ∞, and thus $C = \frac{1}{\lambda}\int_0^\infty g(t)\exp(-t/\lambda)dt$. This integral converges, since g vanishes at infinity and $\exp(-t/\lambda)$ vanishes exponentially fast. Therefore the last displayed equation can be rewritten

$$(I - \lambda A)^{-1}g \;=\; f \qquad \text{where} \qquad f(x) \;=\; \frac{1}{\lambda}\exp\!\left(\frac{x}{\lambda}\right)\int_x^\infty g(t)\exp\!\left(\frac{-t}{\lambda}\right)dt.$$

It is easy to verify that f, defined by the last equation, is indeed a member of $D(A)$ that solves $(I-\lambda A)f = g$; and the preceding computations show that there is no other solution. A further computation shows that $\|f\|_{\sup} \le \|g\|_{\sup}$. Thus

$(I-\lambda A)^{-1}$ is a nonexpansive linear operator defined everywhere on $C_0(\mathbb{R})$.

This is typical of the kind of operator to which the Crandall-Liggett Theorem is applicable — but we emphasize that that theorem applies to much more complicated operators as well.

Exercise. Modifying the computations above, show that $(I+\lambda A)^{-1}$ is also a nonexpansive linear operator defined everywhere on $C_0(\mathbb{R})$, for each $\lambda > 0$.

30.20. Let X be a Banach space, and let $J : X \to \mathcal{P}(X^*)$ be its duality mapping (defined as in 28.44). Let A be some mapping from a subset of X into X. Then the following two conditions are equivalent; if either (hence both) are satisfied, we say A is **dissipative** (or $-A$ is **accretive**):

(A) Whenever $\lambda > 0$, then the mapping $(I-\lambda A) : \mathrm{Dom}(A) \to X$ is injective, and its inverse mapping $(I-\lambda A)^{-1} : \mathrm{Ran}(I-\lambda A) \to \mathrm{Dom}(A)$ is nonexpansive.

(B) Whenever $x_1, x_2 \in \mathrm{Dom}(A)$, then there is some $\varphi \in J(x_1 - x_2)$ such that $\varphi[A(x_1) - A(x_2)] \le 0$.

Proof (following Cioranescu [1990]). Let $y_1 = A(x_1)$ and $y_2 = A(x_2)$. Let $\widehat{x} = x_1 - x_2$ and $\widehat{y} = y_1 - y_2$; then we are to show that

(A′) $\|\widehat{x} - \lambda\widehat{y}\| \ge \|\widehat{x}\|$ for all $\lambda > 0$

if and only if

(B′) there is some $\varphi \in J(\widehat{x})$ such that $\varphi(\widehat{y}) \le 0$.

For (B′) $\Rightarrow$ (A′) we simply compute

$$\|\widehat{x}\|^2 \;=\; \varphi(\widehat{x}) \;\le\; \varphi(\widehat{x}) - \lambda\varphi(\widehat{y}) \;=\; \varphi(\widehat{x} - \lambda\widehat{y}) \;\le\; \|\widehat{x}\|\;\|\widehat{x} - \lambda\widehat{y}\|.$$

The proof of (A′) $\Rightarrow$ (B′) is longer. We may assume $\widehat{x}$ and $\widehat{y}$ are both nonzero (explain), hence $\widehat{x} - \lambda\widehat{y}$ is also nonzero for each $\lambda > 0$. For each $\lambda > 0$ choose some $\xi_\lambda \in J(\widehat{x} - \lambda\widehat{y})$; this vector is also nonzero. Form the unit vector $\eta_\lambda = \xi_\lambda / \|\xi_\lambda\|$. Then

$$\|\widehat{x}\| \;\le\; \|\widehat{x} - \lambda\widehat{y}\| \;=\; \eta_\lambda(\widehat{x} - \lambda\widehat{y}) \;=\; \eta_\lambda(\widehat{x}) - \lambda\eta_\lambda(\widehat{y}) \;\le\; \|\widehat{x}\| - \lambda\eta_\lambda(\widehat{y})$$

from which we conclude both

$$\|\widehat{x}\| \le \eta_\lambda(\widehat{x}) + \lambda\|\widehat{y}\| \qquad \text{and} \qquad \eta_\lambda(\widehat{y}) \le 0. \tag{**}$$

Since the vectors η_λ all lie in the unit ball of X^*, which is weak-star compact by (UF28) in 28.29. the net $(\eta_\lambda : \lambda \downarrow 0)$ has a subnet converging in the weak-star topology to some limit η_0 in that unit ball. Then $\|\eta_0\| \le 1$. Now we may take limits in (**); we obtain

$$\|\widehat{x}\| \le \eta_0(\widehat{x}) \qquad \text{and} \qquad \eta_0(\widehat{y}) \le 0.$$

Since η_0 is in the unit ball, we can conclude $\|\widehat{x}\| = \eta_0(\widehat{x})$ and $\|\eta_0\| = 1$. Then $\varphi = \|\widehat{x}\|\eta_0$ is a member of $J(\widehat{x})$, satisfying $\varphi(\widehat{y}) \leq 0$.

30.21. *A generalization.* Let X be a Banach space, and let $J : X \to \mathcal{P}(X^*)$ be its duality mapping. Let A be a mapping from some subset of X into X, and let ω be a nonnegative number. Then the following three conditions are equivalent (*exercise*); if they are satisfied we say A is **ω-dissipative**:

(A) Whenever $\lambda \in (0, \frac{1}{\omega})$, then the mapping $(I - \lambda A) : \mathrm{Dom}(A) \to X$ is injective, and its inverse mapping

$$R(\lambda) = (I - \lambda A)^{-1} \quad : \quad \mathrm{Ran}(I - \lambda A) \ \to \ \mathrm{Dom}(A)$$

is Lipschitzian with $\big\langle R(\lambda)\big\rangle_{\mathrm{Lip}} \leq (1 - \lambda\omega)^{-1}$.

(B) Whenever $x_1, x_2 \in \mathrm{Dom}(A)$, then there is some $\varphi \in J(x_1 - x_2)$ such that $\varphi[A(x_1) - A(x_2)] \leq \omega\|x_1 - x_2\|^2$.

(C) $A - \omega I$ is dissipative.

30.22. *Remarks.* If A is a Lipschitzian mapping, with $\langle A\rangle_{\mathrm{Lip}} \leq \omega$, then A and $-A$ are both ω-dissipative. For this reason, dissipativeness conditions are sometimes called **one-sided Lipschitz conditions**.

However, that terminology may be misleading. For instance, define A as in 30.19. Then A and $-A$ are both dissipative, but A is not Lipschitzian; in fact, A is not even continuous.

30.23. *Example.* Let X be a real Hilbert space with inner product $\langle\ ,\ \rangle$. Then an operator A is dissipative if and only if it has this property: Whenever $x_1, x_2 \in$ Dom(A), then $\langle x_1 - x_2,\ A(y_1) - A(y_2)\rangle \leq 0$.

If X is one-dimensional — i.e., if X is just the real line — then A is dissipative if and only if $(x_1 - x_2)\Big(A(y_1) - A(y_2)\Big) \leq 0$; that inequality is satisfied if and only if A is a decreasing function.

30.24. *Proposition.* Let C be a subset of a Banach space X, and let S be a semigroup of self-mappings of C. Assume that $\langle S(t)\rangle_{\mathrm{Lip}} \leq e^{\omega t}$ for some constant $\omega \geq 0$ and all $t \geq 0$. Define a mapping from a subset of C into X by

$$A(x) = \lim_{h\downarrow 0} \frac{S(h)x - x}{h}$$

where the domain of the operator A is the set of all $x \in C$ for which the limit exists. Then A is ω-dissipative.

Proof. Fix any $x_1, x_2 \in \mathrm{Dom}(A)$ and $\lambda \in (0, \frac{1}{\omega})$; let $h > 0$. Then

$$\left\|(x_1 - x_2) - \lambda\frac{S(h)x_1 - x_1}{h} + \lambda\frac{S(h)x_2 - x_2}{h}\right\|$$

$$
\begin{aligned}
&= \left\| (1 + \frac{\lambda}{h})(x_1 - x_2) - \frac{\lambda}{h}\left[S(h)x_1 - S(h)x_2\right] \right\| \\
&\geq \left\| (1 + \frac{\lambda}{h})(x_1 - x_2) \right\| - \left\| \frac{\lambda}{h}\left[S(h)x_1 - S(h)x_2\right] \right\| \\
&\geq (1 + \frac{\lambda}{h})\|x_1 - x_2\| - \frac{\lambda}{h} e^{\omega h}\|x_1 - x_2\| \\
&= \left[1 - \lambda \frac{e^{\omega h} - 1}{h} \right] \|x_1 - x_2\|.
\end{aligned}
$$

Take limits as $h \downarrow 0$, to prove

$$\|(x_1 - x_2) - \lambda[A(x_1) - A(x_2)]\| \quad \geq \quad (1 - \lambda\omega)\|x_1 - x_2\|.$$

30.25. *Lemma.* Let A be an ω-dissipative mapping, and let $R(\lambda) = (I - \lambda A)^{-1}$. Then for any numbers $\alpha, \beta \in (0, \frac{1}{\omega})$ and any vectors $u \in \operatorname{Ran}(I - \alpha A)$ and $v \in \operatorname{Ran}(I - \beta A)$, we have

$$(\alpha + \beta - \omega\alpha\beta)\|R(\alpha)u - R(\beta)v\| \ \leq \ \alpha\|R(\alpha)u - v\| + \beta\|u - R(\beta)v\|.$$

Proof. Let $x = R(\alpha)u$ and $y = R(\beta)v$; thus $u = x - \alpha A(x)$ and $v = y - \beta A(y)$. Choose some $\varphi \in J(x - y)$ such that $\varphi[A(x) - A(y)] \leq \omega\|x - y\|^2$. Then

$$
\begin{aligned}
&(\alpha + \beta - \omega\alpha\beta)\|x - y\|^2 \\
&\quad = \quad (\alpha + \beta)\varphi(x - y) - \omega\alpha\beta\|x - y\|^2 \\
&\quad \leq \quad (\alpha + \beta)\varphi(x - y) - \alpha\beta\varphi\{A(x) - A(y)\} \\
&\quad = \quad \varphi\Big\{(\alpha + \beta)(x - y) - \alpha\beta[A(x) - A(y)]\Big\} \\
&\quad = \quad \alpha\varphi\Big\{(x - y) + \beta A(y)\Big\} + \beta\varphi\Big\{(x - y) - \alpha A(x)\Big\} \\
&\quad \leq \quad \alpha\|x - y\|\,\|x - y + \beta A(y)\| + \beta\|x - y\|\,\|x - y - \alpha A(x)\|.
\end{aligned}
$$

Divide through by $\|x - y\|$ to obtain the desired inequality.

