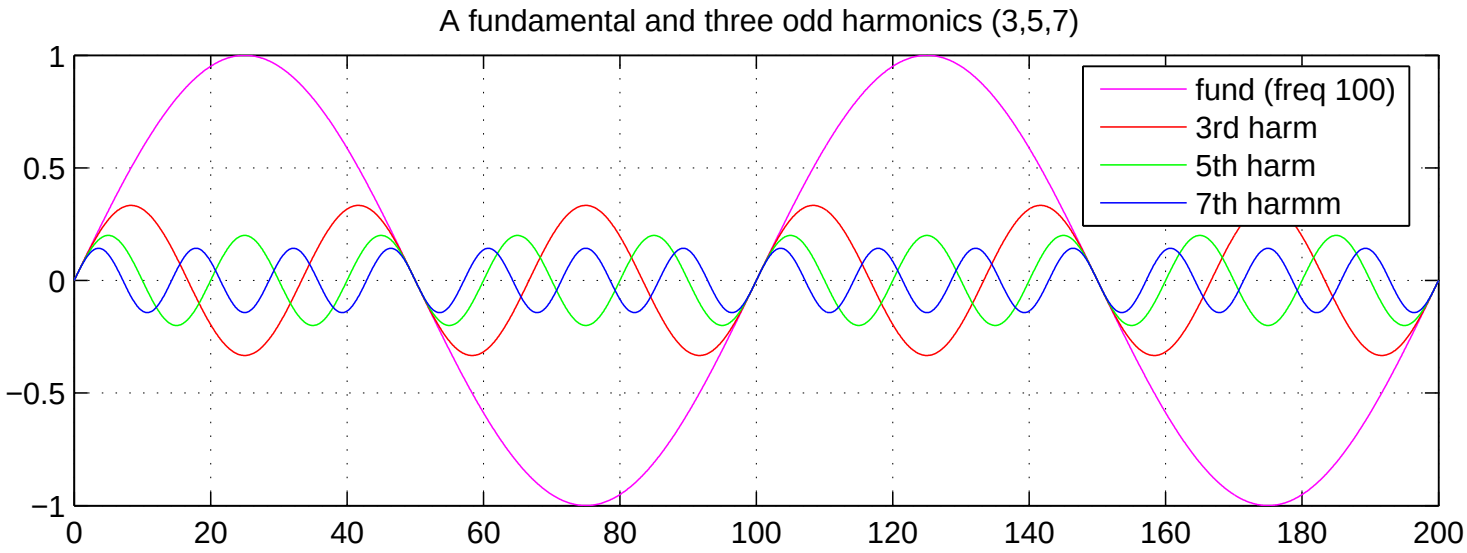
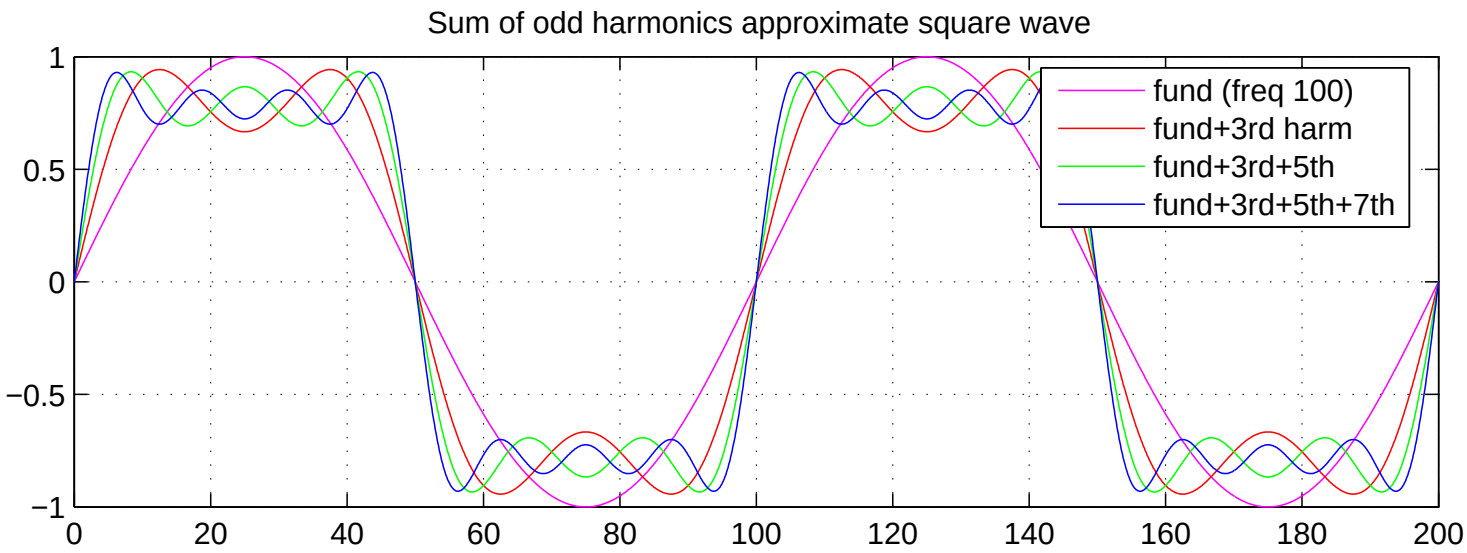
A Tutorial on Fourier Analysis

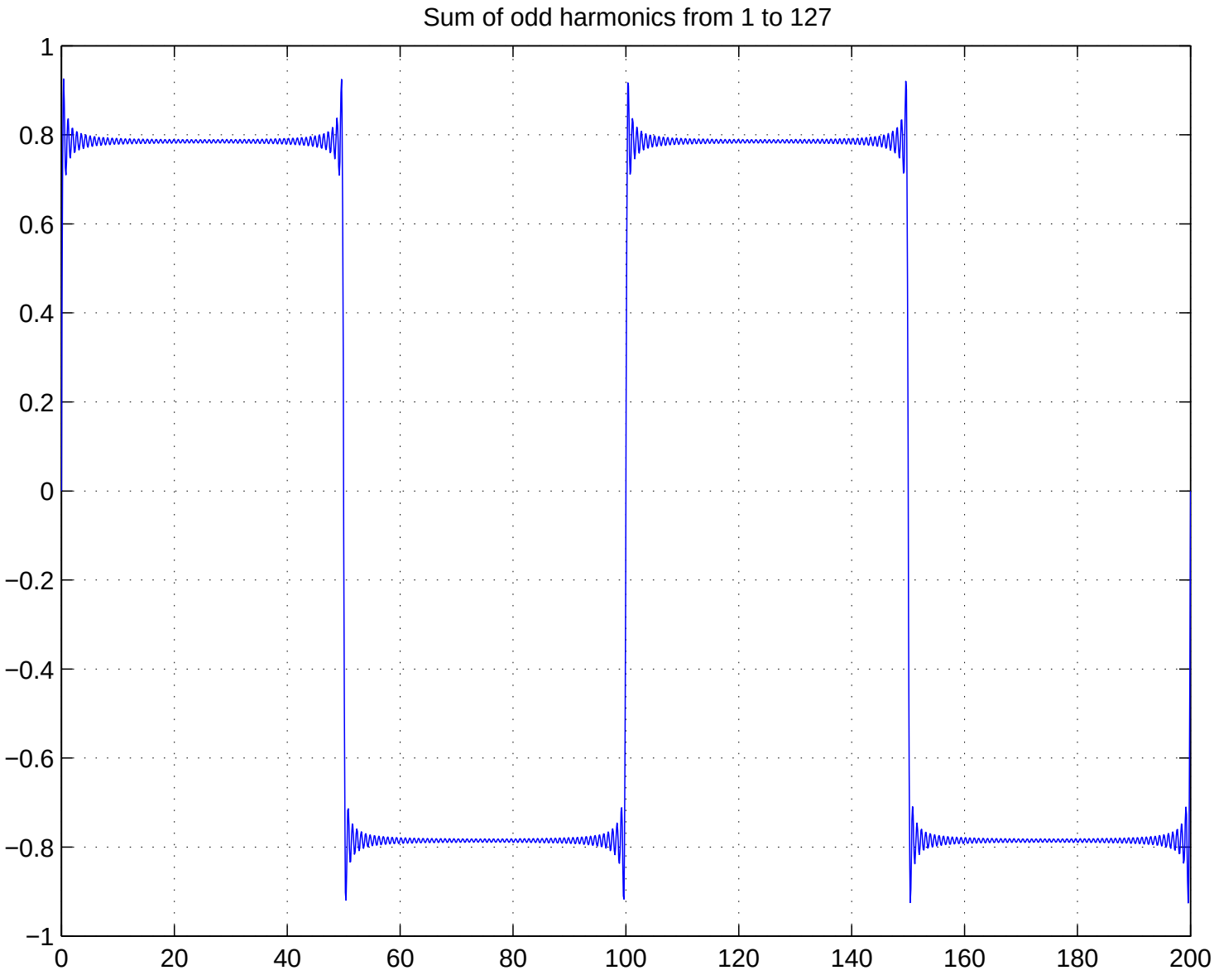
Douglas Eck

University of Montreal

NYU March 2006


A fundamental and three odd harmonics (3,5,7)
fund (freq 100)
3rd harm
5th harm
7th harmm
1
0.5
0
−0.5
−1
0
20
40
60
80
100
120
140
160
180
200


Sum of odd harmonics approximate square wave
fund (freq 100)
fund+3rd harm
fund+3rd+5th
fund+3rd+5th+7th
1
0.5
0
−0.5
−1
0
20
40
60
80
100
120
140
160
180
200


Sum of odd harmonics from 1 to 127
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0
20
40
60
80
100
120
140
160
180
200

Linear Combination

In the interval $[u_1, u_2]$ a function $\Theta(u)$ can be written as a linear combination:

$$\Theta(u) = \sum_{i=0}^{\infty} \alpha_i \psi_i(u)$$

where functions $\psi_i(u)$ make up a set of simple elementary functions. If functions are orthogonal (roughly, perpindicular; inner product is zero)then coefficients α_i are independant from one another.

Continuous Fourier Transform

The most commonly used set of orthogonal functions is the Fourier series. Here is the analog version of the Fourier and Inverse Fourier:

$$X(w) = \int_{-\infty}^{+\infty} x(t) e^{(-2\pi jwt)} \mathrm{d}t$$

$$x(t) = \int_{-\infty}^{+\infty} X(w) e^{(2\pi jwt)} \mathrm{d}w$$

Discrete Fourier and Inverse Fourier Transform

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{-2\pi jnk/N}$$

$$x(k) = \frac{1}{N} \sum_{n=0}^{N-1} X(n) e^{2\pi jnk/N}$$

Taylor Series Expansion

$$f(x) = f(x_0) + \frac{f'(x_0)}{1}(x - x_0) + \frac{f''(x_0)}{1 * 2}(x - x_0)^2 + \frac{f'''(x_0)}{1 * 2 * 3}(x - x_0)^3 + \ldots$$

or more compactly as

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

Taylor series expansion of $e^{j\Theta}$

Since e^x is its own derivative, the Taylor series expansion for $f(x) = e^x$ is very simple:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + ...$$

We can define:

$$e^{j\Theta} == \sum_{n=0}^{\infty} \frac{(j\Theta)^n}{n!} = 1 + j\Theta - \frac{\Theta^2}{2} - j\frac{\Theta^3}{3!} + ...$$

Splitting out real and imaginary parts

All even order terms are real; all odd-order terms are imaginary:

