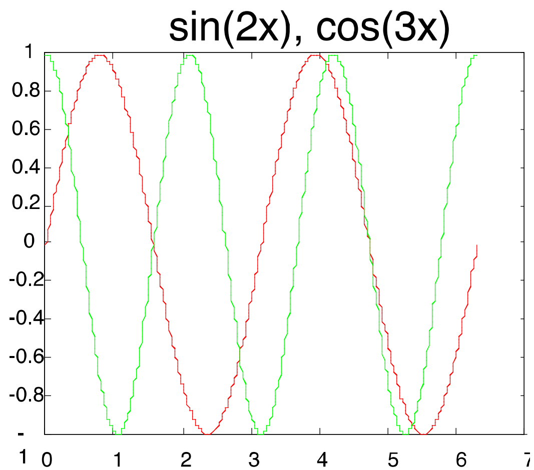
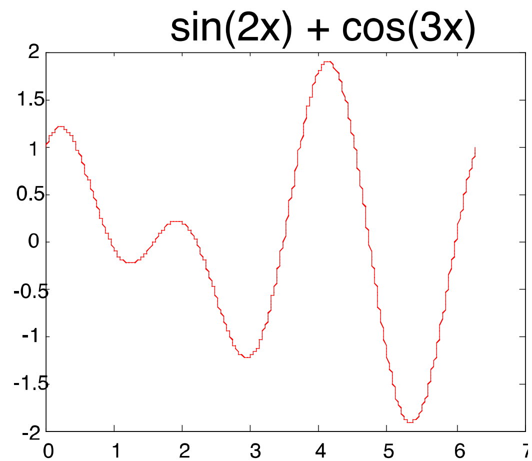
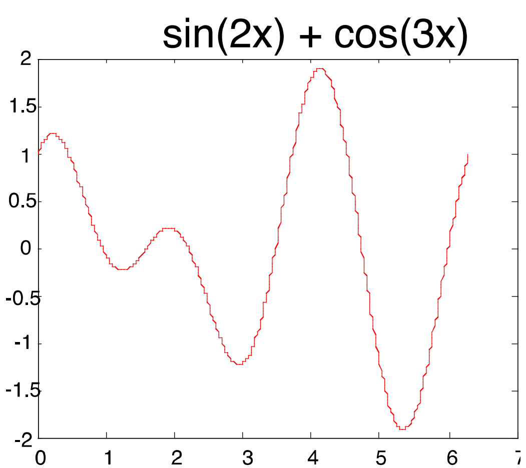
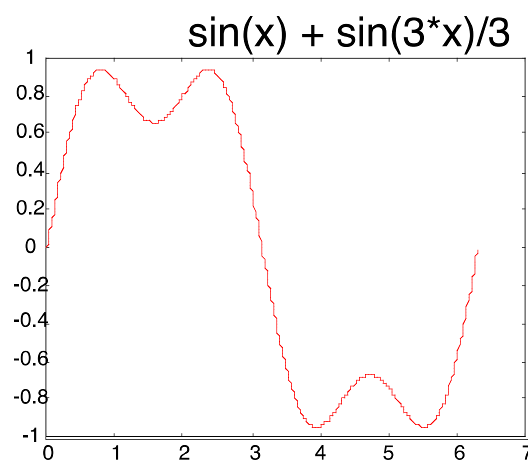
Fourier Analysis



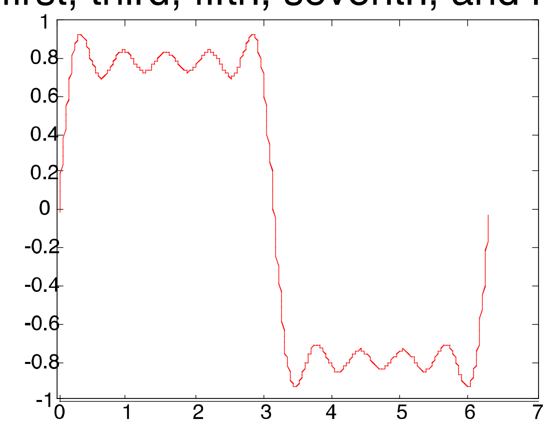


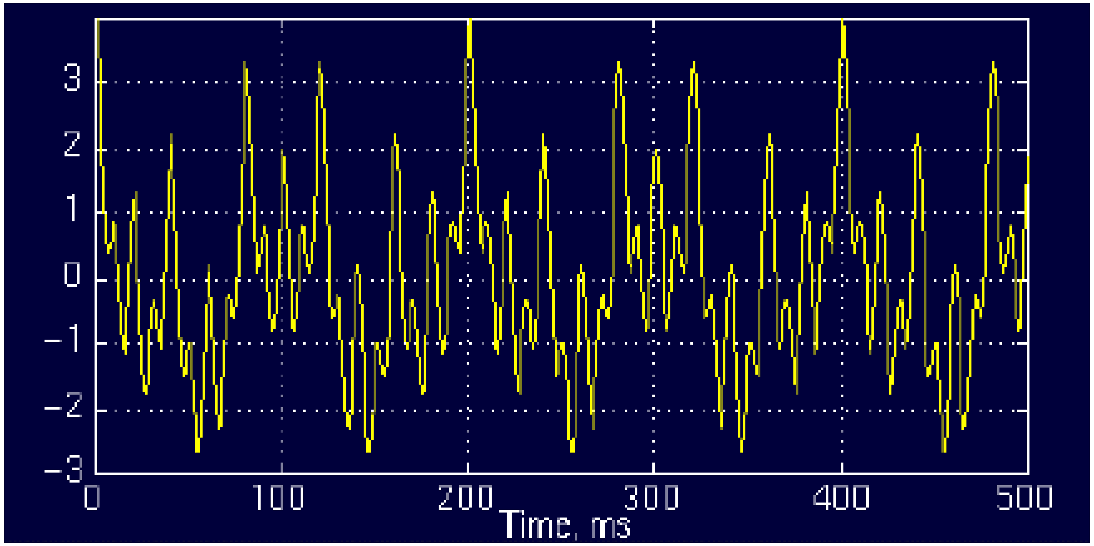




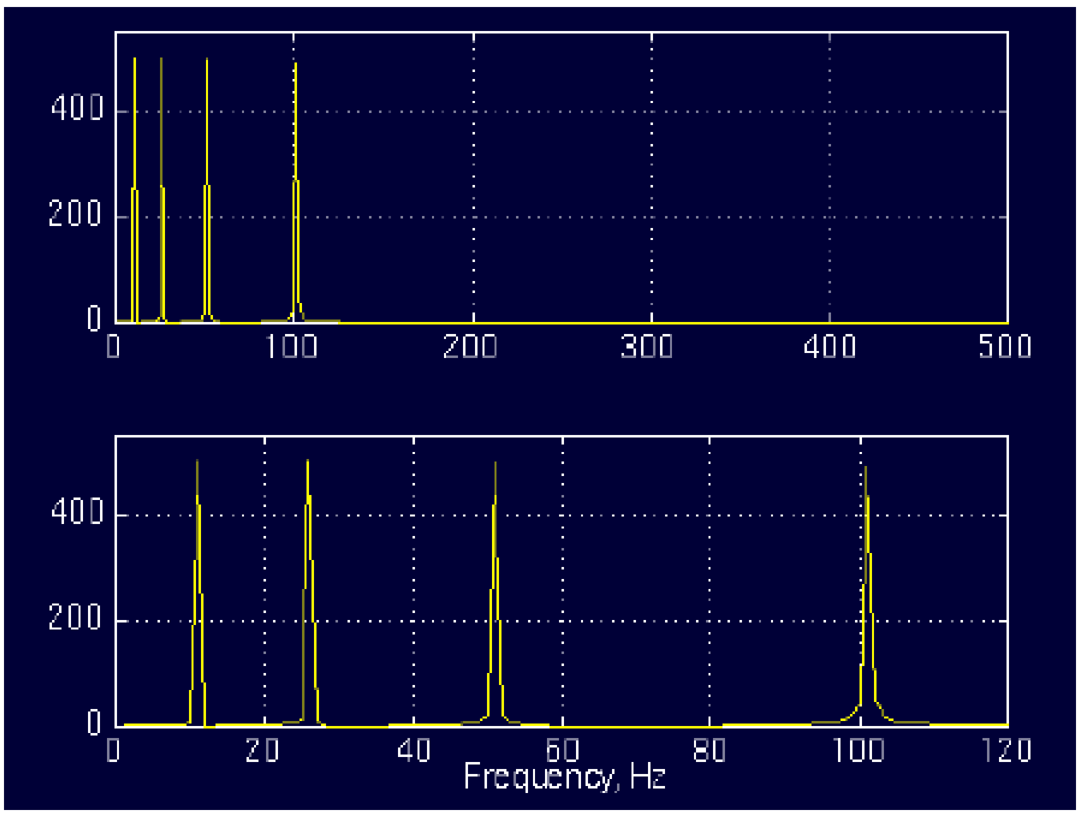
$$y(x) = \sin(x) + \sin(3x)/3 + \sin(7x)/7 + \sin(9x)/9$$

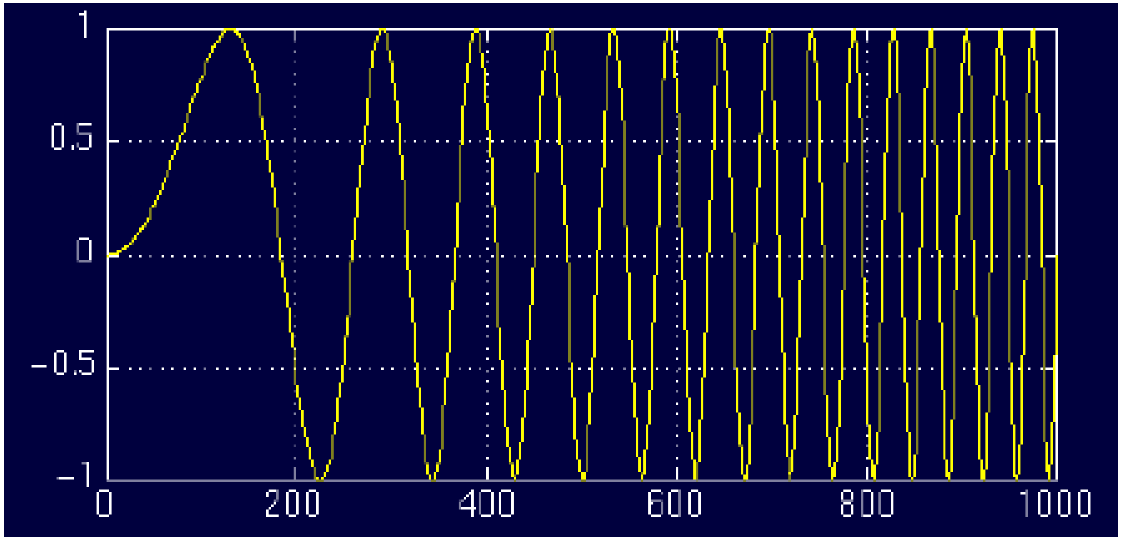
first, third, fifth, seventh, and ninth harmonics



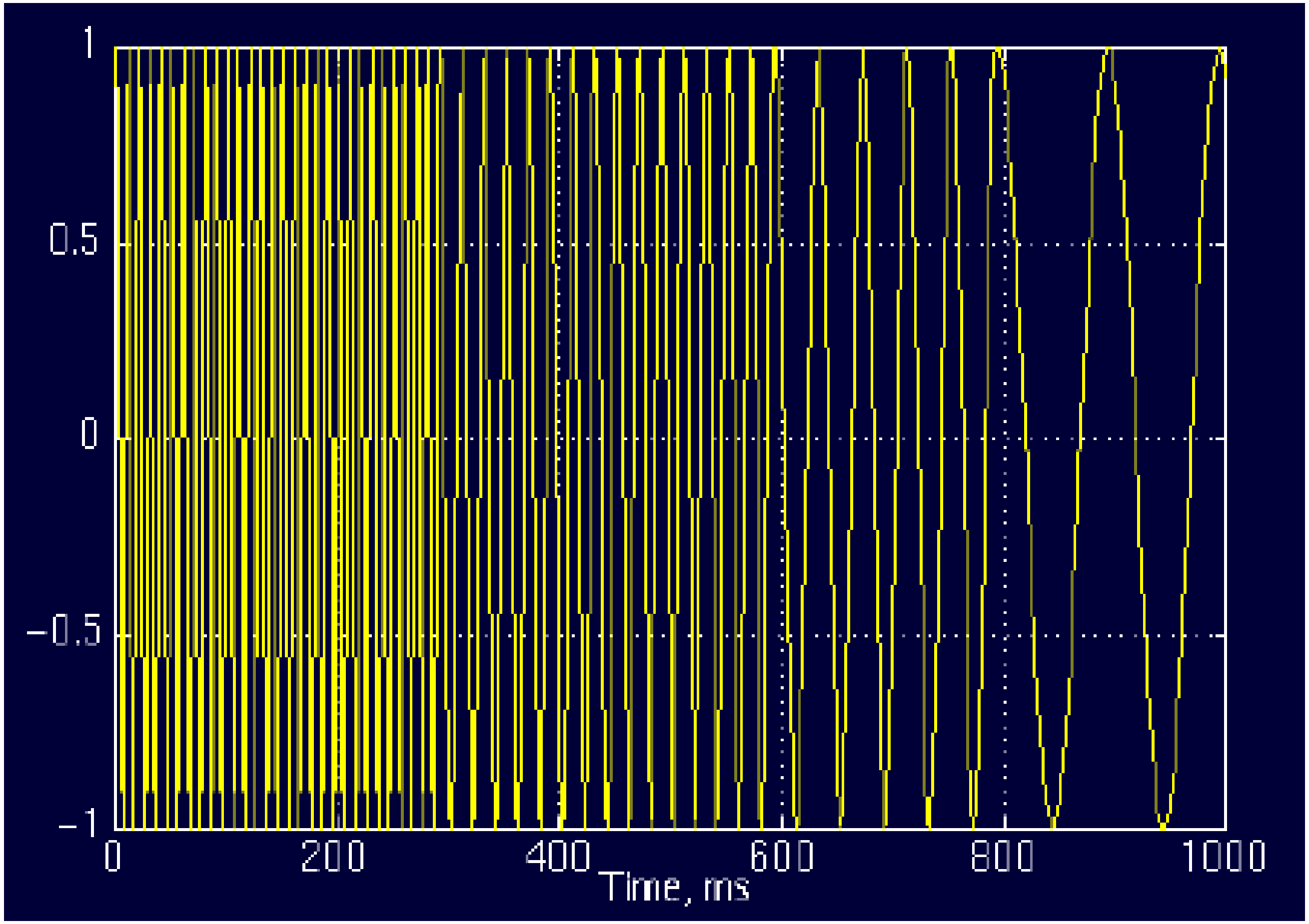


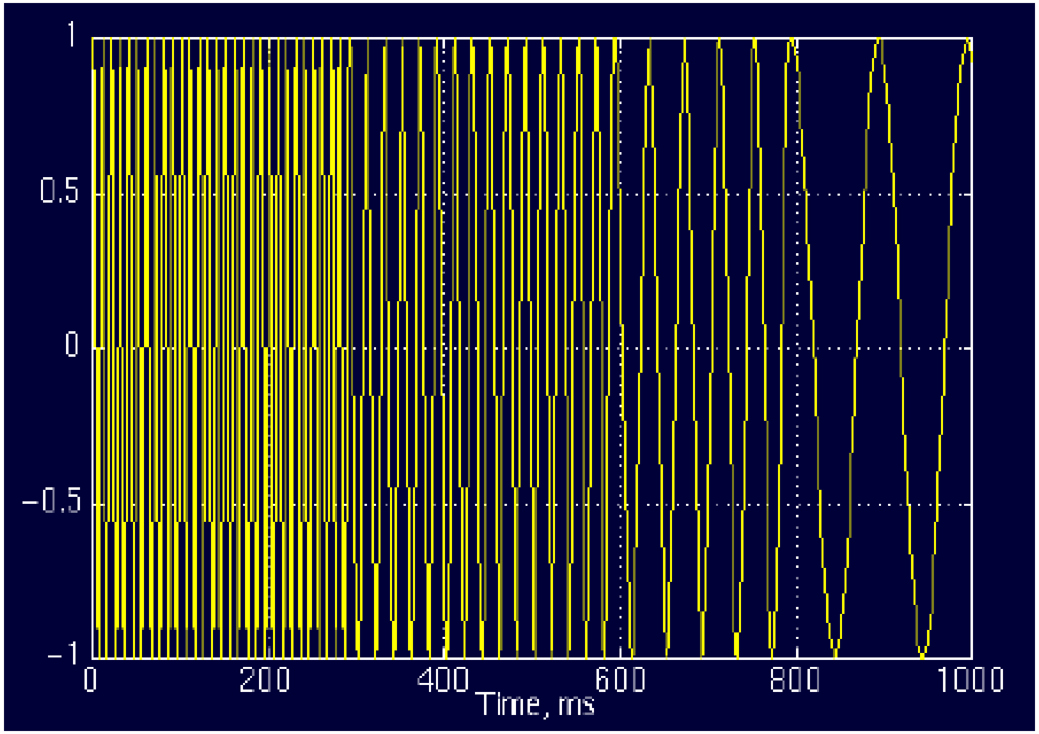
a stationary signal:
frequencies of 10, 25, 50, and 100 Hz are present at any given time instant

