Introduction to Wavelets

Michael Phipps
Vallary Bhopatkar

*Amended from "The Wavelet Tutorial" by Robi Polikar,
http://users.rowan.edu/~polikar/WAVELETS/WTtutoria

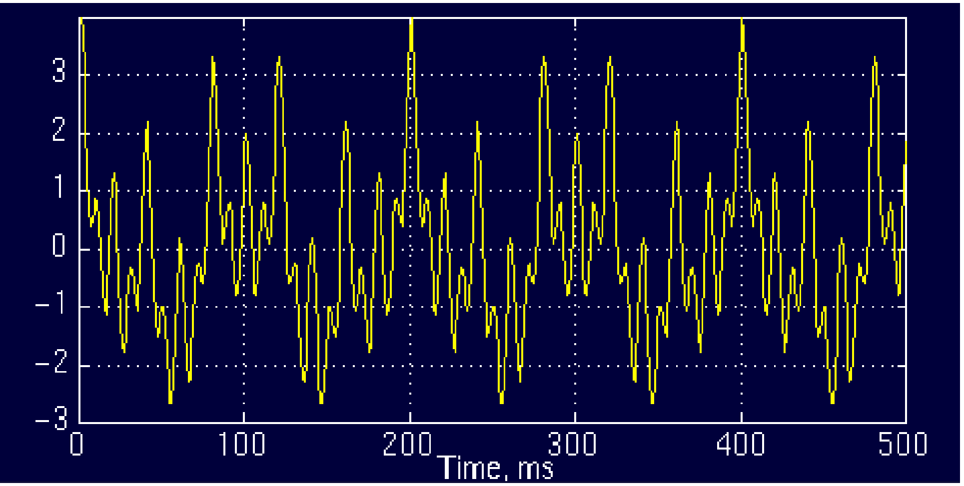
Who can tell me what this means?

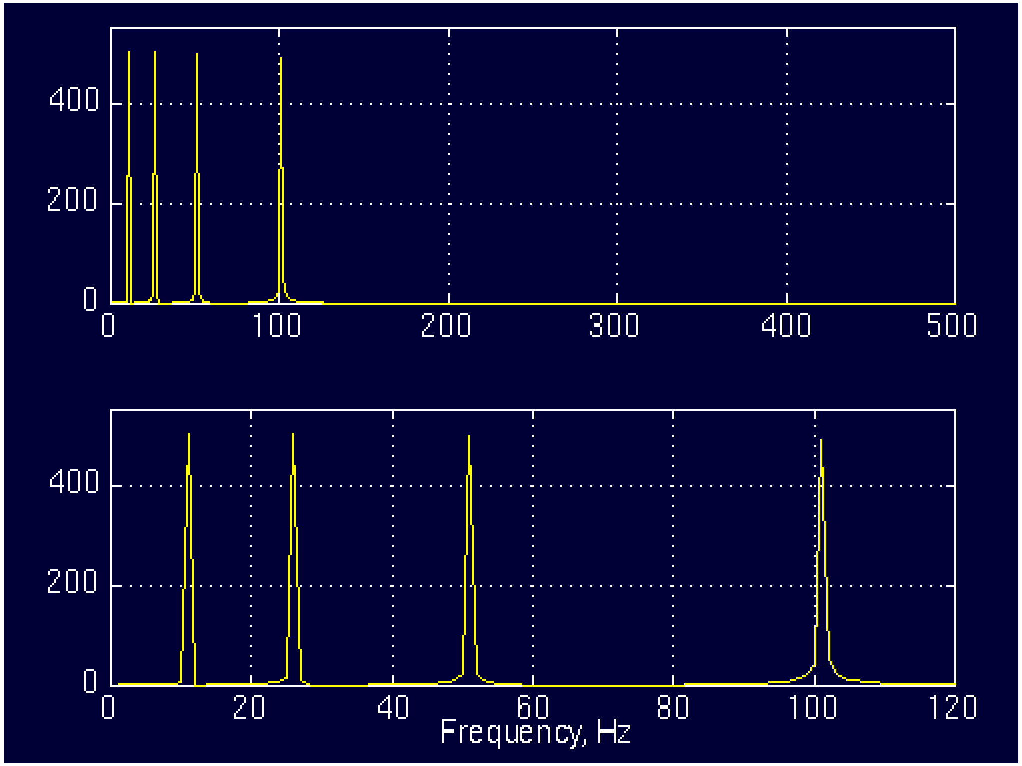
- NR3, pg 700: "What makes the wavelet basis interesting is that, unlike sines and cosines, individual wavelet functions are quite localized in space; simultaneously, like sines and cosines, individual wavelet functions are quite localized in frequency."

FT Motivation

- Time Domain: valuable information suppressed.
- Frequency spectrum shows what frequencies exist in the signal
- Frequency plot tells us how much of each frequency exists in the signal (frequency on the x axis and quantity on the y axis)
- Frequency spectrum of a real valued signal always symmetric. So a 50 Hz signal (from a light bulb) will have a matching frequency at about 950 Hz. Since the symmetric part provides no extra info, usually suppressed
- Frequency plot doesn't tell us when in time the frequency components exist (stationary signal) – ie the frequency of stationary signals doesn't change in time

Stationary signal:
x(t)=cos(2*pi*10*t)+cos(2*pi*25*t) +cos(2*pi*50*t)+cos(2*pi*100*t)