30.26. Rasmussen-Kobayashi Inequalities. Let α and β be positive numbers. Let $c_{j,k}$ be nonnegative real numbers that satisfy

$$c_{j,0} \ \leq \ j\alpha, \qquad c_{0,k} \ \leq \ k\beta, \qquad c_{j+1,k+1} \ \leq \ \frac{\alpha}{\alpha + \beta} c_{j+1,k} + \frac{\alpha}{\alpha + \beta} c_{j,k+1}$$

for all nonnegative integers j, k. Then $c_{j,k} \leq \sqrt{(j\alpha - k\beta)^2 + j\alpha^2 + k\beta^2}$ for all nonnegative integers j, k.

More generally, let $\alpha, \beta > 0$ and $\omega \geq 0$ with $\max\{\omega\alpha, \omega\beta\} < 1$. Let $c_{j,k}$ be nonnegative real numbers that satisfy

$$c_{j,0} \ \leq \ (1 - \omega\alpha)^{-j} j\alpha, \qquad c_{0,k} \ \leq \ (1 - \omega\beta)^{-k} k\beta, \tag{1}$$

$$c_{j+1,k+1} \quad \leq \quad \frac{\alpha c_{j+1,k} + \beta c_{j,k+1}}{\alpha + \beta - \omega\alpha\beta} \tag{2}$$

for all nonnegative integers j, k. Then

$$c_{j,k} \quad \leq \quad (1-\omega\alpha)^{-j}(1-\omega\beta)^{-k}\ \sqrt{(j\alpha-k\beta)^2+j\alpha^2+k\beta^2} \qquad \text{(RK)}$$

for all nonnegative integers j, k.

Remarks. This inequality will be used in 30.27. It shows that $c_{j,k}$ may be small even with j, k large, provided that α, β, and $j\alpha - k\beta$ are small. In a first reading, the reader may wish to concentrate on the special case of $\omega = 0$, stated in the first paragraph of the lemma, since that case is slightly simpler in notation and still contains most of the main ideas.

Outline of proof. First, a few preliminary computations. Show that

$$\begin{aligned} &\alpha\Big\{[j\alpha-(k-1)\beta]^2+j\alpha^2+(k-1)\beta^2\Big\} \\ +\ &\beta\Big\{[(j-1)\alpha-k\beta]^2+(j-1)\alpha^2+k\beta^2\Big\} \\ &\quad = \quad (\alpha+\beta)\Big\{[j\alpha-k\beta]^2+j\alpha^2+k\beta^2\Big\}. \end{aligned} \qquad (3)$$

Also, from $\omega(\alpha+\beta)^2-2(\alpha+\beta)\leq 0\leq\alpha\beta\omega$ we obtain

$$(\alpha+\beta)\left[\alpha(1-\omega\beta)^2+\beta(1-\omega\alpha)^2\right] \quad \leq \quad (\alpha+\beta-\omega\alpha\beta)^2. \qquad (4)$$

Also, by the Cauchy-Bunyakovskiĭ-Schwarz Inequality (2.10),

$$\alpha(1-\omega\beta)\sqrt{p}+\beta(1-\omega\alpha)\sqrt{q}\ \leq\ \sqrt{\alpha(1-\omega\beta)^2+\beta(1-\omega\alpha)^2}\ \sqrt{\alpha p+\beta q} \qquad (5)$$

for any nonnegative numbers p and q.

Now, the Rasmussen-Kobayashi Inequality (RK) is clear from (1) when $j=0$ or $k=0$. The inequality will be proved for larger j and k by double induction. In the computations below, step (Ind) is by the induction hypothesis. Compute

$$\begin{aligned} & (1-\omega\alpha)^j(1-\omega\beta)^k(\alpha+\beta-\omega\alpha\beta)c_{j,k} \\ \overset{(2)}{\leq}\quad & (1-\omega\alpha)^j(1-\omega\beta)^k\left[\alpha c_{j,k-1}+\beta c_{j-1,k}\right] \\ \overset{(\mathrm{Ind})}{\leq}\quad & \alpha(1-\omega\beta)\sqrt{[j\alpha-(k-1)\beta]^2+j\alpha^2+(k-1)\beta^2} \\ & \qquad +\beta(1-\omega\alpha)\sqrt{[(j-1)\alpha-k\beta]^2+(j-1)\alpha^2+k\beta^2} \\ \overset{(5)}{\leq}\quad & \sqrt{\alpha(1-\omega\beta)^2+\beta(1-\omega\alpha)^2} \\ & \qquad \left(\alpha\left\{[j\alpha-(k-1)\beta]^2+j\alpha^2+(k-1)\beta^2\right\}\right. \\ & \qquad \left.+\beta\left\{[(j-1)\alpha-k\beta]^2+(j-1)\alpha^2+k\beta^2\right\}\right)^{1/2} \\ \overset{(3)}{=}\quad & \sqrt{\alpha(1-\omega\beta)^2+\beta(1-\omega\alpha)^2}\sqrt{(\alpha+\beta)\left\{[j\alpha-k\beta]^2+j\alpha^2+k\beta^2\right\}} \\ \overset{(4)}{\leq}\quad & (\alpha+\beta-\omega\alpha\beta)\sqrt{(j\alpha-k\beta)^2+j\alpha^2+k\beta^2}. \end{aligned}$$

This completes the induction step, and thus the proof of (RK).

30.27. *Discussion.* The Crandall-Liggett Theorem is generally viewed as a theorem about differential equations in Banach spaces. The Crandall-Liggett Theorem has no applications except in that setting. However, a large part of the proof can be presented in the simpler setting of a complete metric space. We shall take that approach because it may be conceptually simpler to grasp without the distractions of linear structure, and because it provides an interesting application of metric completeness. It is one of the few cases known to this author where we use Lipschitz mappings *without* using the Contraction Fixed Point Theorem.

In the theorem below, we permit $T = +\infty$ if $\omega = 0$. The computations are slightly simpler in that case and so beginners may wish to concentrate on that case.

Crandall-Liggett Theorem (metric space version). Let (M, ρ) be a complete metric space. Let $T \in (0, +\infty]$ and $\omega \in [0, +\infty)$ with $\omega T < 1$. For each $t \in [0, T)$, let $R(t) : M \to M$ be some Lipschitzian mapping, with

$$\big\langle R(t) \big\rangle_{\text{Lip}} \quad \leq \quad (1 - \omega t)^{-1} \tag{1}$$

and

$$\rho\Big(R(s)x,\ R(t)y\Big) \quad \leq \quad \frac{s\rho\Big(R(s)x,\ y\Big) + t\rho\Big(x,\ R(t)y\Big)}{s + t - \omega st} \tag{2}$$

for all $s, t \in [0, T)$ and $x, y \in M$. Let $R(t/j)^k$ denote the kth iterate of the mapping $R(t/j) : M \to M$. (This is defined for all integers $j > t/T$.) Define the function

$$\Gamma(x) \quad = \quad \sup_{t \in (0,T)} \frac{1 - \omega t}{t}\, \rho\Big(R(t)x, x\Big), \tag{3}$$

and assume that the set $D = \{x \in M : \Gamma(x) < \infty\}$ is dense in M.

Then for each $t \geq 0$, the sequence of functions $R(t/j)^j$ ($j \in \mathbb{N}$, $j > t/T$) converges pointwise on M to a Lipschitzian function $S(t) : M \to M$, with $\Big\langle S(t) \Big\rangle_{\text{Lip}} \leq e^{\omega t}$. In fact, for $x \in D$ we have this estimate of the convergence rate:

$$\rho\left(R\left(\frac{t}{j}\right)^j x,\ S(t)M\right) \quad \leq \quad \frac{t}{\sqrt{j}} \left(1 - \frac{\omega t}{j}\right)^{-j} e^{\omega t}\Gamma(x). \tag{a}$$

The map $(t, x) \mapsto S(t)x$ is jointly continuous from $[0, +\infty) \times M$ into M. It is also Lipschitzian in t for fixed $x \in D$ and bounded t:

$$\rho\Big(S(t)x,\ S(s)x\Big) \quad \leq \quad e^{(t+s)\omega}\, |t - s|\, \Gamma(x). \tag{b}$$

Moreover, the mappings $S(t) : M \to M$ form a *semigroup*:

$$S(0)x = x \qquad \text{and} \qquad S(t + s)x = S(t)S(s)x \tag{c}$$

for $t, s \geq 0$ and $x \in M$.

Outline of proof. Temporarily fix any $\alpha, \beta \in [0, T)$ and any $x \in D$. We may assume $\Gamma(x) > 0$; a separate but easy argument for the case of $\Gamma(x) = 0$ is left as an exercise. We shall apply the Rasmussen-Kobayashi Inequality in 30.26, with $c_{j,k} = \rho(R(\alpha)^j x, R(\beta)^k x)/\Gamma(x)$. Hypotheses (1), (2), (3) of the present theorem imply the hypotheses of the Rasmussen-Kobayashi Inequality; thus we obtain

$$\rho\Big(R(\alpha)^j x,\ R(\beta)^k x\Big) \quad \leq \quad \frac{\sqrt{(j\alpha - k\beta)^2 + j\alpha^2 + k\beta^2}}{(1-\omega\alpha)^j\,(1-\omega\beta)^k}\,\Gamma(x). \tag{4}$$

With $\alpha = t/j$ and $\beta = t/k$ this yields

$$\rho\left(R\left(\frac{t}{j}\right)^j x,\ R\left(\frac{t}{k}\right)^k x\right) \quad \leq \quad \frac{t\sqrt{\frac{1}{j}+\frac{1}{k}}}{\left(1-\frac{\omega t}{j}\right)^j\left(1-\frac{\omega t}{k}\right)^k}\,\Gamma(x),$$

which proves that the sequence $\Big(R(t/j)^j x \,:\, j \in \mathbb{N},\ j > t/T\Big)$ is Cauchy for fixed $t \geq 0$ and $x \in D$; denote its limit by $S(t)x$. Hold j fixed and let $k \to \infty$ to prove the convergence rate (a).

Since $\langle R(s)\rangle_{\mathrm{Lip}} \leq (1-\omega t)^{-1}$ on all of M, it follows that $S(t) = \lim_{j\to\infty} R(t/j)^j$ exists and is Lipschitzian on all of M, with $\langle S(t)\rangle_{\mathrm{Lip}} \leq \lim_{j\to\infty}(1-\frac{\omega t}{j})^{-j} = e^{\omega t}$.

Now apply (4) with $\alpha = t/j$ and $\beta = s/k$, and take limits to prove (b). Since $S(t)$ is Lipschitzian on M, it follows easily that $(t,x) \mapsto S(t)x$ is jointly continuous.

Finally, by induction on n show that $S(s/n)^n x = \lim_{k\to\infty} R(s/kn)^{kn} x = S(s)x$. Use this to prove (c) when t/s is rational; then use continuity to prove (c) for all s and t in $[0,+\infty)$.

