COM3502/4502/6502
SPEECH PROCESSING

Lecture 15
The Fourier Transform

The University Of Sheffield.

1

Phase in Spectral Analysis

- In the last lecture it was shown how the spectrum could be computed by 'cosine correlation' (*Pd Example 14-6*)
- However, this only worked because the target signals all had 'zero phase'
- The correlation between two cosines varies according to the phase difference between them
  - in-phase → maximum correlation
  - 90° phase difference → zero correlation
  - 180° phase difference → maximum negative correlation
- In fact the correlation between two cosines varies as a cosine!

The University Of Sheffield.

2

Phase in Spectral Analysis

Multiplication of two cosines → $\cos(a).\cos(b) = \frac{1}{2}\left(\cos(a-b) + \cos(a+b)\right)$

Multiplication of two cosine waves → $\cos(\omega t).\cos(\omega t) = \frac{1}{2}\left(\cos(0) + \cos(2\omega t)\right)$

$= \frac{1}{2}\left(1 + \cos(2\omega t)\right)$ ← Constant + cosine at twice the frequency

Multiplication of two cosine waves differing in phase → $\cos(\omega t).\cos(\omega t - \phi) = \frac{1}{2}\left(\cos(\phi) + \cos(2\omega t - \phi)\right)$

Cosine of the phase difference + cosine at twice the frequency

The University Of Sheffield.

Phase in Spectral Analysis

Correlation of two cosine waves differing in phase →

$$q = \sum_{i=0}^{N-1} \cos(\omega nT).A\cos(\omega nT - \phi)$$

$$= \sum_{i=0}^{N-1} \left(\frac{A}{2}\left(\cos(\phi) + \cos(2\omega t - \phi)\right) \right)$$

$$= \sum_{i=0}^{N-1} \left(\frac{A}{2}\left(\cos(\phi)\right) \right) + \sum_{i=0}^{N-1} \left(\frac{A}{2}\left(\cos(2\omega t - \phi)\right) \right)$$

$$= \sum_{i=0}^{N-1} \left(\frac{A}{2}\left(\cos(\phi)\right) \right)$$

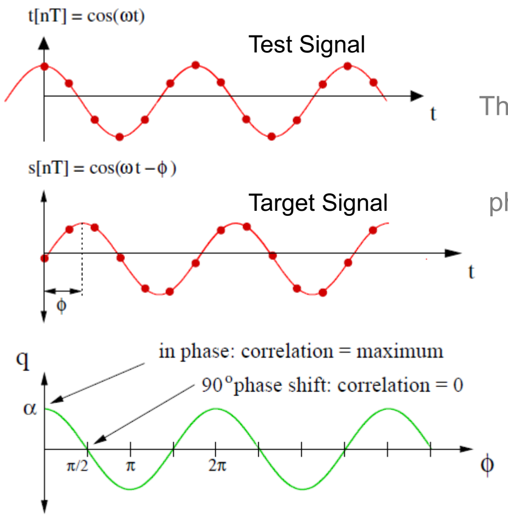
$$= \frac{A.N}{2}\cos(\phi)$$

$$= \alpha\cos(\phi)$$

Cosine of the phase difference

The University Of Sheffield.

Phase in Spectral Analysis



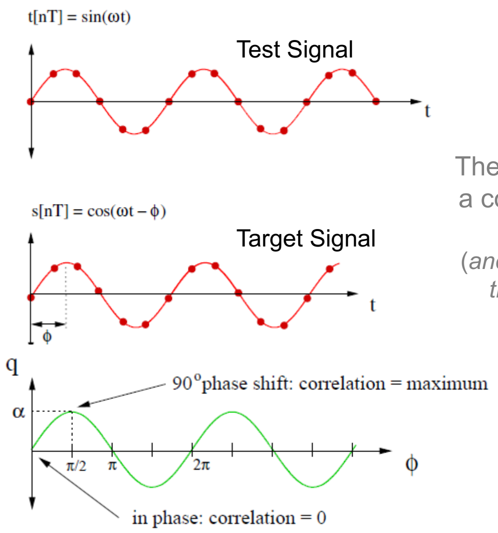
This means that cosine correlation cannot detect a $\pi/2$ ($90°$) phase shift in a target signal!