$$\mathrm{re}e^{j\Theta} = 1 - \Theta^2/2 + \Theta^4/4! - ...$$

$$\mathrm{im}e^{j\Theta} = \Theta - \Theta^3/3! + \Theta^5/5! - ...$$

Fourier Transform as sum of sines and cosines

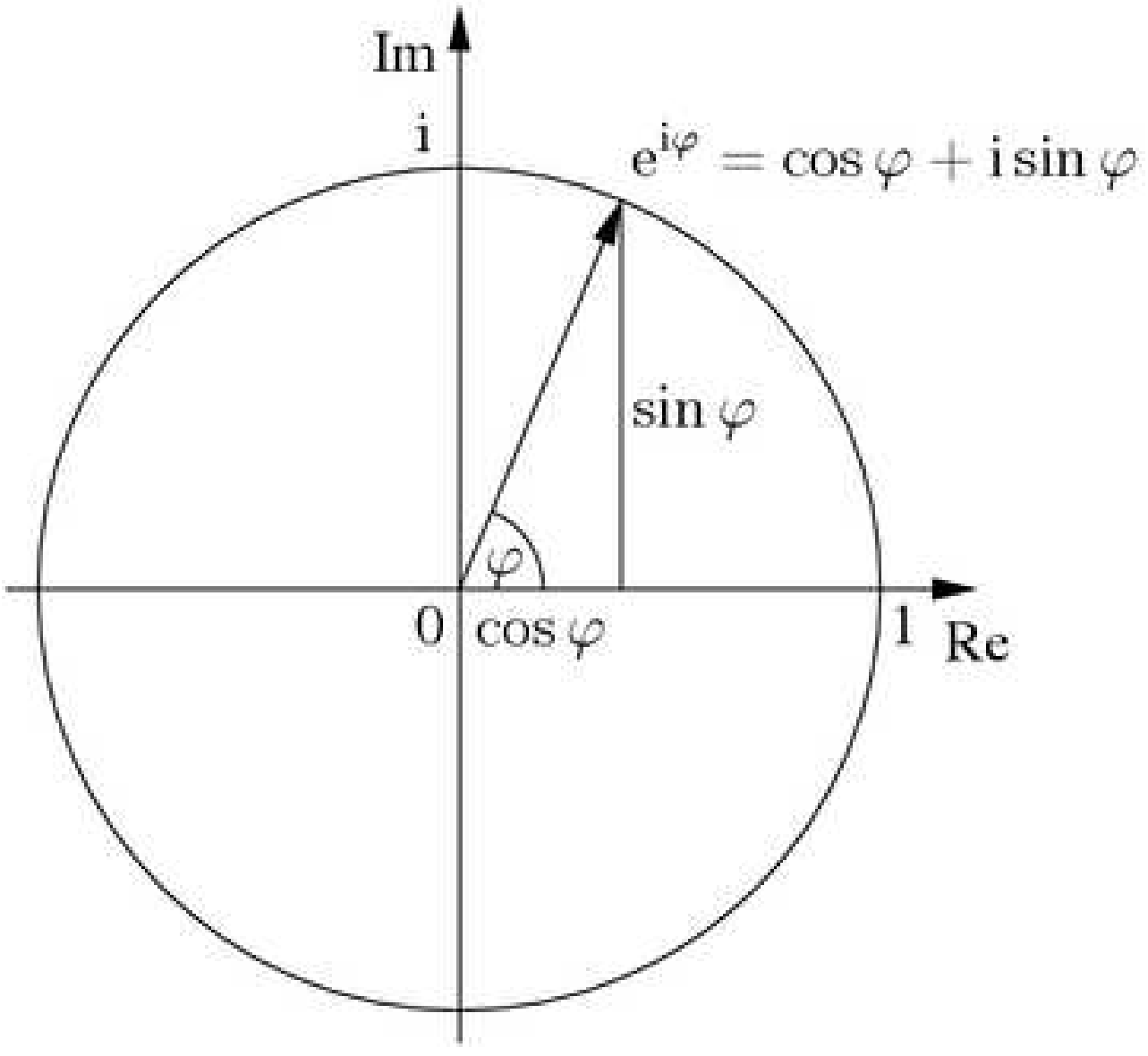
Observe that:

$$\cos(\Theta) = \sum_{n>=0;\text{n is even}}^{\infty} \frac{(-1)^{n/2}}{n!}\Theta^n$$

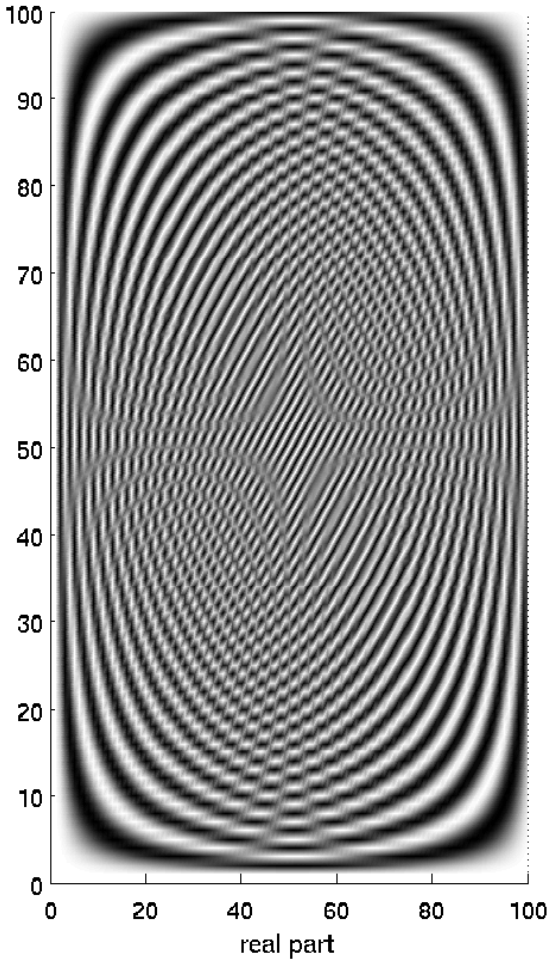
$$\sin(\Theta) = \sum_{n>=0;\text{n is odd}}^{\infty} \frac{(-1)^{(n-1)/2}}{n!}\Theta^n$$

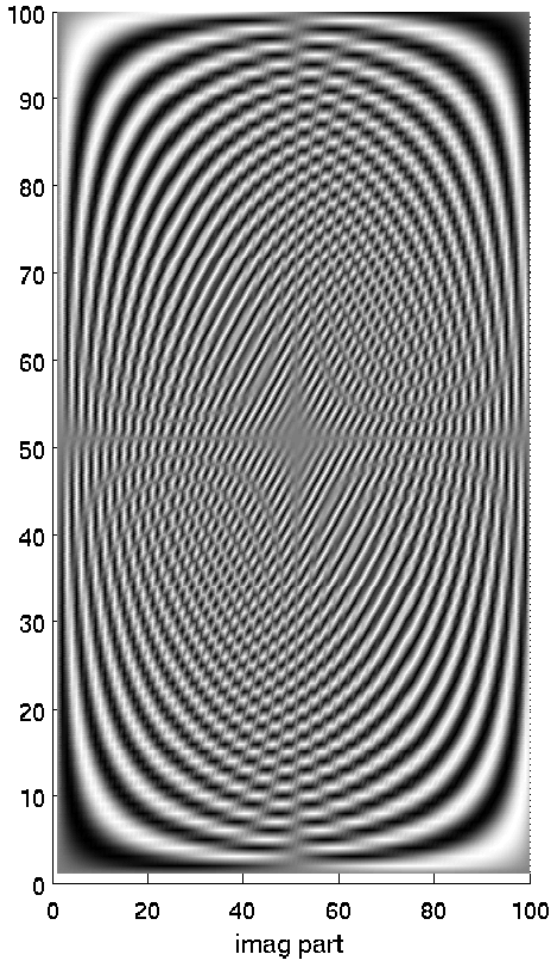
Thus yielding Euler's formula:

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


Im
i
$e^{i\varphi} = \cos\varphi + i\sin\varphi$
$\sin\varphi$
φ
0
$\cos\varphi$
1
Re