30.28. Crandall-Liggett Theorem. Let X be a Banach space, and let A be a mapping from some set $\mathrm{Dom}(A) \subseteq X$ into X. Assume A is ω-dissipative for some $\omega \geq 0$. Also assume this **range condition**:

$$\mathrm{Ran}(I - \lambda A) \ \supseteq\ \mathrm{cl}(\mathrm{Dom}(A)) \qquad \text{for all sufficiently small } \lambda > 0.$$

Then the limit

$$S(t)x \quad = \quad \lim_{n\to\infty}\left(I - \frac{t}{n}A\right)^{-n} x$$

exists for each $x \in \mathrm{cl}(\mathrm{Dom}(A))$ and each $t \geq 0$. In fact, the functions $R(\lambda) = (I-\lambda A)^{-1}$ satisfy the hypotheses and hence the conclusions of 30.27, with $M = \mathrm{cl}(\mathrm{Dom}(A))$.

Proof. Choose $T > 0$ small enough so that the range condition is satisfied for all $\lambda \in (0,T)$, and also so that $\omega T < 1$. It is easy to verify that $\Gamma(x) \leq \|A(x)\|$ for $x \in \mathrm{Dom}(A)$; hence $\mathrm{Dom}(A) \subseteq D$; hence D is indeed dense in M. The other hypotheses of 30.27 follow from properties of ω-dissipative operators developed in the last few pages — see 30.21 and 30.25.

30.29. *Remarks on extensions and generalizations.* For simplicity we have only considered $\omega \geq 0$; with some effort it is possible to generalize so that ω may also take negative values. Actually, much of the literature concerns itself only with the case of $\omega = 0$, because the most interesting ideas are already present in that case and the computations are tidier. We have considered positive values of ω so that beginners may more easily contrast Lipschitz mappings with ω-dissipative mappings.

To avoid burdening beginners with more complicated notation, we have only considered dissipative operators that are mappings $A : \mathrm{Dom}(A) \to X$. However, most of the ideas about dissipative operators developed above can be generalized readily to set-valued mappings $A : \mathrm{Dom}(A) \to \{\text{subsets of } X\}$. The proofs for that generalization are similar to the proofs we have presented above; for the most part, one simply replaces "=" with "$\in$" in appropriate places. Thus, instead of just differential equations $u'(t) = A(u(t))$, it is possible to consider differential inclusions $u'(t) \in A(u(t))$. This greater generality is useful in various ways — e.g., for implicit differential equations $p(u(t), u'(t)) = 0$ or for differential inequalities $p(u(t)) \le u'(t) \le q(u(t))$.

Additional properties of the semigroup $S(t)$ can be proved under additional assumptions about the operator A and/or the Banach space X. When X is a Hilbert space, the resulting theory is particularly elegant; much of it can be found in Brézis [1973]. The book by Haraux [1981] covers some of the Banach space theory but also devotes particular attention to the Hilbert space case.

The Crandall-Liggett Theorem, as we have presented it, extends readily to the differential inclusion $u'(t) \in A(u(t))$. If we strengthen the range condition, and require that $\mathrm{Ran}(I - \lambda A) = X$ for all sufficiently small $\lambda > 0$, then it is possible to prove the existence of solutions to the initial value problem

$$\begin{cases} u'(t) \in A(u(t)) + f(t) & (0 \le t \le T), \\ u(0) = x_0 \end{cases}$$

for any $f \in L^1([0,T], X)$ and $x_0 \in X$. A very elegant theory for problems of this type was developed in Bénilan [1972], Crandall and Evans [1975], Crandall and Pazy [1979], and elsewhere.

Much has also been written about differential inclusions of the form $u'(t) \in A(t, u(t))$, where $A(t, \cdot)$ is a ω-dissipative operator for each fixed t. One reference for this subject is Pavel [1987]; that book also introduces many applications to partial differential equations. This subject's theory is not so elegant, but there is good reason. For maximum applicability to partial differential equations, researchers have been interested in problems where the different operators $A(t, \cdot)$, for different fixed values of t, have different domains, and where $\mathrm{Dom}(A(t, \cdot))$ varies erratically with t. This makes the problem considerably more complicated.

30.30. *Remarks on the lack of a "Grand Unified Theory" of initial value problems.* In the preceding pages we have developed several substantially different theories of initial value problems, using hypotheses of Lipschitz conditions, compactness, isotonicity, and dissipativeness. Historically, these theories developed separately, for different kinds of applications. It is tempting to try to make these theories into special cases of a single, more general theory. Certainly it is possible to prove at least a few weak results in a more general setting — see for instance 30.6.

However, in truth we are very far from a complete or unified theory. The several main subtheories — Lipschitzness, compactness, isotonicity, etc. — are very different in nature; large conceptual gaps lie between them. The literature contains only a handful of examples of nonexistence of solutions, most of them similar to Dieudonné's example 30.4; the examples of nonexistence are not sufficiently diverse to explain the gaps between our theories of existence. Thus, we are very far from a clear understanding of what "really" makes initial value problems work.

More modest than the search for a grand unified theory is the program to solve problems of the form $u'(t) = A(u(t)) + B(u(t))$, where A and B are operators of two different types — e.g., where A satisfies a dissipativeness condition and B satisfies a compactness condition. A theory of this sort would include the dissipativeness and compactness theories as special cases, since we could take $A = 0$ or $B = 0$ (since the operator 0 is both dissipative and compact). This program has met some success, at least when the operators are continuous — for instance, the sum of a continuous dissipative operator, a continuous compact operator, and a continuous isotone operator is known to generate an evolution; see Volkmann [1991]. But without continuity the problem is still open. For the compact plus dissipative problem, some discussions and partial results can be found in Schechter [1987, 1989] and Vrabie [1988].

This Page Intentionally Left Blank

References

Abbreviations of journals are the same as the abbreviations used by *Mathematical Reviews.*

J. F. Aarnes and P. R. Andenæs, On nets and filters, *Math. Scand.* **31** (1972), 285–292.

A. Abian, Calculus must consist of the study of real numbers in their decimal representation and not of the study of an abstract complete ordered field or nonstandard real numbers, *Internat. J. Math. Ed. Sci. Tech.* **12** (1981), 465–472.

P. Aczel, *Non-Well-Founded Sets*, Center for the Study of Language and Information Lecture Notes **14**, Stanford, 1988.

R. A. Adams, *Sobolev Spaces*, Pure and Applied Mathematics **65**, Academic Press, New York, 1975.

N. Adasch, B. Ernst, and D. Keim, *Topological Vector Spaces: The Theory Without Convexity Conditions*, Lecture Notes in Math. **639**, Springer-Verlag, Berlin, 1978.

P. Alexandroff and P. Urysohn, Über nulldimensionale Punktmengen, *Math. Ann.* **98** (1928), 89–106.

C. D. Aliprantis and O. Burkinshaw, *Positive Operators*, Pure Appl. Math. **119**, Academic Press, Orlando, 1985.

M. Arala-Chaves, Spaces where all continuity is uniform, *Amer. Math. Monthly* **92** (1985), 487–489.

K. J. Arrow and F. H. Hahn, *General Competitive Analysis*, Holden-Day, San Francisco, 1971.

B. Artmann, *The Concept of Number: From Quaternions to Monads and Topological Fields*, Ellis Horwood, Chichester, 1988.

B. W. Augenstein, Speculative model of some elementary particle phenomena, *Speculations in Science and Technology* **17** (1994), 21–26.

G. Bachman, *Introduction to p-adic Numbers and Valuation Theory*, Academic Press, New York, 1964.

J.-B. Baillon and S. Simons, Note: Almost-fixed-point and fixed-point theorems for discrete-valued maps, *J. Comb. Theory* (A) **60** (1992), 147–154.

R. N. Ball, Completions of ℓ-groups, pp. 142–174 in: *Lattice-Ordered Groups: Advances and Techniques (workshop in Bowling Green, 1985)*, ed. by A. M. W. Glass and W. C. Holland, Math. Appl. **48**, Kluwer Academic Publishers, Dordrecht, 1989.

S. Banach, Über metrische Gruppen, *Studia Mathematica* **3** (1931), 101–113.

S. Banach, *Théorie des Opérations Linéaires*, Monografje Matematyczne **1**, Warszawa-Lwów, 1932. Later reprints include translation by F. Jellett, *Theory of Linear Operations*, Mathematical Library **38**, North-Holland, Amsterdam, 1987.

J. Banaś, On existence theorems for differential equations in Banach spaces, *Bull. Austral. Math. Soc.* **32** (1985), 73–82.

J. Barwise and J. Etchemendy, *The Liar: An Essay on Truth and Circularity*, Oxford University Press, New York, 1987.

J. E. Baumgartner, Independence proofs and combinatorics, pp. 35–46 in: *Relations Between Combinatorics and Other Parts of Mathematics (conf. Columbus 1978)*, ed. by D. K. Ray-Chaudhuri, Proc. Sympos. Pure Math. **34**, American Mathematical Society, Providence, 1979.

G. Beer and R. Lucchetti, Weak topologies for the closed subsets of a metrizable space, *Trans. Amer. Math. Soc.* **335** (1993), 805–822.

P. R. Beesack, R. Hughes, and M. Ortel, Rotund complex normed spaces, *Proc. Amer. Math. Soc.* **75** (1979), 42–44.

M. J. Beeson, *Foundations of Constructive Mathematics: Metamathematical Studies*, Ergeb. Math. Grenzgeb. (3) **6**, Springer-Verlag, New York, 1985.

M. Behrens, A local inverse function theorem, pp. 34–36 in: *Victoria Symposium on Nonstandard Analysis (University of Victoria 1972)*, ed. by A. Hurd and P. Loeb; Lecture Notes in Math. **369**, Springer-Verlag, Berlin, 1974.

J. L. Bell, *Boolean-Valued Models and Independence Proofs in Set Theory*, 2nd edition, Oxford Logic Guides **12**, Clarendon Press, Oxford, 1985.

J. L. Bell and M. Machover, *A Course in Mathematical Logic*, North-Holland, Amsterdam, 1977.

J. L. Bell and A. B. Slomson, *Models and Ultraproducts: An Introduction*, North-Holland, Amsterdam, 1969, 1974.

P. Bénilan, *Equations d'evolution dans un espace de Banach quelconque et applications*, Thesis, U. Paris XI, Orsay, 1972.

H. L. Bentley, H. Herrlich, and E. Lowen-Colebunders, Convergence, *J. Pure Appl. Algebra* **68** (1970), 27–45.

S. K. Berberian, *Lectures in Functional Analysis and Operator Theory*, Graduate Texts in Math. **15**, Springer-Verlag, New York, 1974.

C. Bessaga, On the converse of the Banach fixed-point principle, *Colloq. Math.* **7** (1959), 41–43.

K. P. S. Bhaskara Rao and M. Bhaskara Rao, *Theory of Charges: A Study of Finitely Additive Measures*, Pure Appl. Math. **109**, Academic Press, London, 1983.

D. C. Biles, Existence of solutions for discontinuous differential equations, *Differential and Integral Equations* **8** (1995), 1525–1532.

P. Billingsley, Van der Waerden's continuous nowhere differentiable function, *Amer. Math. Monthly* **89** (1982), 691.

G. Birkhoff, *Lattice Theory*, Amer. Math. Soc. Colloq. Publ. **25**, Amer. Math. Soc., New York, 1940; 3rd edition in 1967.