The University Of Sheffield.

5

© 2019 The University of Sheffield

Sine Correlation



The correlation between a cosine and a sine is a sine!

(and it varies according to the phase difference between them)

The University Of Sheffield.

6

Sine Correlation

Correlation of sine wave with cosine wave

$$q=\sum_{i=0}^{N-1}\sin(\omega nT).A\cos(\omega nT-\phi)$$

$$=\sum_{i=0}^{N-1}\cos(\omega nT-\pi/2).A\cos(\omega nT-\phi)$$

$$=\sum_{i=0}^{N-1}\left(\frac{A}{2}(\cos(\phi-\pi/2)+\cos(2\omega t-\phi-\pi/2))\right)$$

$$=\sum_{i=0}^{N-1}\left(\frac{A}{2}(\sin(\phi)+\sin(2\omega t-\phi))\right)$$

$$=\sum_{i=0}^{N-1}\left(\frac{A}{2}(\sin(\phi))\right)+\sum_{i=0}^{N-1}\left(\frac{A}{2}(\sin(2\omega t-\phi))\right)$$

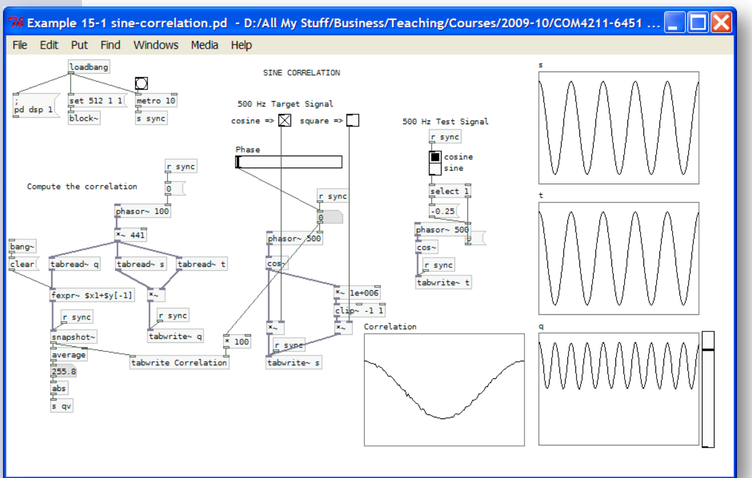
$$=\sum_{i=0}^{N-1}\left(\frac{A}{2}(\sin(\phi))\right)$$

$$=\alpha\sin(\phi)$$

Sine of the phase lag

The University Of Sheffield.

Sine Correlation



The University Of Sheffield.

Sine *and* Cosine Correlation

- Cosine correlation with cosine is a cosine
- Sine correlation with cosine is a sine
- Sine and cosine correlations are 90° ($\pi/2$) out of phase
- So can cosine and sine correlation be combined in some useful way?

The University Of Sheffield.

9

Amplitude and Phase

- Correlate a signal with *both* sines *and* cosines …

$$cosinecorrelation = \alpha.\cos(\phi)$$
$$sinecorrelation = \alpha.\sin(\phi)$$

- Note …

$$(\alpha.\cos(\phi))^2 + (\alpha.\sin(\phi))^2 = \alpha^2$$

- Hence the amplitude of the sinusoidal component independent of phase is given by …

$$\alpha = \sqrt{(cosinecorrelation)^2 + (sinecorrelation)^2}$$

- The phase of this component is given by …

$$\tan(\phi) = \alpha.\sin(\phi)/\alpha.\cos(\phi)$$

$$\therefore \phi = \tan^{-1}(sinecorrelation/cosinecorrelation)$$

The University Of Sheffield.