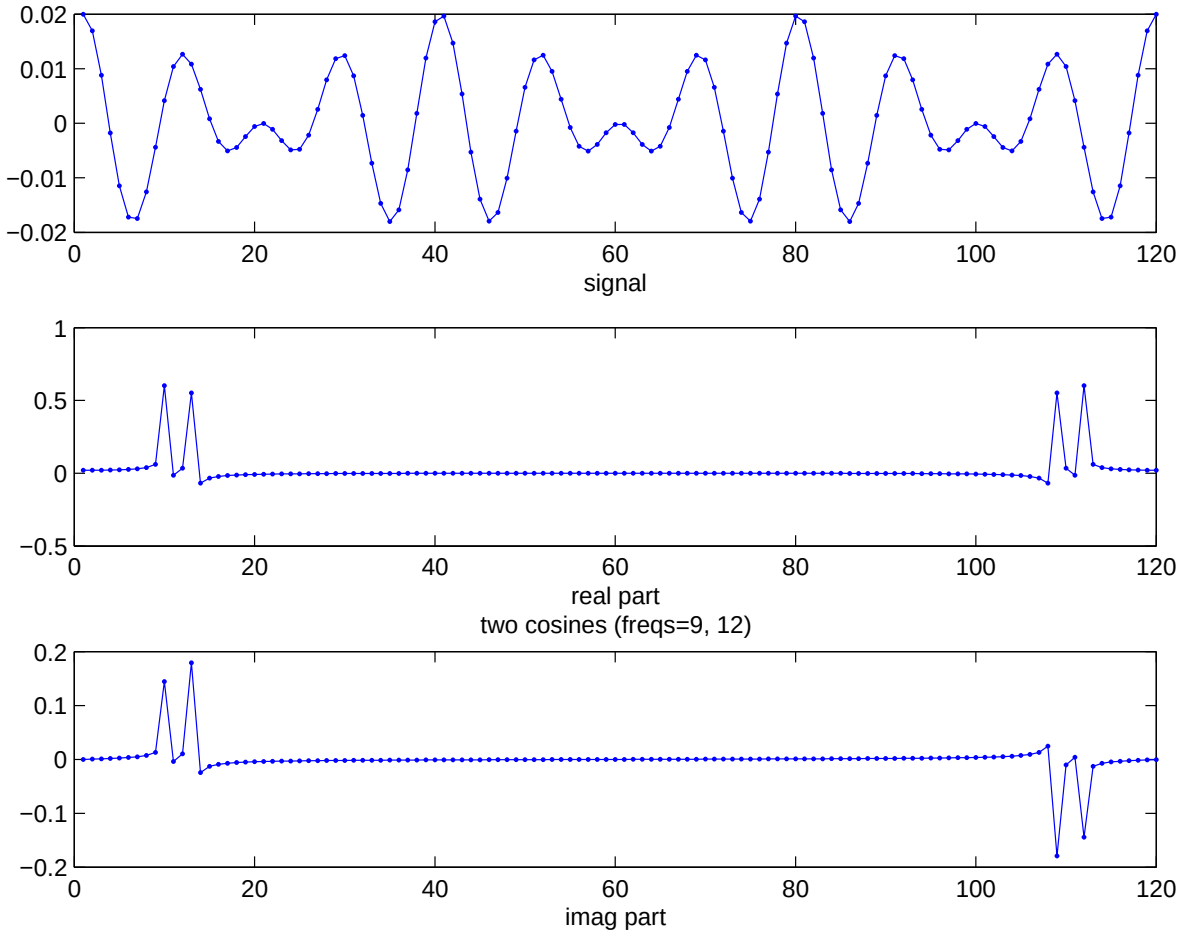
Fourier transform as kernel matrix





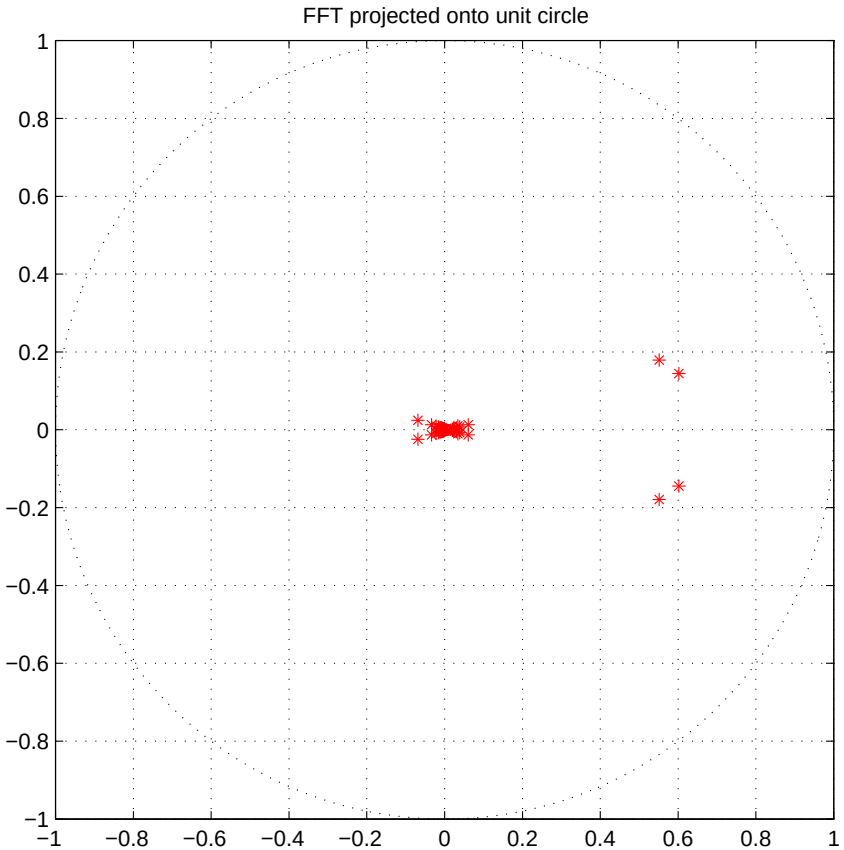
Example

Sum of cosines with frequencies 12 and 9, sampling rate = 120

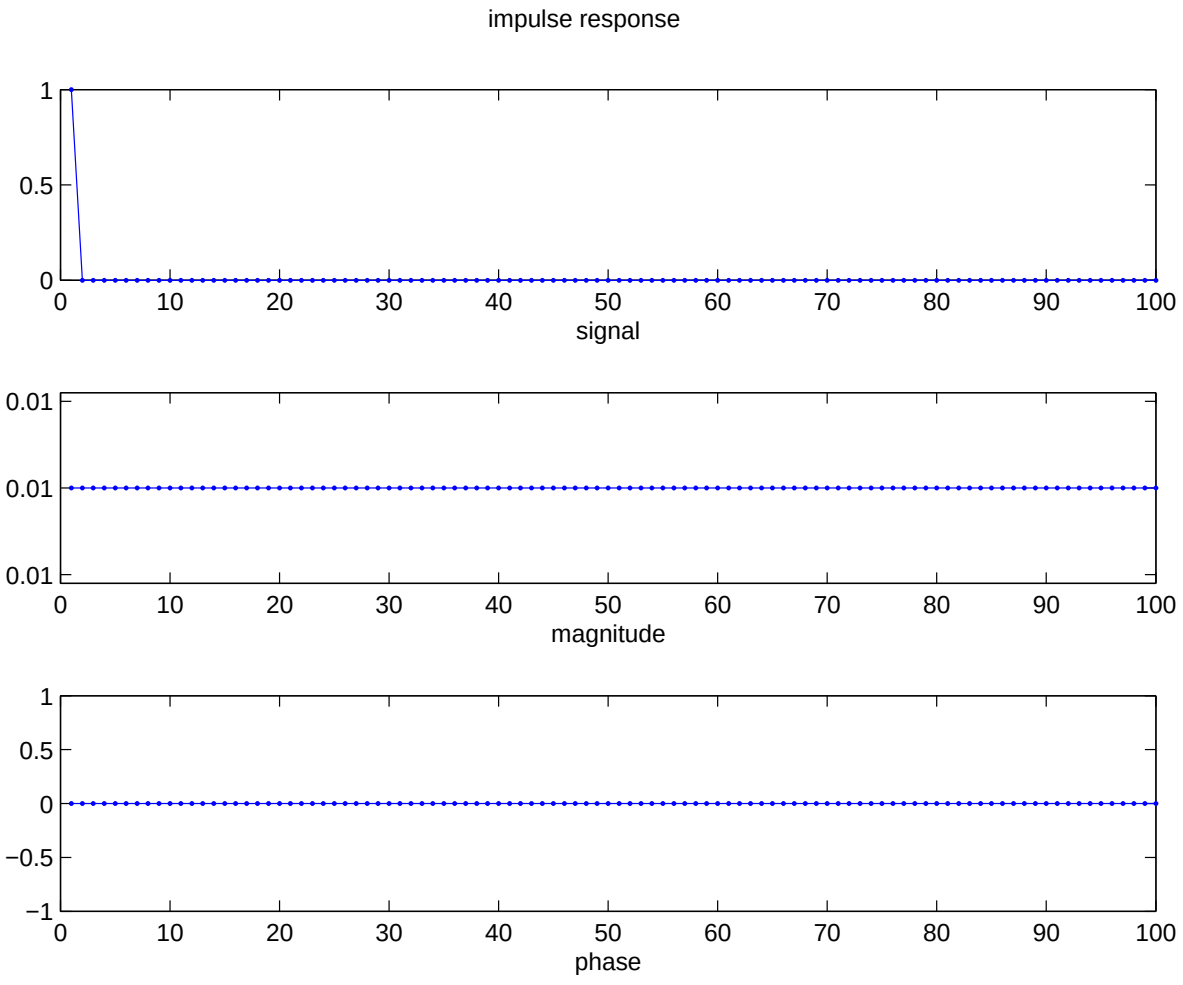


Example

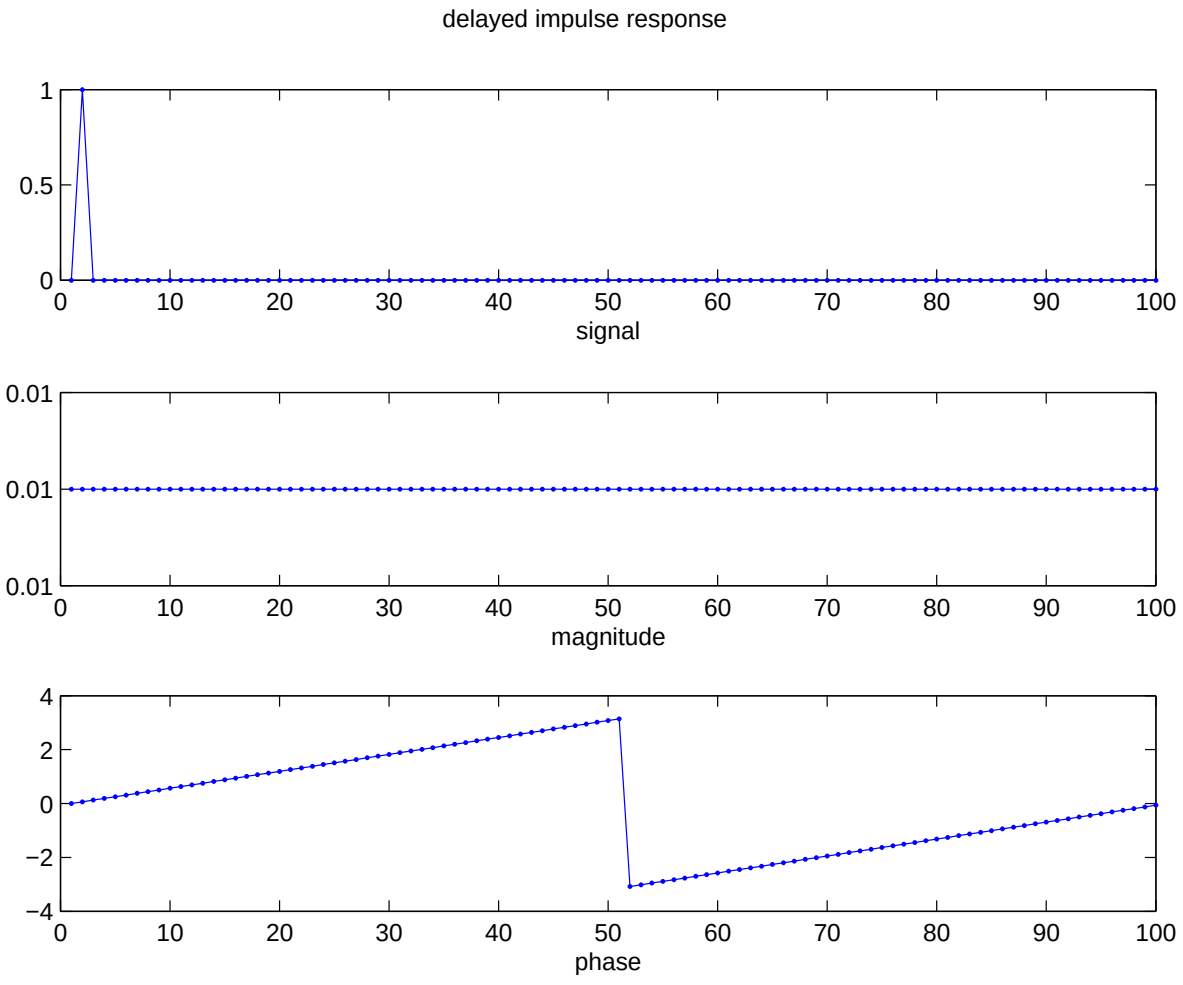
FFT coefficients mapped onto unit circle



Impulse response

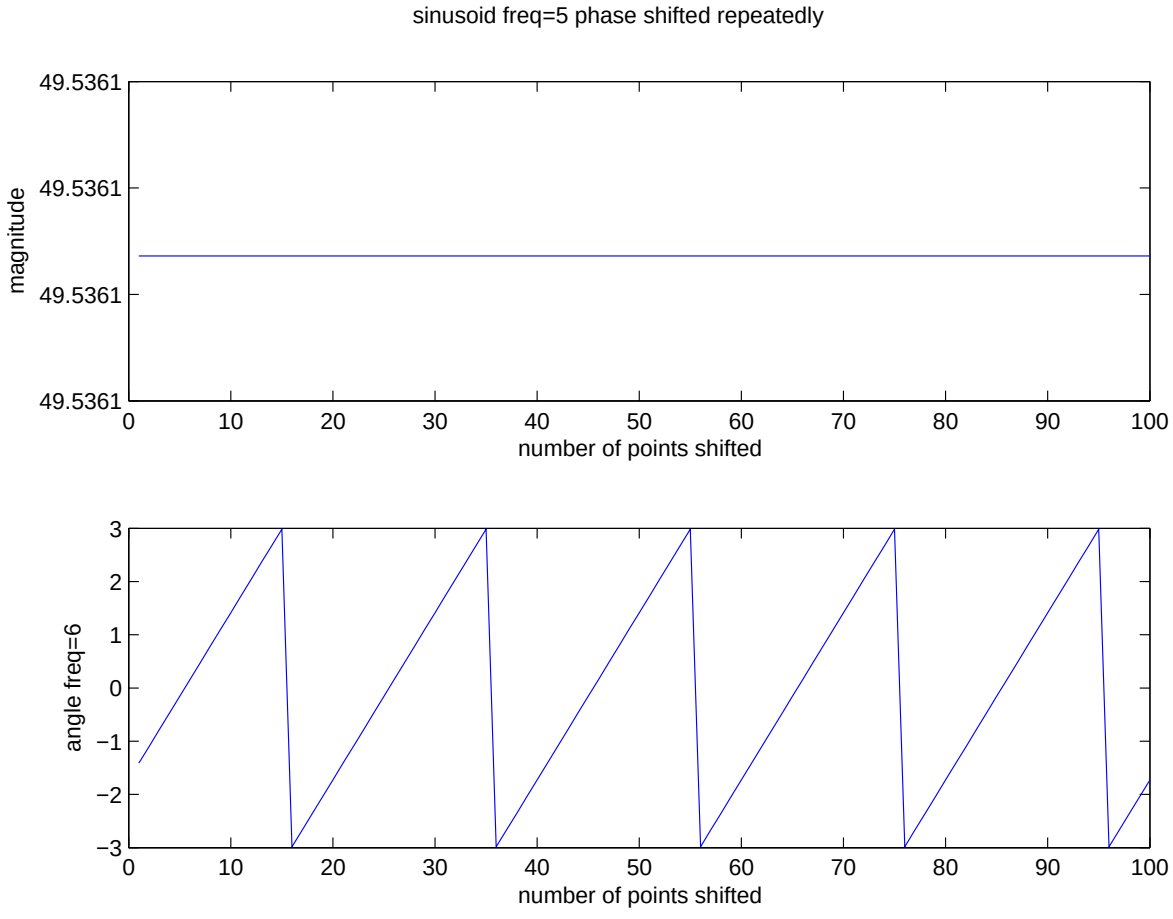


Impulse response

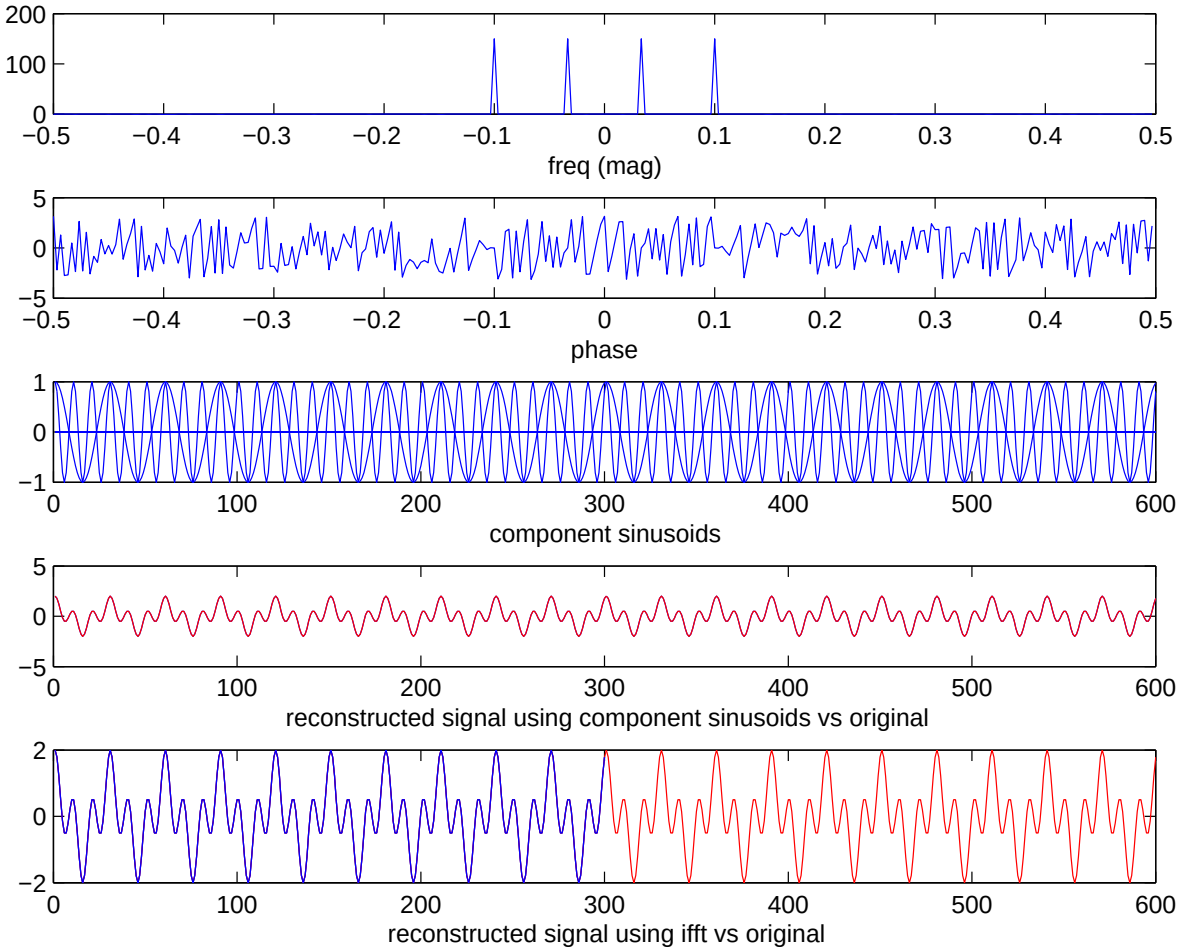


A look at phase shifting

Sinusoid frequency=5 phase shifted multiple times.