E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967. (Later revised as *Constructive Analysis* by Bishop and Bridges.)

E. Bishop, "Schizophrenia in Contemporary Mathematics," talk given in 1973. Subsequently published in *Errett Bishop: Reflections on Him and His Research*, ed. by Murray Rosenblatt, Contemp. Math. **39**, American Mathematical Society, Providence, 1985.

E. Bishop and D. Bridges, *Constructive Analysis*, Grundlehren Math. Wiss. **279**, Springer-Verlag, Berlin, 1985. (Revision of Bishop's *Foundations of Constructive Analysis.*)

C. E. Blair, The Baire category theorem implies the principle of dependent choices, *Bull. Acad. Polon. Sci.* **25** (1977), 933–934.

A. Blass, Existence of bases implies the Axiom of Choice, pp. 31–33 in: *Axiomatic Set Theory*, ed. by J. E. Baumgartner, D. A. Martin, and S. Shelah, Contemp. Math. **31**, Amer. Math. Soc., Providence, 1984.

J. L. Bona, private communication, 1977.

K. C. Border, *Fixed Point Theorems with Applications to Economics and Game Theory*, Cambridge University Press, Cambridge, 1985.

N. Bourbaki, *General Topology, Part 2*, Addison-Wesley, Reading, 1966. English translation of *Éléments de Mathématique, Topologie Générale*, Hermann, Paris.

D. W. Boyd and J. S. Wong, On nonlinear contractions, *Proc. Amer. Math. Soc.* **20** (1969), 458–464.

R. C. Bradley, An elementary treatment of the Radon-Nikodym derivative, *Amer. Math. Monthly* **96** (1989), 437–440.

H. Brézis, *Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Mathematics Studies **5**, Notas de Matemática **50**, North-Holland, Amsterdam, 1973.

D. S. Bridges, *Constructive Functional Analysis*, Res. Notes Math. **28**, Pitman Publishing, London, 1979.

D. Bridges and R. Mines, What is constructive mathematics?, *Mathematical Intelligencer* **6** (1984), 32–38.

D. Bridges and F. Richman, *Varieties of Constructive Mathematics*, London Math. Soc. Lecture Note Ser. **97**, Cambridge University Press, Cambridge, 1987.

A. Brønsted, On a lemma of Bishop and Phelps, *Pacific J. Math.* **55** (1974), 335–341.

F. E. Browder, *Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces*, Proc. Sympos. Pure Math. **18** part 2, Amer. Math. Soc., Providence, 1976.

R. A. Brualdi, Comments and complements, *Amer. Math. Monthly* **84** (1977), 803–807.

G. Bruns and J. Schmidt, Zur Äquivalenz von Moore-Smith-Folgen und Filtern, *Math. Nachr.* **13** (1955), 169–186.

N. Brunner, Amorphe Potenzen kompakter Raume, *Arch. Math. Logik Grundlag.* **24** (1984), 119–135. (*Math. Reviews* **86k**:03043.)

N. Brunner, Garnir's dream spaces with Hamel bases, *Arch. Math. Logik Grundlag.* **26** (1987), 123–126.

N. Brunner, Topologische Maximalprinzipien, *Z. Math. Logik Grundlag. Math.* **33** (1987), 135–139.

P. S. Bullen, P. Y. Lee, J. L. Mawhin, P. Muldowney, and W. F. Pfeffer (editors), *New Integrals (Proceedings of the Henstock Conference held in Coleraine, Northern Ireland, August 9–12, 1988)*, Lecture Notes in Mathematics **1419**, Springer-Verlag, Berlin, 1990.

J. P. Burgess, Forcing, pp. 403–452 in: *Handbook of Mathematical Logic*, ed. by J. Barwise, Stud. Logic Found. Math. **90**, North-Holland, Amsterdam, 1977.

G. Buskes, *The Hahn-Banach Theorem Surveyed*, Dissertationes Math. (Rozprawy Mat.) **327**, Polish Acad. Sci., Warsaw, 1993.

G. J. H. M. Buskes and A. C. M. van Rooij, Riesz spaces and the ultrafilter theorem I, *Compositio Math.* **83** (1992), 311–327.

H. Cartan, Filtres et ultrafiltres, *C. R. Acad. Sci. Paris* **205** (1937), 777–779.

S. D. Chatterji, Martingale convergence and the Radon-Nikodym Theorem in Banach spaces, *Mathematica Scandinavica* **22** (1968), 21–41.

I. Cioranescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*, Mathematics and its Applications **62**, Kluwer Academic Publishers, Dordrecht, 1990.

F. H. Clarke, On the inverse function theorem, *Pacific J. Math.* **64** (1976), 97–102.

P. J. Cohen, *Set Theory and the Continuum Hypothesis*, W. A. Benjamin, New York, 1966.

D. L. Cohn, *Measure Theory*, Birkhäuser, Boston, 1980.

P. M. Cohn, *Universal Algebra*, Harper and Row, New York, 1965.

E. Coleman, Banach space ultraproducts, *Bull. Irish Math. Soc.* **18** (1987), 30–39.

J. F. Colombeau, *Elementary Introduction to New Generalized Functions*, North-Holland Math. Stud. **113** / Notas Mat. **103**, North-Holland, Amsterdam, 1985.

W. W. Comfort, A theorem of Stone-Cech type, and a theorem of Tychonoff type, without the axiom of choice, and their realcompact analoguse, *Fund. Math.* **63** (1968), 97–110.

J. B. Conway, The inadequacy of sequences, *Amer. Math. Monthly* **76** (1969), 68–69.

C. H. Cook and H. R. Fischer, Regular convergence spaces, *Math. Ann.* **174** (1967), 1–7.

R. H. Cox, A proof of the Schroeder-Bernstein theorem, *Amer. Math. Monthly* **75** (1968), 508.

M. G. Crandall and L. C. Evans, On the relation of the operator $\partial/\partial s + \partial/\partial\tau$ to evolution governed by accretive operators, *Israel J. Math.* **21** (1975), 261–278.

M. G. Crandall and A. Pazy, An approximation of integrable functions by step functions with an application, *Proc. Amer. Math. Soc.* **76** (1979), 74–80.

D. van Dalen, *Logic and Structure*, Springer-Verlag, Berlin, 1983.

D. van Dalen, H. C. Doets, and H. de Swart, *Sets: Naive, Axiomatic, and Applied*, Internat. Ser. Pure Appl. Math. **106**, Pergamon Press, Oxford, 1978.

H. G. Dales and W. H. Woodin, *An Introduction to Independence for Analysts*, London Math. Soc. Lecture Note Ser. **115**, Cambridge Univ. Press, Cambridge, 1987.

L. Danzer, B. Grünbaum, and V. Klee, Helly's theorem and its relatives, pp. 101–180 in: *Convexity (proc. conf. Seattle 1961)*, ed. by V. Klee, Proc. Sympos. Pure Math. **7**, Amer. Math. Soc., Providence, 1963.

A. S. Davis, Indexed systems of neighborhoods for general topological spaces, *Amer. Math. Monthly* **68** (1961), 886–893.

M. Davis, *Applied Nonstandard Analysis*, Wiley, New York, 1977.

M. M. Day, Normed linear spaces, Ergeb. Math. Grenzgeb. **21**, Springer-Verlag, Berlin, 1973.

J. D. DePree and C. W. Swartz, *Introduction to Real Analysis*, Wiley, New York, 1988.

R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*, 2nd edition, Addison-Wesley, Redwood City, 1989.

R. Deville, G. Godefroy, and V. Zizler, *Smoothness and Renormings in Banach Spaces*, Pitman Monographs and Surveys in Pure and Appl. Math. **64**, Longman Scientific and Technical, 1993.

P. Dienes, *The Taylor Series: An Introduction to the Theory of Functions of a Complex Variable*, Dover Publications, New York, 1957.

J. Diestel, *Geometry of Banach Spaces: Selected Topics*, Lecture Notes in Math. **485**, Springer-Verlag, Berlin, 1975.

J. Diestel, *Sequences and Series in Banach Spaces*, Graduate Texts in Math. **92**, Springer-Verlag, New York, 1984.

J. Diestel and J. J. Uhl, *Vector Measures*, Mathematical Surveys **15**, Amer. Math. Soc., Providence, 1977.

J. Dieudonné, Deux examples singuliers d'équations différentielles, *Acta Sci. Math. (Szeged)* **12 B** (1950), 38–40.

H. C. Doets, Cantor's paradise, *Nieuw Arch. Wisk.* (4) **1** (1983), 290–344.

S. Dolecki and G. H. Greco, Cyrtologies of convergences I, *Math. Nachr.* **126** (1986), 327–348.

J. Dugundji, *Topology*, Allyn & Bacon, Boston, 1966.

J. Dugundji and A. Granas, *Fixed Point Theory, Volume I*, Monografie Matematyczne **61**, Polska Akademia Nauk, Instytut Matematyczny, Polish Scientific Publishers (PWN), Warszawa, 1982.

N. Dunford and J. T. Schwartz, *Linear Operators, Part I: General Theory*, Pure Appl. Math. **7**, Wiley Interscience, New York, 1957.

R. E. Edwards, *Fourier Series: A Modern Introduction, Vol. 1*, Holt Rinehart Winston, New York, 1967.

H. B. Enderton, Elements of recursion theory, pp. 527–566 in: *Handbook of Mathematical Logic*, ed. by J. Barwise, Stud. Logic Found. Math. **90**, North-Holland, Amsterdam, 1977.

H. B. Enderton, *Elements of Set Theory*, Academic Press, New York, 1977.

R. Engelking, *General Topology*, Monografie Matematyczne **60**, PWN, Warszawa, 1977.

P. Erdös, L. Gillman, and M. Henriksen, An isomorphism theorem for real-closed fields, *Ann. of Math.* **61** (1955), 542–554.

J. J. M. Evers and H. van Maaren, Duality principles in mathematics and their relations to conjugate functions, *Nieuw Archief voor Wiskunde (4)* **3** (1985), 23–68.

R. Fefferman, lectures, University of Chicago, 1977.

I. Fleischer, Infinitesimals, *Nordisk Matematisk Tidskrift* **15** (1967), 151–155.

K. Floret, *Weakly Compact Sets*, Lecture Notes in Math. **801**, Springer-Verlag, Berlin, 1980.

S. R. Foguel, On a theorem by A. E. Taylor, *Proc. Amer. Math. Soc.* **9** (1958), 325.

D. Freedman, *Brownian Motion and Diffusion*, Holden-Day, San Francisco, 1971.

D. H. Fremlin, *Topological Riesz Spaces and Measure Theory*, Cambridge University Press, London, 1974.

L. Fuchs, *Partially Ordered Algebraic Systems*, Internat. Ser. Monographs Pure Appl. Math. **28**, Pergamon Press, Oxford, 1963.