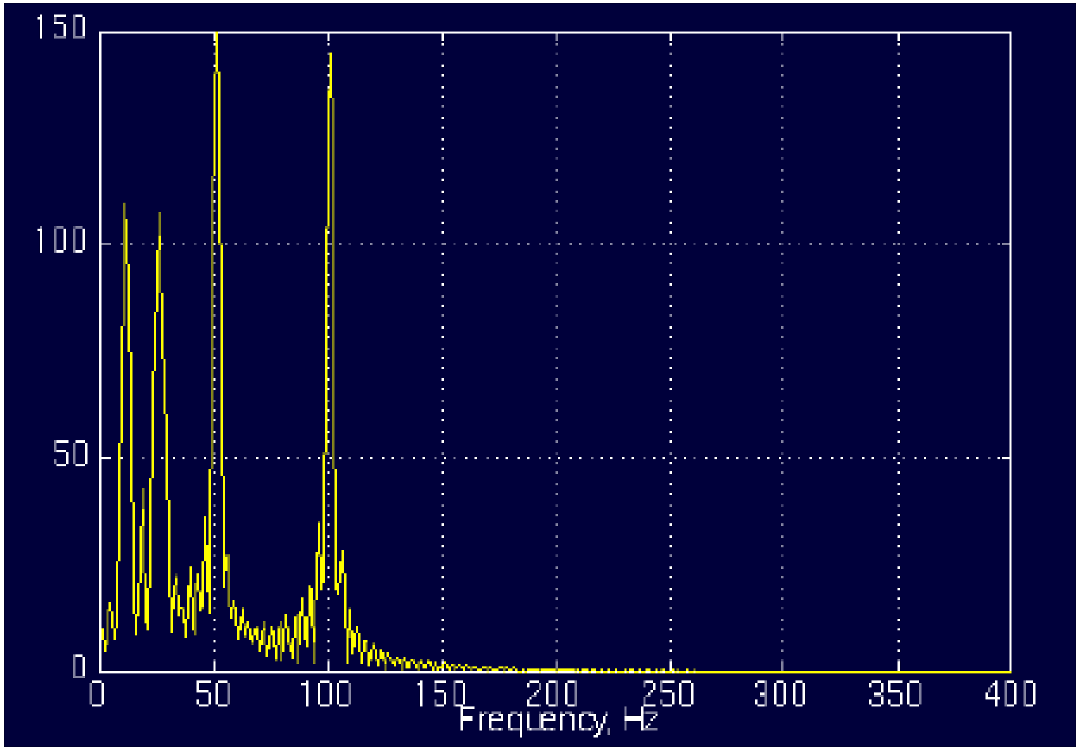




This signal is known as the "chirp" signal. This is a non-stationary signal.







Note that the amplitudes of higher frequency components are higher than those of the lower frequency ones. This is due to fact that higher frequencies last longer (300 ms each) than the lower frequency components (200 ms each). (The exact values of the amplitudes are not important).

At what times (or time intervals), do these frequency components occur?

At all times! Remember that in stationary signals, all frequency components that exist in the signal, exist throughout the entire duration of the signal. There is 10 Hz at all times, there is 50 Hz at all times, and there is 100 Hz at all times.

Fourier transform

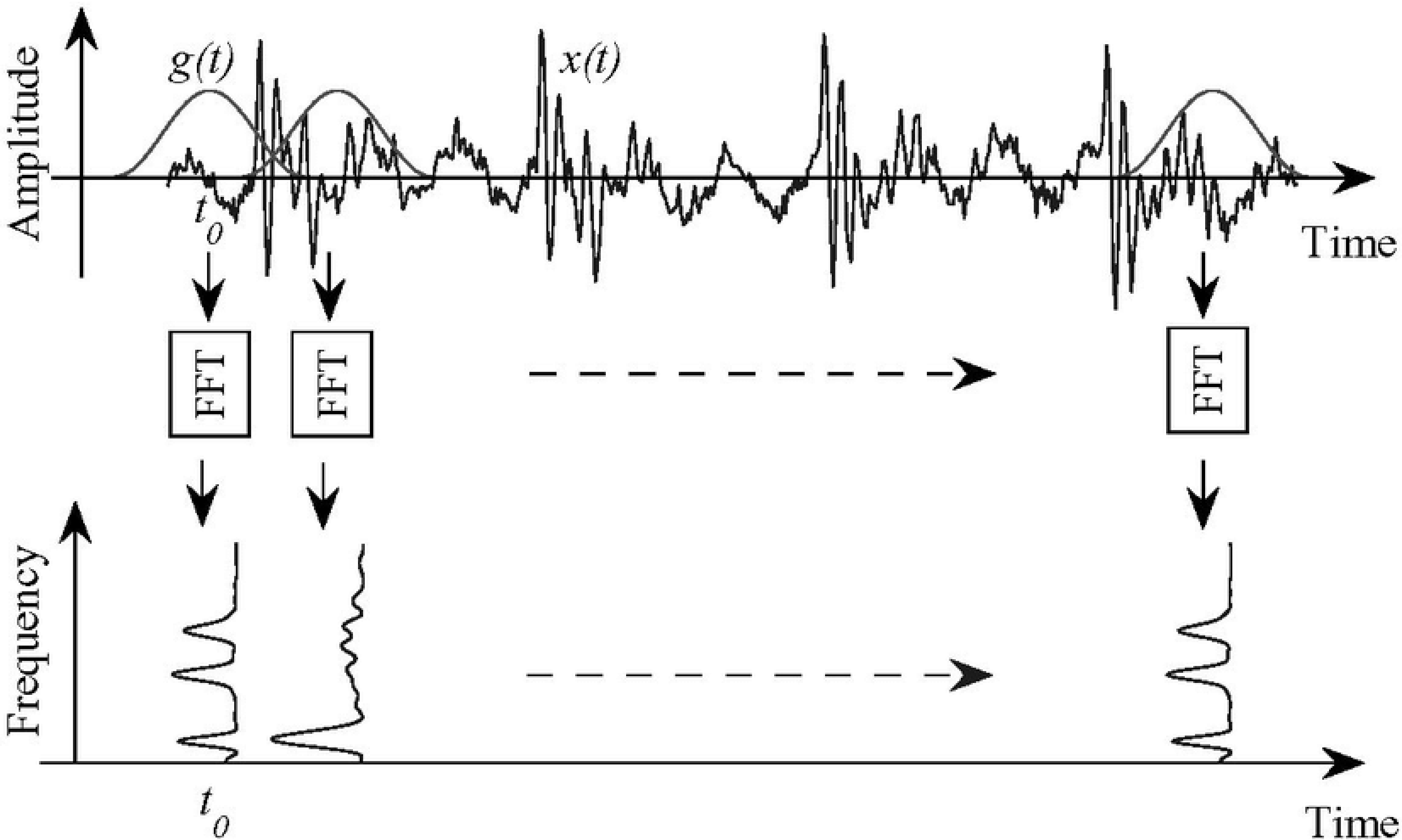
$$F(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-j\omega t) dt. \tag{4.9}$$

As we know, Fourier transform convert signal in time domain to frequency domain by integrating over the whole time axis. However, if the signal is not stationary, that is, the frequency composition is a function of time, we cannot tell when a certain frequency rises.

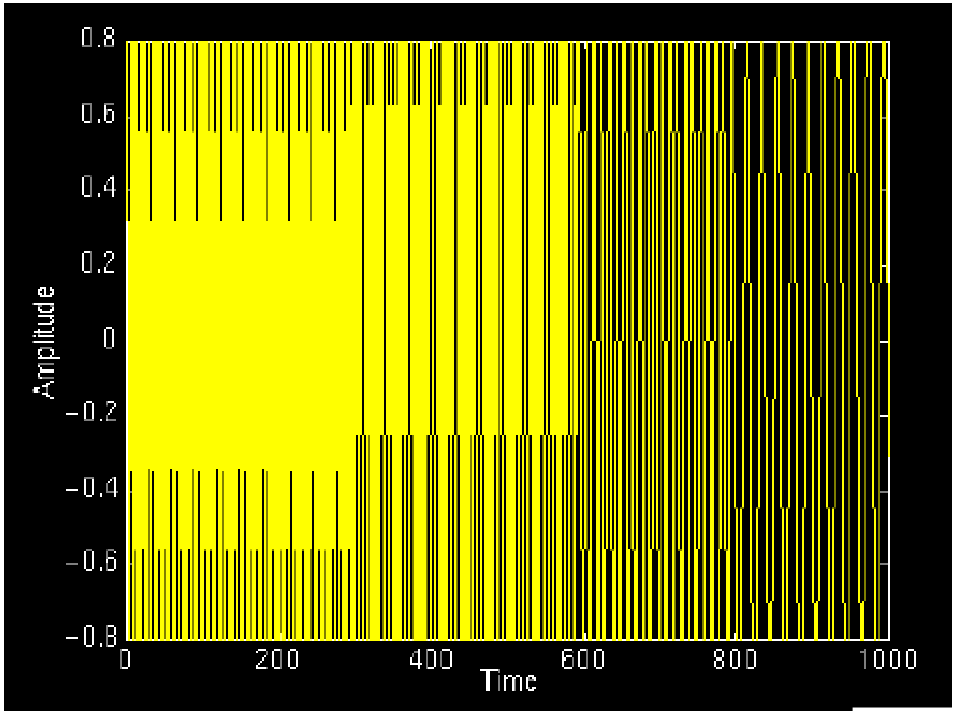
STFT

$$Sf(u,\xi) = \int_{-\infty}^{\infty} f(t)w(t-u) \exp(-j\xi t) dt. \tag{4.10}$$

The STFT tries to solve the problem in Fourier transform by introducing a sliding window $w(t-u)$. The window is designed to extract a small portion of the signal $f(t)$ and then take Fourier transform. The transformed coefficient has two independent parameters. One is the time parameter τ, indicating the instant we concern. The other is the frequency parameter ξ, just like that in the Fourier transform. However another problem rises. The very low frequency component cannot be detected on the spectrum. It is the reason that we use the window with fixed size. Suppose the window size is 1. If there is a signal with frequency 0.1Hz, the extracted data in 1 second look like flat (DC) in the time domain.



Amplitude
g(t)
x(t)
t0
Time
FFT
FFT
FFT
Frequency
t0
Time

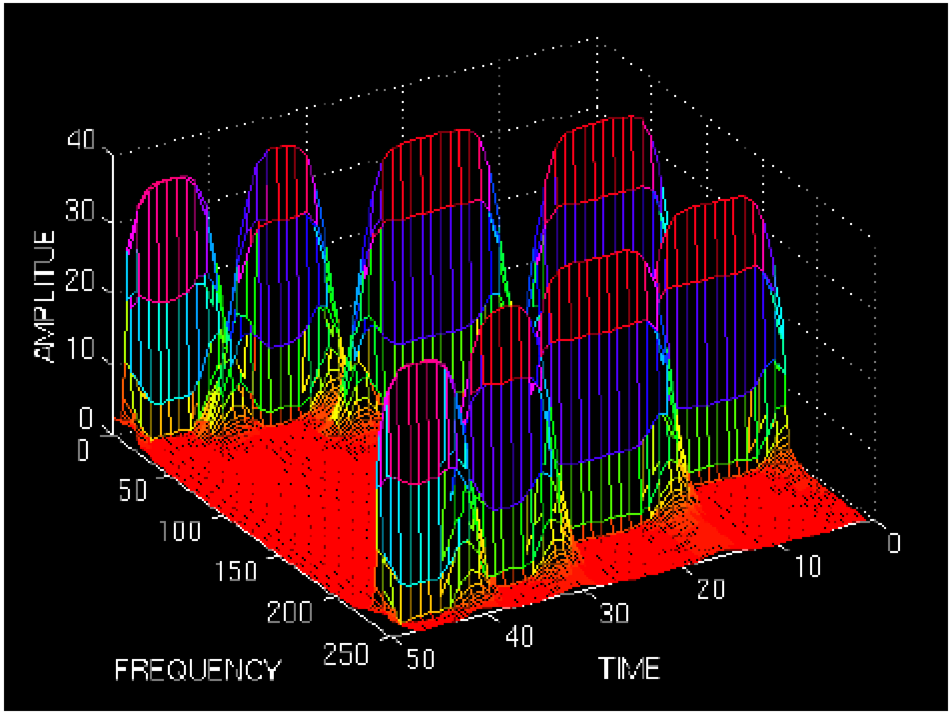


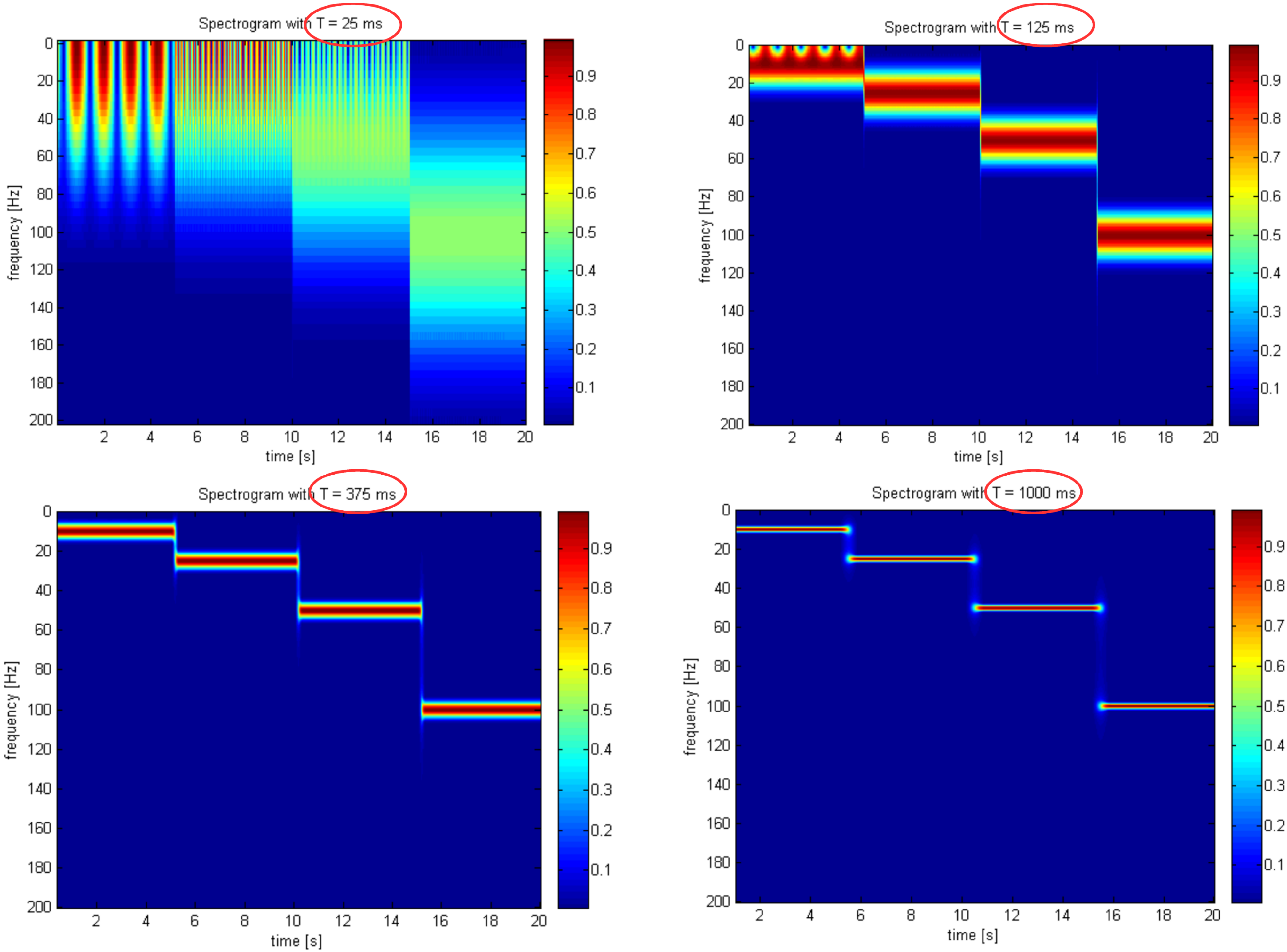
In this signal, there are four frequency components at different times. The interval 0 to 250 ms is a simple sinusoid of 300 Hz, and the other 250 ms intervals are sinusoids of 200 Hz, 100 Hz, and 50 Hz, respectively

The STFT →>

As expected, this is a two dimensional plot (3 dimensional, if you count the amplitude too). The "x" and "y" axes are time and frequency, respectively

Now we have a true time-frequency representation of the signal. We not only know what frequency components are present in the signal, but we also know where they are located in time.




Spectrogram with T = 25 ms
Spectrogram with T = 125 ms
Spectrogram with T = 375 ms
Spectrogram with T = 1000 ms
frequency [Hz]
time [s]
0 20 40 60 80 100 120 140 160 180 200
2 4 6 8 10 12 14 16 18 20
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9

In FT, the kernel function, allows us to obtain perfect frequency resolution, because the kernel itself is a window of infinite length. In STFT is window is of finite length, and we no longer have perfect frequency resolution. You may ask, why don't we make the length of the window in the STFT infinite, just like as it is in the FT, to get perfect frequency resolution? Well, than you loose all the time information, you basically end up with the FT instead of STFT.