10

General Spectral Analysis Algorithm

- Recall that any periodic signal can be expressed as the sum of a fundamental sinusoid and its harmonics (*Fourier*)
- The individual components at a frequency $\Omega = p\omega$ can be found by correlating $s(nT)$ with $\cos(\Omega nT)$ and $\sin(\Omega nT)$
- Let $c(\Omega)$ be the cosine correlation and $s(\Omega)$ the sine correlation …

$$c(\Omega) = \sum_{n=0}^{N-1} s(nT).\cos\left(\frac{2\pi np}{N}\right) \quad p = 0, 1, \ldots, N\text{-}1$$

$$s(\Omega) = \sum_{n=0}^{N-1} s(nT).\sin\left(\frac{2\pi np}{N}\right) \quad p = 0, 1, \ldots, N\text{-}1$$

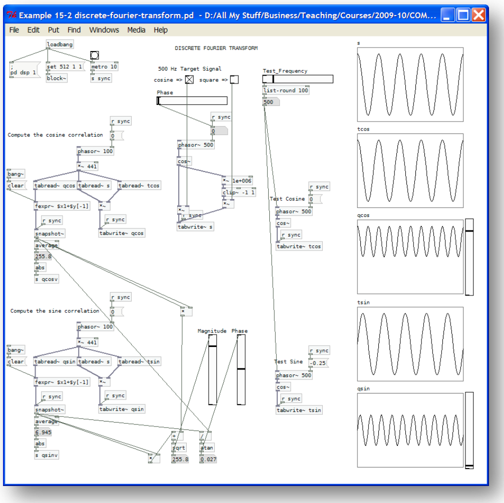
$$\mathrm{a_p} = \sqrt{c(\Omega)^2 + s(\Omega)^2}$$

$$\phi_p = \tan^{-1}\left(\frac{s(\Omega)}{c(\Omega)}\right)$$

- This is the '**Discrete Fourier Transform**' (*DFT*)

The University Of Sheffield.

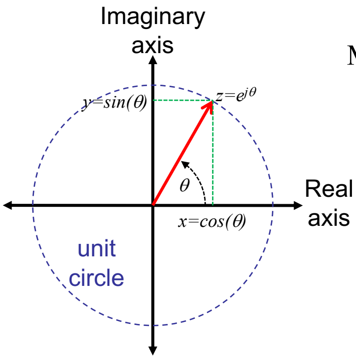
Discrete Fourier Transform



The University Of Sheffield.

Complex Numbers: *A Reminder*

$$z = x + jy$$



$$\text{Magnitude } |z| = \sqrt{x^2 + y^2}$$

$$\text{Phase } \theta = \tan^{-1}\left(\frac{y}{x}\right)$$

$$e^{j\theta} = \cos(\theta) + j\sin(\theta)$$

The University Of Sheffield.

13

© 2019 The University of Sheffield

Complex Formulation of the DFT

- The DFT is often expressed using 'complex number notation'
- The cosine and sine correlations are associated with the real and imaginary parts of a complex number …

$$S_p = \sum_{n=0}^{N-1} s(nT).\cos\left(\frac{2\pi np}{N}\right) - j\sum_{n=0}^{N-1} s(nT).\sin\left(\frac{2\pi np}{N}\right)$$

$$= \sum_{n=0}^{N-1} s(nT)\left[\cos\left(\frac{2\pi np}{N}\right) - j\sin\left(\frac{2\pi np}{N}\right)\right]$$

$$= \sum_{n=0}^{N-1} s(nT).e^{-j\left(\frac{2\pi np}{N}\right)}$$

The University Of Sheffield.

14

Complex Formulation of the DFT

- Hence the DFT can be expressed as …

$$S_p = \sum_{n=0}^{N-1} s(nT).e^{-j\left(\frac{2\pi np}{N}\right)} \qquad p = 0, 1, \ldots, N-1$$

… where S_p is a complex number whose magnitude and phase correspond to that of the spectrum of $s(nT)$ at a frequency p/NT

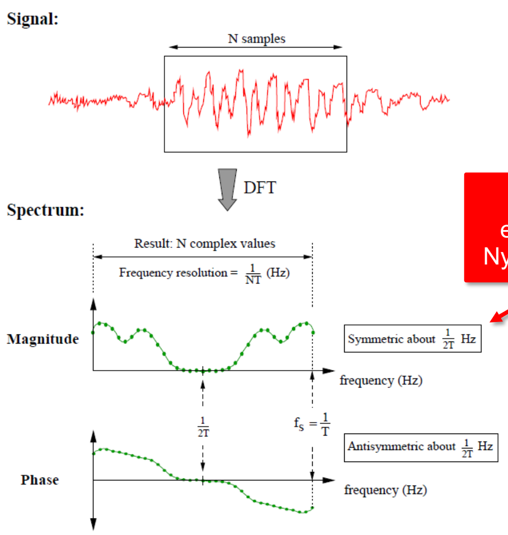
- Note also the 'inverse DFT' …

$$s(nT) = \frac{1}{N} \sum_{p=0}^{N-1} S_p.e^{j\left(\frac{2\pi np}{N}\right)} \qquad n = 0, 1, \ldots, N-1$$

The University Of Sheffield.