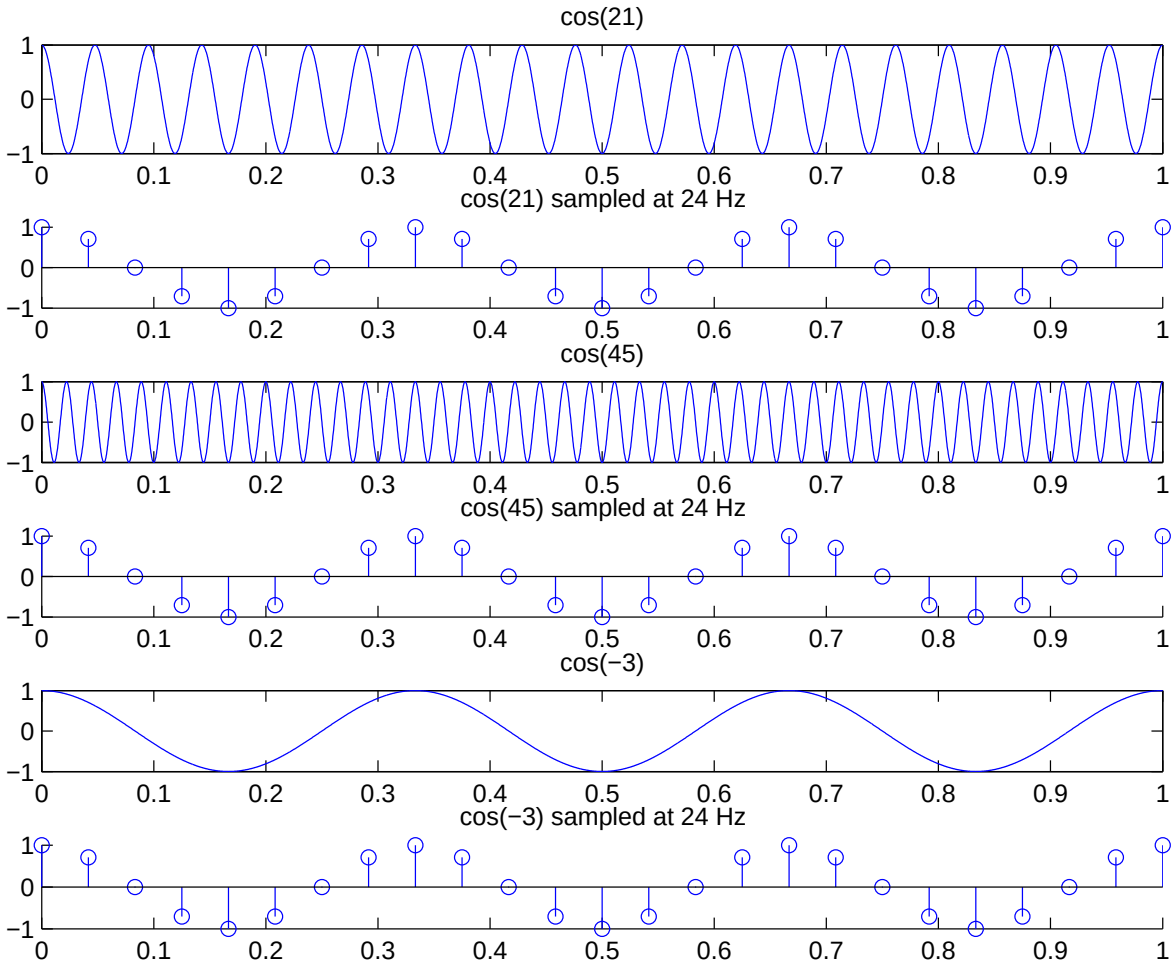


Sin period 10 + period 30



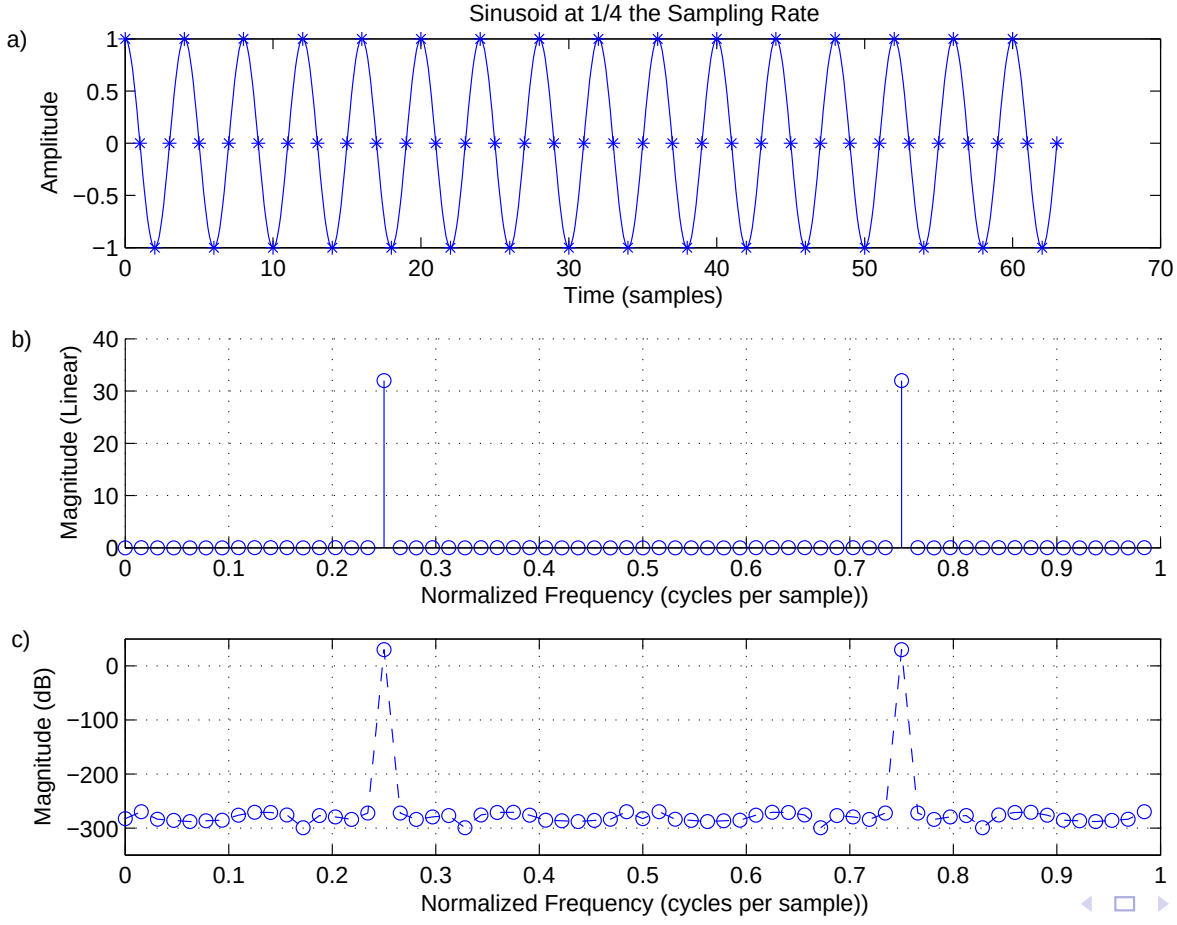
Aliasing

The useful range is the "Nyquist frequency" ($fs/2$)



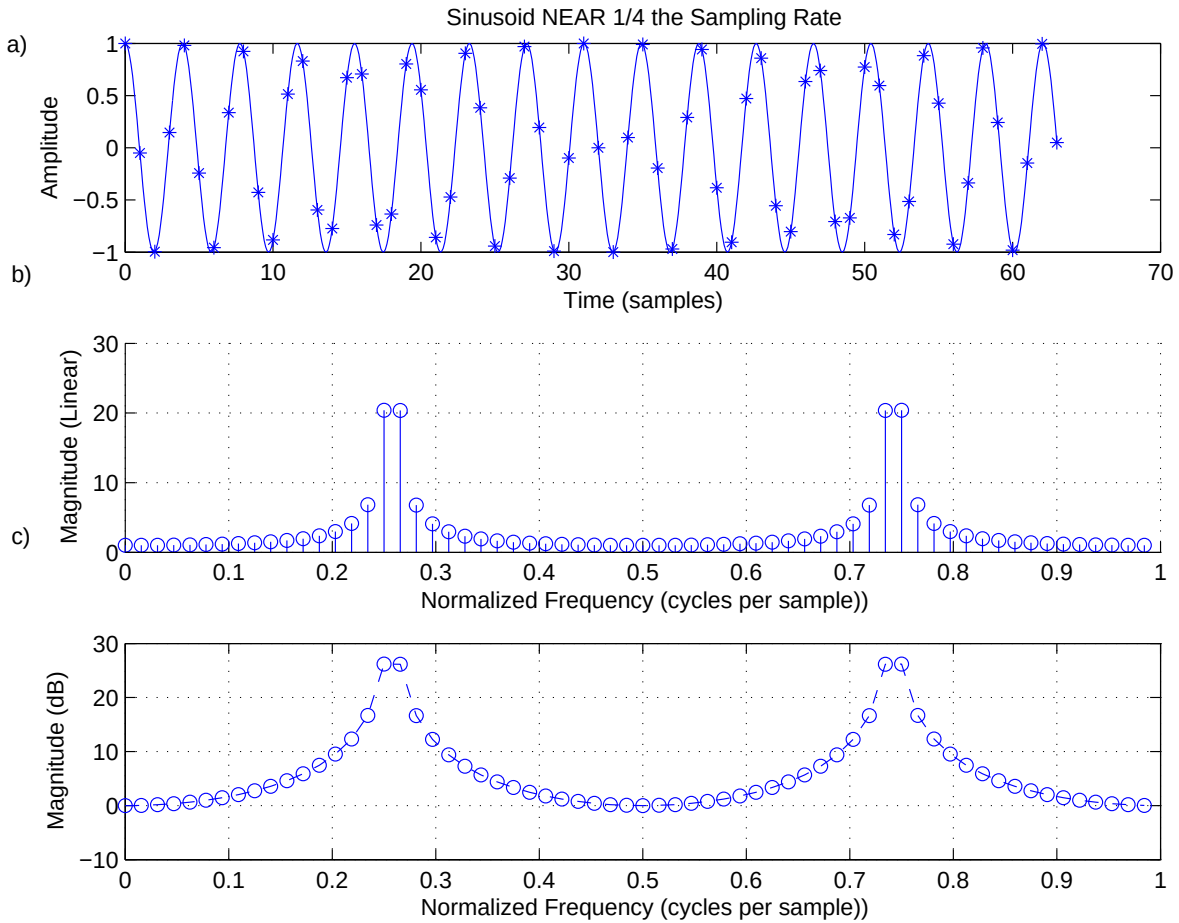
Leakage

Even below Nyquist, when frequencies in the signal do not align well with sampling rate of signal, there can be "leakage". First consider a well-aligned exampl (freq = .25 sampling rate)



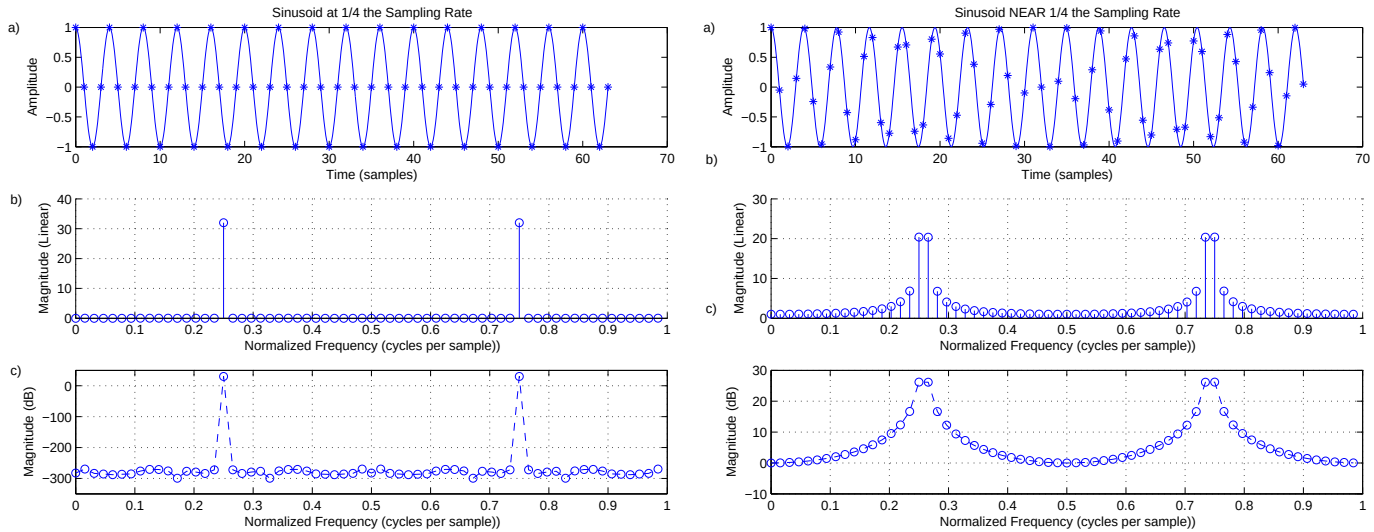
Leakage

Now consider a poorly-aligned example (freq = (.25 + .5/N) * sampling rate)