To make a long story real short, we are faced with the following dilemma:

If we use a window of infinite length, we get the FT, which gives perfect frequency resolution, but no time information. Furthermore, in order to obtain the stationarity, we have to have a short enough window, in which the signal is stationary. The narrower we make the window, the better the time resolution, and better the assumption of stationarity, but poorer the frequency resolution:

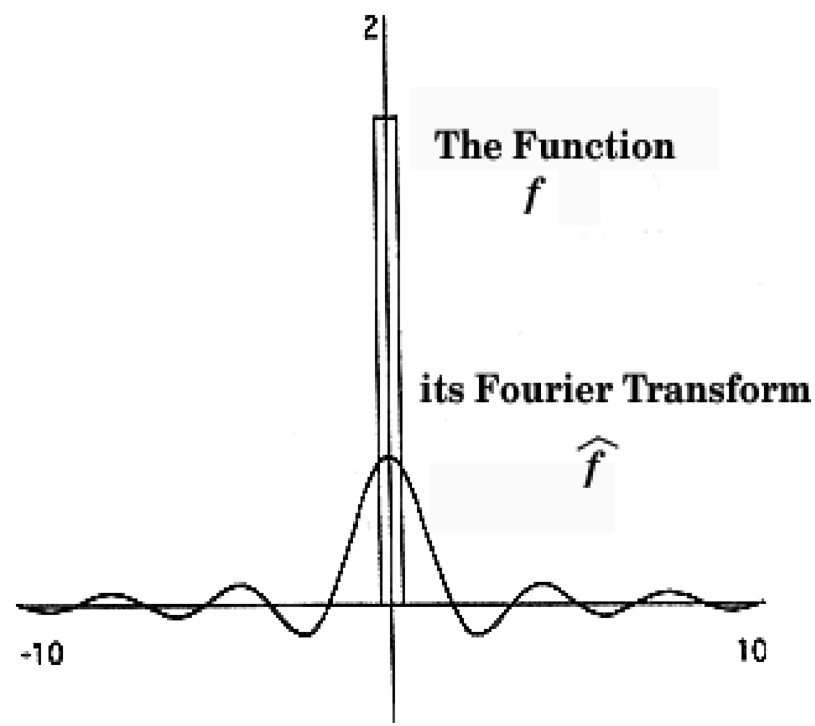
Narrow window ===>good time resolution, poor frequency resolution.

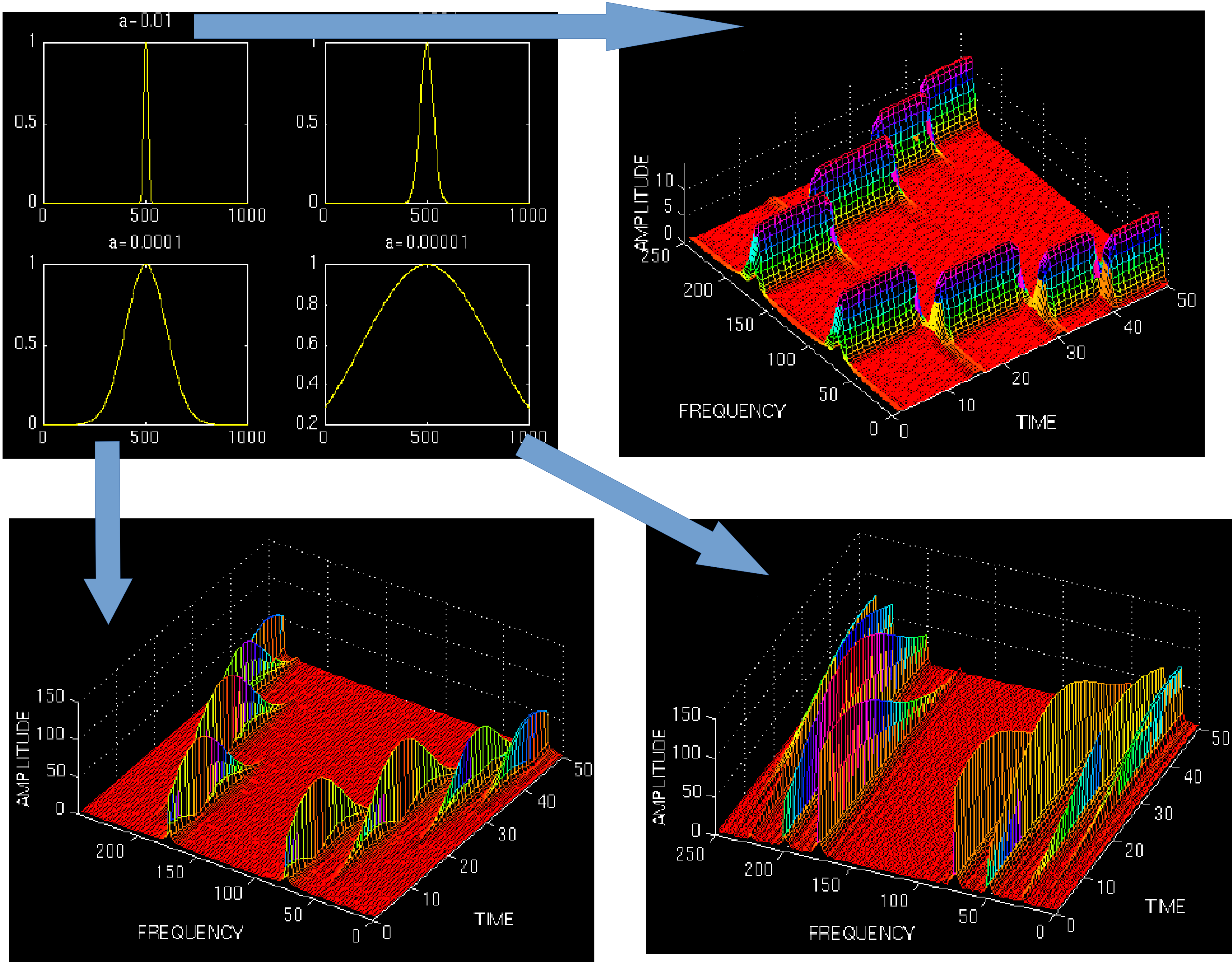
Wide window ===>good frequency resolution, poor time resolution.

Fourier: A single window is used for all frequencies

Let's examine the coverage in the time-frequency plane of a simple function, in this case, a sharply concentrated function around t=0.

If you want precise information about time, you have to accept vague information about frequency, if you want precise information about frequency, you have to accept vagueness about time. However we can be smarter about how we cover the time-frequency plane.



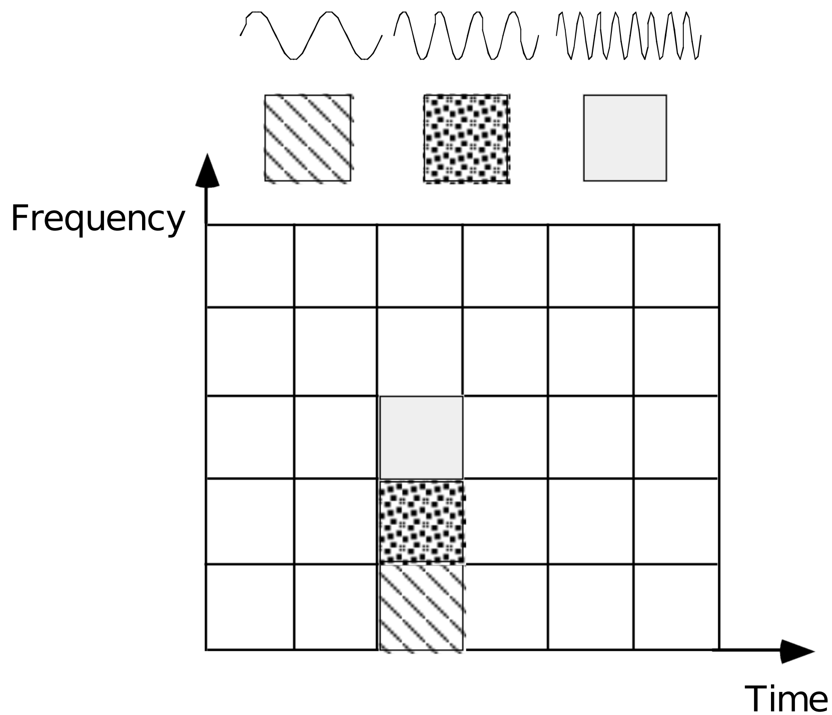


a- 0.01
a=0.0001
a=0.00001
AMPLITUDE
FREQUENCY
TIME
AMPLITUDE
FREQUENCY
TIME
AMPLITUDE
FREQUENCY
TIME

Fourier Time-Space

Fourier: A single window is used for all frequencies

The figure below shows a *windowed Fourier transform*, where the window is simply a square wave. The square wave window truncates the sine or cosine function to fit a window of a particular width. Because a single window is used for all frequencies in the windowed Fourier transform, the resolution of the analysis is the same at all locations in the time-frequency plane. With a windowed Fourier transform, a small window looks at short intervals of time at the cost of being vague about frequency. A big window is less precise about time, but more precise about frequency. But the width of the box (window) remains fixed.



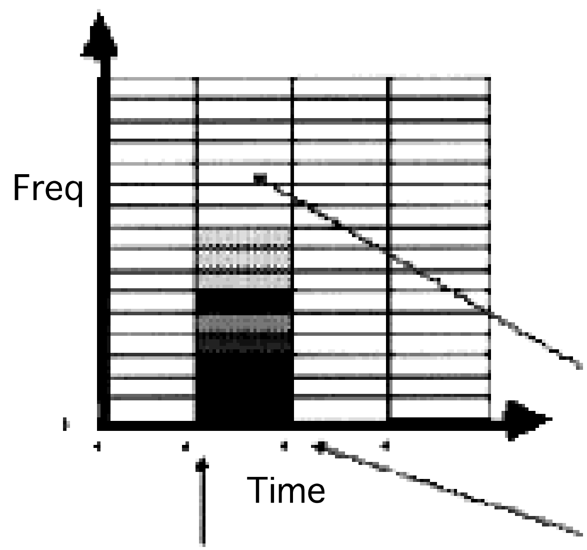
Fourier Time-Space

Fourier: A single window is used for all frequencies

In a windowed Fourier transform, it is the number of oscillations that varies. A small window is "blind" to low frequencies, which are too large for the window. But if one uses a large window, information about a brief change ("discontinuity") will be lost in the information concerning the entire interval corresponding to the window.

<u>Fourier Analysis of a sharp transient</u>

1) each graph represent one coefficient of a Fourier transform
2) darker shading represents a higher numerical magnitude for that coefficient
3) the equal area of each rectangle in both windows can be seen as an interpretation of the Heisenberg uncertainty principle.



There are many windows which have nonzero average values for relatively long periods of time

The click happened quickly in time here

...but the analysis still registers a reaction here, a long time after the click occurred.

THE CONTINUOUS WAVELET TRANSFORM

The continuous wavelet transform was developed as an alternative approach to the short time Fourier transform to overcome the resolution problem. The wavelet analysis is done in a similar way to the STFT analysis, in the sense that the signal is multiplied with a function, th*e wavelet*, similar to the window function in the STFT, and the transform is computed separately for different segments of the time-domain signal. However, there are two main differences between the STFT and the CWT:

1. The Fourier transforms of the windowed signals are not taken, and therefore single peak will be seen corresponding to a sinusoid, i.e., negative frequencies are not computed.

2. The width of the window is changed as the transform is computed for every single spectral component, which is probably the most significant characteristic of the wavelet transform.

Wavelet transform

$$Wf(s,u) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt. \tag{4.11}$$

Wavelet transform overcomes the previous problem. The wavelet function is designed to strike a balance between time domain (finite length) and frequency domain (finite bandwidth). As we dilate and translate the mother wavelet, we can see very low frequency components at large s while very high frequency component can be located precisely at small s.

The continuous wavelet transform is reversible if $C\psi<\infty$, even though the basis functions are in general may not be orthonormal.

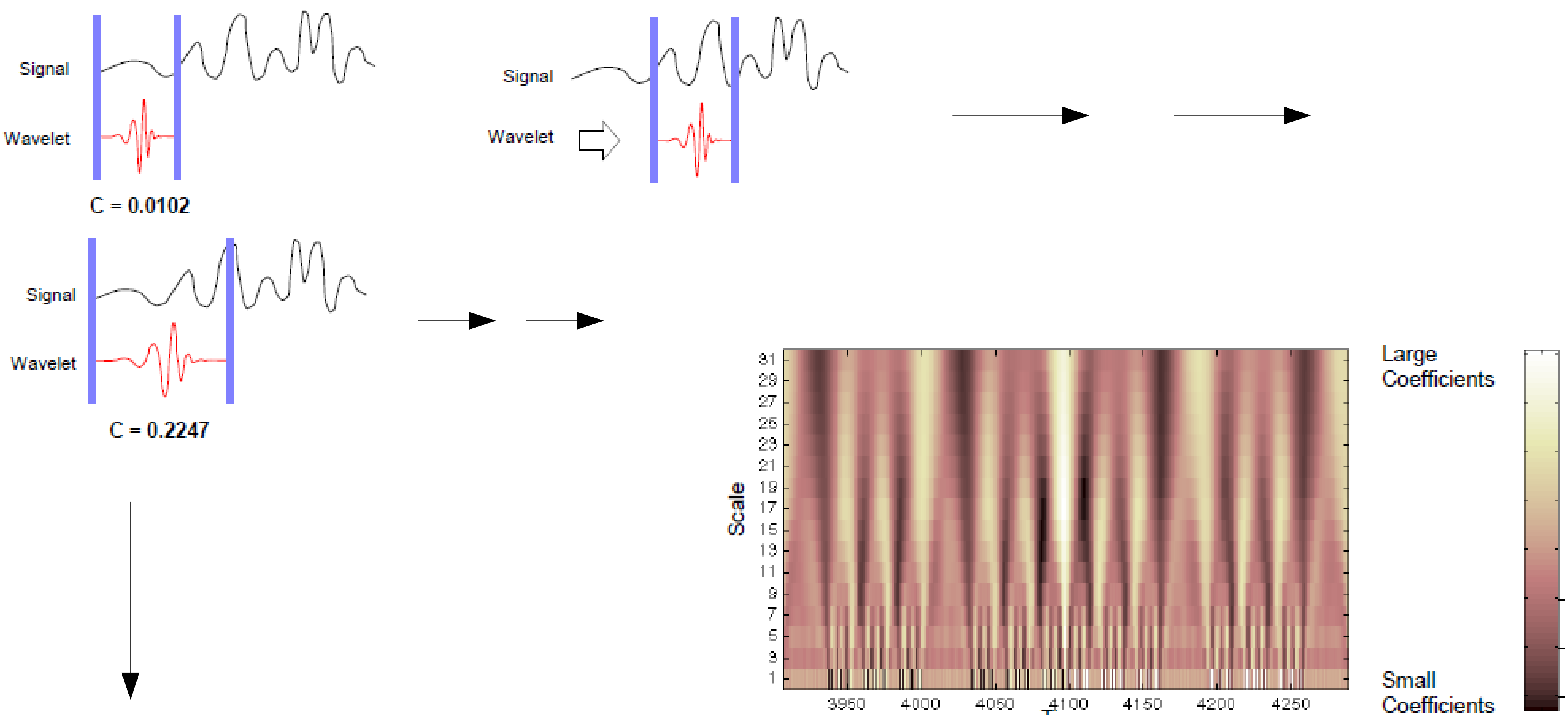
$$f(t) = \frac{1}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} Wf(s,u) \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) du \frac{ds}{s^2},$$

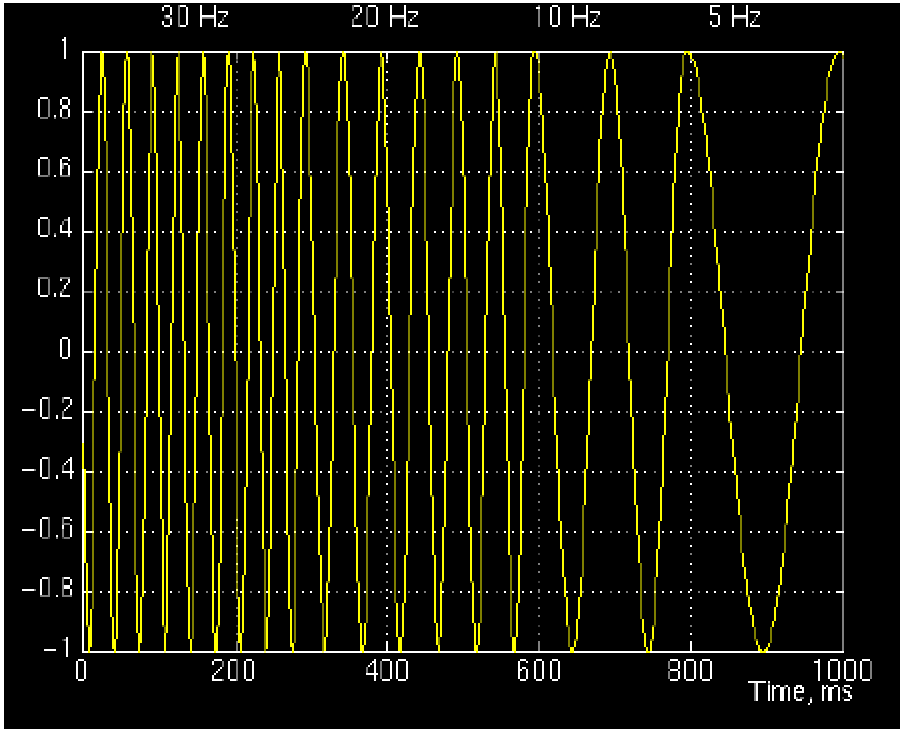
$$C_\psi = \int_0^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty.$$

Wavelet transform

$$Wf(s,u) = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{s}}\psi^*\left(\frac{t-u}{s}\right)dt. \tag{4.11}$$

Wavelet transform overcomes the previous problem. The wavelet function is designed to strike a balance between time domain (finite length) and frequency domain (finite bandwidth). As we dilate and translate the mother wavelet, we can see very low frequency components at large s while very high frequency component can be located precisely at small s.