15

Using the DFT



This is the explanation for Nyquist's theorem

The University Of Sheffield.

16

Time *vs.* Frequency Resolution

Frequency resolution = $1/NT$ *(Hz)*

- Increasing the analysis frame N
  - decreases the spacing between the spectral components
  - reduces the ability to respond to changes in the signal
- Hence *large* N leads to 'narrowband analysis'
  - good spectral resolution
  - poor time resolution

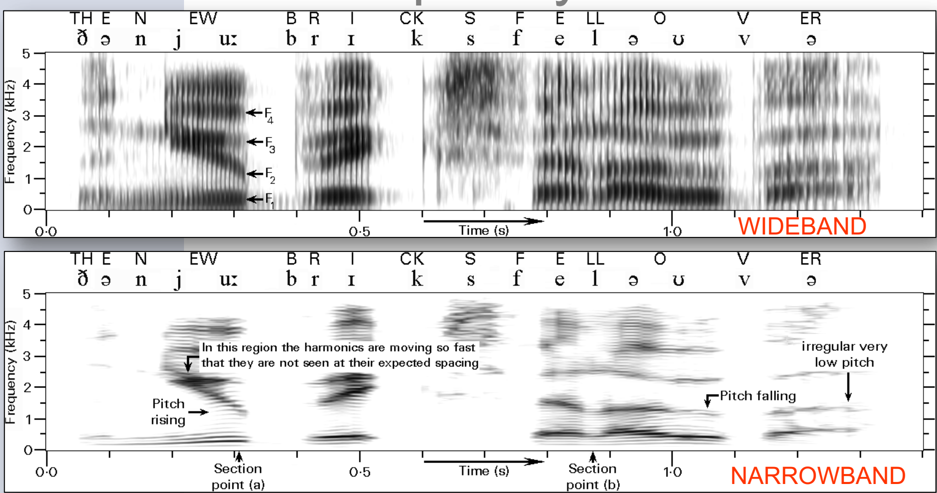
- Decreasing the analysis frame N
  - increases the spacing between the spectral components
  - increases the ability to respond to changes in the signal
- Hence *small* N leads to 'wideband analysis'
  - good time resolution
  - poor spectral resolution

This is the *time-frequency trade-off* we saw in Lecture 4

The University Of Sheffield.

17

Time *vs.* Frequency Resolution



Taken from: Holmes, J. N., & Holmes, W. (2002). *Speech Synthesis and Recognition*: Taylor & Francis.

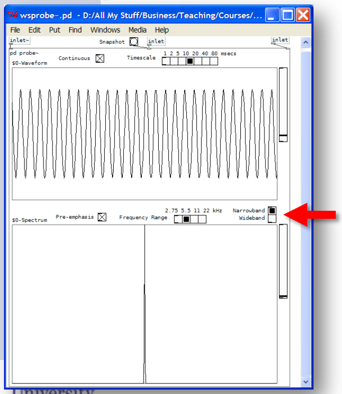
The University Of Sheffield.

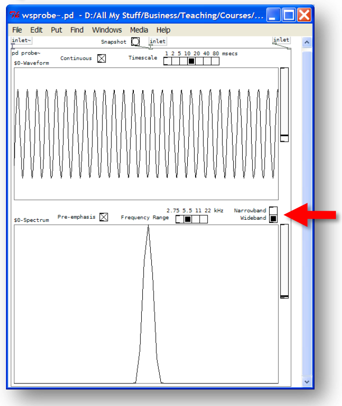
18

Time *vs.* Frequency Resolution

In [wsprobe~] …

- narrowband analysis uses a frame/block size of 8192 samples (*~190 msecs*)
- wideband analysis uses a frame/block size of 512 samples (*~12 msecs*)





The University Of Sheffield.