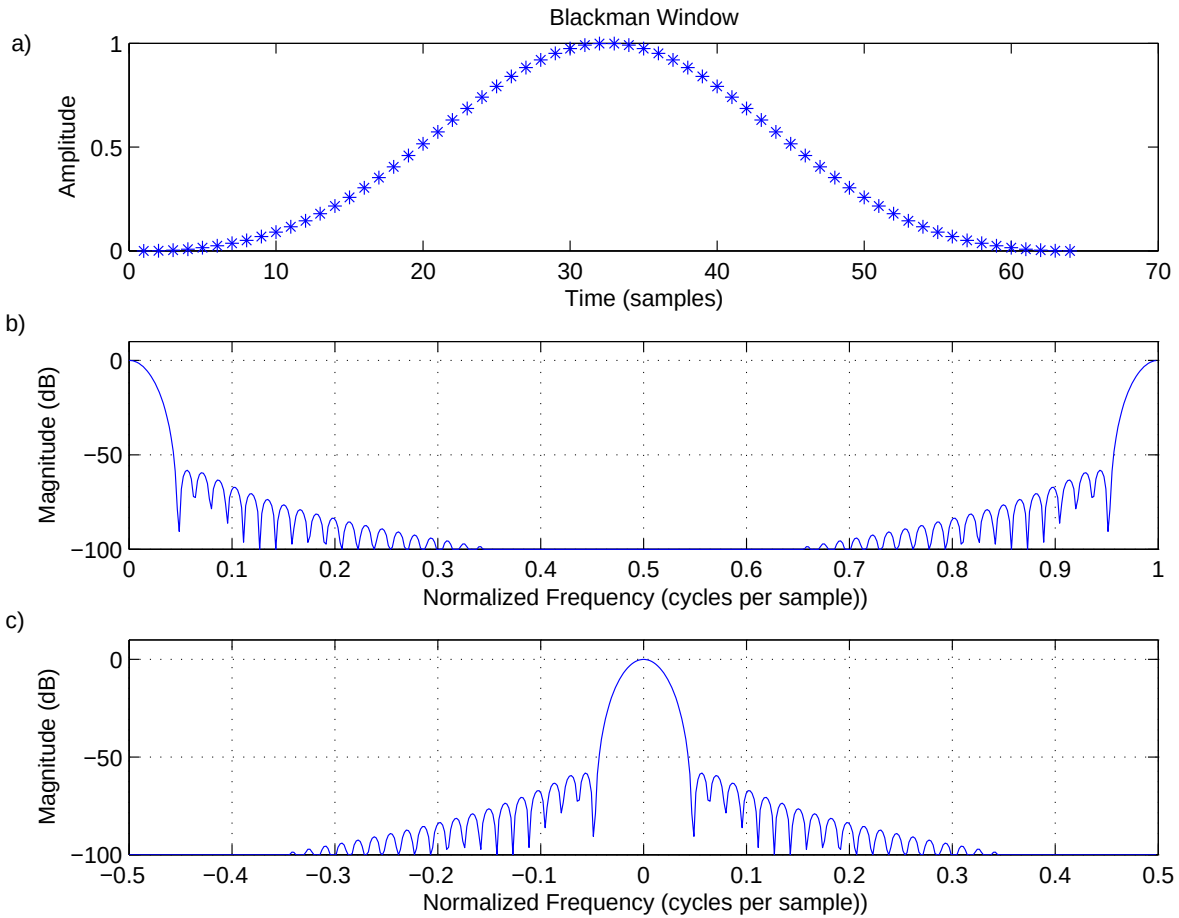
Leakage

Comparison:



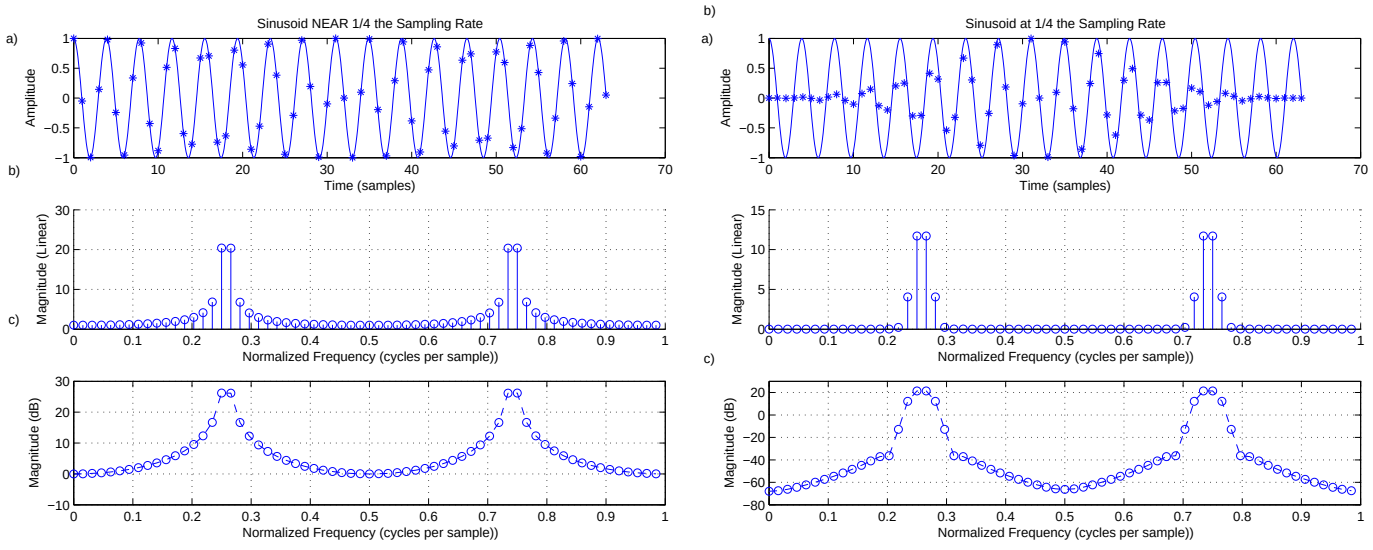
Windowing can help

Can minimize effects by multiplying time series by a window that diminishes magnitude of points near signal edge:



Leakage Reduced

Comparison:



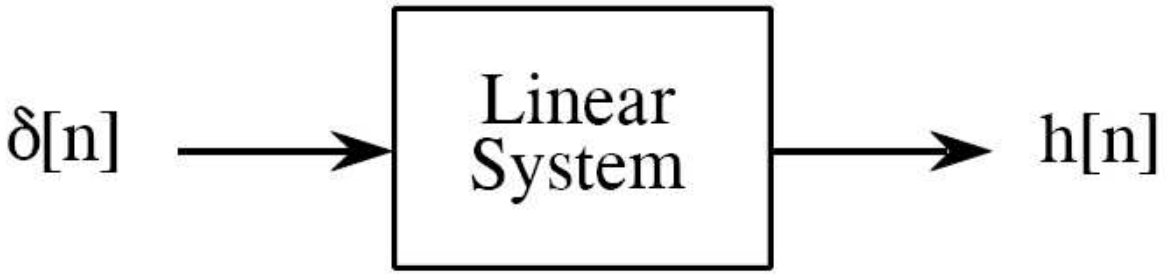
Convolution theorem

This can be understood in terms of the Convolution Theorem. Convolution in the time domain is multiplication in the frequency domain via the Fourier transform ($\mathcal{F}$).

$$\mathcal{F}(f * g) = \mathcal{F}(f) \times \mathcal{F}(g)$$

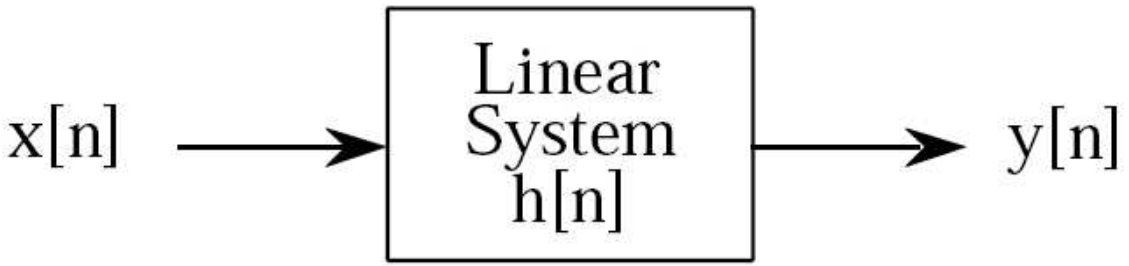
Computing impulse response

The impulse response $h[n]$ is the response of a system to the unit impulse function.



Using the impulse response

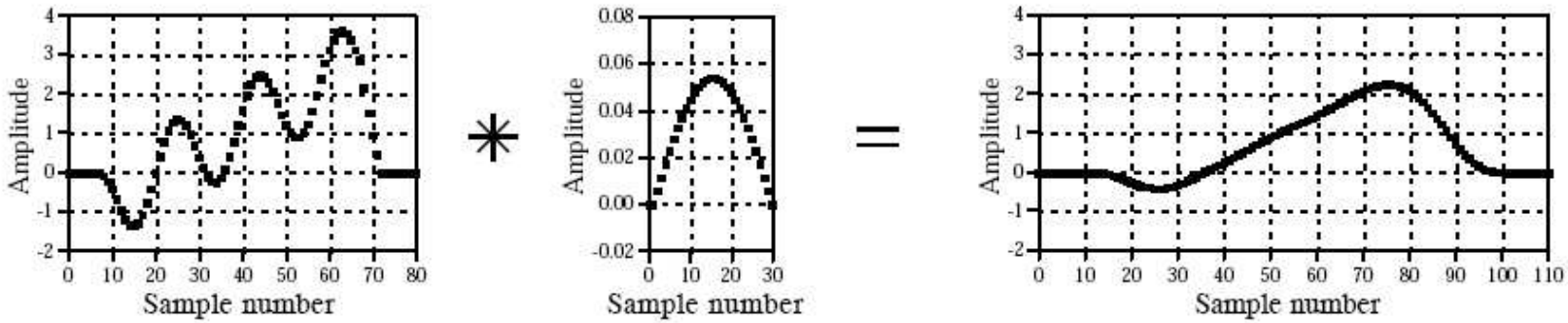
Once computed, the impulse response can be used to filter any signal $x[n]$ yielding $y[n]$.



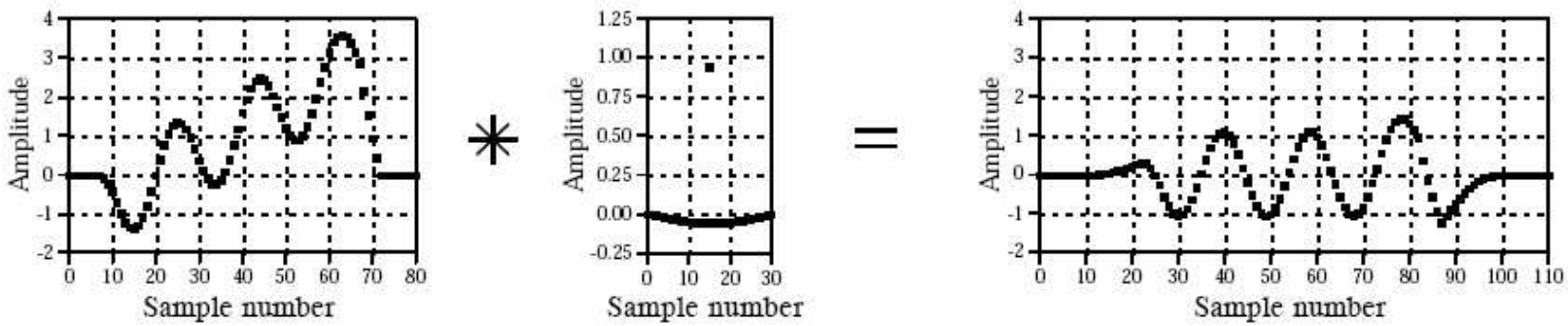
$$x[n] * h[n] = y[n]$$

Examples

a. Low-pass Filter



b. High-pass Filter



Filtering using DFT

- Goal is to choose good impulse response $h[n]$

Filtering using DFT

- Goal is to choose good impulse response $h[n]$
- Transform signal into frequency domain

Filtering using DFT

- Goal is to choose good impulse response $h[n]$
- Transform signal into frequency domain
- Modify frequency properties of signal via *multiplication*

Filtering using DFT

- Goal is to choose good impulse response $h[n]$
- Transform signal into frequency domain
- Modify frequency properties of signal via *multiplication*
- Transform back into time domain

Difficulties (Why not a perfect filter?)

- You *can* have a perfect filter(!)

Difficulties (Why not a perfect filter?)