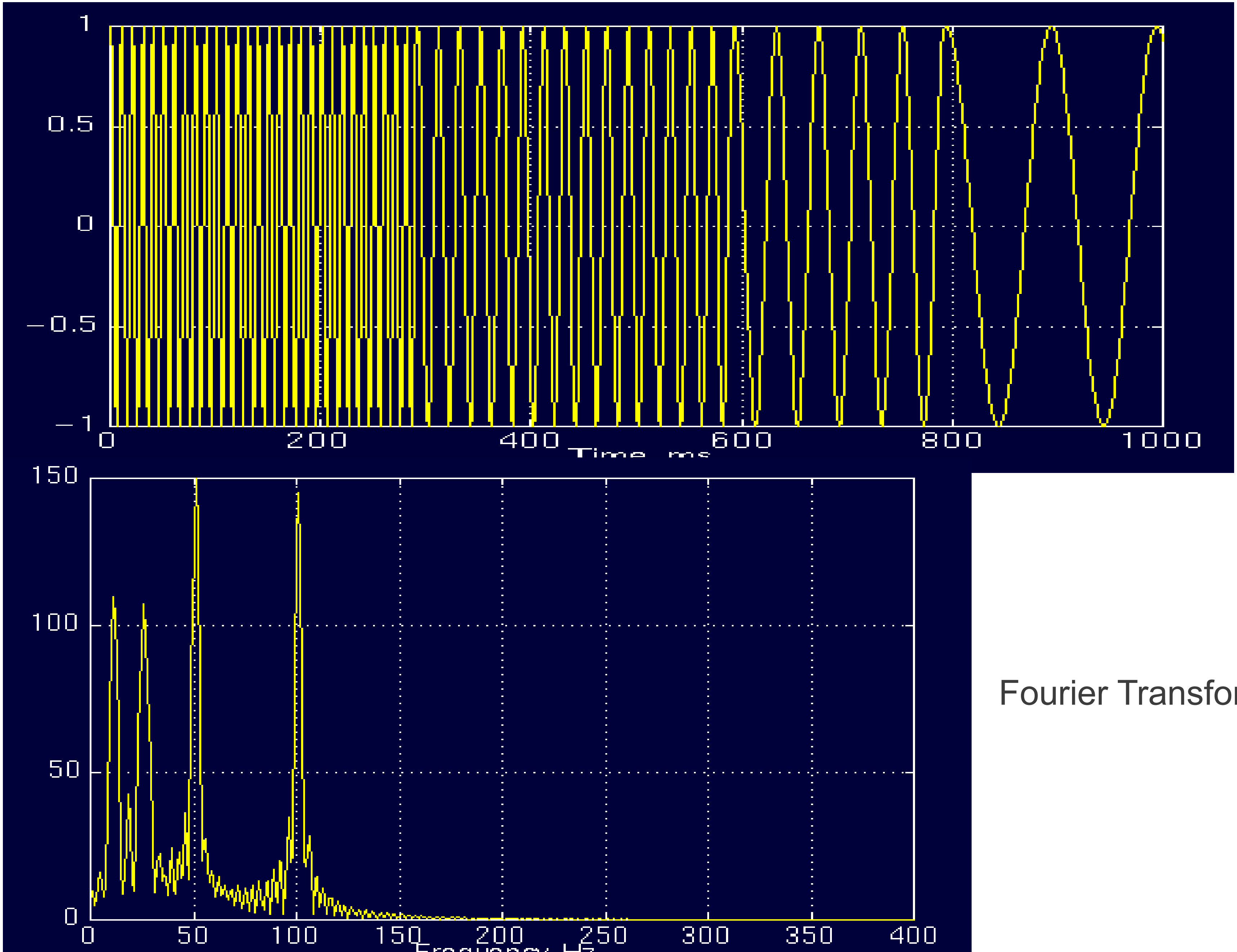


Fourier Transform

Non-stationary signals

- Biological signals: ECG (heart), EEG (brain), EMG (muscles)
- On an EEG you would like to know the time intervals between spectral components
- Example: turn on and off a flashlight. What's the time lapse of the brain between the stimulus and response?

Non-stationary signal



Fourier Transform

FT not able to distinguish two signals of same frequency

Not suitable for non-stationary signals (ie frequency varying in time)

If you don't care about time, FT can be a good tool

What is the Fourier Transform Actually Doing?

- Any periodic function can be expressed as infinite sum of periodic complex exponential functions
- Techniques generalized to non-periodic functions
- Signal x(t) mutliplied by exponential term at some frequency and integrated over ALL TIMES
- What we're actually doing is multiplying original signal by complex expression of sines and cosines and integrated
- Sinusoidal term has a frequency f, so if the signal has a high amplitude coefficient at that frequency their product will result in a large value, ie) a major spectra component of f. On the other hand, if the signal doesn't have that component it will warrant zero
- Integral runs from -infinity to + infinity in time. All times added and no matter where in time the frequency occurs it will affect the final product equally.
- THIS IS WHY THE FOURIER TRANSFORM NOT SUITABLE FOR A TIME VARYING FREQUENCY!!! IE) NON-STATIONARY SIGNAL

Alternatives

- Many potential transforms: Hilbert transform, short-time Fourier transform, Wigner distributions, Radon Transform, Wavelet transform
- Short-time Fourier, Wigner and Wavelet give time-frequency representation of signal
- WT developed to fix resolution problems of STFT

Short Term Fourier Transform

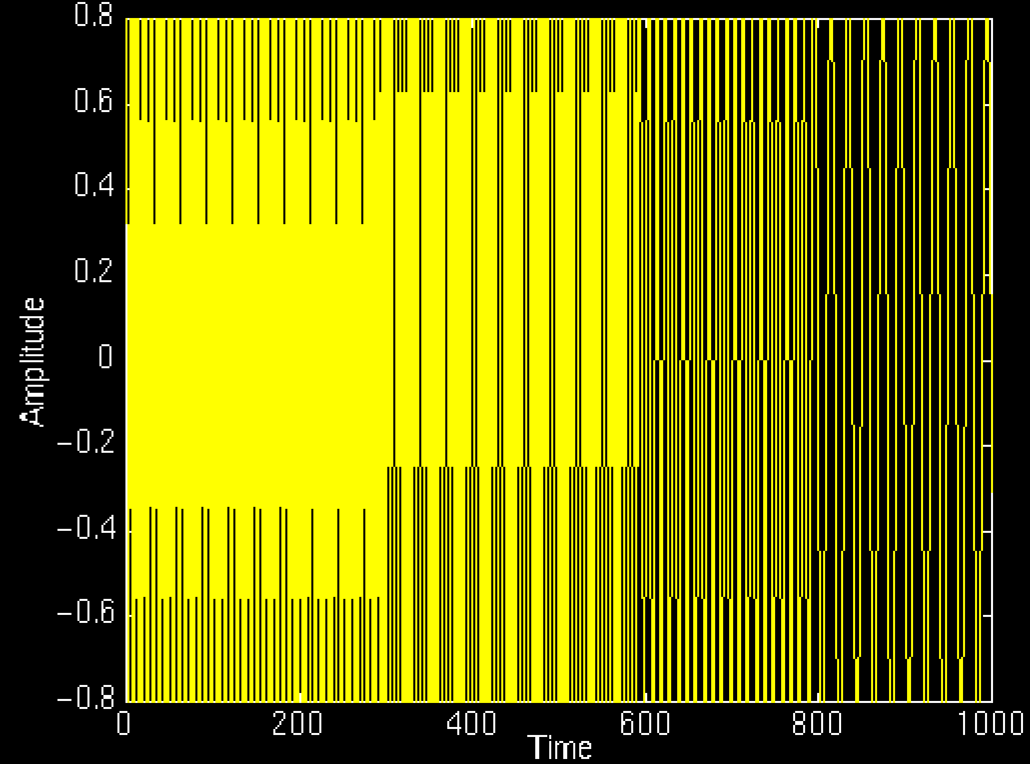
- First solution to time varying frequency signals and the short-comings of FT
- Even non-stationary signals have portions in which they are stationary – see example every 250 time units
- So the solution was to break the signal up into narrow, stationary portions of the signal
- Difference between FT and STFT is that a window function is needed to designate the width of these windows of information
- Narrow window → good time resolution, poor frequency resolution
- Large window → good frequency resolution, poor time resolution

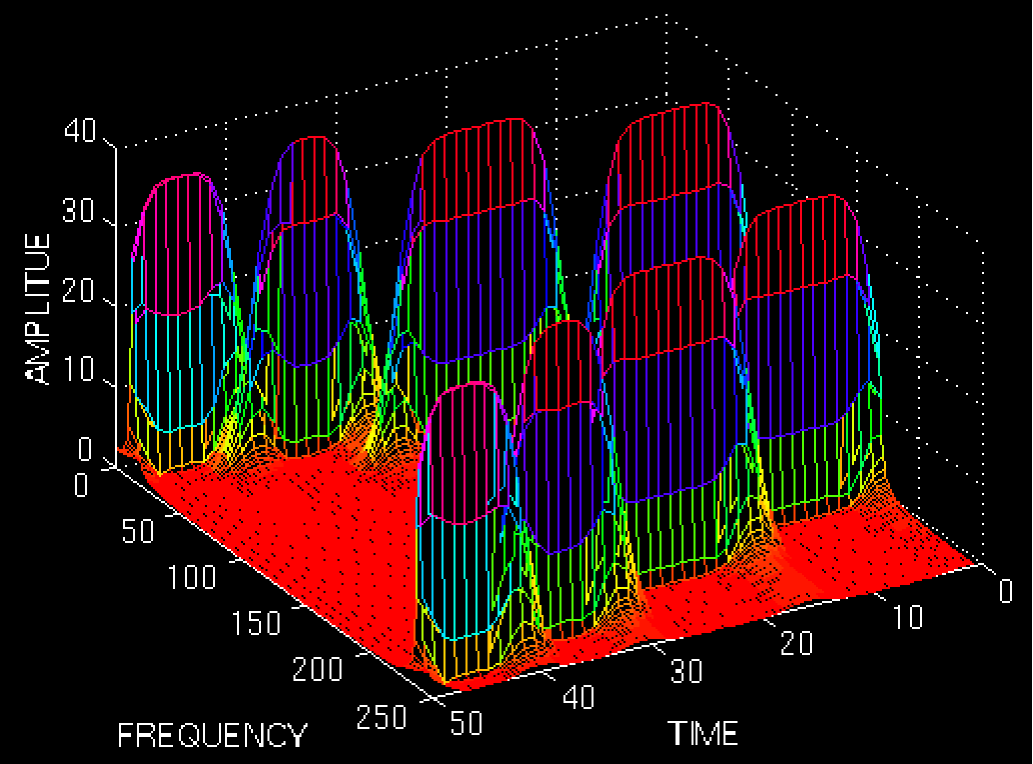
$$\mathbf{STFT}\{x(t)\}(\tau,\omega) \equiv X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}\,dt$$