19

© 2019 The University of Sheffield

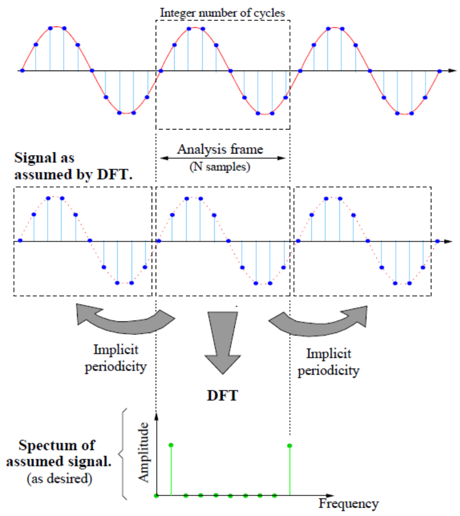
Implicit Periodicity in the DFT

- The DFT computes the spectrum at N evenly spaced discrete frequencies
- The algorithm assumes periodicity outside the analysis frame (*with a period equal to the frame length*)
- This means that discontinuities will arise for …
  - periodic signals with a *non-integer* number of cycles in the analysis frame
  - all aperiodic signals
  - all stochastic signals
- Such discontinuities give rise to unwanted spectral components

The University Of Sheffield.

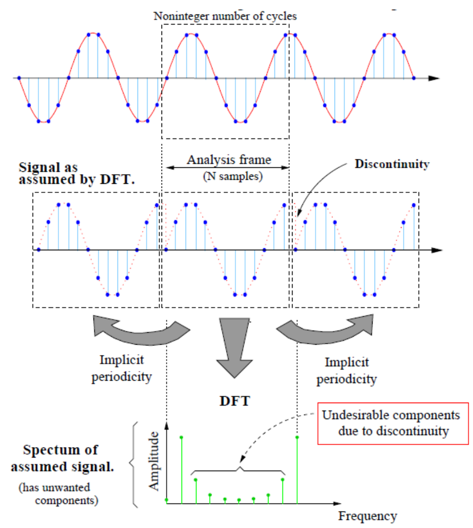
20

Implicit Periodicity in the DFT



21

Implicit Periodicity in the DFT



22

Windowing

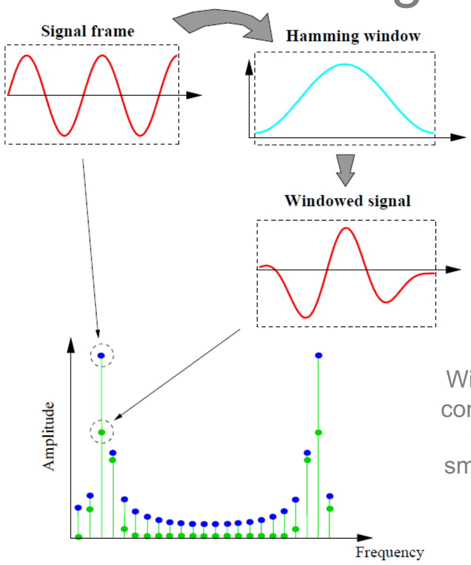
- The discontinuities arising from segmenting the signal into frames distorts the spectrum
- The distortion can be reduced by multiplying each signal frame with a 'window function'
- The most common window function is the 'Hamming Window' (*proposed by Richard W. Hamming*) …

$$w(nT) = 0.54 - 0.46.\cos\left(\frac{2\pi n}{N-1}\right)$$

The University Of Sheffield.