- You *can* have a perfect filter(!)
- Need long impulse response function in both directions

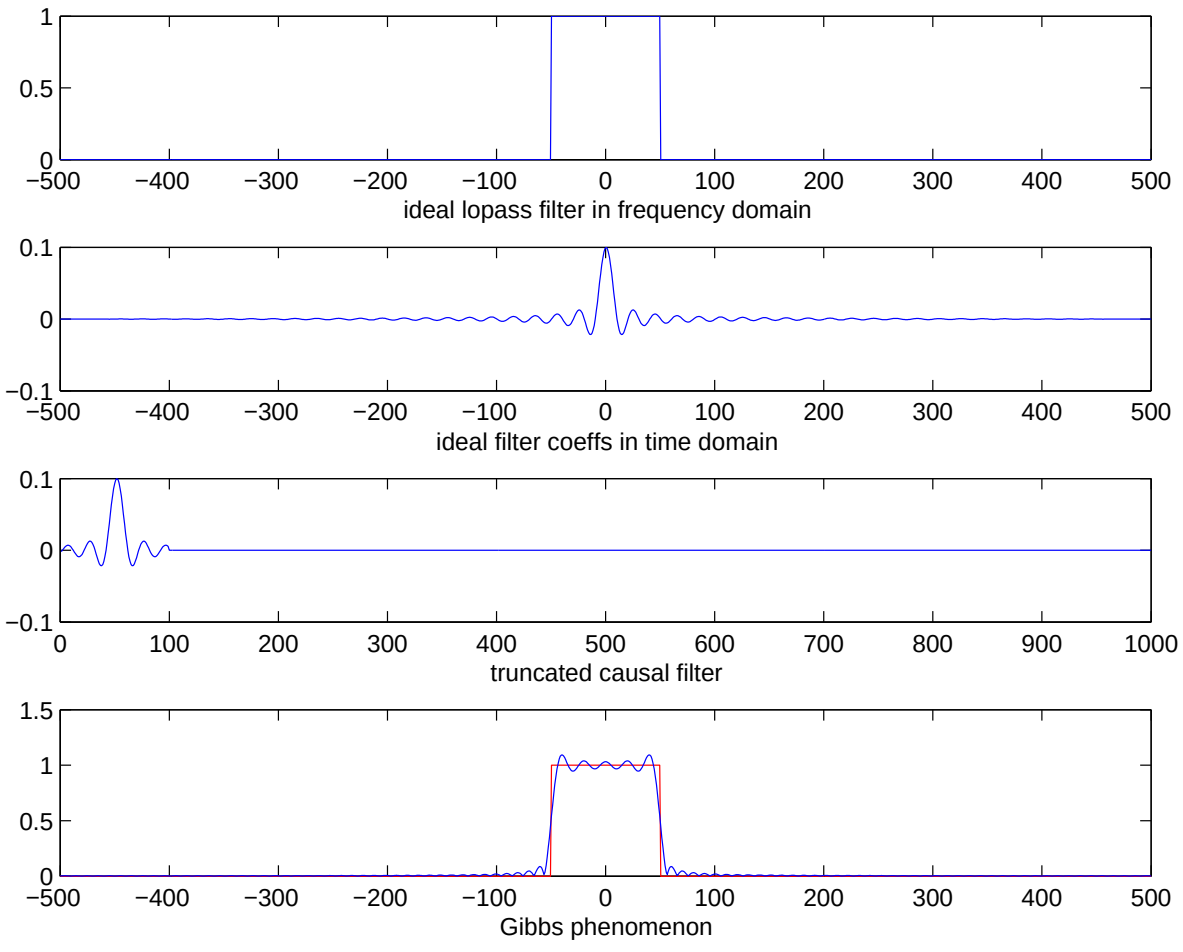
Difficulties (Why not a perfect filter?)

- You *can* have a perfect filter(!)
- Need long impulse response function in both directions
- Very non causal

Difficulties (Why not a perfect filter?)

- You *can* have a perfect filter(!)
- Need long impulse response function in both directions
- Very non causal
- In generating causal version, challenges arise

Gibbs Phenomenon



Spectral Analysis

- Often we want to see spectral energy as a signal evolves over time

Spectral Analysis

- Often we want to see spectral energy as a signal evolves over time
- Compute Fourier Transform over evenly-spaced frames of data

Spectral Analysis

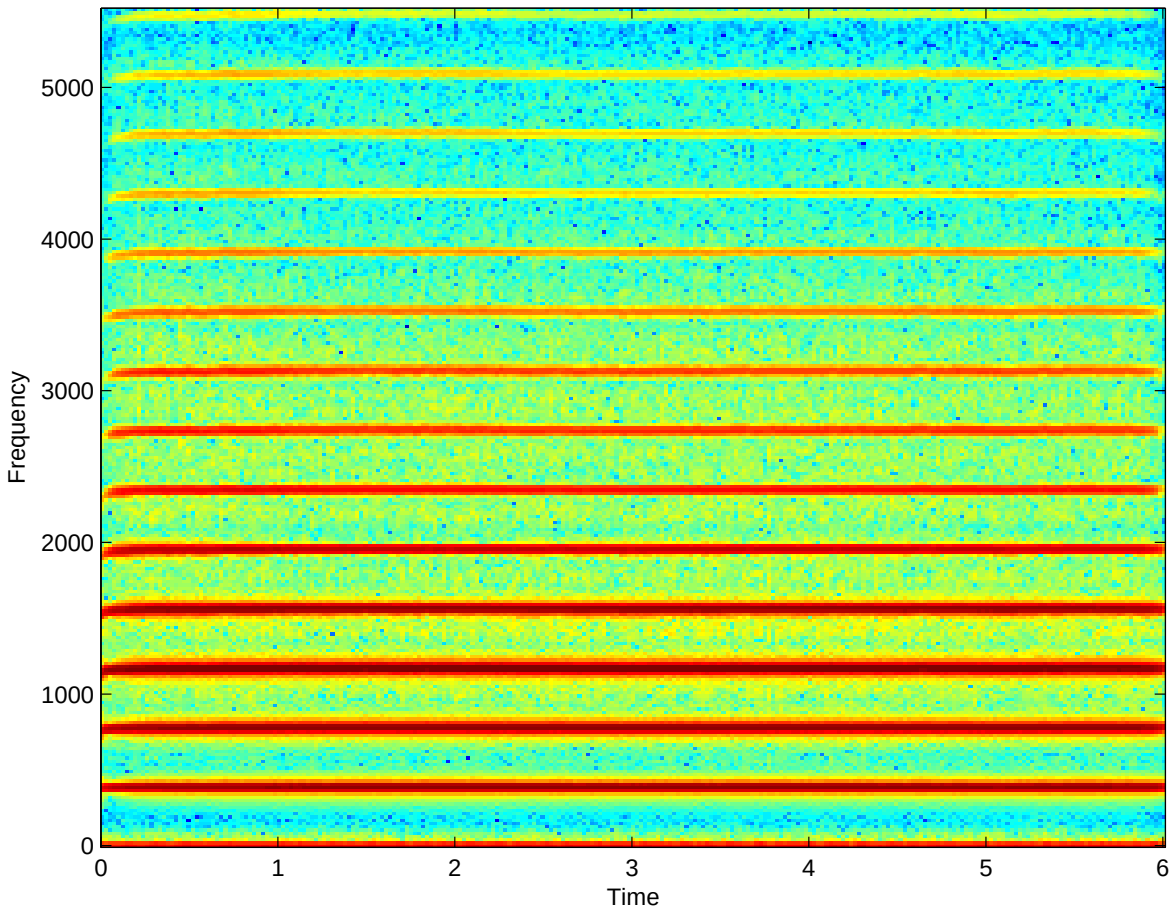
- Often we want to see spectral energy as a signal evolves over time
- Compute Fourier Transform over evenly-spaced frames of data
- Apply window to minimize edge effects

Short-Timescale Fourier Transform (STFT)

$$X(m,n) = \sum_{k=0}^{N-1} x(k)w(k-m)e^{-2\pi jnk/N}$$

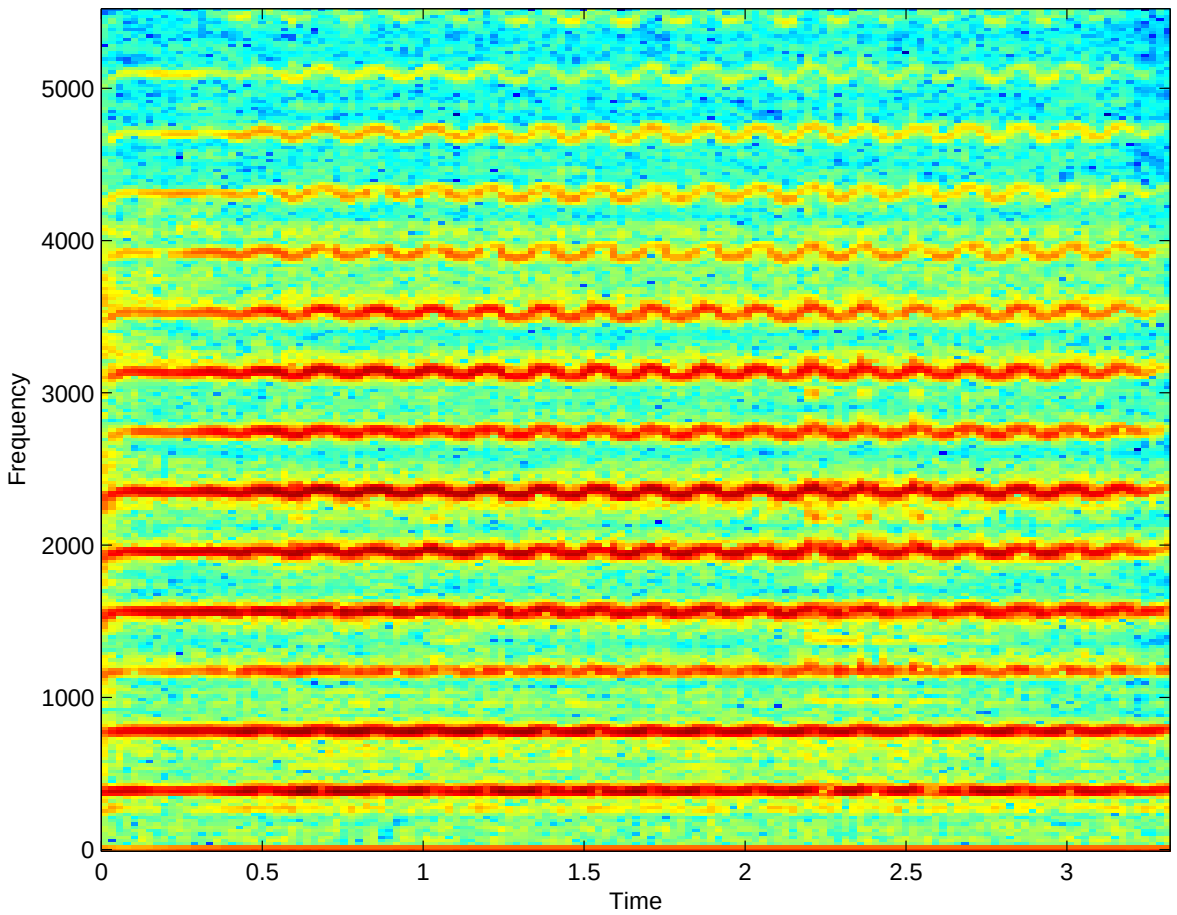
Where w is some windowing function such as Hanning or gaussian centered around zero.
The spectrogram is simply the squared magnitude of these STFT values

Trumpet (G4)



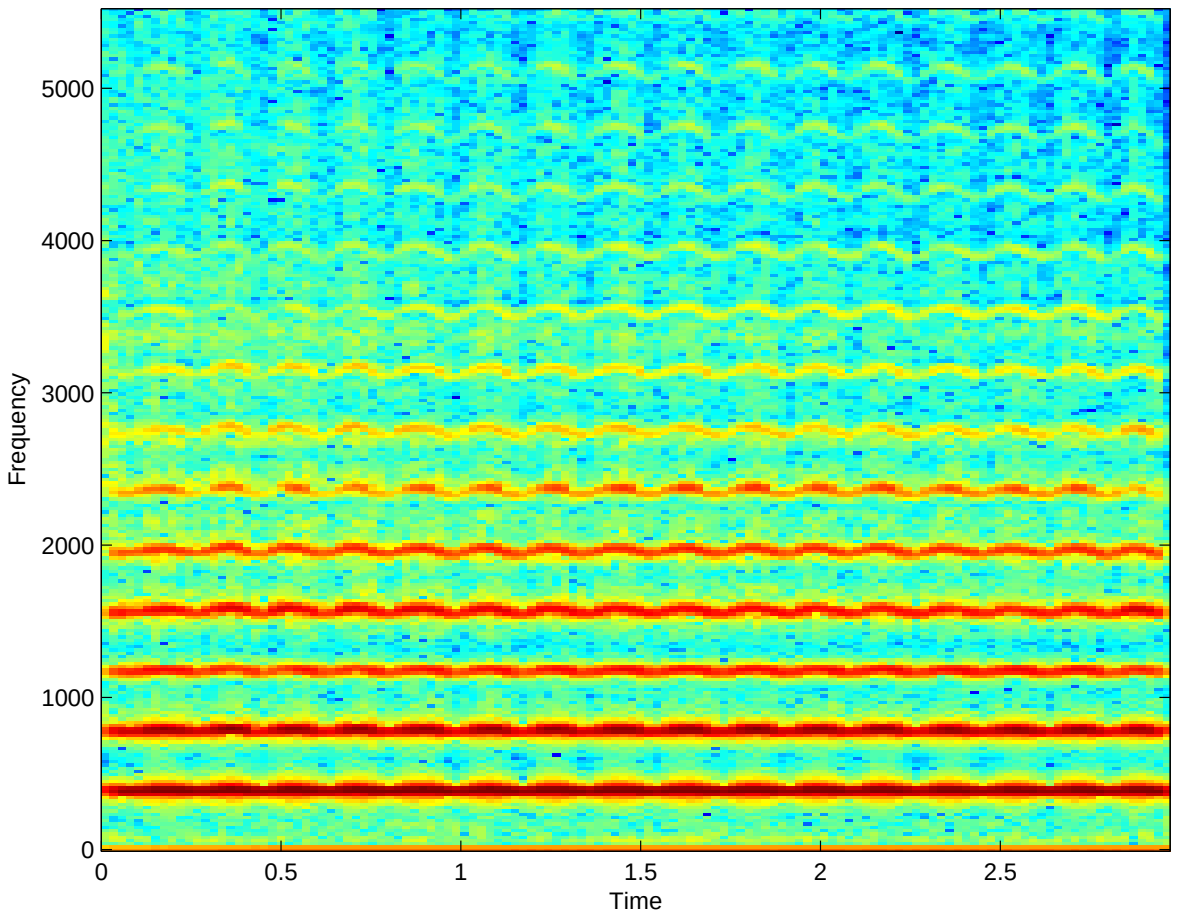
[Listen]