B. Fuchssteiner, Exposed fixpints in order-structures, pp. 359–376 in: *Aspects of Mathematics and its Applications (proc. conf. Rio de Janeiro, 1982)*, ed. by J. A. Barroso, North-Holland Math. Library **34**, North-Holland, Amsterdam, 1986.

W. Gähler, *Grundstrukturen der Analysis I*, Birkhäuser Verlag, Basel, 1977.

W. Gähler, General convergence and applications, pp. 90–101 in: *Convergence, Topology, and Measure (Proceedings of the Conference on Topology and Measure IV, Trassenheide, 1983)*, ed. by J. Flachsmeyer, Z. Frolik, S. Lotz, Y. M. Smirnov, F. Terpe, and F. Topsøe, Wissenschaftliche Beiträge der Ernst-Moritz-Arndt-Universität, Greifswald, 1984.

G. García Márquez, "One Day After Saturday," story appearing in the collection *Big Mama's Funeral*, translated from the Spanish by S. J. Bernstein, originally published as *Los Funerales de la Mama Grande*, Universidad Veracruzana, Vera Cruz, 1962.

H. G. Garnir, Solovay's axiom and functional analysis, pp. 189–204 in: *Functional Analysis and its Applications (International Conference, Madras, 1973)*, ed. by H. G. Garnir, K. R. Unni, and J. H. Williamson, Lecture Notes in Math. **399**, Springer-Verlag, Berlin, 1974.

H. G. Garnir, M. De Wilde, and J. Schmets, *Analyse Fonctionnelle: Théorie constructive des espaces linéaires à semi-normes*, Birkhäuser Verlag, Basel, 1968.

A. O. Gelfond, *Transcendental and Algebraic Numbers*, Dover Publications, New York, 1960.

R. Gherman, Remark on axioms defining a topology by convergent filters, *Analele Universitatii din Graiova, Seria Matematica, Fizica-Chimie* **8** (1980), 43–46.

L. Gillman and M. Jerison, *Rings of Continuous Functions*, Van Nostrand, Princeton, 1960.

D. Gluschankof and M. Tilli, On some extension theorems in functional analysis and the theory of Boolean algebras, *Rev. Un. Mat. Argentina* **33** (1987), 44–54.

A. N. Godunov, Peano's theorem in Banach spaces (Russian), *Funktsional'nyi Analiz i Ego Prilozheniya* **9** (1975), 59–60; translated into English in: *Functional Analysis and its Applications* **9** (1975), 53–55.

K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Non-expansive Mappings*, Monographs Textbooks Pure Appl. Math. **83**, Dekker, New York, 1984.

N. Goodman and J. Myhill, Choices implies excluded middle, *Z. Math. Logik Grundlag. Math.* **24** (1978), 461.

R. Gordon, Riemann integration in Banach spaces, *Rocky Mountain J. Math.* **21** (1991), 923–949.

R. A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron, and Henstock*, Amer. Math. Soc., Providence, 1994.

D. H. Griffel, *Applied Functional Analysis*, Ellis Horwood, New York, 1981.

A. Grothendieck, *Topological Vector Spaces*, Gordon and Breach, New York, 1973.

L. Haddad, Condorcet et les ultrafiltres, pp. 343–360 in: *Mathématiques Finitaires et Analyse Non Standard, tome 2*, ed. by M. Diener and G. Wallet, Publ. Math. Univ. Paris VII no. 31, U. F. R. de Mathématiques, Paris, 1989.

M. Hall, Jr., Distinct representatives of subsets, *Bull. Amer. Math. Soc.* **54** (1948), 922–926.

M. Hall, Jr., A survey of combinatorial analysis, pp. 35–104 in: *Some Aspects of Analysis and Probability*, Surveys in Applied Mathematics **4**, Wiley, New York, 1958.

P. Hall, On representatives of subsets, *J. London Math. Soc.* **10** (1935), 26–30.

P. R. Halmos, *Naive Set Theory*, Van Nostrand, Princeton, 1960.

P. R. Halmos, *Lectures on Boolean Algebras*, Van Nostrand, Princeton, 1963.

P. R. Halmos, *I Want to be a Mathematician: An Automathography*, Springer-Verlag, New York, 1985.

P. R. Halmos and H. E. Vaughan, The marriage problem, *Amer. J. Math.* **72** (1950), 214–215.

J. D. Halpern, Bases for vector spaces and the axiom of choice, *Proc. Amer. Math. Soc.* **17** (1966), 670–673.

J. D. Halpern and A. Lévy, The Boolean prime ideal theorem does not imply the axiom of choice, pp. 83–134 in: *Axiomatic Set Theory, Part 1 (proc. conf. Los Angeles 1967)*, ed. by D. S. Scott, Proc. Sympos. Pure Math. **13** part 1, AMS, Providence, 1971.

A. G. Hamilton, *Logic for Mathematicians*, Cambridge University Press, Cambridge, 1978.

O. Hanner, On the uniform convexity of L^p and ℓ^p, *Arkiv för Matematik* **3** (1955), 239–244.

A. Haraux, *Nonlinear Evolution Equations — Global Behavior of Solutions*, Lecture Notes in Math. **841**, Springer-Verlag, Berlin, 1981.

W. S. Hatcher, Calculus is algebra, *Amer. Math. Monthly* **89** (1982), 362–370.

H.-P. Heinz, On the behaviour of measures of noncompactness with respect to differentiation and integration of vector-valued functions, *Nonlinear Anal.* **7** (1983), 1351–1371.

L. A. Henkin, Boolean representation through propositional calculus, *Fund. Math.* **41** (1954), 89–96.

J. Hennefeld, A nontopological proof of the uniform boundedness theorem, *Amer. Math. Monthly* **87** (1980), 217.

C. W. Henson and H. J. Kiesler, On the strength of nonstandard analysis, *J. Symbolic Logic* **51** (1986), 377–386.

R. Henstock, *Lectures on the Theory of Integration*, Series in Real Analysis **1**, World Scientific, Singapore, 1988.

R. Henstock, *The General Theory of Integration*, Clarendon Press, Oxford, 1991.

H. Herrlich and G. E. Strecker, *Category Theory*, Sigma Ser. Pure Math. **1**, Heldermann Verlag, Berlin, 1973, 1979.

T. H. Hildebrandt, *Introduction to the Theory of Integration*, Pure and Applied Mathematics **13**, Academic Press, New York, 1963.

E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, Amer. Math. Soc. Colloq. Publ. **31**, AMS, Providence, 1957.

J. R. Hindley, B. Lercher, and J. P. Seldin, *Introduction to Combinatory Logic*, London Math. Soc. Lecture Notes **7**, Cambridge Univ. Press, Cambridge, 1972.

M. W. Hirsch, Review of *Realism in Mathematics* by P. Maddy, *Bull. Amer. Math. Soc.* **32** (1995), 137–148.

J. G. Hocking and G. S. Young, *Topology*, Addison-Wesley, Reading, 1961.

D. R. Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid*, Random House, New York, 1979.

R. B. Holmes, *Geometric Functional Analysis and its Applications*, Graduate Texts in Math. **24**, Springer-Verlag, New York, 1975.

J. Horvath, *Topological Vector Spaces and Distributions*, Addison-Wesley, Reading, 1966.

P. E. Howard, Rado's selection lemma does not imply the Boolean prime ideal theorem, *Z. Math. Logik Grundlag. Math.* **30** (1984), 129–132.

P. E. Howard, private communication, 1992.

P. E. Howard, Variations of Rado's lemma, *Math. Logic Quarterly* **39** (1993), 353–356.

P. Howard and J. E. Rubin, The Boolean Prime Ideal Theorem + Countable Choice do not imply Dependent Choice, to appear, *Math. Logic Quarterly*, 1996.

P. Howard and J. E. Rubin, *Weak Forms of the Axiom of Choice*, book in preparation. (Currently available on-line at http://www.math.purdue.edu/~jer/Papers/papers.html)

B. R. Hunt, T. Sauer, and J. A. Yorke, Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces, *Bull. Amer. Math. Soc.* **27** (1992), 217–238.

B. R. Hunt, T. Sauer, and J. A. Yorke, Prevalence: an addendum, *Bull. Amer. Math. Soc.* **28** (1993), 306–307.

R. A. Hunt, On the convergence of Fourier series, pp. 235–255 in: *Orthogonal Expansions and Their Continuous Analogues (Proc. Conf. Edwardsville, 1967)*, ed. by D. T. Haimo,

Southern Illinois University Press, Carbondale, 1968.
A. D. Ioffe, Nonsmooth analysis and the theory of fans, pp. 93–117 in: *Convex Analysis and Optimization (proc. conf. London 1980)*, ed. by J.-P. Aubin and R. B. Vinter, Res. Notes Math. **57**, Pitman, Boston, 1982.
J. R. Isbell, *Uniform Spaces*, Mathematical Surveys **12**, Amer. Math. Soc., Providence, 1964.
H. Ishihara, Constructive reflexivity of a uniformly convex Banach space, *Proc. Amer. Math. Soc.* **104** (1988), 735–740.
V. I. Istrăţescu, *Strict Convexity and Complex Strict Convexity: Theory and Applications*, Lecture Notes in Pure and Appl. Math. **89**, Dekker, New York, 1984.
A. J. Izzo, A functional analysis proof of the existence of Haar measure on locally compact abelian groups, *Proc. Amer. Math. Soc.* **115** (1992), 581–583.
I. M. James, *Topological and Uniform Spaces*, Springer-Verlag, New York, 1987.
R. C. James, A non-reflexive Banach space isometric with its second conjugate space, *Proc. Nat. Acad. Sci. U.S.A.* **37** (1951), 174–177.
R. C. James, Bases in Banach spaces, *Amer. Math. Monthly* **89** (1982), 625–640.
L. Janos, A converse of Banach's contraction theorem, *Proc. Amer. Math. Soc.* **18** (1967), 287–289.
T. J. Jech, *The Axiom of Choice*, Stud. Logic Found. Math. **75**, North-Holland, Amsterdam, 1973.
T. J. Jech, About the Axiom of Choice, pp. 345–370 in: *Handbook of Mathematical Logic*, ed. by J. Barwise, Stud. Logic Found. Math. **90**, North-Holland, Amsterdam, 1977.
P. T. Johnstone, The point of pointless topology, *Bull. Amer. Math. Soc.* **8** (1983), 41–53.
P. T. Johnstone, *Notes on Logic and Set Theory*, Cambridge Math. Textbooks, Cambridge University Press, Cambridge, 1987.
M. Josephy, Composing functions of bounded variation, *Proc. Amer. Math. Soc.* **83** (1981), 354–356.
N. J. Kalton, N. T. Peck, and J. W. Roberts, *An F-space Sampler*, London Math. Soc. Lecture Note Ser. **89**, Cambridge University Press, Cambridge, 1984.
I. Kaplansky, *Set Theory and Metric Spaces*, Allyn & Bacon, Boston, 1972; Chelsea, New York, 1977.
J. L. Kelley, The Tychonoff theorem implies the axiom of choice, *Fund. Math.* **37** (1950), 75–76.
J. L. Kelley, *General Topology*, Van Nostrand, New York, 1955; reprinted by Springer-Verlag, New York, 1975.
J. L. Kelley, and I. Namioka, *Linear Topological Spaces*, Graduate Texts in Math. **36**, Springer-Verlag, New York, 1963, 1976.
D. C. Kent and G. D. Richardson, Cauchy spaces and their completions, pp. 103–112 in: *Convergence Structures and Applications II (proc. conf. Schwerin, 1983)*, ed. by S. Gähler, W. Gähler, and G. Kneis, Akademie-Verlag, Berlin, 1984.
A. P. Kirman and D. Sondermann, Arrow's theorem, many agents, and invisible dictators, *J. Econom. Theory* **5** (1972), 267–277.
F. Klein, *Lectures on the Ikosahedron and the Solution of Equations of the Fifth Degree*, Trübner and Co., London, 1888.
M. Kline, *Mathematics: The Loss of Certainty*, Oxford University Press, New York, 1980.

M. Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, New York, 1972, 1990.

G. T. Kneebone, *Mathematical Logic and the Foundations of Mathematics: An Introductory Survey*, Van Nostrand, London, 1963.

H. König, Theory and applications of superconvex spaces, pp. 79–118 in: *Aspects of Positivity in Functional Analysis (proc. conf. Tubingen 1985)*, ed. by R. Nagel, U. Schlotterbeck, and M. P. H. Wolff, Notas Mat. **108**, North-Holland, Amsterdam, 1986.

G. Köthe, *Topological Vector Spaces I*, Grundlehren Math. Wiss. **237**, Springer-Verlag, Berlin, 1969.

R. Kopperman, All topologies come from generalized metrics, *Amer. Math. Monthly* **95** (1988), 89–97.

G. Kreisel, What have we learnt from Hilbert's second problem?, pp. 93–130 in: *Mathematical Developments Arising from Hilbert Problems*, ed. by F. Browder, Proc. Sympos. Pure Math. **28**, Amer. Math. Soc., Providence, 1976.

J.-L. Krivine, *Introduction to Axiomatic Set Theory*, Reidel, Dordrecht, 1971.

H. W. Kuhn, Some combinatorial lemmas in topology, *IBM J. Res. Develop.* **4** (1960), 518–524.

H. Kuo, *Gaussian Measures in Banach Spaces*, Lecture Notes in Math. **463**, Springer-Verlag, Berlin, 1975.

K. Kunen, *Set Theory: An Introduction to Independence Proofs*, Stud. Logic Found. Math. **102**, North-Holland, Amsterdam, 1980.

K. Kuratowski, *Topology, Volume 1*, Academic Press, New York, 1966. Translation and revision of *Topologie, Vol. 1*, Monagrafie Matematycnzne **20**, Warsaw, 1948.

A. G. Kurosh, *Lectures in General Algebra*, Pergamon Press, Oxford, 1965.

J. Kurzweil, Generalized ordinary differential equations and continuous dependence on a parameter, *Czech. Math. Journal* **7** (82) (1957), 418–446.

E. Landau, *Darstellung und Begründung Einiger Neuerer Ergebnisse der Funktionentheorie*, 1916, second edition 1929, reprinted in *Das Kontinuum und Andere Monographien*, Chelsea, New York.

S. Lang, *Real Analysis*, Addison-Wesley, Reading, 1983.

A. Lasota and J. Yorke, The generic property of existence of solutions of differential equations in Banach space, *J. Differential Equations* **13** (1973), 1–12.

P. D. Lax, *Hyperbolic systems of conservation laws and the theory of shock waves*, Regional Conference Series in Applied Math. **11**, Soc. Indus. Appl. Math., Philadelphia, 1973.

S. Leader, A topological characterization of Banach contractions, *Pacific J. Math.* **69** (1977), 461–466.

J. Lembcke, Two extension theorems effectively equivalent to the axiom of choice, *Bull. London Math. Soc.* **11** (1979), 285–288.

A. Lévy, Definability in axiomatic set theory I, pp. 127–151 in: *Logic, Methodology and Philosophy of Science (proc. conf. Jerusalem 1964)*, ed. by Y. Bar-Hillel, North-Holland, Amsterdam, 1965.

R. Lidl and H. Niederreiter, *Finite Fields*, Encyclopedia Math. Appl. **20**, Addison-Wesley, Reading, 1983.

A. H. Lightstone and A. Robinson, *Nonarchimedean Fields and Asymptotic Expansions*, North-Holland Math. Library **13**, North-Holland, Amsterdam, 1975.

B. V. Limaye, *Functional Analysis*, Halsted Press, New York, 1981.

J. Lindenstrauss, Some useful facts about Banach spaces, pp. 185–200 in: *Geometric Aspects of Functional Analysis (Proc. Israel GAFA Seminar 1986–87)*, Lecture Notes in Math. **1317**, Springer-Verlag, New York, 1988.

J. Lindenstrauss and L. Tzafriri, On the complemented subspaces problem, *Israel J. Math.* **9** (1971), 263–269.

J. Loś, Sur le théorème de Gödel pour les théories indénombrables, *Bull. de l'Acad. Polon. des Sci.* Cl. III, 2 (1954), 319–320.

J. Loś and C. Ryll-Nardzewski, Effectiveness of the representation theory for Boolean algebras, *Fund. Math.* **41** (1954), 49–56.

E. Lowen-Colebunders, *Function Classes of Cauchy Continuous Maps*, Monographs Textbooks Pure Appl. Math. **123**, Dekker, New York, 1989.

D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.

R. Lutz and M. Goze, *Nonstandard Analysis: A Practical Guide with Applications*, Lecture Notes in Math. **881**, Springer-Verlag, Berlin, 1981.

W. A. J. Luxemburg, A remark on a paper by N. G. de Bruijn and P. Erdös, *Nederl. Akad. Wetensch.* (A) **65** (1962), 343–345.

W. A. J. Luxemburg, Reduced powers of the real number system and equivalents of the Hahn-Banach extension theorem, pp. 123–137 in: *Applications of Model Theory to Algebra, Analysis, and Probability (symposium, CalTech, 1967)*, ed. by W. A. J. Luxemburg, Holt Rinehart Winston, New York, 1969.

W. A. J. Luxemburg, Arzela's dominated convergence theorem for the Riemann integral, *Amer. Math. Monthly* **78** (1971), 970–979.

W. A. J. Luxemburg, Non-standard analysis, pp. 107–119 in: *Logic, Foundations of Mathematics, and Computability Theory (part 1 of proc. conf. London, Ontario, 1975)*, ed. by R. E. Butts and J. Hintikka, Univ. Western Ontario Ser. Philos. Sci. **9**, Reidel, Dordrecht, 1977.

H. van Maaren, Generalized pivoting and coalitions, pp. 155–176 in: *The Computation and Modelling of Economic Equilibria*, ed. by D. Talman and G. van der Laan, Contributions to Economic Analysis **167**, North-Holland, Amsterdam, 1987.

M. Machover and J. Hirschfeld, *Lectures on Non-Standard Analysis*, Lecture Notes in Math. **94**, Springer-Verlag, Berlin, 1969.

S. Mac Lane, *Categories for the Working Mathematician*, Graduate Texts in Math. **5**, Springer-Verlag, New York, 1971.

S. Mac Lane, *Mathematics: Form and Function*, Springer-Verlag, New York, 1986.

S. Mac Lane and G. Birkhoff, *Algebra*, MacMillan, New York, 1967.

J. Malitz, *Introduction to Mathematical Logic*, Springer-Verlag, New York, 1979.

Y. I. Manin, *A Course in Mathematical Logic*, Graduate Texts in Math. **53**, Springer-Verlag, New York, 1977.

R. Mánka, Some forms of the axiom of choice, *Jahrb. Kurt-Gödel Ges.* **1** (1988), 24–34.

R. Mánka, private communication, 1992.

E. Marchi and J.-E. Martínez-Legaz, Some results on approximate continuous selections, fixed points and minimax inequalities, pp. 327–342 in: *Fixed Point Theory and Applications*, ed. by M. A. Théra and J.-B. Baillon, Pitman Research Notes in Mathematics **252**, Longman Scientific and Technical, New York, 1991.

S. Mazurkiewicz, Teoria zbiorów G_δ (Theory of G_δ sets), *Wektor* **6** (1917–1918), 129–185.

R. N. McKenzie, G. F. McNulty, and W. F. Taylor, *Algebras, Lattices, Varieties: Volume I*, Wadsworth & Brooks/Cole, Monterey, 1987.

R. M. McLeod, *The Generalized Riemann Integral*, Carus Math. Monographs **20**, Math. Assoc. America, Washington, D.C., 1980.

E. J. McShane, Partial orderings and Moore-Smith limits, *Amer. Math. Monthly* **59** (1952), 1–11.

E. J. McShane, *Unified Integration*, Pure and Appl. Math. **107**, Academic Press, Orlando, 1983.

E. Mendelson, *Introduction to Mathematical Logic*, Van Nostrand, Princeton, 1964.

P. R. Meyers, A converse to Banach's contraction theorem, *J. of Research of the Nat'l. Bureau of Standards* **71B** (1967), 73–76.

H. I. Miller and B. Živaljević, Remarks on the zero-one law, *Math. Slovaca* **34** (1984), 375–384.

L. Mirsky, *Transversal Theory: An Account of Some Aspects of Combinatorial Mathematics*, Math. Sci. Engrg. **75**, Academic Press, New York/London, 1971.

H. Mönch and G.-F. von Harten, On the Cauchy problem for ordinary differential equations in Banach spaces, *Arch. Math.* **39** (1982), 153–160.

J. D. Monk, ed., *Handbook of Boolean Algebras*, North-Holland, Amsterdam, 1989.

E. H. Moore, Introduction to a form of general analysis, pp. 1–150 in: *The New Haven Mathematical Colloquium (Proc. of the 5th Colloq. of the A.M.S., New Haven, 1906)*, Yale Univ. Press, New Haven, 1910.

E. H. Moore and H. L. Smith, A general theory of limits, *Amer. J. Math.* **44** (1922), 102–121.

G. H. Moore, *Zermelo's Axiom of Choice: Its Origins, Development, and Influence*, Stud. Hist. Math. Phys. Sci. **8**, Springer-Verlag, New York, 1982.

G. H. Moore, Lebesgue's measure problem and Zermelo's axiom of choice: the mathematical effects of a philosophical dispute, pp. 129–154 in: *History and Philosophy of Science: Selected Papers*, ed. by J. W. Dauben and V. S. Sexton, Ann. New York Acad. Sci. **412**, New York Acad. Sci., New York, 1983.

J. C. Morgan II, *Point Set Theory*, Dekker, New York, 1990.

M. Morillon, Divers axiomes de choix, Exp. no. 4 and 5 (22 pages) in: *Seminaire d'analyse 1985–1986 Université de Clermont-Ferrand II*, Aubiere, France, 1986.

M. G. Murdeshwar, *General Topology*, Wiley, New York, 1983.

J. Mycielski, Two remarks on Tychonoff's product theorem, *Bull. Acad. Polon. Sci. Ser. Sci. Math. Astr. Phys.* **12** (1964), 439–441.