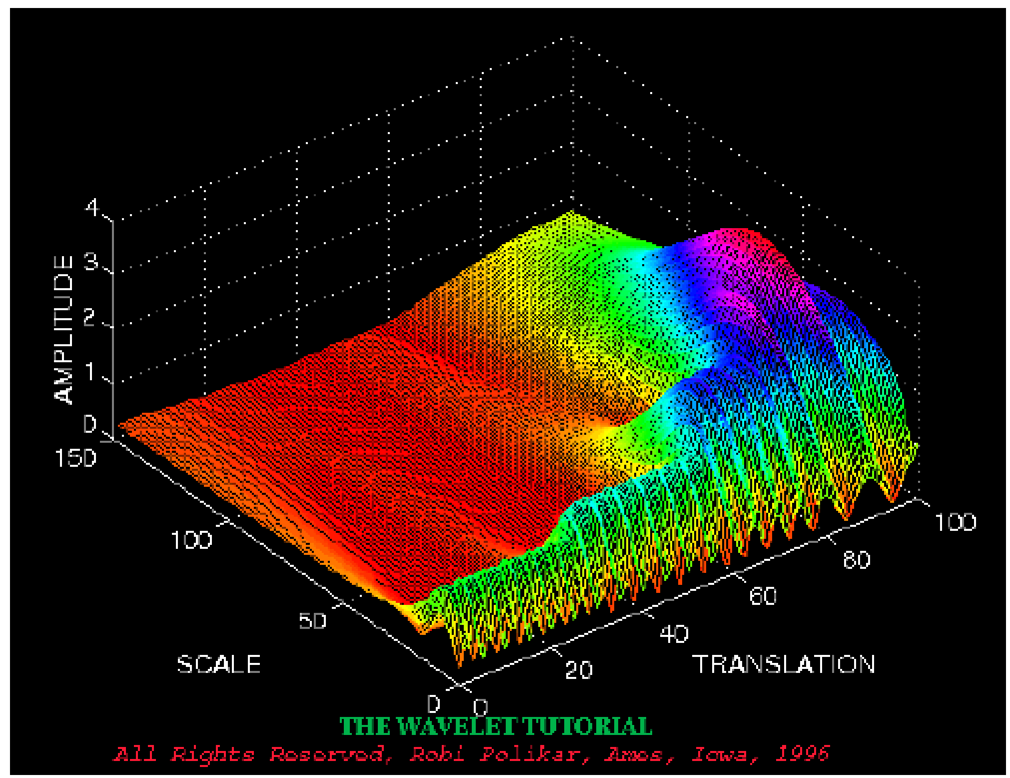




- Low scale $a \Rightarrow$ Compressed wavelet $\Rightarrow$ Rapidly changing details $\Rightarrow$ High frequency ω.
- High scale $a \Rightarrow$ Stretched wavelet $\Rightarrow$ Slowly changing, coarse features $\Rightarrow$ Low frequency ω.

Note that that **smaller scales** correspond to **higher frequencies**, i.e., frequency decreases as scale increases, therefore, that portion of the graph with scales around zero, actually correspond to highest frequencies in the analysis, and that with high scales correspond to lowest frequencies.

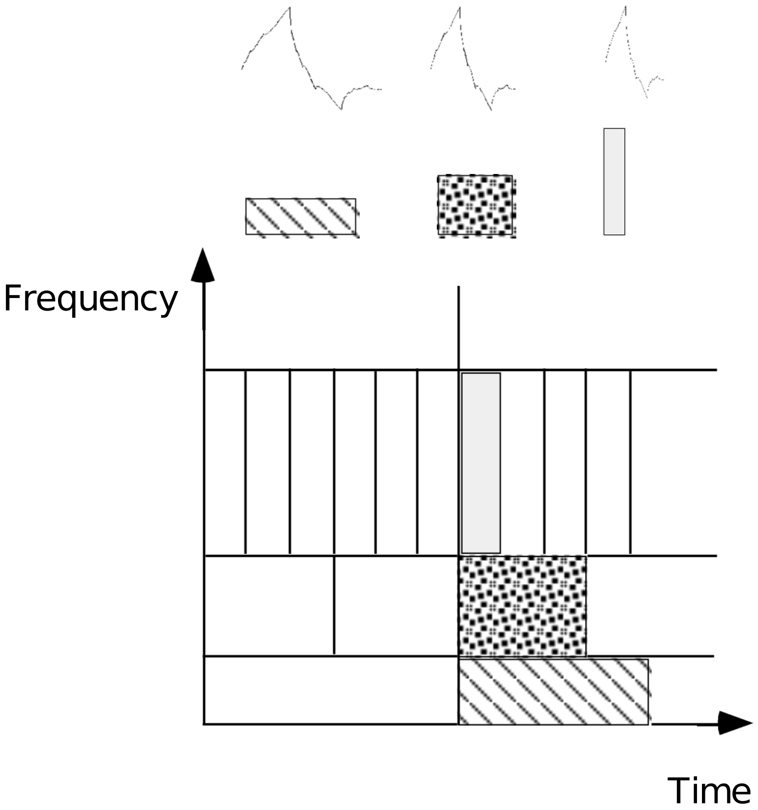
Remember that the signal had 30 Hz (highest frequency) components first, and this appears at the lowest scale at a translations of 0 to 30. Then comes the 20 Hz component, second highest frequency, and so on. The 5 Hz component appears at the end of the translation axis (as expected), and at higher scales (lower frequencies) again as expected.



Wavelet Time-Space

Wavelets: The windows vary for different frequencies.

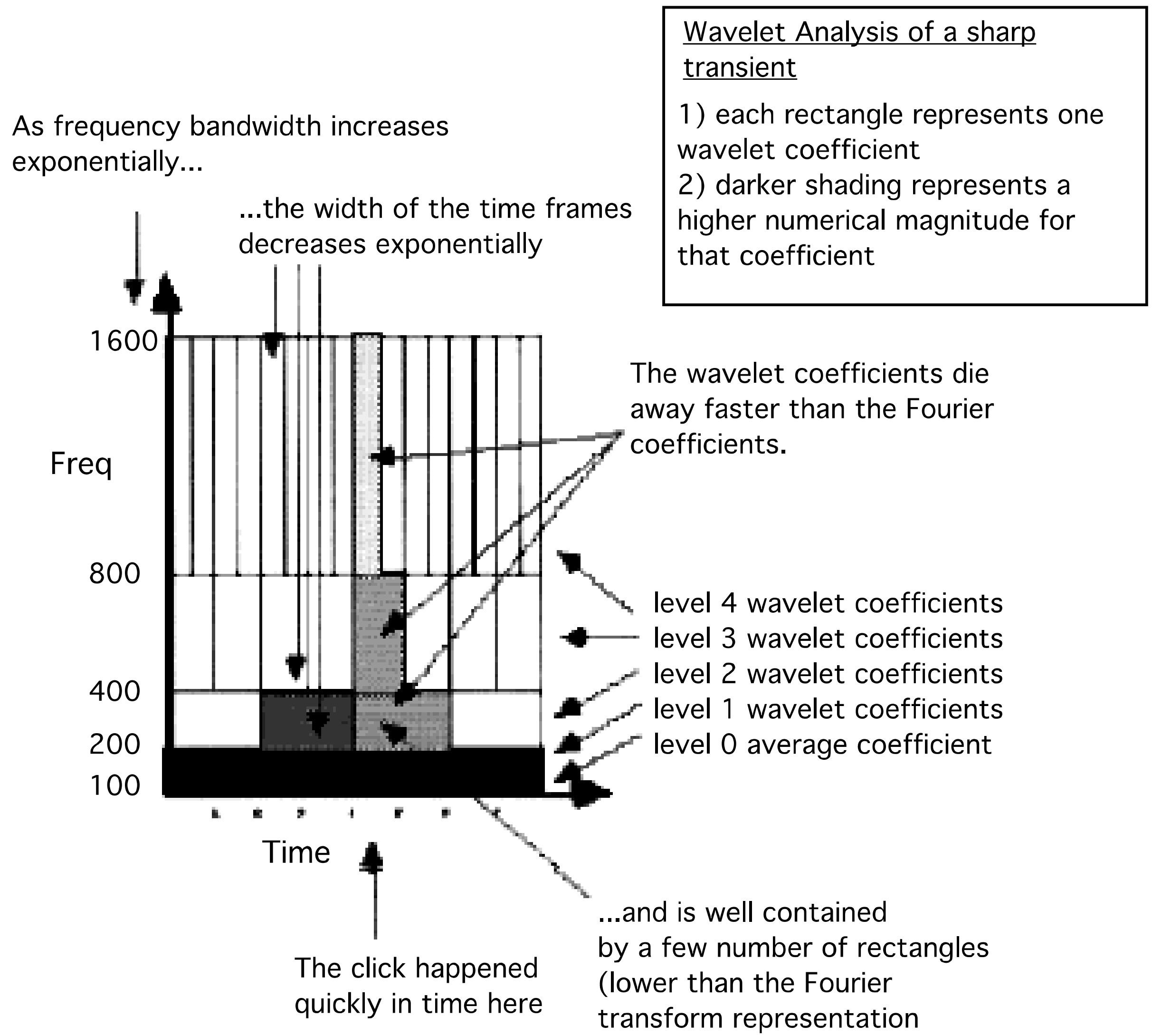
In order to isolate signal discontinuities, one would like to have some very short basis functions. At the same time, in order to obtain detailed analysis for low frequencies, one would like to have some very long basis functions. A way to achieve this is to have short high-frequency basis functions and long low-frequency functions.



This happy medium is exactly what you get with wavelet transforms.

Wavelet Time-Space

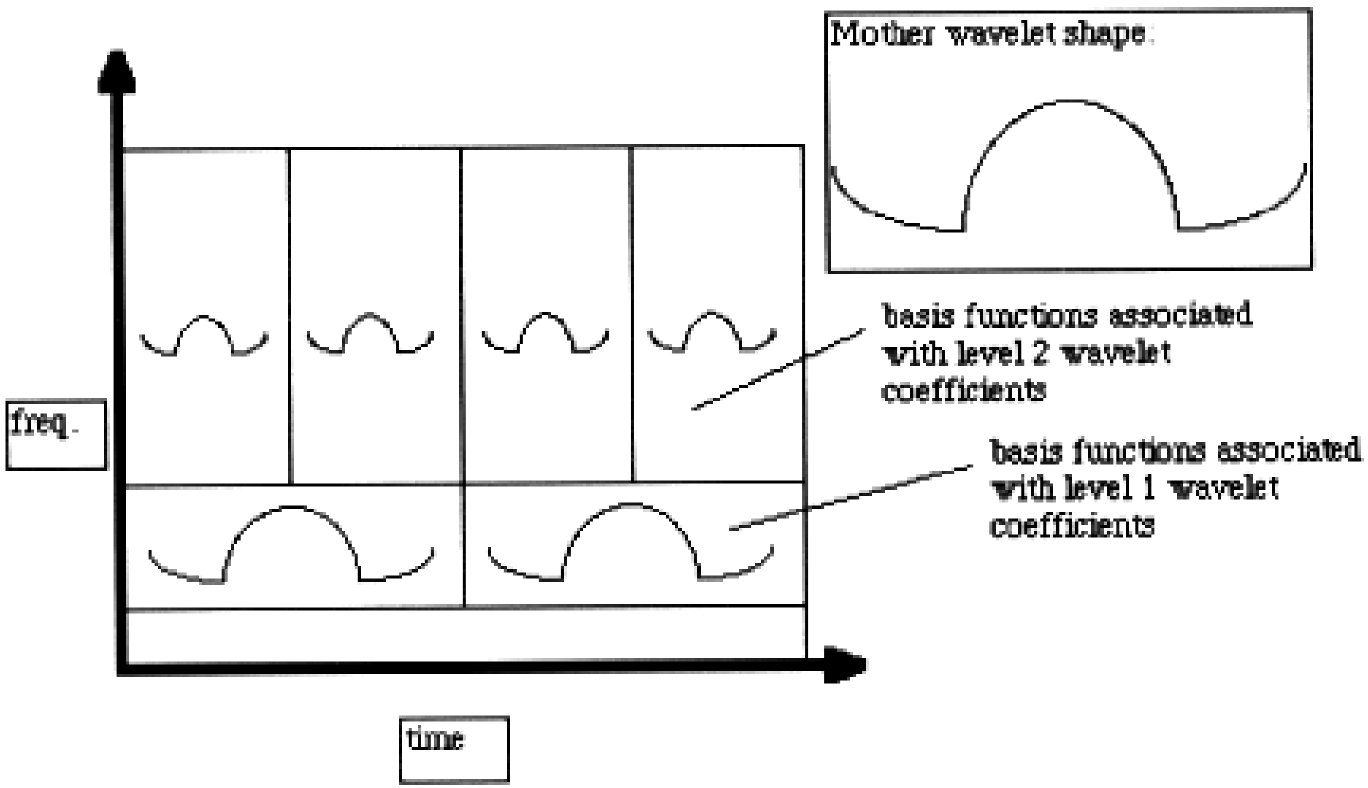
Wavelet analysis divides the time-frequency plane in a non-uniform manner. Frequency resolution is finer than time resolution at low frequencies, while time resolution is finer than frequency resolution at higher frequencies.



Wavelet Time-Space

Wavelets: The windows vary for different frequencies.

In a wavelet transform, a **Mother wavelet** is stretched or compressed to change the size of the window. This makes it possible to analyze a signal at different scales.



The wavelet transform is sometimes called a "mathematical microscope": big wavelets give an approximate image of the signal, while smaller and smaller wavelets zoom in on small details.

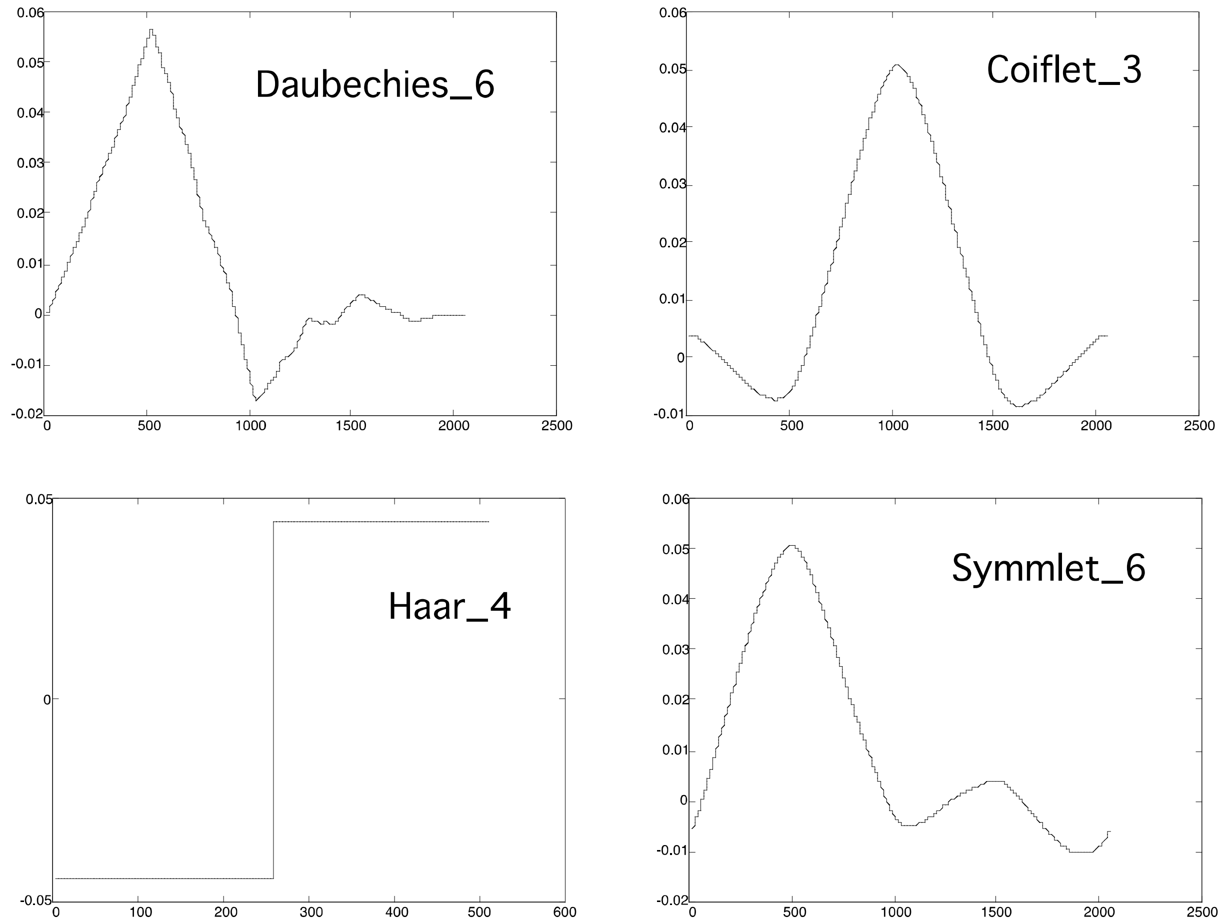
Some Wavelets

Wavelets comprise an infinite set. The different wavelet families make different trade-offs between how compactly the basis functions are localized in space and how smooth they are.