Violin (G4)



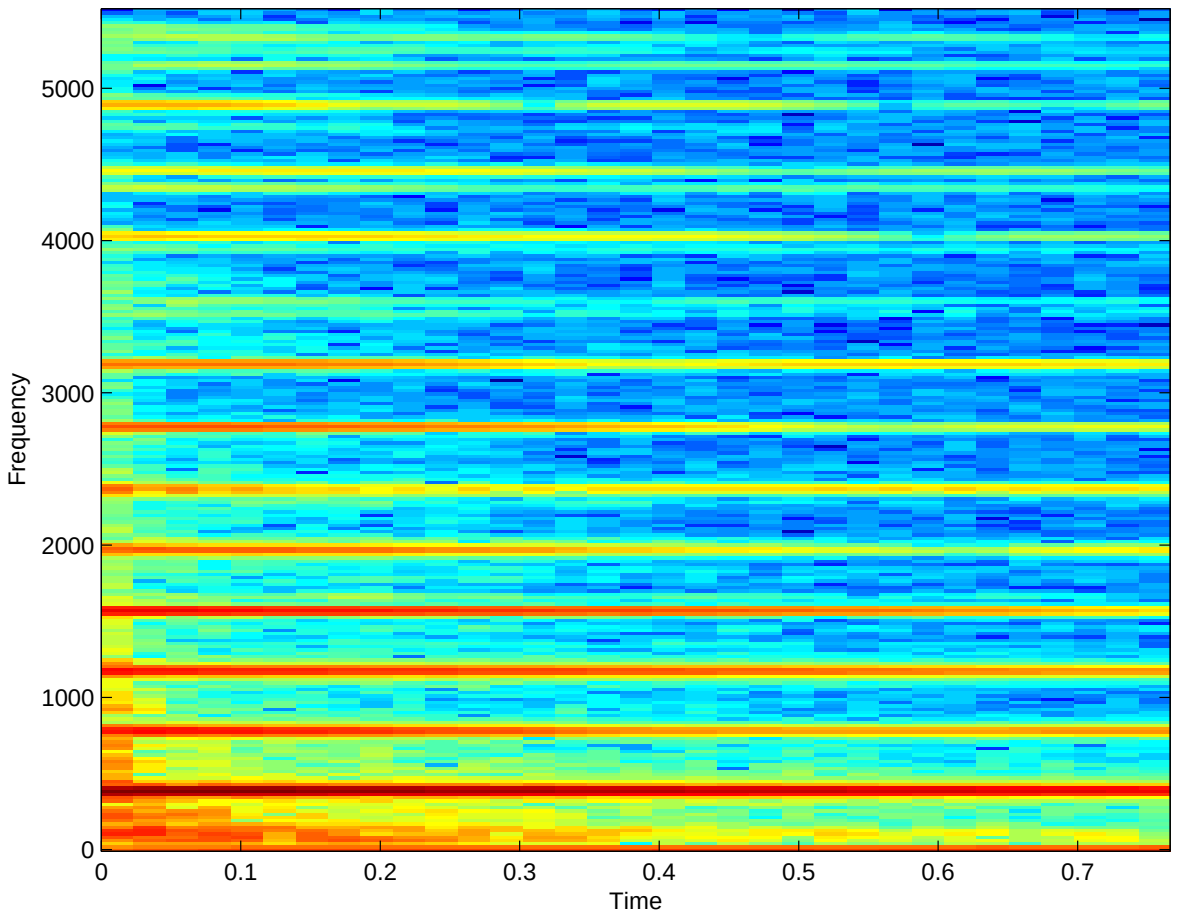
[Listen]

Flute (G4)



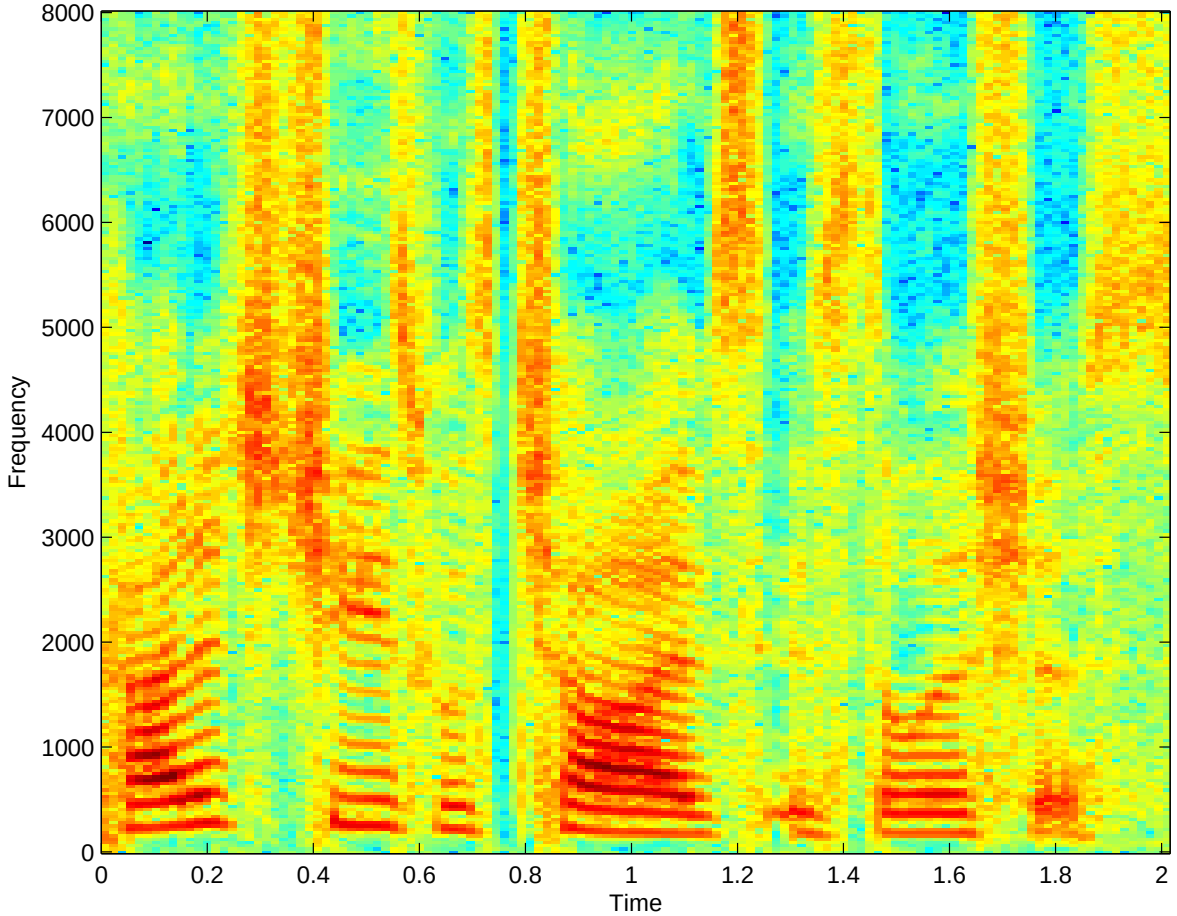
[Listen]

Piano (G4)



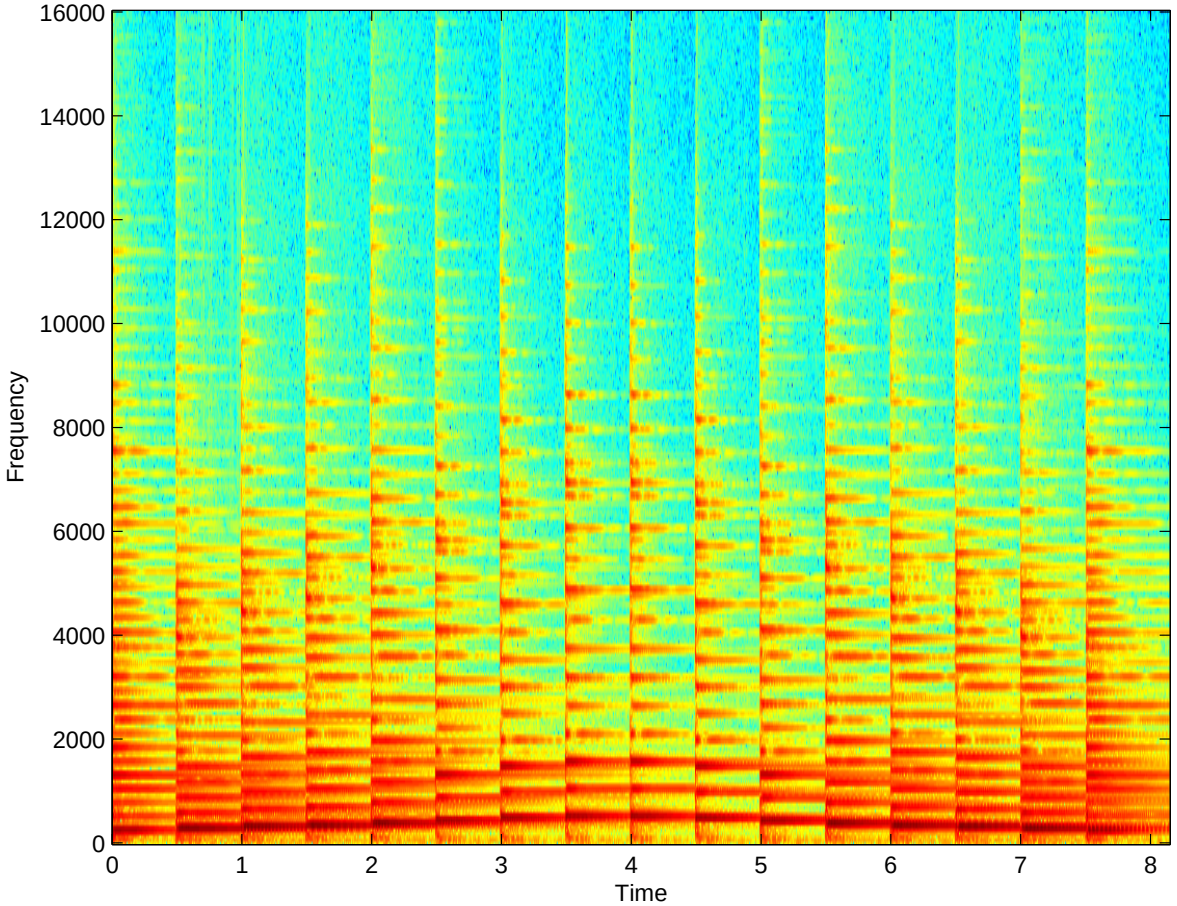
[Listen]

Voice



[Listen]

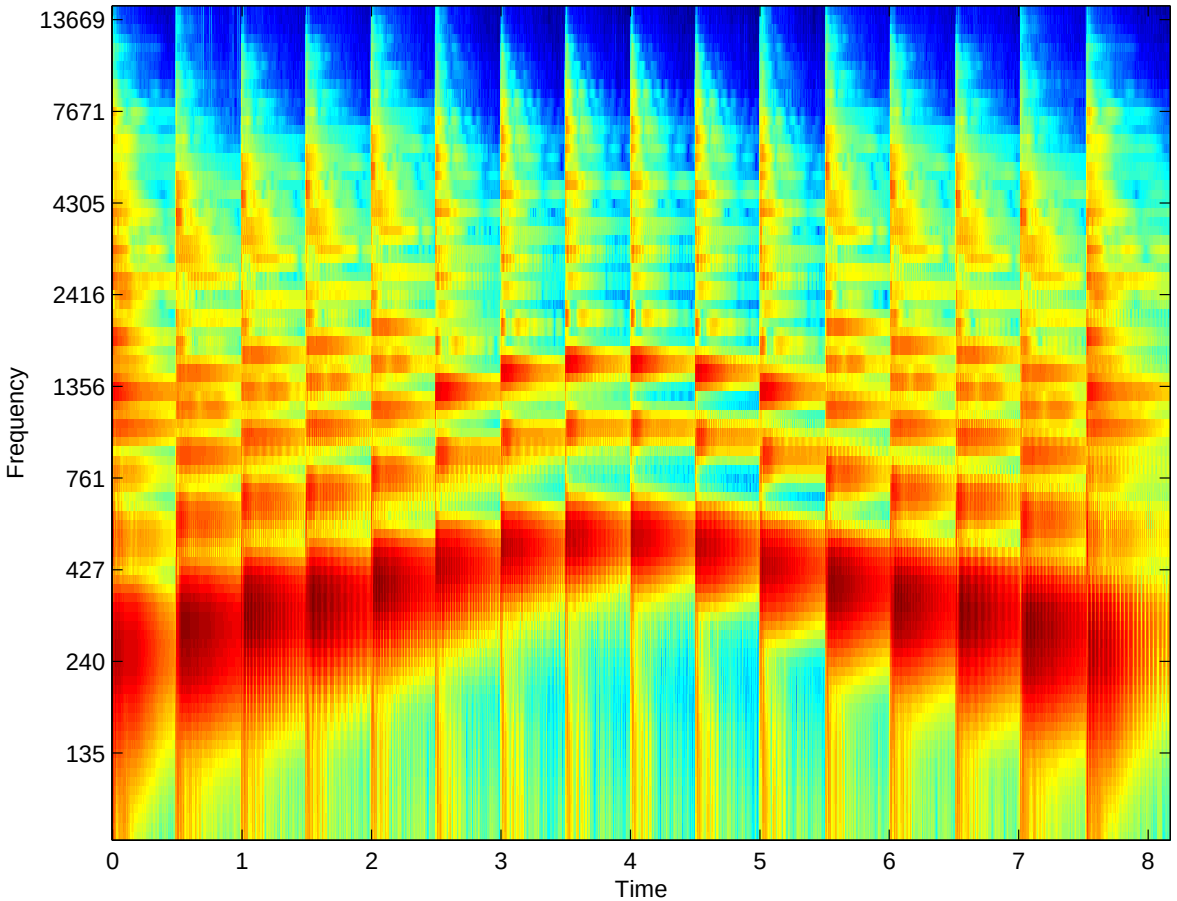
C Major Scale (Piano)