Uncertainty Principle

- Originally formulated by Heisenberg in reference to Quantum Mechanics: Momentum and position of moving particle cannot be known simultaneously
- In reference to our problem . . . We can't know exactly what frequency exists at what time
- But we can know what frequency bands exist at what time intervals
- Problem of resolution and the main reason for the switch from STFT to WT
- STFT gives fixed resolution at all times while WT gives variable resolution . . . higher frequencies are better resolved in time while lower frequencies are better resolved in frequency, ie) less error

Example

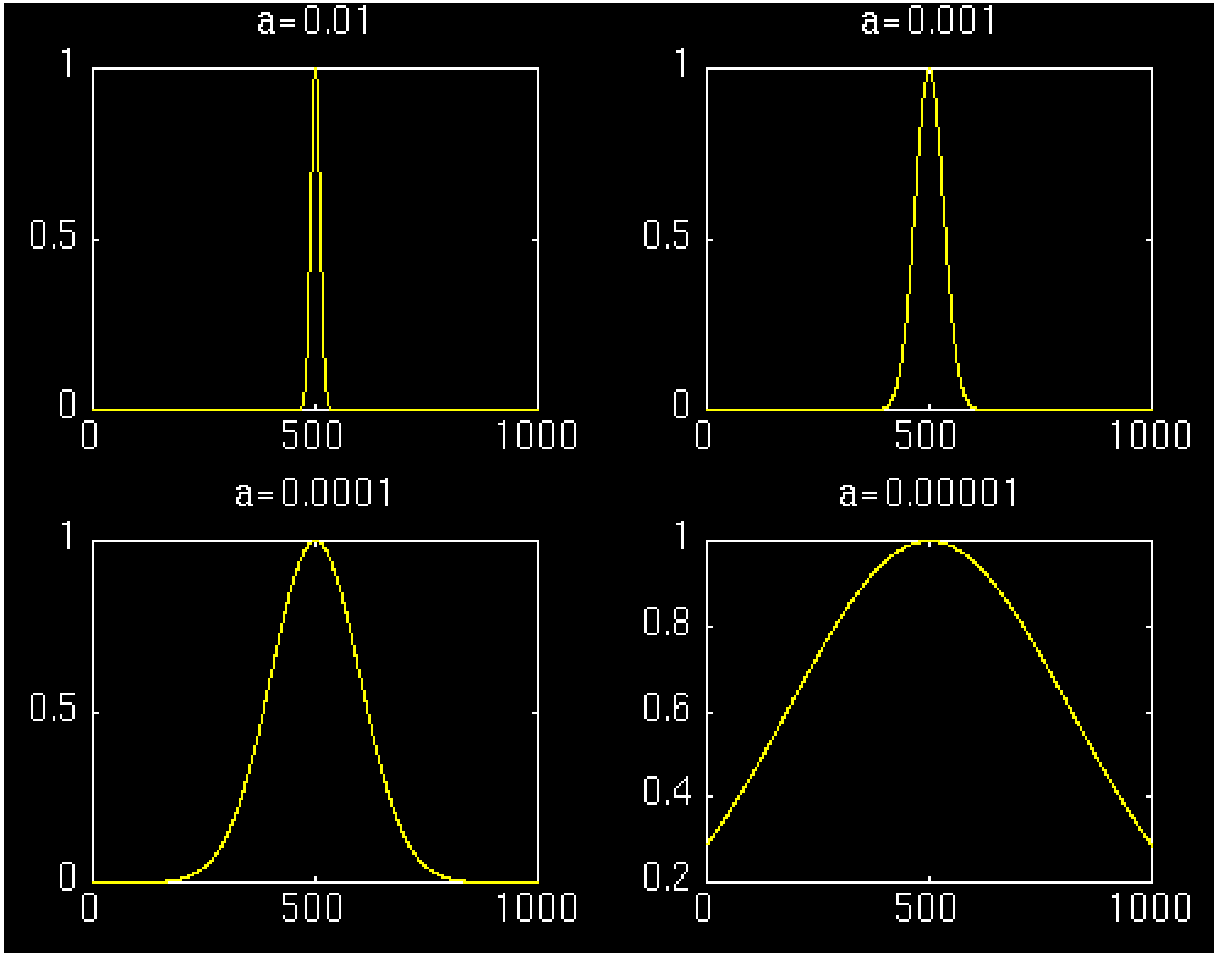




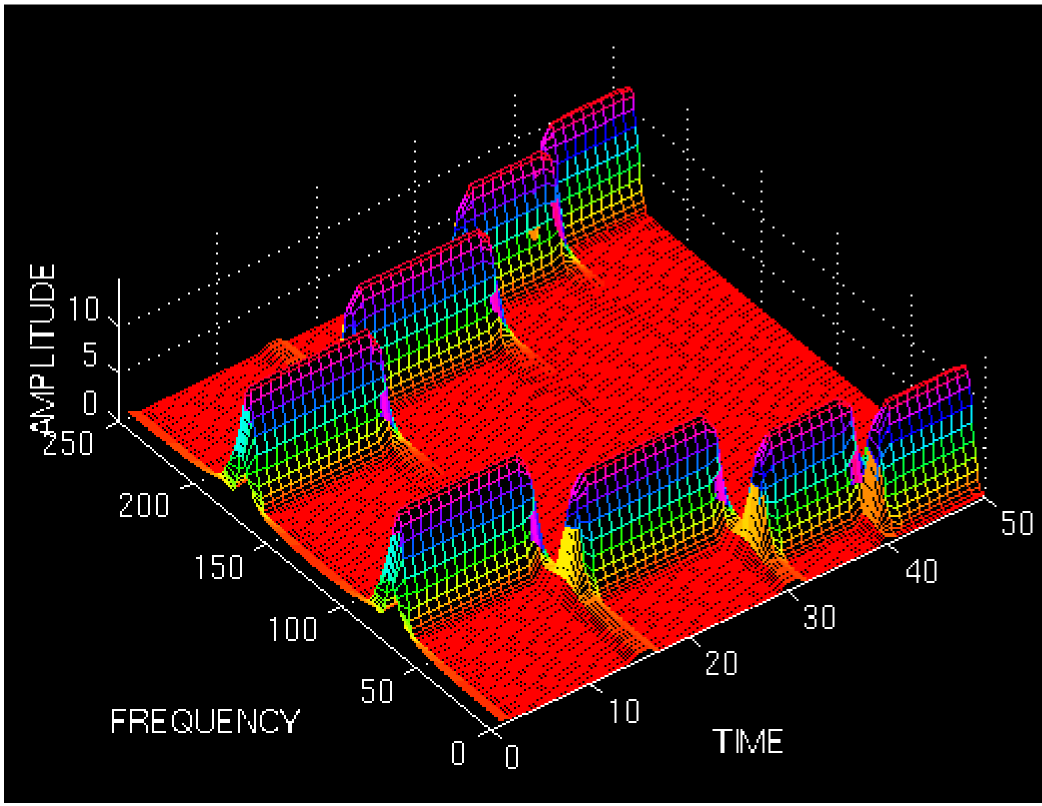
4 Frequency components at different times: Sine curves at 250 ms intervals: 300 Hz, 200 Hz, 100 Hz, 50 Hz