Wavelets are classified within a family most often by the *number of vanishing moments*.

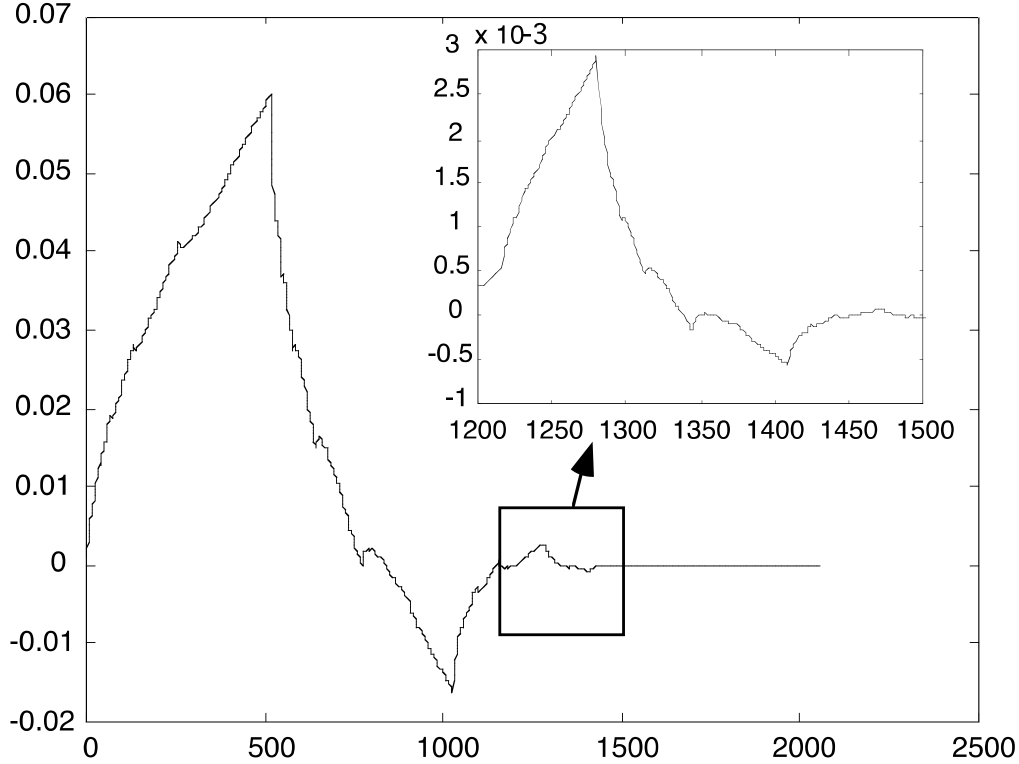
Haar -- the first wavelet; a square-wave wavelet
Daubechies -- the first continuous, compactly supported orthonormal wavelet family
Coiflet -- orthonormal wavelets system where both father and mother have special vanishing moments properties
Symmlet -- smooth orthogonal wavelet of compact support with 6 vanishing moments.

What do some wavelet look like?



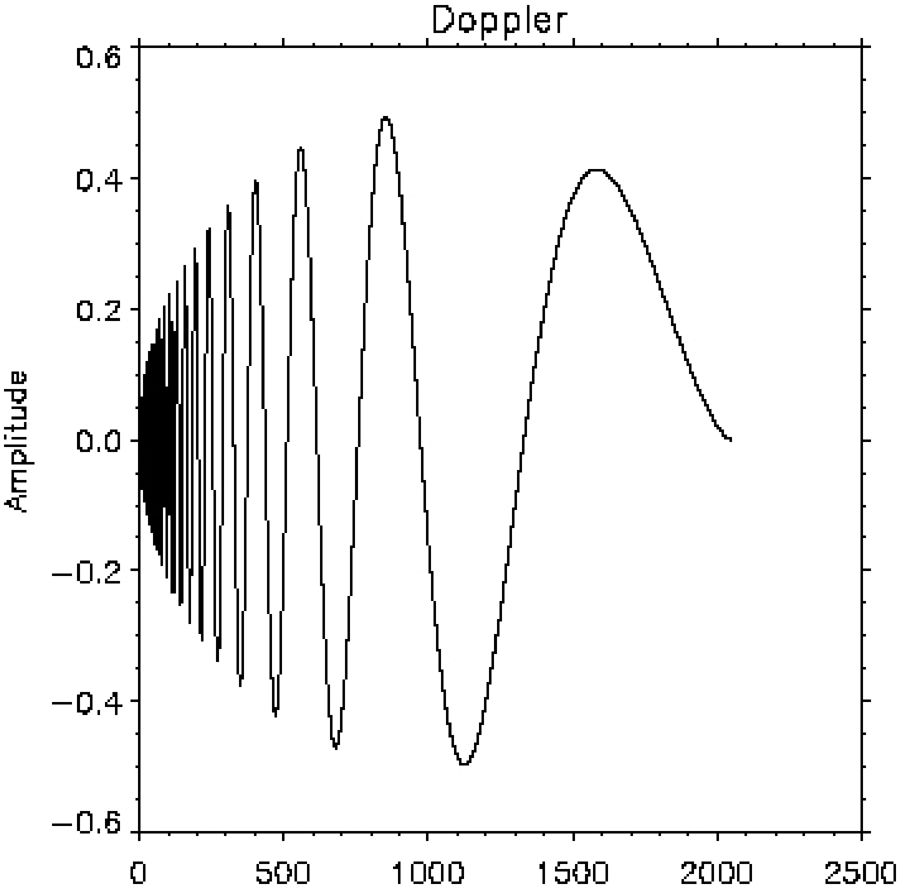
Some Wavelets

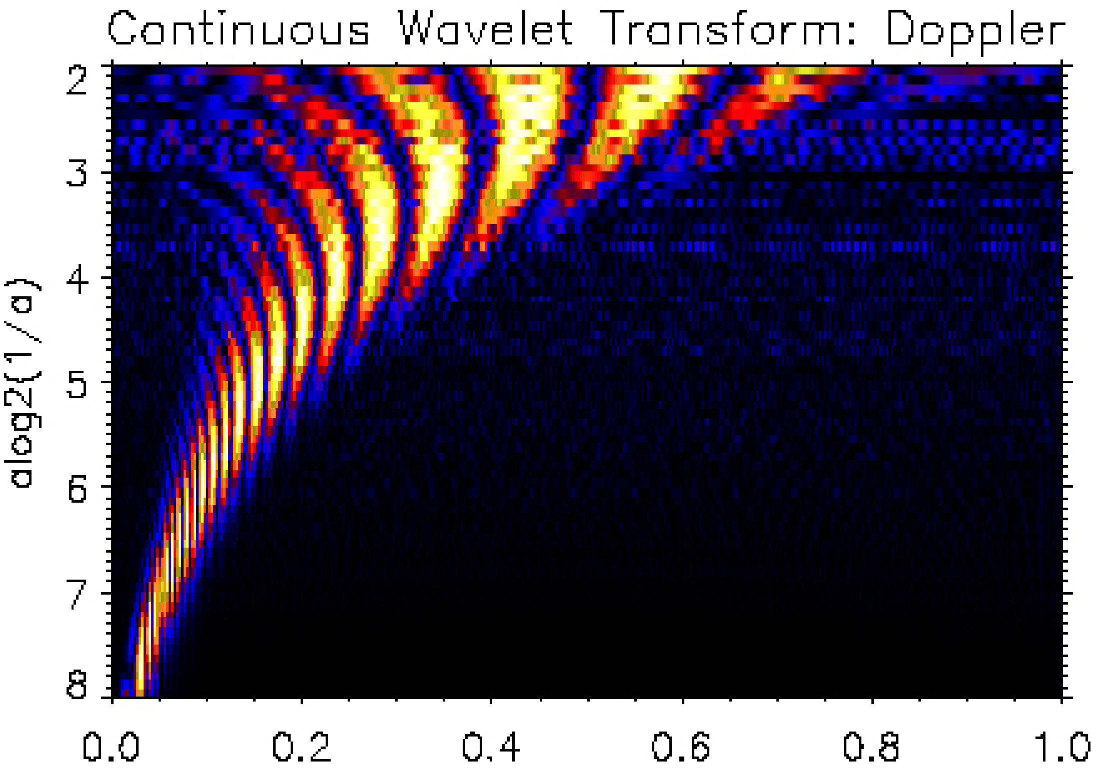
Some of the wavelets have fractal structure. The Daubechies wavelet family is one example:



Continuous Wavelet Transform

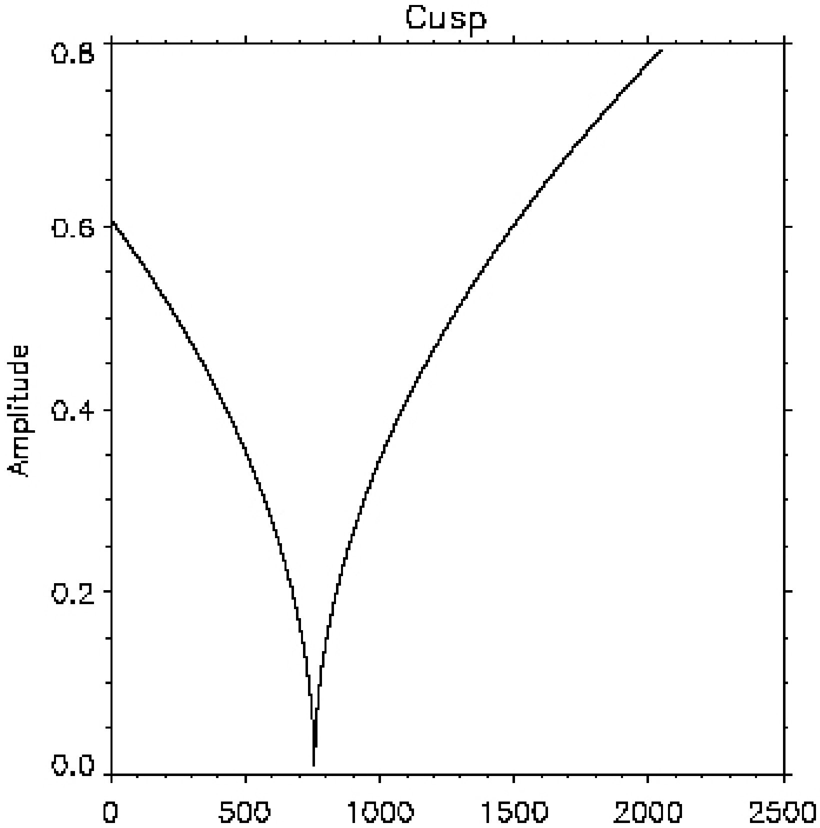
Doppler Signal

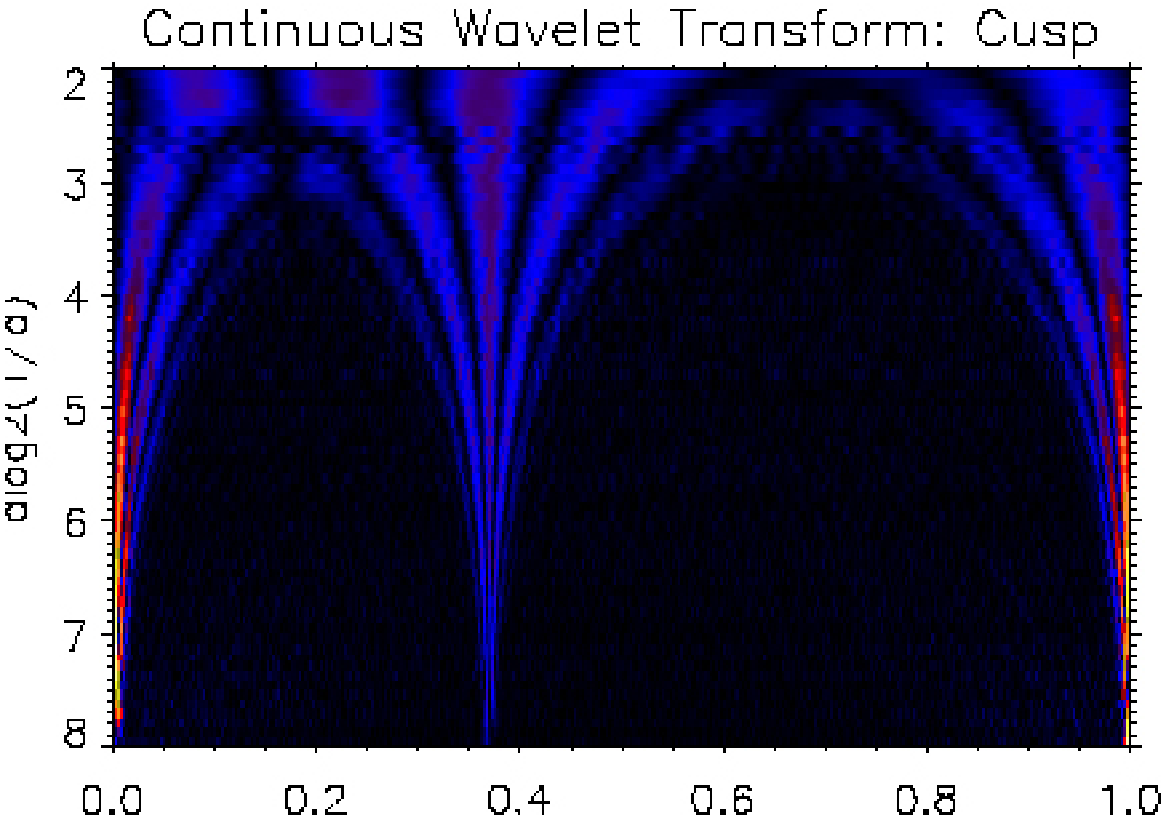




Continuous Wavelet Transform

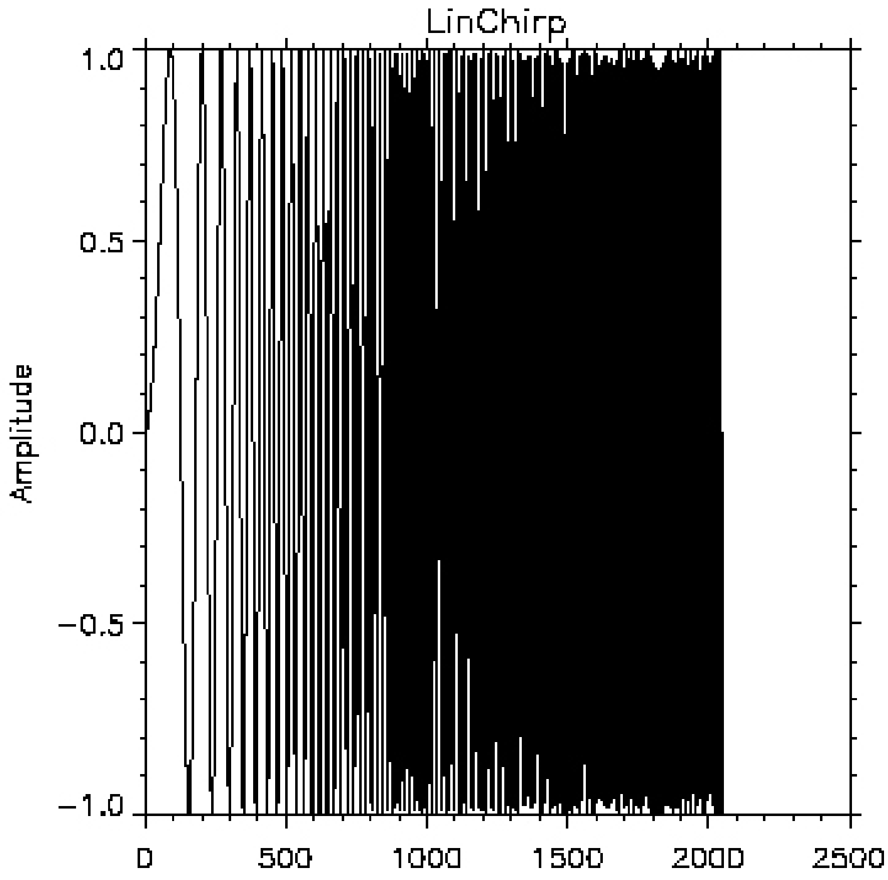
Cusp Signal

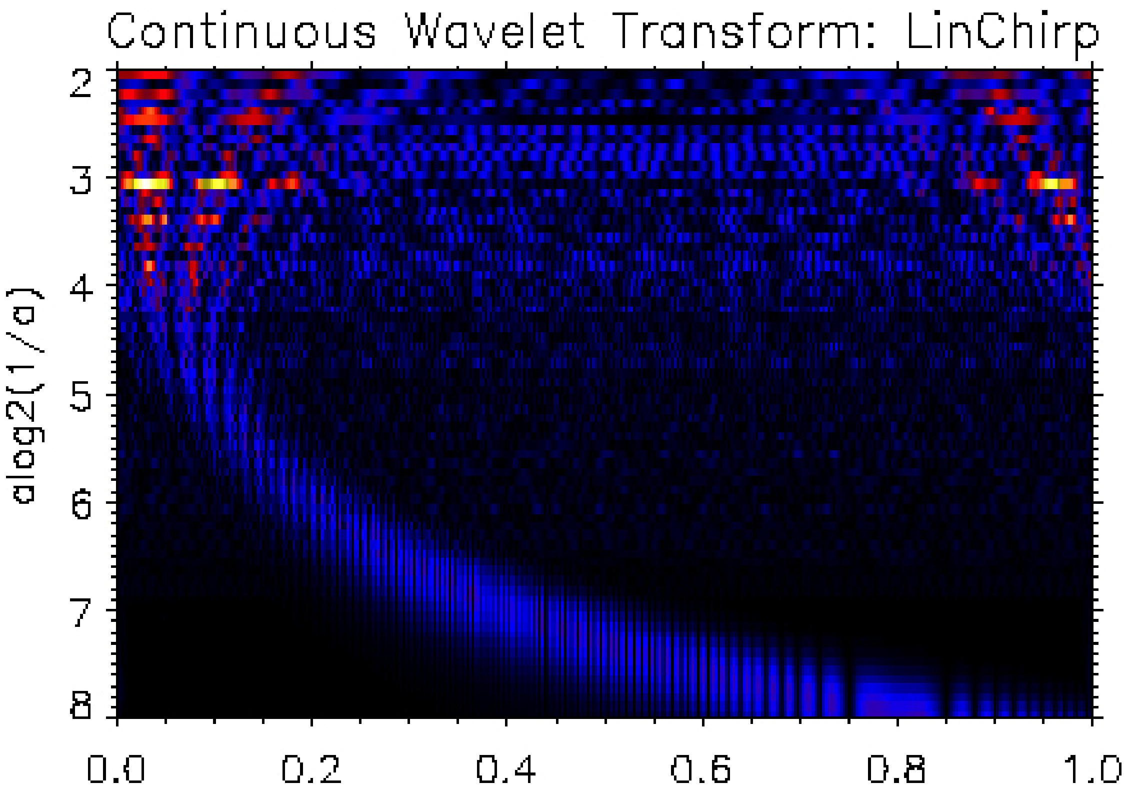




Continuous Wavelet Transform

Linear Chirp Signal





Wavelet Shifting and Scaling

Dilations and translations of the $\Psi(x)$ **Mother function**, or **analyzing wavelet** define an orthogonal basis, our wavelet basis:

$$\Psi_{(s,l)}(x) = 2^{-s/2}\Psi\left(2^{-s}x - l\right)$$

The variables s and l are integers that scale and dilate the mother function to generate wavelets, such as a Daubechies wavelet family. The scale index $\boldsymbol{s}$ indicates the wavelet's width, and the location index $\boldsymbol{l}$ gives its position. Notice that the mother functions are rescaled, or "dilated" by powers of two, and translated by integers. What makes wavelet bases especially interesting is the self-similarity caused by the scales and dilations. Once we know about the mother functions, we know everything about the basis.

To span our data domain at different resolutions, the mother wavelet is used in a scaling equation:

$$W(x) = \sum_{k=-1}^{N-2} (-1)^k c_{k+1} \Psi(2x + k)$$

where ***W(x)*** is the scaling function (or **Father function**) for the mother function and ***c_k*** are the wavelet coefficients.

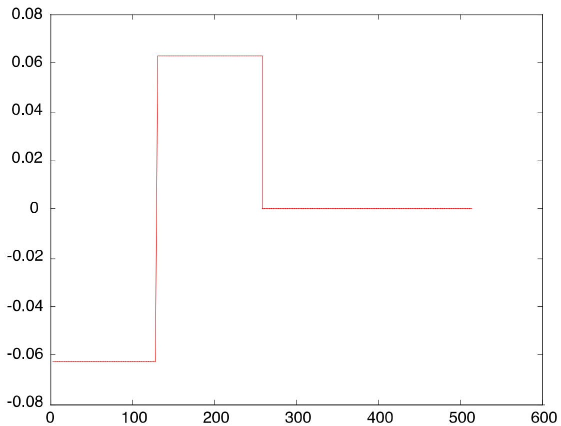
$$\sum_{k=0}^{N-1} c_k = 2 \qquad \sum_{k=0}^{N-1} c_k c_l = 2\delta_{l,0}$$

where ***d*** is the delta function and ***l*** is the location index.

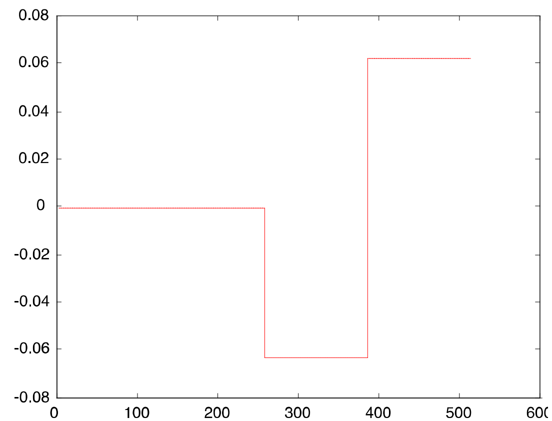
Wavelet Shifting and Scaling

"Wavelets comprise an infinite set of scale-varying, translated basis functions."

First Level scaled

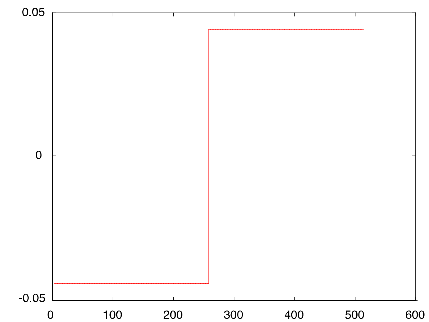


First Level scaled and shifted

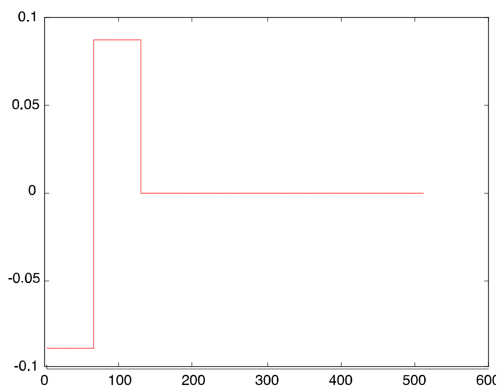


Haar Wavelet:

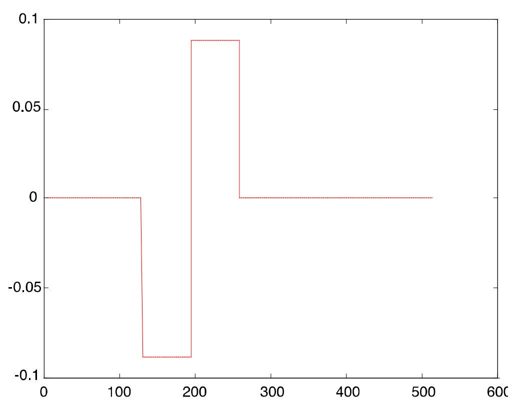
Mother Wavelet



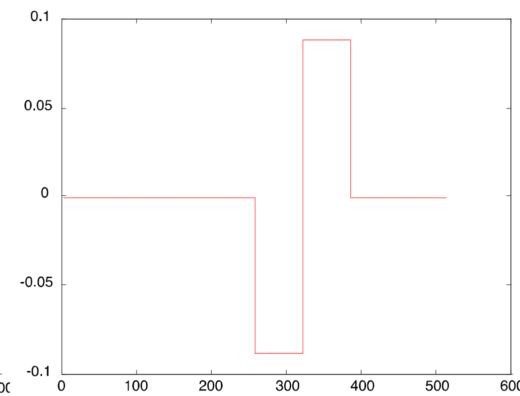
Second Level scaled



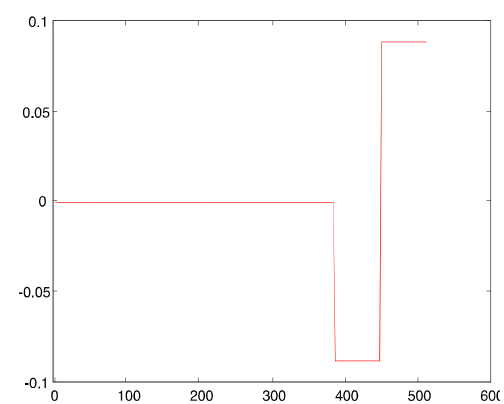
Second Level scaled and shifted



Second Level scaled and shifted



Second Level scaled and shifted



Discretization of the Continuous Wavelet Transform: The Wavelet Series

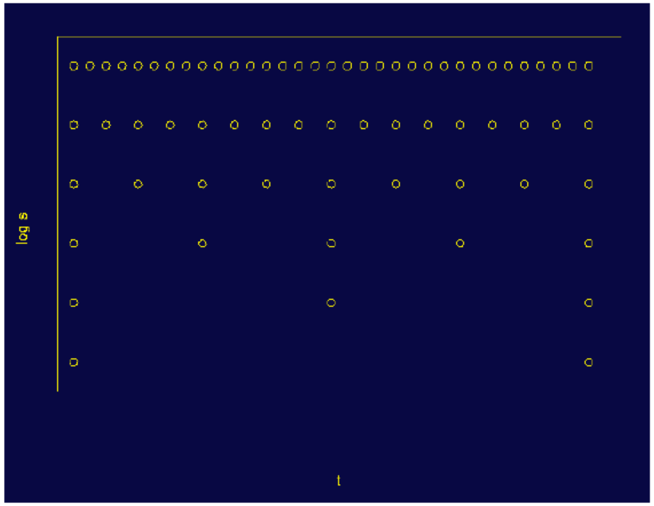
At higher scales (lower frequencies), the sampling rate can be decreased, according to Nyquist's rule. In other words, if the time-scale plane needs to be sampled with a sampling rate of N_1 at scale s_1, the same plane can be sampled with a sampling rate of N_2, at scale s_2, where, s_1 < s_2 (corresponding to frequencies f1>f2) and N_2 < N_1. The actual relationship between N_1 and N_2 is

$$N_2 = \frac{s_1}{s_2} N_1$$

In other words, at lower frequencies the sampling rate can be decreased which will save a considerable amount of computation time.

As mentioned earlier, the wavelet $\psi(\tau,s)$ satisfying $C\psi<\infty$, allows reconstruction of the signal. However, this is true for the continuous transform.
The question is: can we still reconstruct the signal if we discretize the time and scale parameters?
The answer is ``yes", under certain conditions (as they always say in commercials: certain restrictions apply!!!).

The scale parameter s is discretized first on a logarithmic grid. The time parameter is then discretized with respect to the scale parameter, i.e., a different sampling rate is used for every scale.
In other words, the sampling is done on the dyadic sampling grid:

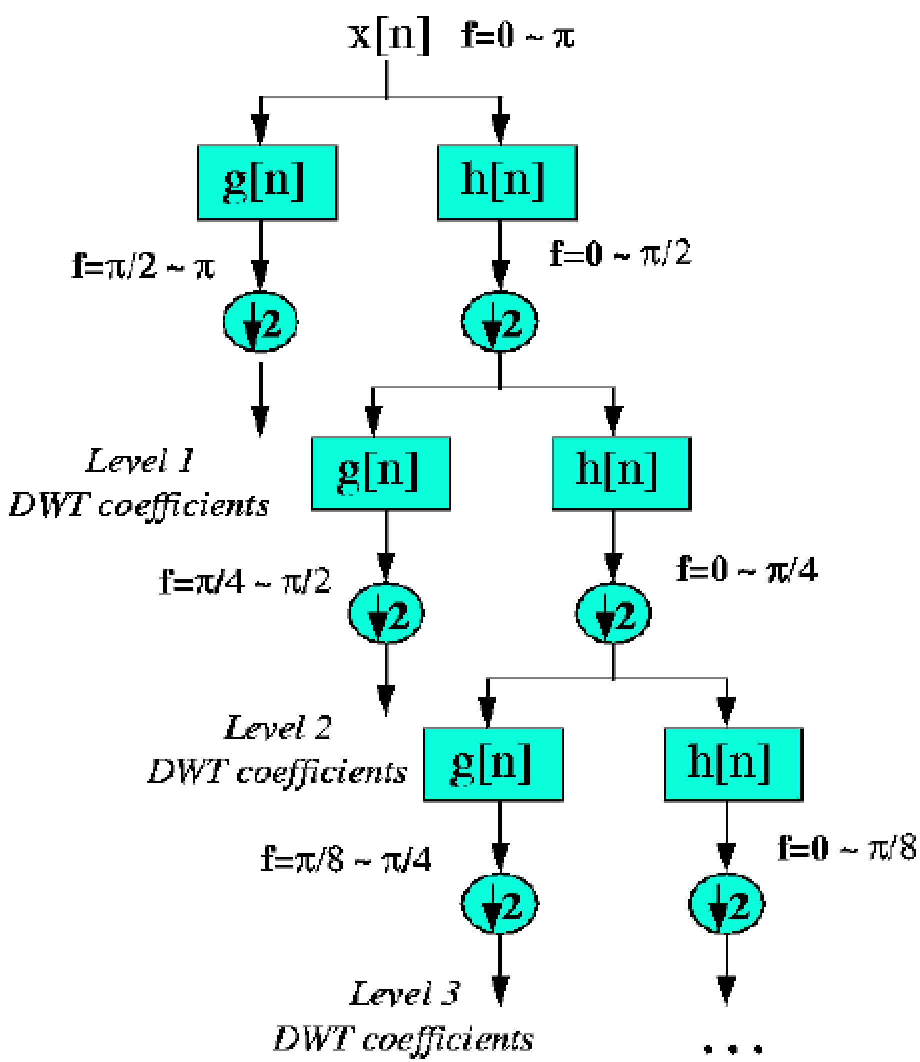


MULTIRESOLUTION ANALYSIS:
THE DISCRETE WAVELET TRANSFORM

Although the discretized continuous wavelet transform enables the computation of the continuous wavelet transform by computers, it is not a true discrete transform.
As a matter of fact, the wavelet series is simply a sampled version of the CWT, and the information it provides is highly redundant as far as the reconstruction of the signal is concerned.
This **redundancy**, on the other hand, requires a significant amount of computation time and resources.

The discrete wavelet transform (DWT), on the other hand, provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time.
The DWT is considerably easier to implement when compared to the CWT

The DWT analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information.
DWT employs two sets of functions, called **scaling functions** and **wavelet functions**, which are associated with **low pass** and **highpass** filters, respectively.
The decomposition of the signal into different frequency bands is simply obtained by successive highpass and lowpass filtering of the time domain signal. The original signal x[n] is first passed through a halfband highpass filter g[n] and a lowpass filter h[n].
After the filtering, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency of p/2 radians instead of p. The signal can therefore be subsampled by 2, simply by discarding every other sample.
The above procedure, which is also known as the subband coding, can be repeated for further decomposition.
At every level, the filtering and subsampling will result in half the number of samples (and hence half the time resolution) and half the frequency band spanned (and hence double the frequency resolution).

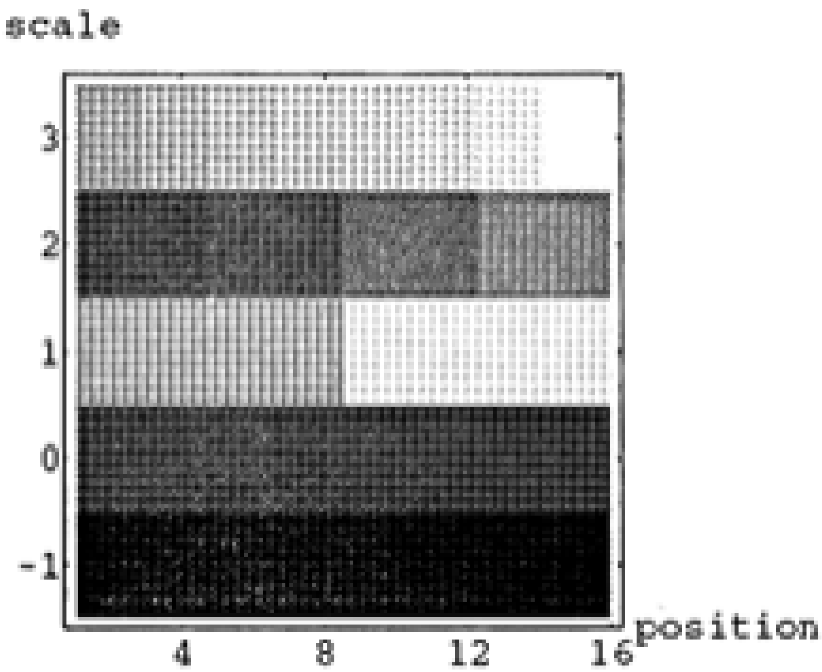

x[n] f=0 ~ π
g[n]
h[n]
f=π/2 ~ π
f=0 ~ π/2
↓2
↓2
Level 1
DWT coefficients
g[n]
h[n]
f=π/4 ~ π/2
f=0 ~ π/4
↓2
↓2
Level 2
DWT coefficients
g[n]
h[n]
f=π/8 ~ π/4
f=0 ~ π/8
↓2
↓2
Level 3
DWT coefficients
. . .

Spectral Density Representation

The **discrete wavelet transform** of a vector of length 2^n is another vector of length 2^n. The output contains both scale (frequency) and spatial information.