23

Windowing

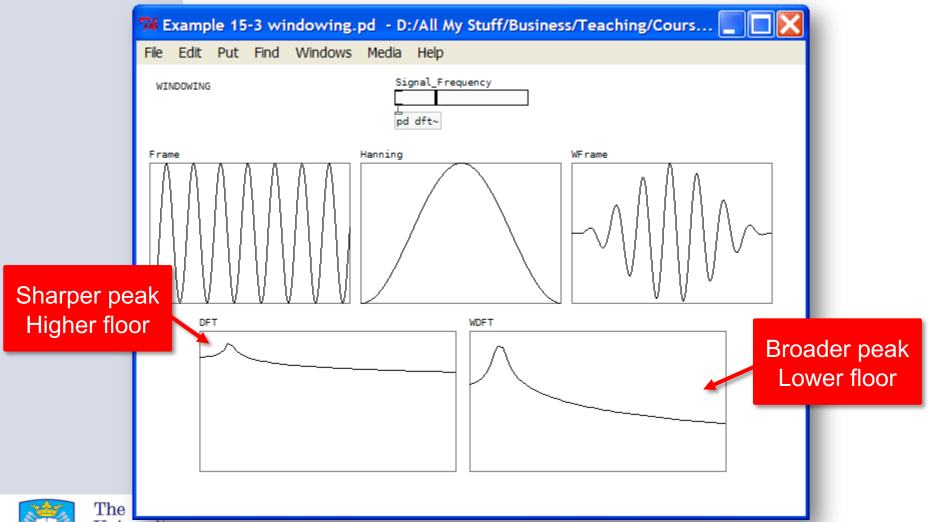


Windowing attenuates the components caused by the discontinuity, but also smears the spectral peaks

The University Of Sheffield.

24

Windowing



The University Of Sheffield.

25

The Fast Fourier Transform (FFT)

- Implementation of the DFT requires the order of N^2 multiply-add operations
- By exploiting symmetry, it is possible to devise an algorithm that requires only $N\log_2 N$ multiply-add operations
  - e.g. for $N=2048$, the result is ~100x faster
- This more efficient algorithm is the … '**Fast Fourier Transform**' (*FFT*)
- The FFT requires that the window/frame should be a power of 2 in size
- This can be achieved by …
  - choosing the appropriate analysis frame size, and/or
  - zero-padding a frame to the nearest power of 2

The University Of Sheffield.

26

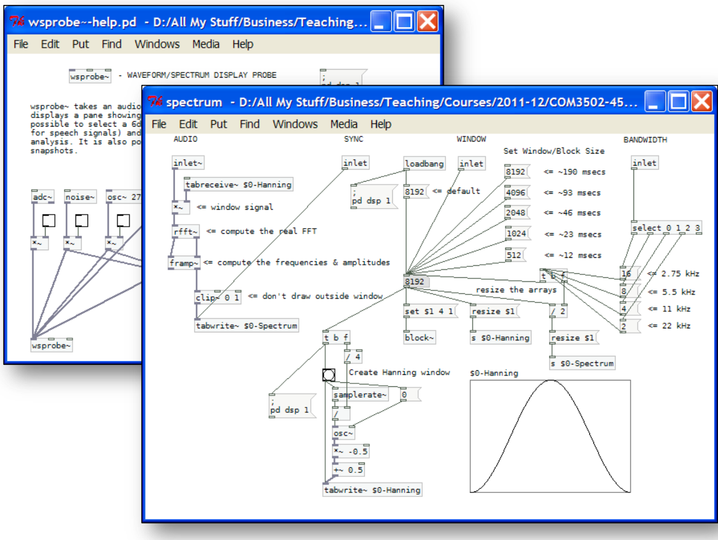
The FFT in Pure Data

- There are two key FFT objects available in Pd …
  - [rfft~] computes the forward transform
  - [ifft~] computes the inverse transform
- We saw both [rfft~] and [ifft~] in action in Example 12-3
- The magnitude spectrum can be computed by squaring the real and imaginary outputs, then taking the square root [sqrt~]
- Alternatively, the Pd object [framp~] outputs frequencies and amplitudes directly
- [rfft~] followed by [framp~] is used in [wsprobe~]

The University Of Sheffield.

27

The FFT in Pure Data



The University Of Sheffield.

28

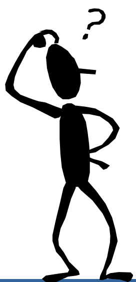
This lecture has covered …

- Phase in spectral analysis
- Sine and cosine correlation
- The discrete Fourier transform (*DFT*)
- Time versus frequency resolution
- Windowing
- The fast Fourier transform (*FFT*)

The University Of Sheffield.

© 2019 The University of Sheffield

Any Questions ?



The University Of Sheffield.

Next time …

The 'Z' Transform