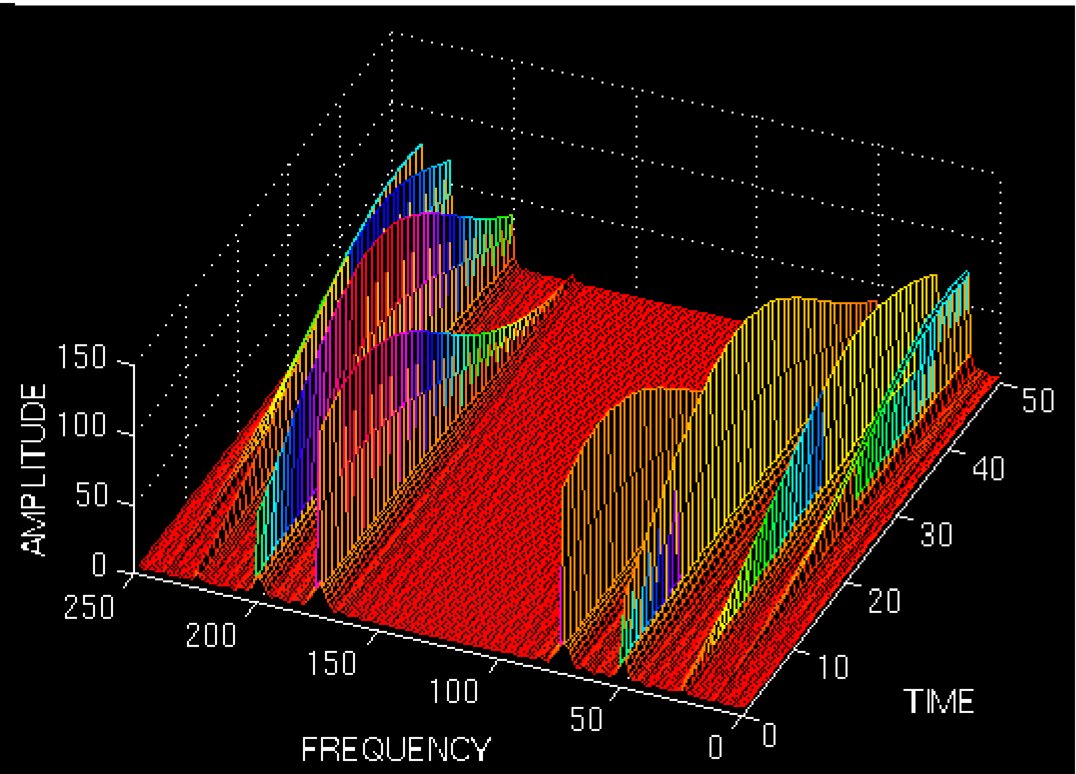
STFT (symmetric since FT is also symmetric)

Problem with STFT
Example: 4 window sizes



STFT of Window Lengths





Window 1: narrow window → Good time resolution, poor frequency resolution

Window 2: Wide window → Good frequency resolution, poor time resolution

Problem with STFT: Resolution

- In FT, there's no resolution problem. In time domain, we know exactly the time and none of the frequency. In frequency, we know exactly the frequency and none of the time
- Happens because the window in FT is infinite the e^iwt function. While in STFT our window is finite length
- Resolution dilemma
- WT solves it, to some extent

Wavelet Transform

- HUP exists regardless of the transform employed
- But possible to analyze the signal with a different approach called multiresolution analysis (MRA)
- Analyzes the signal at different frequencies and resolutions
- Every frequency component not resolved equally
- At high frequencies: Good time resolution; poor frequency resolution
- At low frequencies: Good frequency resolution; poor time resolution
- Makes sense when signal has high frequency for short durations and low frequency for long durations
- Signals encountered often of this type

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi} \, dx$$

ξ :frequency

$$X(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \overline{\Psi\left(\frac{t-b}{a}\right)} x(t) \, dt$$

a:scaling b:time

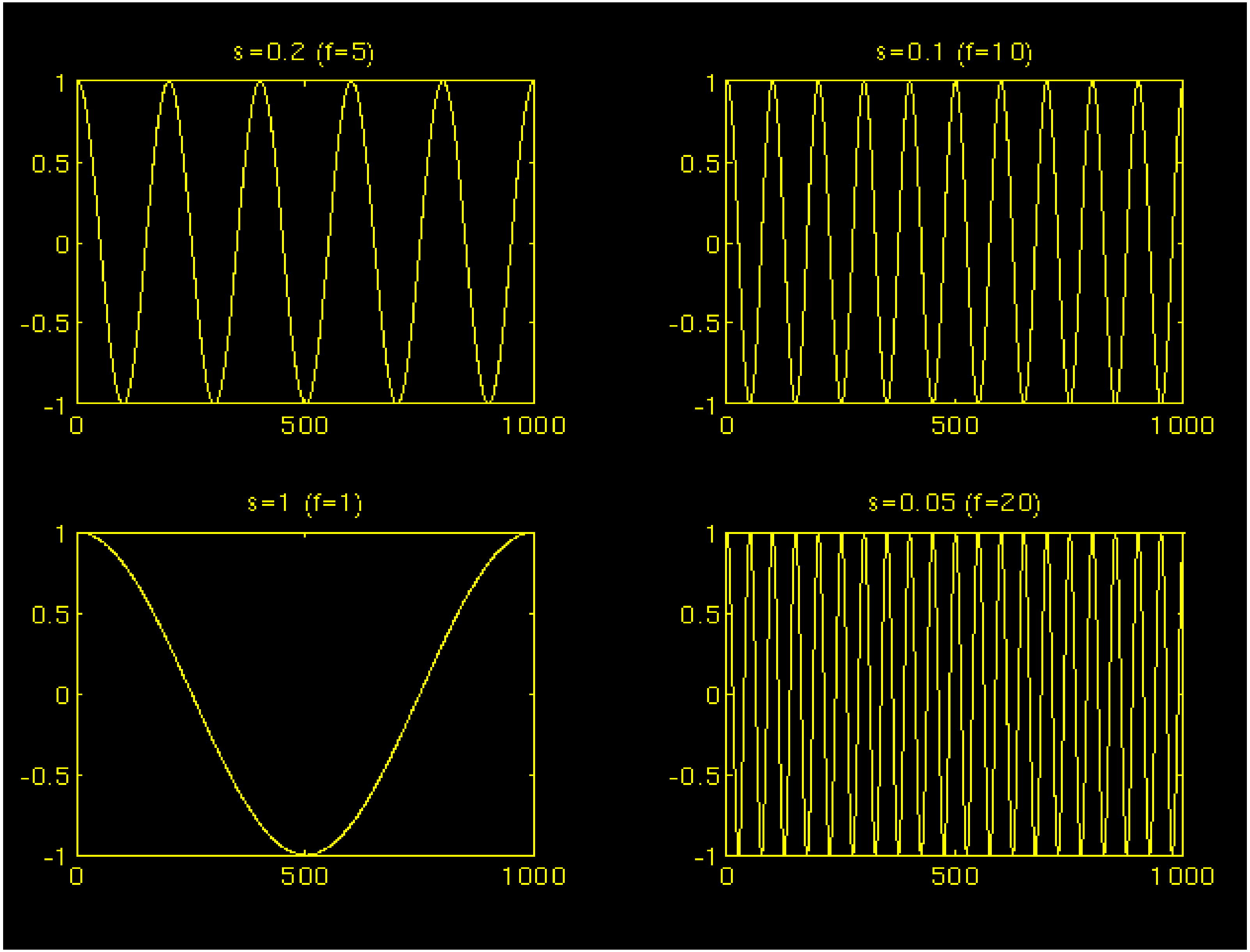
WT

- Similar to STFT, the signal multiplied by a function (a wavelet) and computed separately for different segments of time signal
- Two differences
  - FT of window signals not taken
  - Width of window changed as transform is taken for every spectral component