Represent the wavelet transformation of a one dimensional vector as a two-dimensional plane, with one axis corresponding to scale and one to spatial location:

Spectral Density Plot (SDP)



Scale "N":

The spectral density plot shows the wavelet coefficients, with the gray scale representing the magnitude of the wavelet coefficient at a given position and scale.

Scale "-1" corresponds to the Father wavelet, it is the scalar product of the sampled function (signal) with the the Father wavelet (sometimes called the "DC" term).

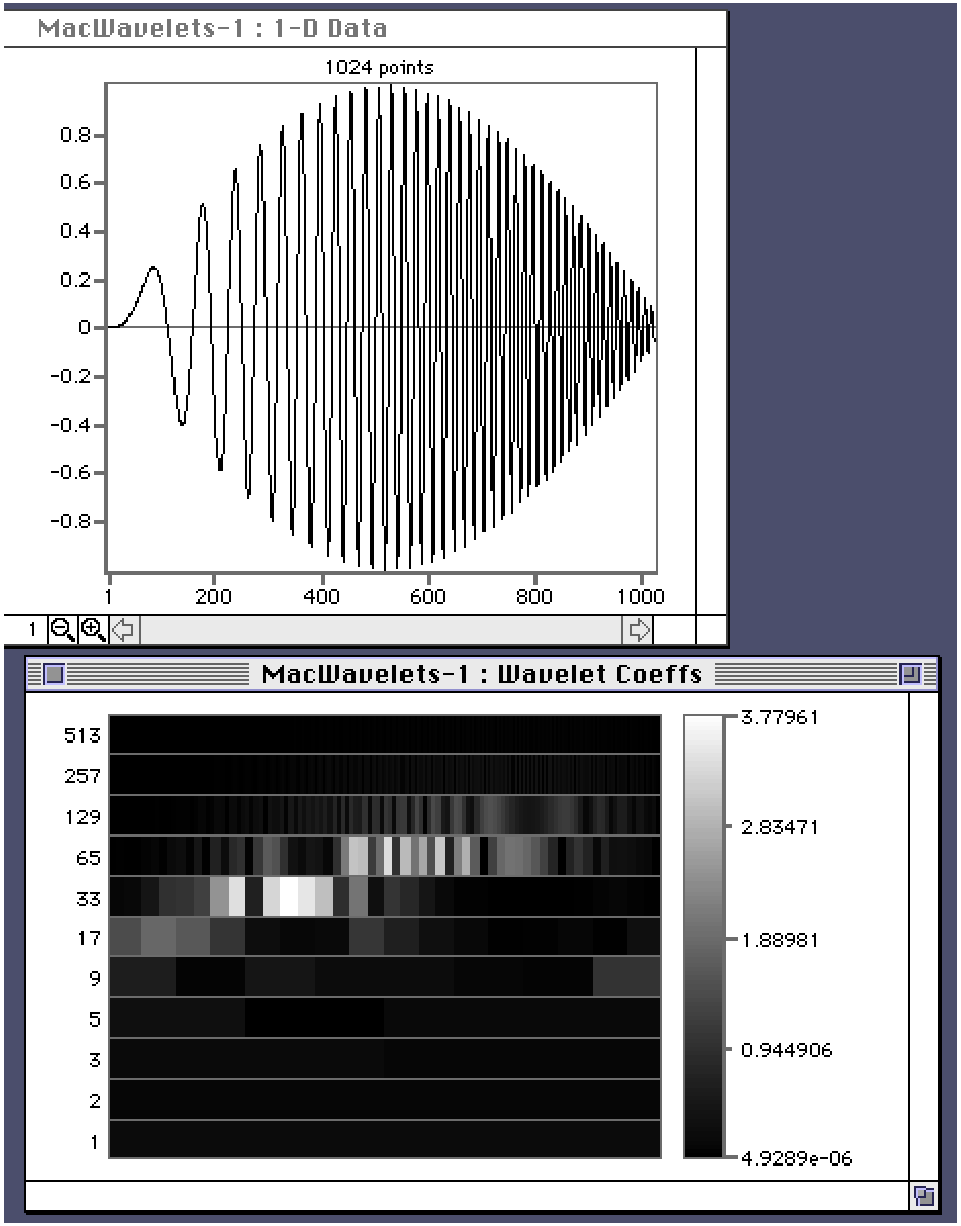
Scale "0" corresponds to the Mother wavelet, it is the scalar product of the sampled function with the Mother wavelet.

At Scale "1", the two allowable basis functions are wavelets constructed by contracting the Mother wavelet by a factor of 2.

This process continues to the finest level, where each wavelet will be defined over 2^{n-1} points along the spatial axis.

Wavelet Decomposition of Functions

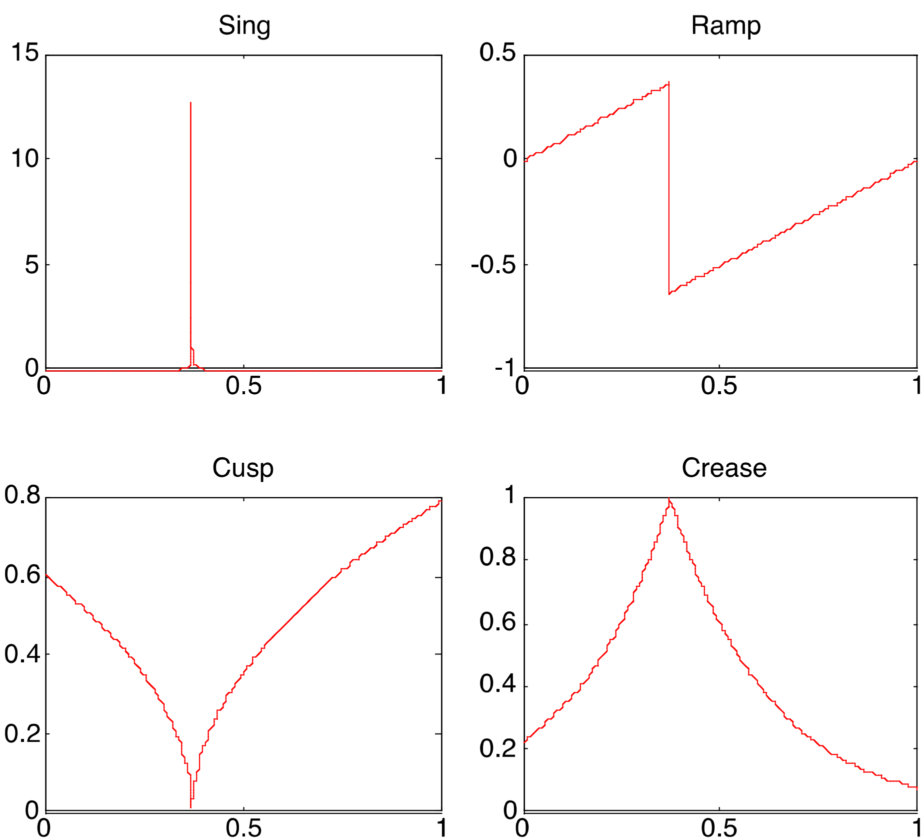
Example function: a **chirp function**; A sine decomposed in discrete wavelet transform using a Daubechies-4 filter. [from MacWavelets]

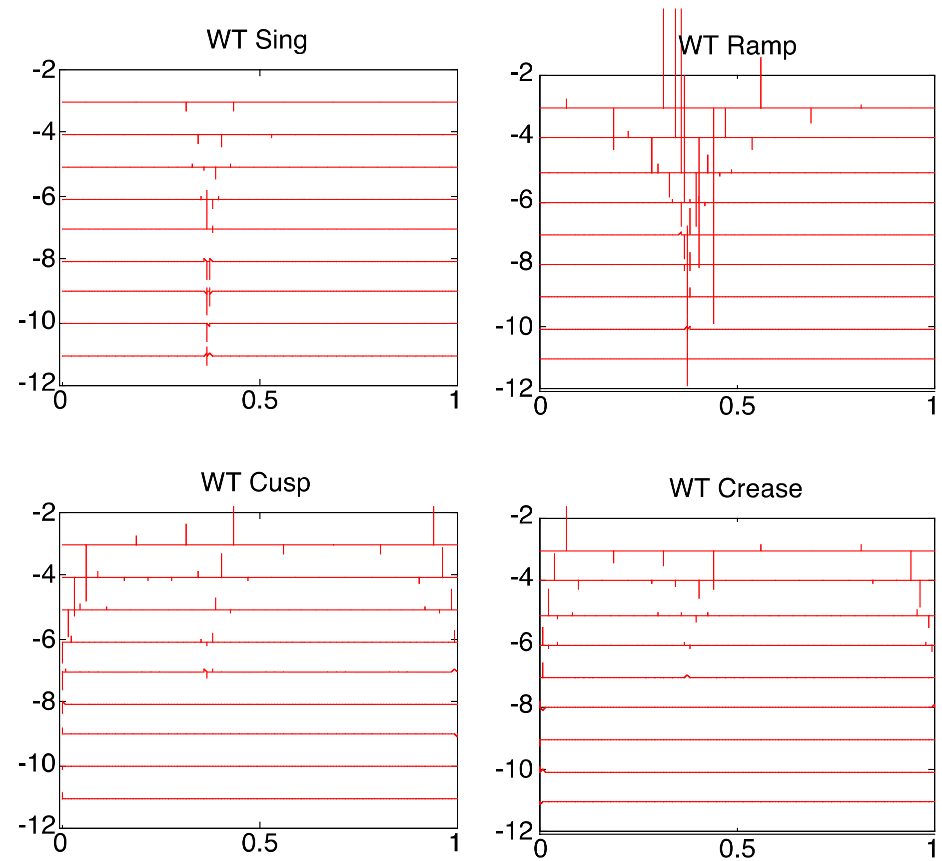


Wavelet Decomposition of Functions

Wavelet Decomposition of Singularities.
[from Wavelab]

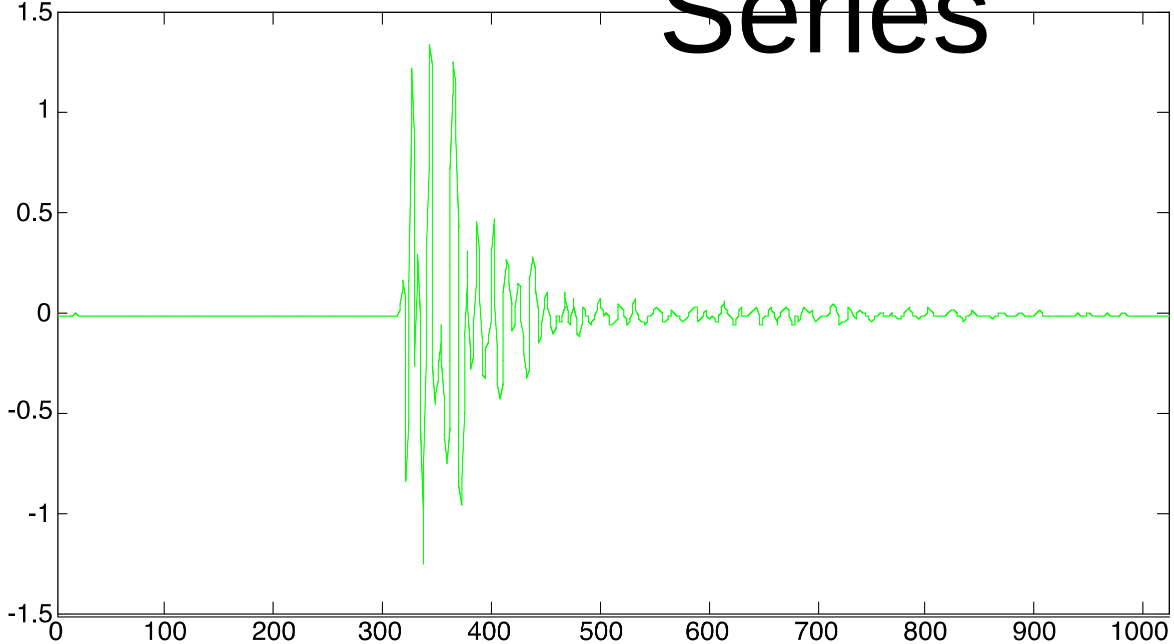
The filter used for these decompositions was a Coiflet 3 wavelet.





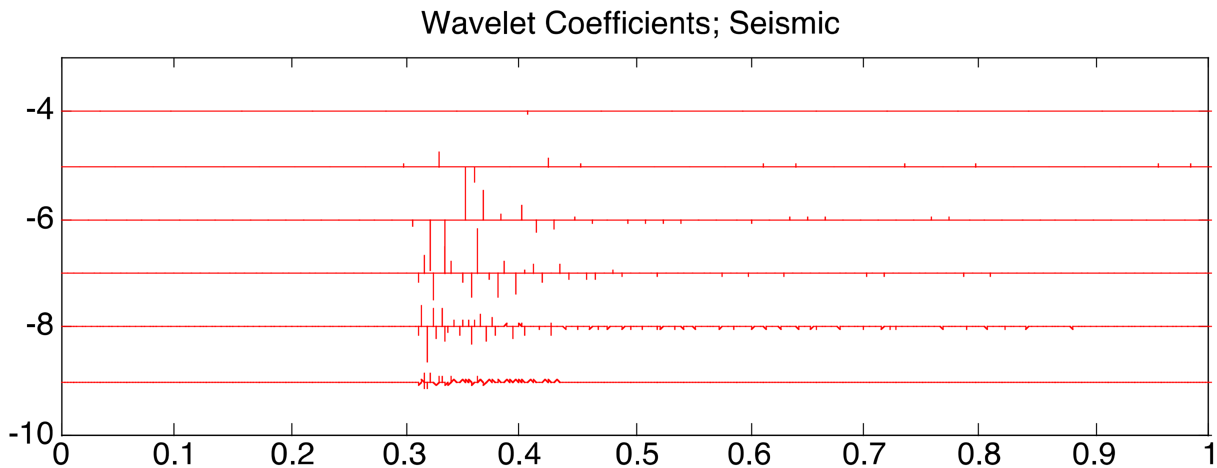
Notice that away from the singularities, the wavelet coefficients decay rapidly.

Wavelet Decomposition of Time Series



Seismic data.

[from Wavelab]



The filter used for these decompositions was a Daub-4 wavelet.

Wavelet Decomposition of Images

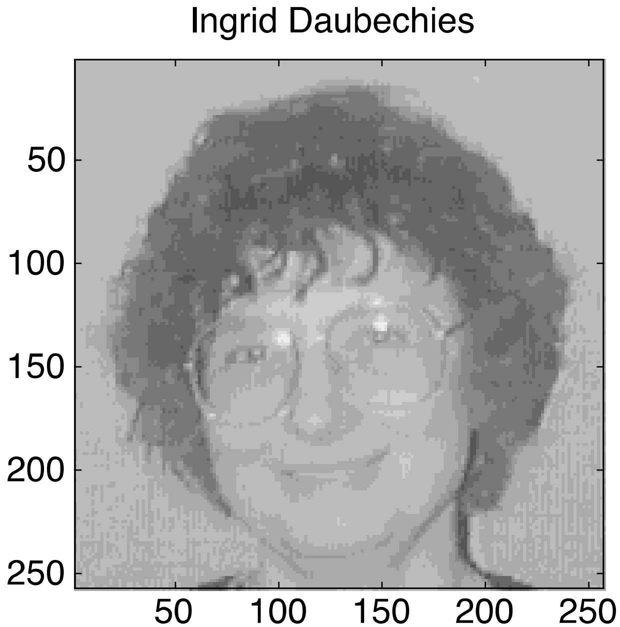
The discrete wavelet transform (DWT) can be applied to 2-D and greater data-sets.

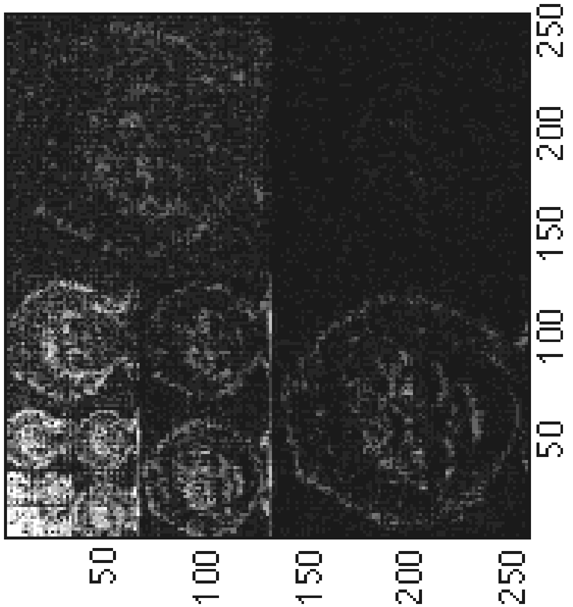
The DWT operates in the following way.
For each scale, it performs a high-pass downsampling (throw away every other data point) and a low-pass smoothing operation of the other half of the data. The result is a square with squares-within-squares of low-pass (L) operations and high-pass operations (H). The filter operations always work in pairs (Quadrature Mirror Filters), the filters being the wavelets you've chosen (say the Daubechies wavelet).