J. Myjak, *Orlicz Type Category Theorems for Functional and Differential Equations*, Dissertationes Math. (Rozprawy Mat.) **206**, Polish Acad. Sci., Warsaw, 1983.

L. Nachbin, *The Haar Integral*, Van Nostrand, Princeton, 1965.

E. Nagel and J. R. Newman, *Gödel's Proof*, New York University Press, New York, 1958.

H. Nakano, On an extension theorem, *Proc. Japan Acad.* **35** (1959), 127.

L. Narici and E. Beckenstein, On non-Archimedean analysis, pp. 107–115 in: *Mathematical Vistas: Papers from the Mathematics Section*, ed. by J. Malkevitch and D. McCarthy, Ann. New York Acad. Sci. **607**, New York, 1990.

J. Nedoma, Note on generalized random variables, pp. 139–141 in: *Transactions of the First Prague Conference on Information Theory, Statistical Decision Functions, Random Pro-*

cesses, held at Liblice, 1956, Publishing House of the Czechoslovak Academy of Sciences, Prague, 1957.

E. Nelson, Internal set theory: a new approach to nonstandard analysis, *Bull. Amer. Math. Soc.* **83** (1977), 1165–1198.

M. Neumann, Automatic continuity of linear operators, pp. 269–296 in: *Functional Analysis: Surveys and Recent Results II (proceedings of the Conference on Functional Analysis, Paderborn, Germany, 1979)*, ed. by K.-D. Bierstedt and B. Fuchssteiner, North-Holland Math. Stud. **38** / Notas Mat. **68**, North-Holland, Amsterdam, 1980.

M. M. Neumann, Uniform boundedness and closed graph theorems for convex operators, *Math. Nachr.* **120** (1985), 113–125.

M. M. Neumann, Sandwich theorems for operators of convex type, *J. Math. Anal. Appl.* **188** (1994), 759–773.

A. Nijenhuis, Strong derivatives and inverse mappings, *Amer. Math. Monthly* **81** (1974), 969–980.

A. Nijenhuis, Addendum to "Strong derivatives and inverse mappings," *American Math. Monthly* **83** (1976), 22.

L. Nirenberg, *Functional Analysis, 1960–1961*, lecture notes based on notes taken by Lesley Sibner, Courant Institute of Mathematical Sciences, New York University, New York, 1975.

J. C. Oxtoby, *Measure and Category: A Survey of the Analogies Between Topological and Measure Spaces*, 2nd edition, Graduate Texts in Math. **2**, Springer-Verlag, New York, 1980.

N. H. Pavel, *Nonlinear Evolution Operators and Semigroups; Applications to Partial Differential Equations*, Lecture Notes in Math. **1260**, Springer-Verlag, Berlin, 1987.

J. Pawlikowski, The Hahn-Banach theorem implies the Banach-Tarski paradox, *Fund. Math.* **138** (1991), 21–22.

B. J. Pettis, On continuity and openness of homomorphisms in topological groups, *Annals of Math.* **52** (1950), 293–308.

A. L. Peressini, *Ordered Topological Vector Spaces*, Harper and Row, New York, 1967.

W. J. Pervin, Quasi-uniformization of topological spaces, *Math. Ann.* **147** (1962), 316–317.

P. B. Pierce, On the preservation of certain properties of functions under composition, thesis, Syracuse University, Syracuse, 1994.

D. Pincus, Independence of the prime ideal theorem from the Hahn-Banach theorem, *Bull. Amer. Math. Soc.* **78** (1972), 766–770.

D. Pincus, The strength of the Hahn-Banach theorem, pp. 203–248 in: *Victoria Symposium on Nonstandard Analysis (University of Victoria 1972)*, ed. by A. Hurd and P. Loeb; Lecture Notes in Math. **369**, Springer-Verlag, Berlin, 1974.

D. Pincus, Adding dependent choice to the prime ideal theorem, pp. 547–565 in: *Logic Colloquium 76 (Proceedings of a conference held in Oxford in July 1976)*, ed. by R. O. Gandy and J. M. E. Hyland, North-Holland, Amsterdam/New York, 1977.

D. Pincus and R. M. Solovay, Definability of measures and ultrafilters, *J. Symbolic Logic* **42** (1977), 179–190.

J. D. Pryce, Weak compactness in locally convex spaces, *Proc. Amer. Math. Soc.* **17** (1966), 148–155.

M. M. Rao, *Measure Theory and Integration*, Wiley, New York, 1987.

M. M. Rao and Z. D. Ren, *Theory of Orlicz Spaces*, Monographs Textbooks Pure Appl. Math. **146**, Dekker, New York, 1991.

H. Rasiowa, *An Algebraic Approach to Non-Classical Logics*, Studies in Logic and the Foundations of Mathematics **78**, North-Holland, Amsterdam, 1974.

H. Rasiowa and R. Sikorski, A proof of the completeness theorem of Gödel, *Fundamenta Mathematicae* **37** (1951), 193–200.

H. Rasiowa and R. Sikorski, *The Mathematics of Metamathematics*, Monografie Matematyczne **41**, Panstwowe Wydawnictwo Naukowe, Warsaw, 1963.

Y. Rav, Variants of Rado's selection lemma and their applications, *Math. Nachr.* **79** (1977), 145–165.

M. Reed and B. Simon, *Functional Analysis*, Methods of Modern Mathematical Physics **1**, Academic Press, New York, 1972.

P. F. Reichmeider, *The Equivalence of Some Combinatorial Matching Theorems*, Polygonal Publishing House, Washington, New Jersey, 1984.

I. L. Reilly, Quasi-gauge spaces, *J. London Math. Soc.* **6** (1973), 481–487.

N. M. Rice, A general selection principle with applications in analysis and algebra, *Canad. Math. Bull.* **11** (1968), 573–584.

F. Richman, private communication, 1994.

C. Riley, A note on strict convexity, *Amer. Math. Monthly* **88** (1981), 198–199.

J. R. Ringrose, A note on uniformly convex spaces, *J. London Math. Soc.* **34** (1959), 92.

J. F. Ritt, *Theory of Functions*, King's Crown Press, New York, 1946.

J. W. Robbin, *Mathematical Logic: A First Course*, Benjamin, New York, 1969.

A. Robert, *Non-Standard Analysis*, Wiley, New York, 1988.

A. W. Roberts and D. E. Varberg, *Convex Functions*, Academic Press, New York, 1973.

A. W. Roberts and D. E. Varberg, Another proof that convex functions are locally Lipschitz, *Amer. Math. Monthly* **81** (1974), 1014–1016.

A. Robinson, Non-standard analysis, *Kon. Nederl. Akad. Wetensch. Amsterdam Proc.* **A64** (= *Indag. Math.* **23**) (1961), 432–440.

A. Robinson and E. Zakon, A set theoretical characterization of enlargements, pp. 109–122 in: *Applications of Model Theory to Algebra, Analysis, and Probability (symposium, CalTech, 1967)*, ed. by W. A. J. Luxemburg, Holt Rinehart Winston, New York, 1969.

R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

C. A. Rogers, A less strange version of Milnor's proof of Brouwer's fixed-point theorem, *Amer. Math. Monthly* **87** (1980), 525–527.

S. Rolewicz, *Metric Linear Spaces*, Reidel, Dordrecht, 1985.

A. B. Romanowska and J. D. H. Smith, *Modal Theory: An Algebraic Approach to Order, Geometry, and Convexity*, Res. Exp. Math. **9**, Heldermann-Verlag, Berlin, 1985.

U. Rønnow, On integral representation of vector valued measures, *Math. Scand.* **21** (1967), 45–53.

A. C. M. van Rooij, *Non-Archimedean Functional Analysis*, Monographs Textbooks Pure Appl. Math. **51**, Dekker, New York, 1978.

E. E. Rosinger, *Non-Linear Partial Differential Equations: An Algebraic View of Generalized Solutions*, North-Holland Math. Stud. **164**, North-Holland, Amsterdam, 1990.

J. B. Rosser, *Logic for Mathematicians*, McGraw-Hill, New York, 1953. Second edition, Chelsea, New York, 1978.

J. B. Rosser, An informal exposition of proofs of Gödel's theorems and Church's theorem, *Journal of Symbolic Logic* **4** (1939), 53–60; also reprinted in pp. 223–230 of: *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*, ed. by M. Davis, Raven Press, Hewlett, New York, 1965.

H. Rubin and J. E. Rubin, *Equivalents of the Axiom of Choice II*, Stud. Logic Found. Math. **116**, North-Holland, Amsterdam, 1985.

H. Rubin and D. Scott, Some topological theorems equivalent to the Boolean prime ideal theorem, Preliminary report, *Bull. Amer. Math. Soc.* **60** (1954), 389.

M. E. Rudin, A new proof that metric spaces are paracompact, *Proc. Amer. Math. Soc.* **20** (1969), 603.

W. Rudin, *Fourier Analysis on Groups*, Interscience Tracts in Pure and Applied Mathematics **12**, Wiley, New York, 1960.

D. Rüthing, Some definitions of the concept of function from Joh. Bernoulli to N. Bourbaki, *Mathematical Intelligencer* **6** (1984), 72–77.

C. H. Sah, review of R. J. Gardner and S. Wagon's paper "At long last, the circle has been squared" (*Not. Amer. Math. Soc.* **10** (1989), 1338–1343), in *Math. Reviews* **90j:51023**, 1990.

H. H. Schaefer, *Topological Vector Spaces*, Graduate Texts in Math. **3**, Springer-Verlag, New York, 1971.

E. Schechter, Existence and limits of Carathéodory-Martin evolutions, *Nonlin. Anal.* **5** (1981), 897–930.

E. Schechter, Compact perturbations of linear m-dissipative operators which lack Gihman's property, pp. 142–161 in: *Nonlinear Semigroups, Partial Differential Equations and Attractors (proc. conference at Howard University, Washington D.C., 1985)*, ed. by T. L. Gill and W. W. Zachary, Lecture Notes in Math. **1248**, Springer-Verlag, Berlin, 1987.

E. Schechter, A survey of local existence theories for abstract nonlinear initial value problems, pp. 136–184 in: *Nonlinear Semigroups, Partial Differential Equations, and Attractors (proc. sympos. Washington 1987)*, ed. by T. L. Gill and W. W. Zachary, Lecture Notes in Math. **1394**, Springer-Verlag, Berlin, 1989.

E. Schechter, Two topological equivalents of the axiom of choice, *Z. Math. Logik Grundlag. Math.* **38** (1992), 555–557.

E. Schechter, submitted for publication, 1996.

E. Schechter, K. Ciesielski, and J. S. Norden, private communication, 1993.

M. Schechter, *Principles of Functional Analysis*, Academic Press, New York, 1971.

S. Schwabik, *Generalized Ordinary Differential Equations*, Series in Real Analysis **5**, World Scientific Publishing, Singapore, 1992.