Wavlet Jargon

- Function of two variables: translation and scale
- Psi(t) is the transforming function (mother wavelet)
- Wavelet means small wave. ie) window function is of finite length
- Wave means window function is oscillatory
- Mother means the functions used in different regions all derive from one common function
- Mother wavelet is the prototype for other window functions
- Translation refers to the location of the window. The window is shifted through the signal . . . corresponds to time info
- Scale defined as inverse frequency
- Scale similar to scale on a map. A large scale tells us that we have a non-detailed global view. In terms of frequency, a low frequency (high scale) means we see a global, non-detailed image
- Low scales usually don't last for entire duration of signal. High scales do

Scale

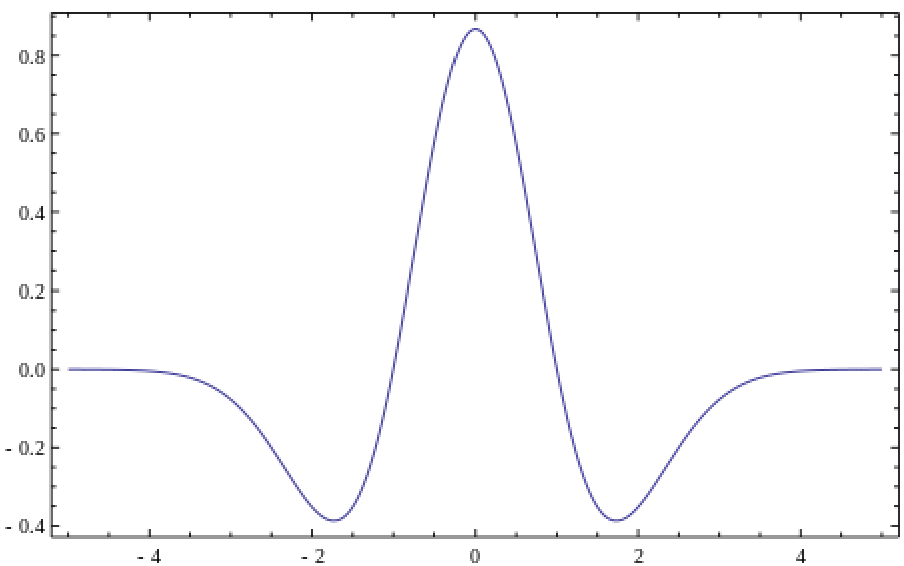


Mother wavelet

- All windows used are dilated/shifted versions of the mother wavelet (scale/transformation)
- Many candidates: Morlet wavelet and Mexican hat function are two possibilities