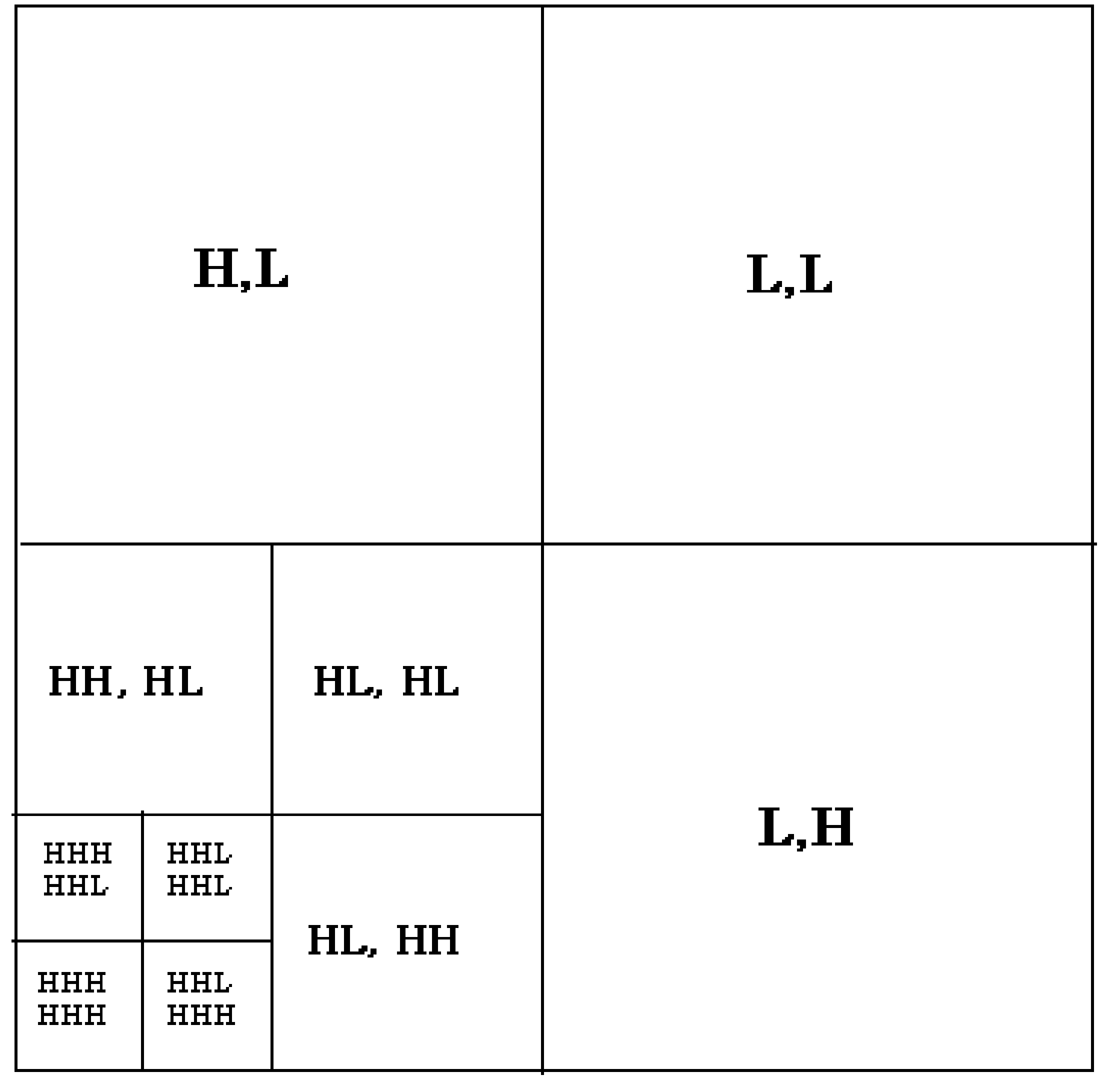
Next is an ASCII sketch of some output filter operations.

The sketch shows three scale levels of smoothing, down-sampled operations. The square labeled all LL will be the most bland, smoothed result. The square with all HH will be the tiniest representation of your original image (just down-sampled, with every other point thrown away several times). The others will show you different things. Some combinations of LH will show the fine detail in your image. Other combinations will show the sharp contrast portions of your image (those are the combinations that make wavelets good edge detectors.

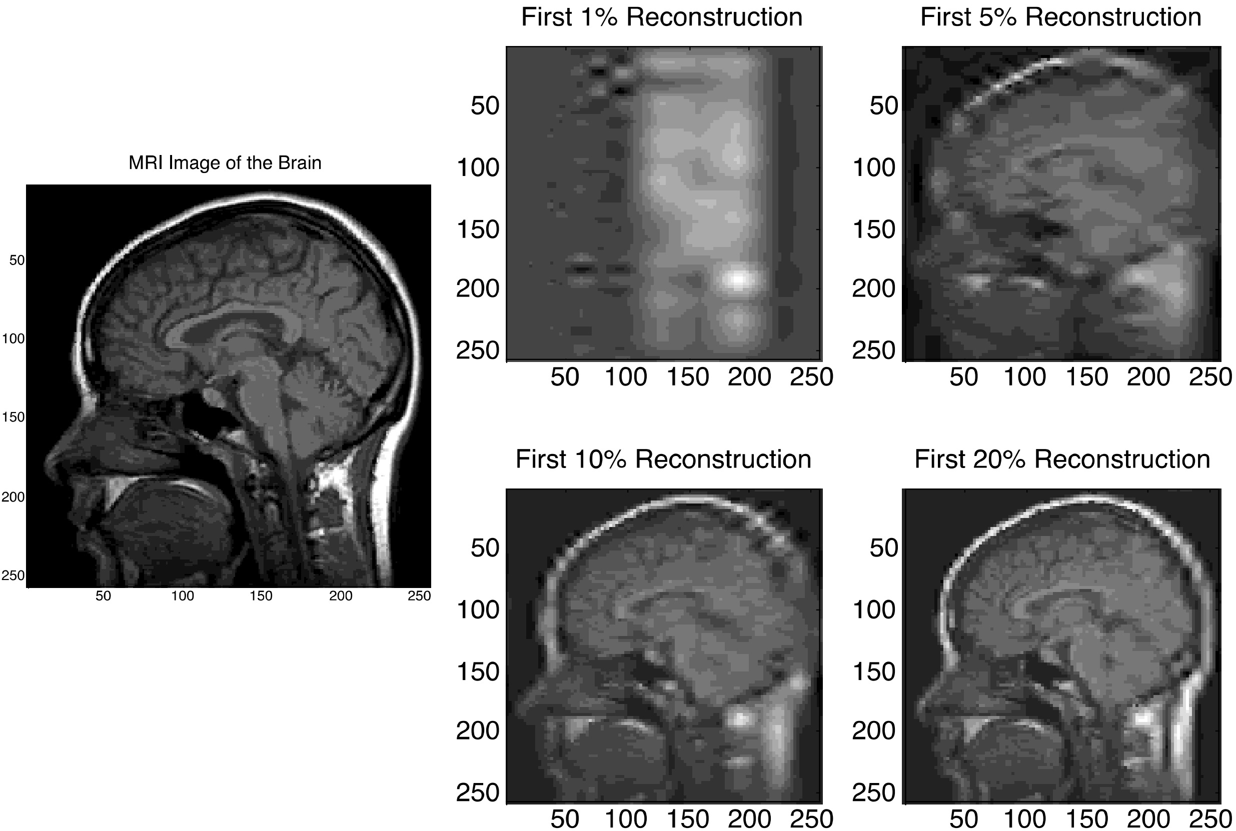
Wavelet Decomposition of Images







Wavelet Synthesis/Reconstruction



- The traditional approach to remove noise is based on simple low-pass filtering, or neighbor averaging. Wavelet representations permit a more efficient noise removal while preserving high frequencies.

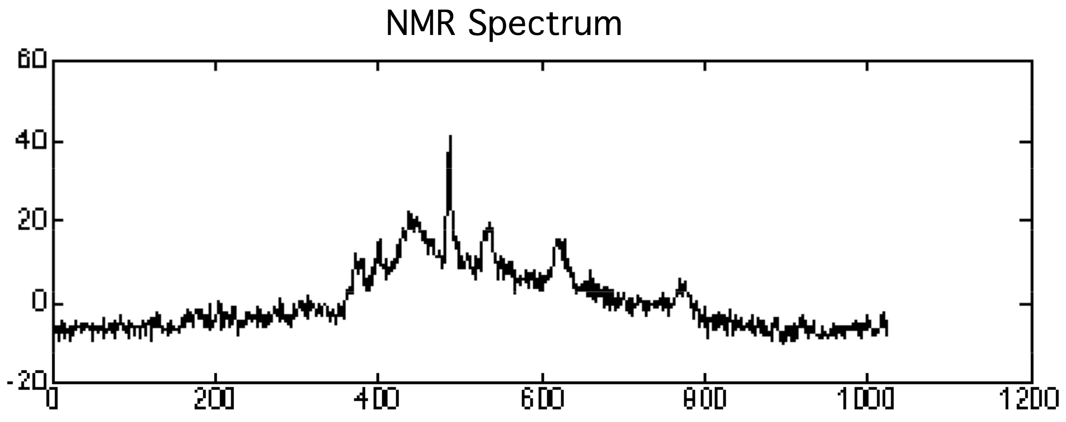
- When decomposing a data set, one can omit details without substantially affecting the main features of the data set.

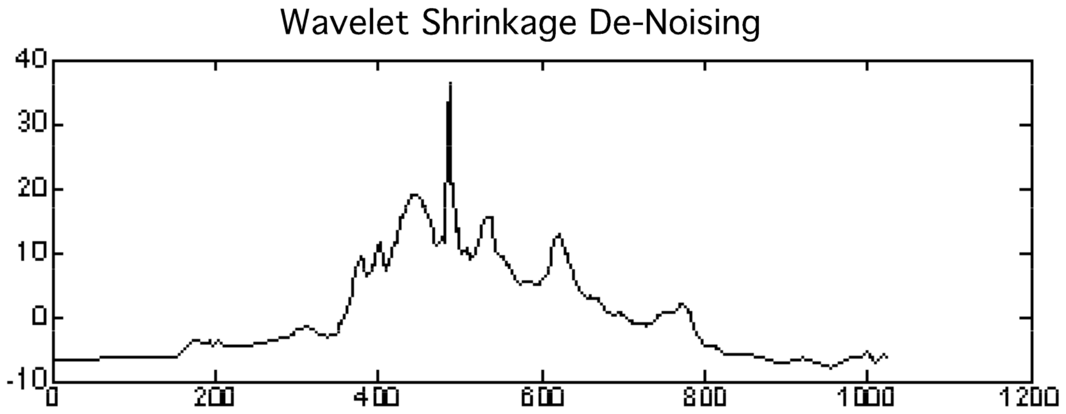
- "Thresholding" sets to zero all coefficients that are less than a particular threshold. Then the coefficients above the threshold are used in reconstructing the data set through an inverse wavelet transformation.

Wavelet Denoising

This NMR signal was denoised using the following procedure:

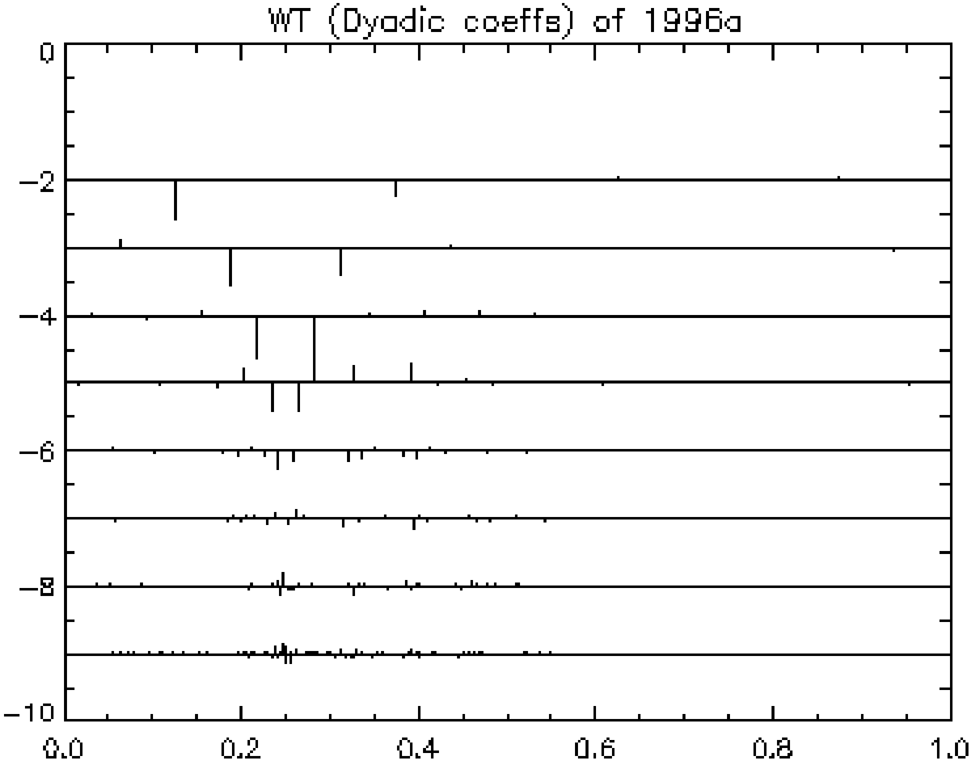
1. The signal was transformed to the wavelet domain using Coiflets with three vanishing moments.
2. A threshold was applied at two standard deviations.
3. The image was inverse transformed to the signal domain.



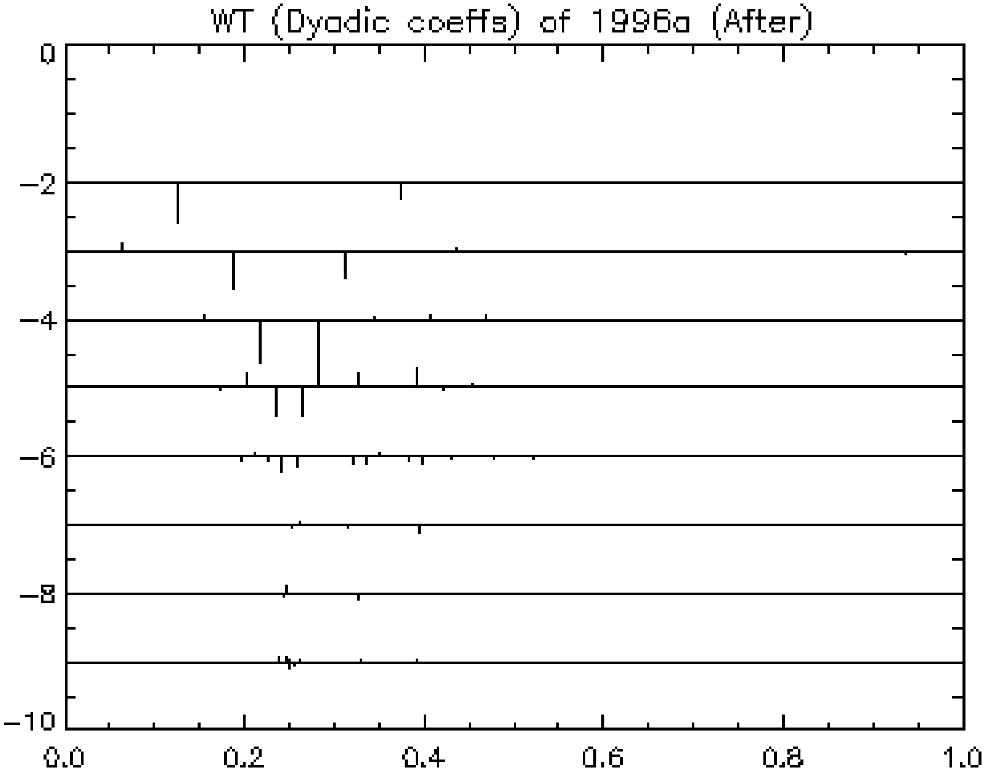


Wavelet Denoising

The wavelet transform coefficients **before** the threshold was applied.



The wavelet transform coefficients **after** the threshold was applied.



Wavelet References

Graps, Amara L., "An Introduction to Wavelets", IEEE Computational Sciences and Engineering, Volume 2, Number 2, Summer 1995, pp 50-61.

[Hubbard] Barbara Burke Hubbard, _The World According to Wavelets: The Story of a Mathematical Technique in the Making_, A.K. Peters, 1995.

[Newland] Newland, D.E., _An Introduction to Random Vibrations, Spectral and Wavelet Analysis_, New York, John Wiley, 1993.

[Robinson] Robinson, Sam L. and Ryczek, Peter F., "Wavelets," _The Mathematica Journal_, Volume 5, Issue 1, Winter 1995, pp. 74-81.

[Rowe] Rowe, Alistair C.H. and Abbott, Paul C., "Daubechies Wavelets and Mathematica," _Computers in Physics_, Volume 9, Issue 6, November/December 1995, pp. 635-648.

[StrangA] Gilbert. Strang, "Wavelets," American Scientist, Vol. 82, 1994, pp. 250-255.

[Strang] Gilbert Strang and Truong Nguyen, _Wavelets and Filter Banks_, Wellesley-Cambridge Press, 1996.

[Stollnitz] E.J. Stollnitz and T.D. DeRose and D.H. Salesin, _Wavelets for Computer Graphics: Theory and Application_, Morgan Kaufmann Publishers,
1996.

[Torrence] Torrence, Christopher and Gilbert P. Compo, "A Practical Guide to Wavelet Analysis, Bull. Amer. Met. Soc, June 21, 1997.