D. Scott, Axiomatizing set theory, pp. 207–214 in: *Axiomatic Set Theory, Part 2 (Los Angeles 1967)*, ed. by T. Jech, Proc. Sympos. Pure Math. **13**, part 2, AMS, Providence, 1974.

N. A. Shanin, On separation in topological spaces, *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS* **38** (1943), 110–113.

I. A. Shashkin, *Fixed Points*, Mathematical World **2**, Amer. Math. Soc., Providence, 1991.

S. Shelah, Can you take Solovay's inaccessible away? *Israel J. Math.* **48** (1984), 1–47.

S. Shelah, On measure and category, *Israel J. Math.* **52** (1985), 110–114.

S. Shelah, private communication, 1991.

J. R. Shoenfield, *Mathematical Logic*, Addison-Wesley, Reading, 1967.

J. R. Shoenfield, Axioms of set theory, pp. 321–344 in: *Handbook of Mathematical Logic*, ed. by J. Barwise, North-Holland, Amsterdam, 1977.

J. Shurman, textbook in progress, Wiley, 1996.

W. Sierpiński, Fonctions additives non complètement additives et fonctions non mesurables, *Fund. Math.* **30** (1938), 96–99.

R. Sikorski, On an analogy between measures and homomorphisms, *Annales de la Societé Polonaise de Mathematique = Rocznik Polskiego Tow. Matematyycznego* **23** (1950), 1–20.

R. Sikorski, *Boolean Algebras*, Ergeb. Math. Grenzgeb. **25**, Springer-Verlag, Berlin, 1st edition 1960, 2nd edition 1964.

S. Simons, A convergence theorem with boundary, *Pacific J. Math.* **40** (1972), 703–708.

S. Simons, An eigenvector proof of Fatou's lemma for continuous functions, *Math. Intelligencer* **17** (1995), 67–70.

I. Singer, *Bases in Banach Spaces I*, Grundlehren Math. Wiss. **154**, Springer-Verlag, Berlin, 1970.

T. Skolem, Über die Nicht-charakterisierbarkeit der zahlenreihe Mittels endlich odor abzahlbar unendlich vieler Aussagen mit aussachliesslich Zahlenvariablen, *Fund. Math.* **23** (1934), 150–161.

D. R. Smart, *Fixed Point Theorems*, Cambridge University Press, Cambridge, 1974.

M. F. Smiley, Filters and equivalent nets, *Amer. Math. Monthly* **64** (1957), 336–338.

R. M. Solovay, unpublished notes, 1964–1965.

R. M. Solovay, A model of set theory in which every set of reals is Lebesgue measurable, *Ann. of Math.* **92** (1970), 1–56.

Z.-F. Song, Existence of generalized solutions for ordinary differential equations in Banach spaces, *J. Math. Anal. Appl.* **128** (1987), 405–412.

R. M. Starr, Quasi-equilibria in markets with non-convex preferences, *Econometrica* **37** (1969), 25–38.

L. A. Steen and J. A. Seebach, Jr., *Counterexamples in Topology*, Holt Rinehart Winston, New York, 1970.

H. G. Steiner, Equivalente Fassungen des Vollständigkeitsaxioms für die Theorie der reelen Zahlen, *Math. Phys. Sem. Ber.* **13** (1966), 180–201.

J. Stern, Generic extensions which do not add random reals, pp. 395–407 in: *Methods in Mathematical Logic*, ed. by C. A. Di Prisco, Lecture Notes in Math. **1130**, Springer-Verlag, Berlin, 1985.

J. Stillwell, Eisenstein's footnote, *Mathematical Intelligencer* **17** (1995), 58–62.

J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions I*, Grundlehren Math. Wiss. **163**, Springer-Verlag, New York, 1970.

R. R. Stoll, *Set Theory and Logic*, W. H. Freeman, San Francisco, 1963.

A. H. Stone, Paracompactness and product spaces, *Bull. AMS* **54** (1948), 977–982.

A. H. Stone, Topology and measure theory, pp. 43–48 in: *Measure Theory (Proc. Conf. Oberwolfach 1975)*, ed. by A. Bellow and D. Kolzow, Lecture Notes in Math. **541**, Springer-Verlag, Berlin, 1976.

C. Swartz, *An Introduction to Functional Analysis*, Dekker, New York, 1992.

Y. Takeuchi, Estructuras topologicas de $\mathbb{R}^*$, *Boletín de Matemáticas* **18** (1984), 36–72.

G. Takeuti and W. M. Zaring, *Introduction to Axiomatic Set Theory*, 2nd edition, Graduate Texts in Math. **1**, Springer-Verlag, New York, 1982.

A. Tarski, Ideale in vollständigen Mengenkörpern, *Fund. Math.* **32** (1939), 45–63.

A. Tarski, Prime ideals for set algebras and the axiom of choice (preliminary report), *Bull. Amer. Math. Soc.* **60** (1954), 391.

R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math.* **141** (1995), 553–572.

J. Thierfelder, Separation theorems for sets in product spaces and equivalent assertions, *Kybernetika* **27** (1991), 522–534.

C. Thomassen, Infinite graphs, pp. 129–160 in: *Selected Topics in Graph Theory 2*, ed. by L. W. Beineke and R. J. Wilson, Academic Press, London/New York, 1983.

H. Thurston, Math bite: a simple proof that every sequence has a monotone subsequence, *Amer. Math. Monthly* **67** (1994), 344.

H. Tietze, *Famous Problems of Mathematics*, Graylock Press, New York, 1965.

R. Tijdeman, Hilbert's seventh problem: on the Gel'fond-Baker method and its applications, pp. 241–268 in: *Mathematical Developments Arising from Hilbert Problems*, ed. by F. Browder, Proc. Sympos. Pure Math. **28**, Amer. Math. Soc., Providence, 1976.

F. Treves, *Topological Vector Spaces, Distributions and Kernels*, Pure Appl. Math. **25**, Academic Press, New York, 1967.

A. S. Troelstra and D. van Dalen, *Constructivism in Mathematics: An Introduction, vol. I*, Stud. Logic Found. Math. **121**, North-Holland, Amsterdam, 1988.

C. Tsinakis, lecture notes, 1993.

A. M. Turing, Systems of logic based on ordinals, *Proc. London Math. Soc.*, Ser. 2, **45** (1939), 161–228.

H. Tuy, *Convex Inequalities and the Hahn-Banach Theorem*, Dissertat. Math. (Rozprawy Mat.) **97**, Polish Acad. Sci., Warsaw, 1972.

J. E. Vaughan, *The Irrational Numbers: Topology and Set Theory*, textbook in preparation, 1988.

M. L. J. van de Vel, *Theory of Convex Structures*, North-Holland Mathematical Library **50**, North-Holland, Amsterdam, 1993.

G. Vidossich, Existence, uniqueness, and approximation of fixed points as a generic property, *Bol. Sci. Brasil. Math.* **5** (1974), 17–29.

A. Villani, Another note on the inclusion $L^p(\mu) \subset L^q(\mu)$, *Amer. Math. Monthly* **92** (1985), 485–487.

P. Volkmann, Quelques résultats récents sur les equations différentielles dans les espaces de Banach, pp. 447–452 in: *Fixed Point Theory and Applications*, ed. by M. A. Théra and J.-B. Baillon, Pitman Research Notes in Math. Series **252**, Longman Scientific and Technical, Essex, 1991.

I. I. Vrabie, *Compactness Methods for Nonlinear Evolutions*, Pitman Monographs and Surveys in Pure and Applied Math. **32**, Longman Scientific and Technical, Essex, 1988.

B. Z. Vulikh, *Introduction to the Theory of Partially Ordered Spaces*, Wolters-Noordhdoff Scientific Publicns., Groningen, 1967.

R. Výborný, Kurzweil-Henstock absolute integrable means McShane integrable, *Real Analysis Exchange* **20** (1994/95), 363–366.

L. Waelbroeck, *Topological Vector Spaces and Algebras*, Lecture Notes Math. **230**, Springer-Verlag, Berlin, 1971.
S. Wagon, *The Banach-Tarski Paradox*, Encyclopedia Math. Appl. **24**, Cambridge Univ. Press, Cambridge, 1985.
J. S. Weber, An elementary proof of the conditions for a generalized Condorcet paradox, *Public Choice* **77** (1993), 415–419.
A. J. Weir, *Integration and Measure, Volume Two: General Integration and Measure*, Cambridge University Press, Cambridge, 1974.
A. Wilansky, *Topology for Analysis*, Ginn and Co., Waltham, 1970.
A. Wilansky, *Modern Methods in Topological Vector Spaces*, McGraw-Hill International Book Co., London, 1978.
S. Willard, *General Topology*, Addison-Wesley, Reading, 1970.
Wolfram Research, *Solving the Quintic with Mathematica* (a wall-poster), 1994.
E. S. Wolk, On the inadequacy of cofinal subnets and transfinite sequences, *Amer. Math. Monthly* **89** (1982), 310–311.
L. A. Wolsey, Cubical Sperner lemmas as applications of generalized complementary pivoting, *J. Combin. Theory* (A) **23** (1977), 78–87.
Y.-C. Wong and K.-F. Ng, *Partially Ordered Topological Vector Spaces*, Clarendon Press, Oxford, 1973.
J. D. M. Wright, On the continuity of mid-point convex functions, *Bull. London Math. Soc.* **7** (1975), 89–92.
J. D. M. Wright, Functional analysis for the practical man, pp. 283–290 in: *Functional Analysis: Surveys and Recent Results (proceedings of the Conference on Functional Analysis, Paderborn, Germany, 1976)*, ed. by K.-D. Bierstedt and B. Fuchssteiner, North-Holland Math. Stud. **27** / Notas Mat. **63**, North-Holland, Amsterdam, 1977.
J. W. A. Young, *Monographs on Topics of Modern Mathematics Relevant to the Elementary Field*, Longmans, Green, New York, 1911; reprinted by Dover, New York, 1955.
K. Yosida, *Functional Analysis*, Grundlehren Math. Wiss. **123**, Springer-Verlag, New York, various editions, 1964 and later.
W. Yuan, ed., *Goldbach Conjecture*, World Scientific Publishing, Singapore, 1984.
J. Zowe, Sandwich theorems for convex operators with values in an ordered vector space, *J. Math. Anal. Applic.* **66** (1978), 282–296.

Index and Symbol List

Symbols are listed at the end of this index.

LIST OF SYMBOLS

The Greek alphabet:

A, α	alpha
B, β	beta
Γ, γ	gamma
Δ, δ	delta
E, ε	epsilon
Z, ζ	zeta
H, η	eta
Θ, θ	theta
I, ι	iota
K, κ	kappa
Λ, λ	lambda
M, μ	mu
N, ν	nu
Ξ, ξ	xi
O, o	omicron
Π, π	pi
P, ρ	rho
Σ, σ	sigma
T, τ	tau
Υ, υ	upsilon
Φ, φ	phi
X, χ	chi
Ψ, ψ	psi
Ω, ω	omega

Sets of Numbers:

This Page Intentionally Left Blank