Mexican Hat

$$w(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-t^2}{2\sigma^2}}$$

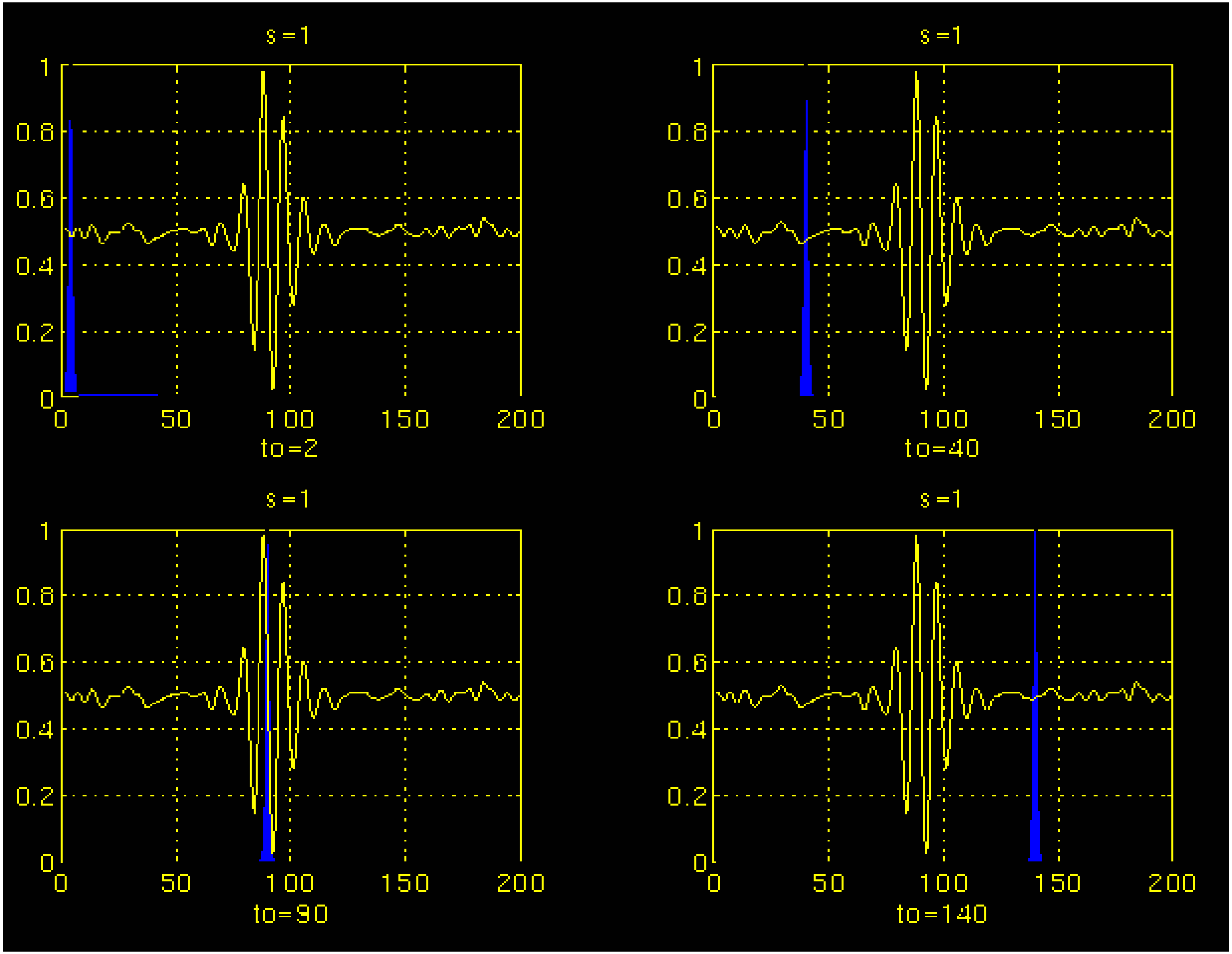


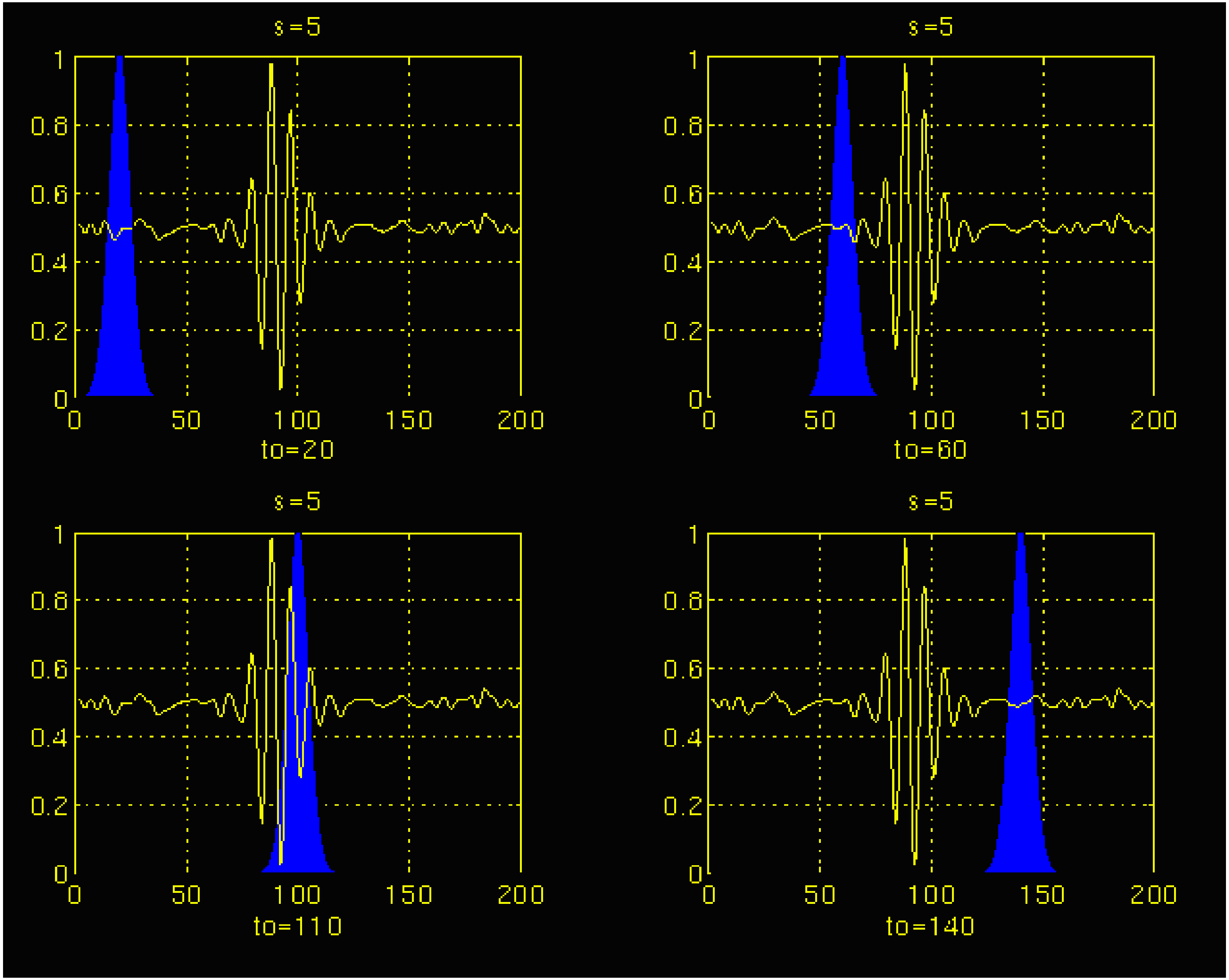
Computation of Continuous Wavelet Transform

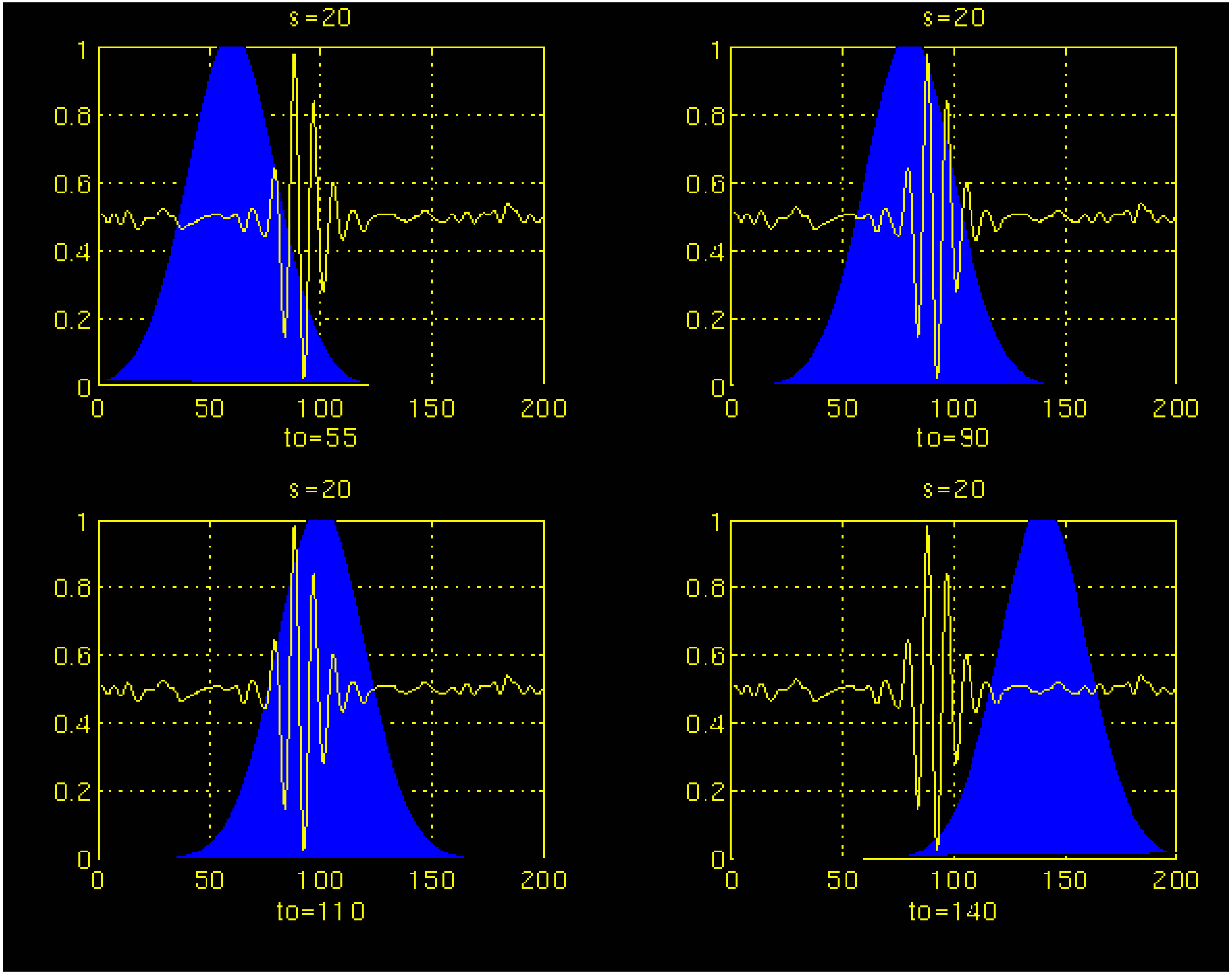
- After choosing the mother, computation starts with s= 1 and CWT taken for all values of s greater and smaller than 1
- Wavelet placed at beginning of signal (t=0)
- Wavelet at scale 1 multiplied by signal/integrated over all times. Result multiplied by 1/sqrt{s} for normalization. This gives the value at tau=0, s=1. Wavelet at s=1 then translated by tau and value found for tau=tau, s=1. Continued till end of signal when s is incremented
- Product at each step is nonzero only when signal falls in region of support of the wavelet. ie) signal must have spectral component that corresponds to value of s to yield a product
- Shifting wavelet in time, localizes signal in time. Changing value of s, localizes signal in scale (frequency)

- For every scale and every time, one point in time-scale plane computed
- Translation related to time. Scale shows inverse frequency
- Time-Frequency plot consists of squares in STFT
- In WT it consists of rectangles of different dimensions
- HUP still holds. But smarter, targeted analysis
- CWT can be thought of as inner product of test signal with basis function psi . . . Gives a measure of similarity between the basis functions (wavelets) and the signal itself
- CWT is a reversible transform as long as it meats non-restrictive admissibility condition:

Example: different scale and translation values

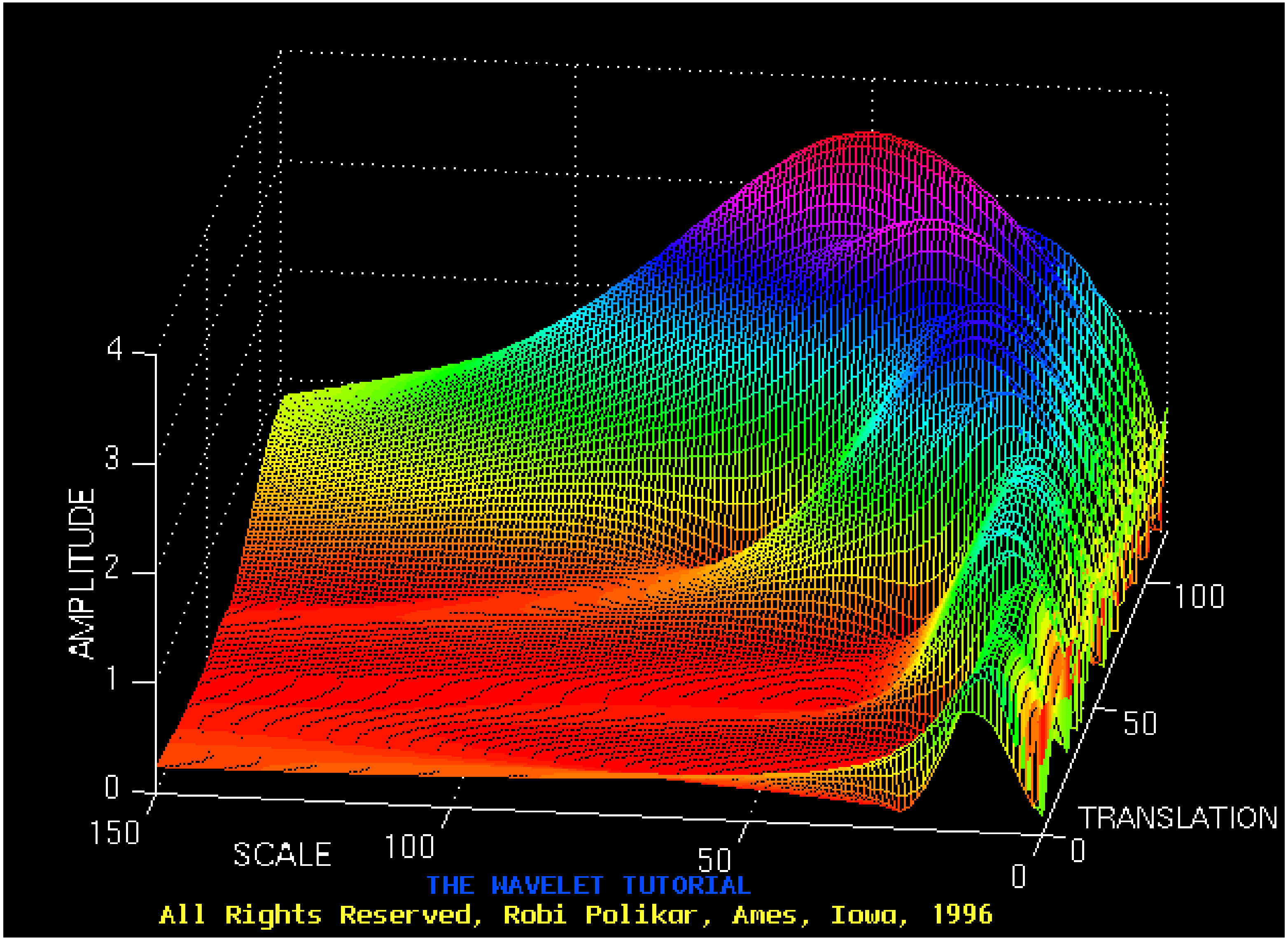



s=5
to=20
s=5
to=60
s=5
to=110
s=5
to=140


s=20
to=55
s=20
to=90
s=20
to=110
s=20
to=140

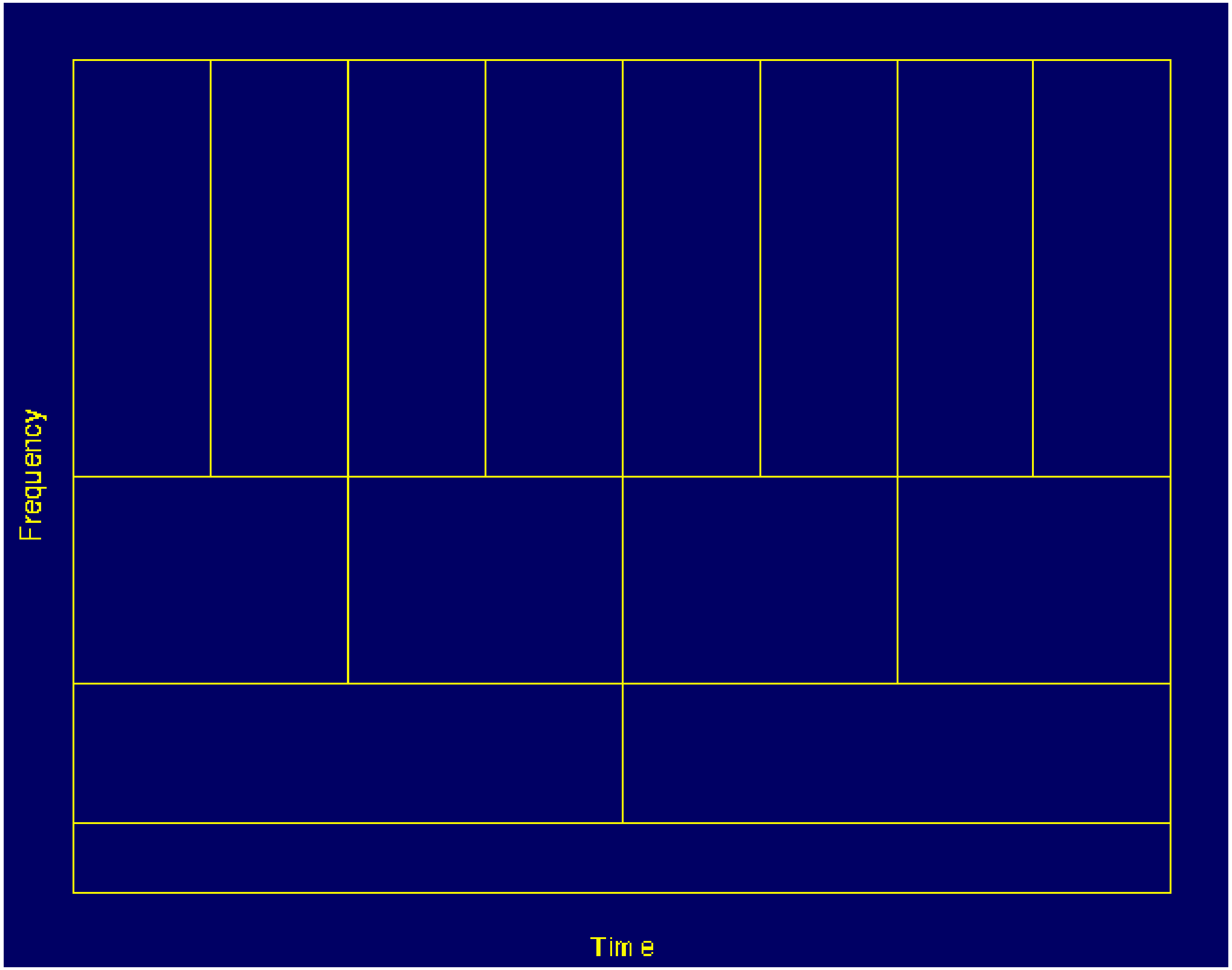
Example

30 Hz
20 Hz
10 Hz
5 Hz
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0
200
400
600
800
1000
Time, ms



4
3
2
1
0
AMPLITUDE
150
SCALE
100
50
0
TRANSLATION
0
50
100
THE WAVELET TUTORIAL
All Rights Reserved, Robi Polikar, Ames, Iowa, 1996

CWT Resolution



DWT

- Filters of different cutoff frequencies used to analyze signal at different scales
- Rather than taking samples at all scales and translations, signal passed through series of high pas filters to analyze high frequencies and low pass filters to analyze low frequencies
- Filtering means convolving signal with filter response
- Scale changed by adding or removing data samples (subsampling by two means removing every two samples)
- Sampling frequency taken in radians (2pi) . . .
- highest frequency component will be pi radians, if signal sampled at Nyquist's rate (2x the max frequency in signal).
- Hz not appropriate for discrete signals (can be used in discussion but not application)
- Resolution affected by filtering operations, since related to amount of information in signal
- Low pass filtering halves the resolution but leaves scale unchanged