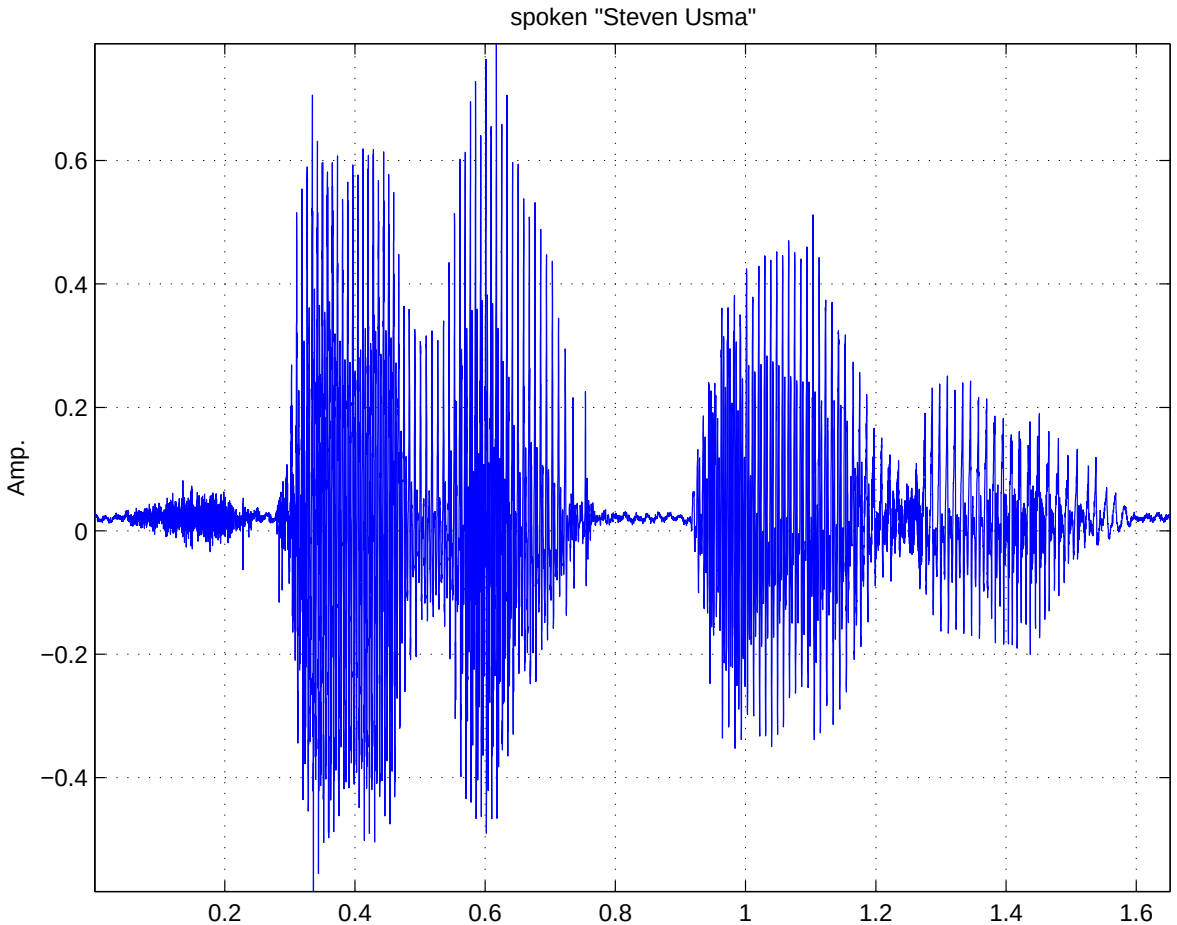
[Listen]

C Major Scale (Piano)

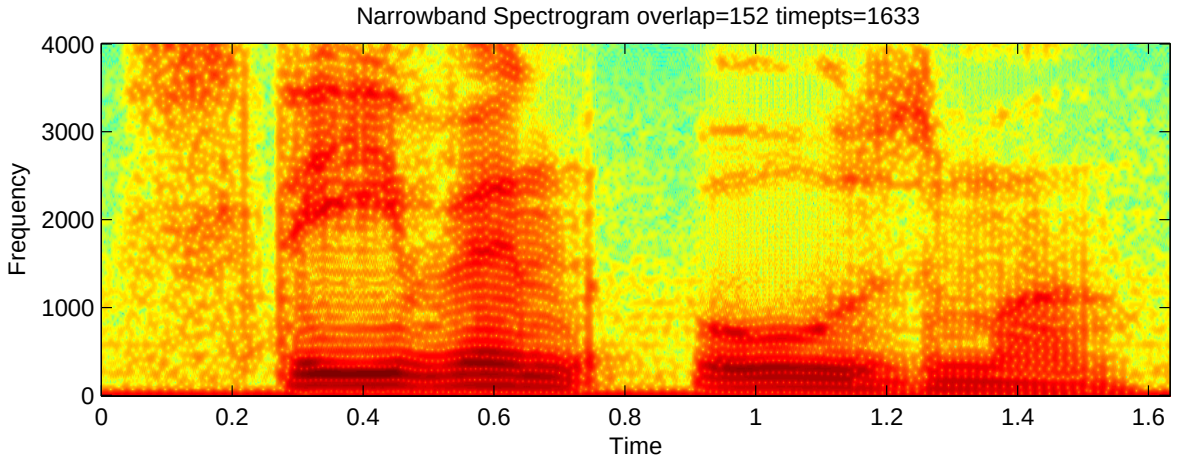
Log Spectrogram (Constant-Q Transform) reveals low-frequency structure

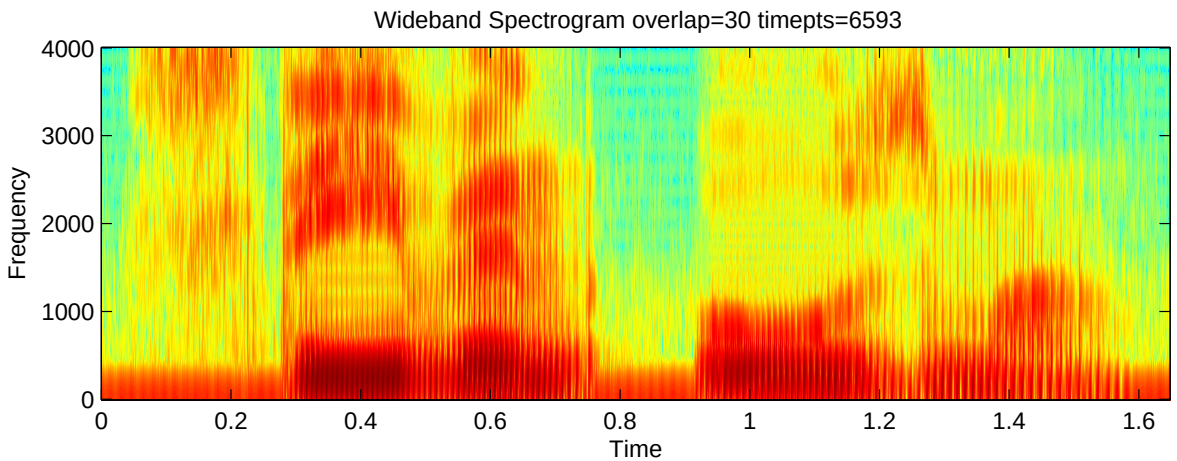


Time-Space Tradeoff



Time-Space Tradeoff





Auto-correlation and meter

Autocorrelation long used to find meter in music (Brown 1993)
Lag k auto-correlation $a(k)$ is a special case of cross-correlation where a signal x is correlated with itself:

$$a(k) = \frac{1}{N} \sum_{n=k}^{N-1} x(n)\, x(n-k)$$

Auto-correlation and meter

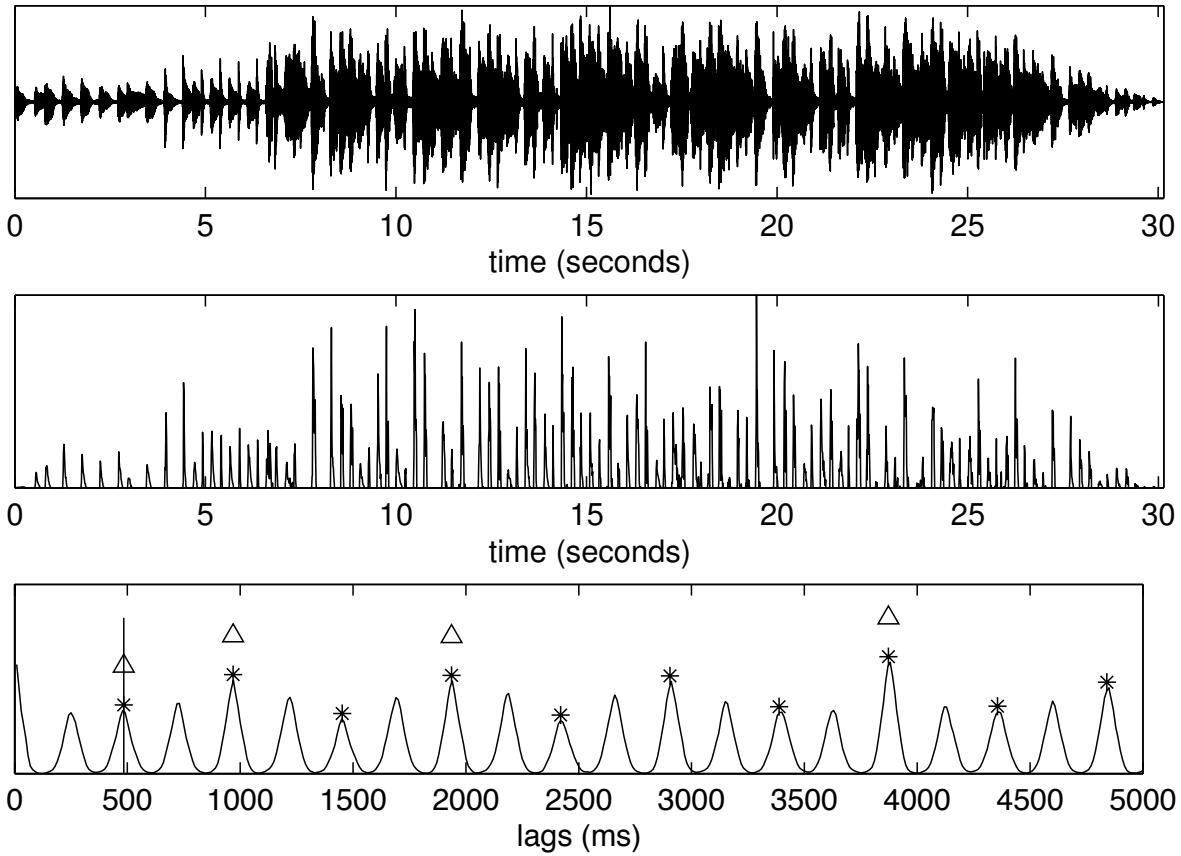
Autocorrelation long used to find meter in music (Brown 1993)
Lag k auto-correlation $a(k)$ is a special case of cross-correlation where a signal x is correlated with itself:

$$a(k) = \frac{1}{N} \sum_{n=k}^{N-1} x(n)\, x(n-k)$$

Autocorrelation can also be defined in terms of Fourier analysis

$$a = \mathcal{F}^{-1}(|\mathcal{F}(x)|)$$

where $\mathcal{F}$ is the Fourier transform, $\mathcal{F}^{-1}$ is the inverse Fourier transform and $||$ indicates taking magnitude from a complex value.



Time series (top), envelope (middle) and autocorrelation (bottom) of excerpt from ISMIR 2004 Tempo Induction contest (Albums-Cafe_Paradiso-08.wav). A vertical line marks the actual tempo (484 msec, 124bpm). Stars mark the tempo and its integer multiples. Triangles mark levels in the metrical hierarchy.

Fast Fourier Transform

- Fourier Transform $O(N^2)$
- Fast version possible $O(NlogN)$
- Size must be a power of two
- Strategy is decimation in time or frequency
- Divide and conquer
  - Rearrange the inputs in bit reversed order
  - Output transformation (Decimation in Time)