- Decomposition of signal into different frequency bands done by successive high and lowpass filtering of time domain signal
- Nyquist's rule says after filtering half the samples can be removed
- Decomposition of the signal into low and high pass components, halves the time resolution, since you have half the # of samples but doubles the frequency resolution, since you have frequency band spans only half the previous band
- Can be repeated indefinitely

Example - Subbanding

- Original signal has 512 sample points with frequency band from zero to p rad/s
- First decomposition and subsampling of 2, leaves the highpass filter output with 256 points and frequencies from p/2 to p rad/s
- Lowpass filter output left with 256 points with frequencies from 0 to p rad/s
- Process continues until only 2 samples are left (8 levels of decomposition in this example)
- DWT found by concatenating all coefficients from last level of decomposition. DWT left with same number of coefficients as original signal

- Most prominent frequencies in original signal have high amplitudes in that region of DWT with those frequencies
- Unlike FT, the time frequency not lost!
- But resolution depends on where the signal lies
- If it lies in high frequency, as it usually does, time localization is precise . . . since many samples are used to express this
- If it lies in low frequency it is imprecise, since few samples compose this measurement
- ie) good time resolution at high frequencies. Good frequency resolution at low frequencies
- Data reduction from discarding low amplitude data
- In the low frequency region, only first 64 samples carry relevant info and rest can be discarded

- Ideal filters that always provide perfect reconstruction not possible, but under certain conditions they can provide perfect reconstruction
- Most famous are the Daubechies wavelets
- Length of signal must be power of 2
- DWT coefficients of each level are concatenated starting with last level

Example

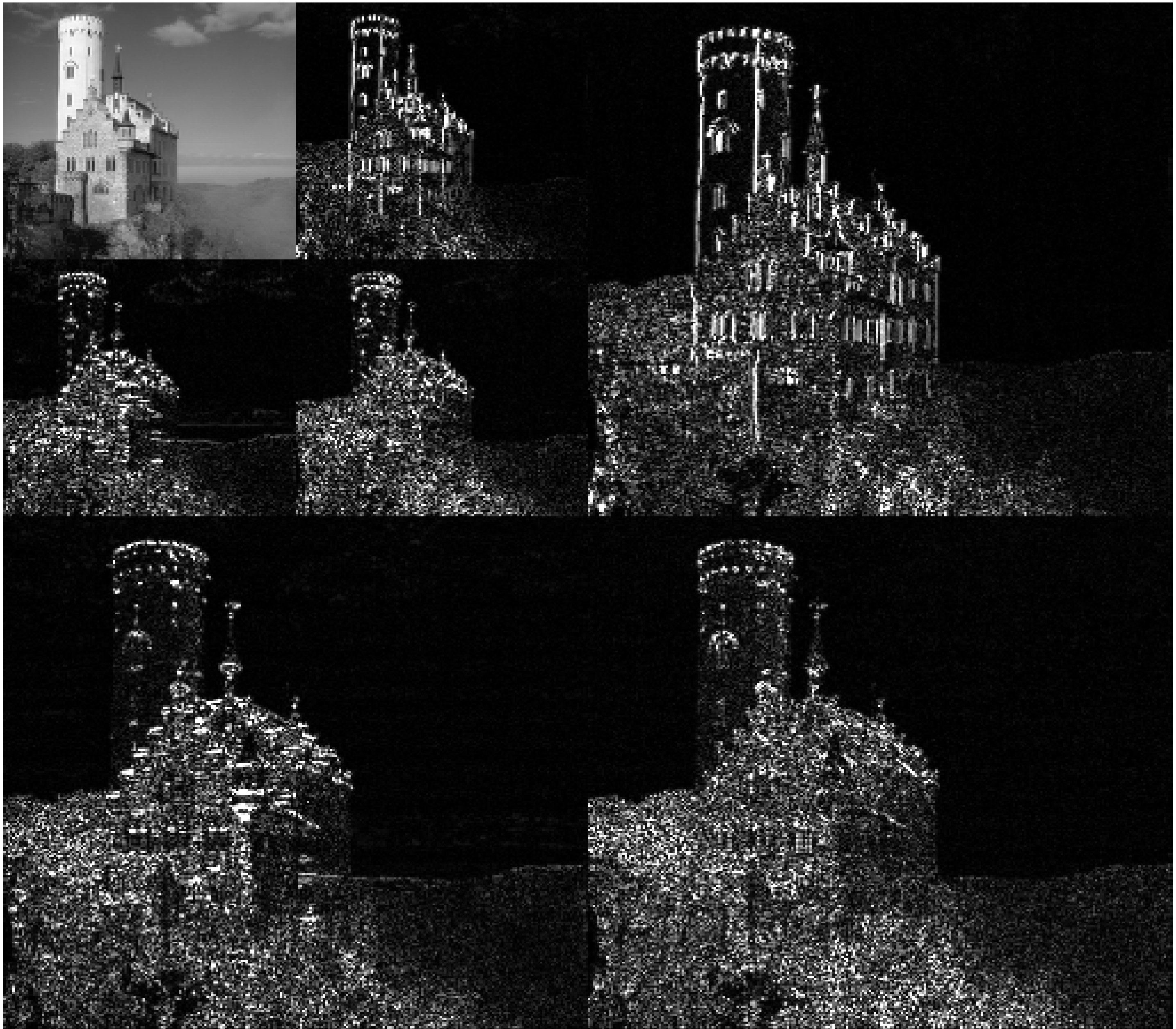
- Suppose a 256 sample signal sampled at 10 MHz and we are looking for DWT coefficients
- Since we have a 10 MHz signal, the highest frequency component in the signal can be 5 MHz
- High and low pass filters applied at each level. 9 levels total

Coefficient levels: 256 samples

- Level 1: 256 coefficients
- Level 2: 128 coefficients
- Level 3: 64 coefficients
- Level 4: 32 coefficients
- Level 5: 16 coefficients
- Level 6: 8 coefficients
- Level 7: 4 coefficients
- Level 8: 2 coefficients
- Level 9: 1 coefficient

Wavelet Tutorial: Robi Polikar, Rowan University

- Part I: http://users.rowan.edu/~polikar/WAVELETS/WTpart1.h
- Part II: http://users.rowan.edu/~polikar/WAVELETS/WTpart2.h
- Part III: http://users.rowan.edu/~polikar/WAVELETS/WTpart3.h
- Part IV: http://users.rowan.edu/~polikar/WAVELETS/WTpart4.h



- DWT used for data compression for previously sampled signal
- CWT used for signal analysis
- DWT common in engineering and computer science while CWT common in scientific research
- JPEG 2000 is a compression standard that uses biorthogonal wavelets
- Smoothing/denoising data using wavelet coefficient threshold … threshold wavelet coefficients to smooth undesired frequency components
- Wavelet transform is representation of function by wavelets
- Wavelets are scaled/translated copies (daughter wavelets) of a finite-length/fast-decaying oscillating waveform (mother wavelet)
- Unlike Fourier transform, can represent functions with discontinuities/non-smooth functions and deconstruct finite/non-periodic signals. In fact smooth, periodic signals may be better compressed with Fourier methods
- CWT operate over every possible scale and translation while wavelets use a specific subset of scale/translation values

- Nyquist's rule tells us at lower frequencies, sampling rate can be decreased . . . saving us significant computational time
- How low can sampling rate be and still reconstruct the signal? Main question of optimization
- Discretized CWT is just a sampled version of CWT that allows computer processing (not the same as DWT) – information is highly redundant
- Discretized continuous wavelet transform can take long time depending on size of signal and desired resolution. DWT much faster