

WAVELETS: “BEYOND COMPARISON” - D. L. FUGAL

Wavelets are used extensively in Signal and Image Processing, Medicine, Finance, Radar, Sonar, Geology and many other varied fields. They are usually presented in mathematical formulae, but can actually be understood in terms of simple comparisons with your data.

As a background, we first look at the Discrete Fourier Transform (DFT) or it’s faster and more famous cousin, the Fast Fourier Transform (FFT). These transforms can be thought of as a series of comparisons with your data, which we will call for now a “signal” for consistency. Signals that are simple waves of constant frequencies can be processed with ordinary DFT/FFT methods.

Real-world signals, however, often have frequencies that can change over time or have pulses, anomalies, or other “events” at certain specific times. This type of signal can tell us where something is located on the planet, the health of a human heart, the position and velocity of a “blip” on a Radar screen, stock market behavior, or the location of underground oil deposits. For these signals, we will often do better with wavelets. We now demonstrate both the Fourier and Wavelet Transforms of a simple pulse signal.

THE DISCRETE FOURIER TRANSFORM/FAST FOURIER TRANSFORM (DFT/FFT).

We start with a point-by-point comparison of the pulse signal (D) with a high frequency wave or “sinusoid” of constant frequency (A) as shown in Figure 1. We obtain a single “goodness” value from this comparison (a correlation value) which indicates how much of that particular sinusoid is found in our own pulse signal.

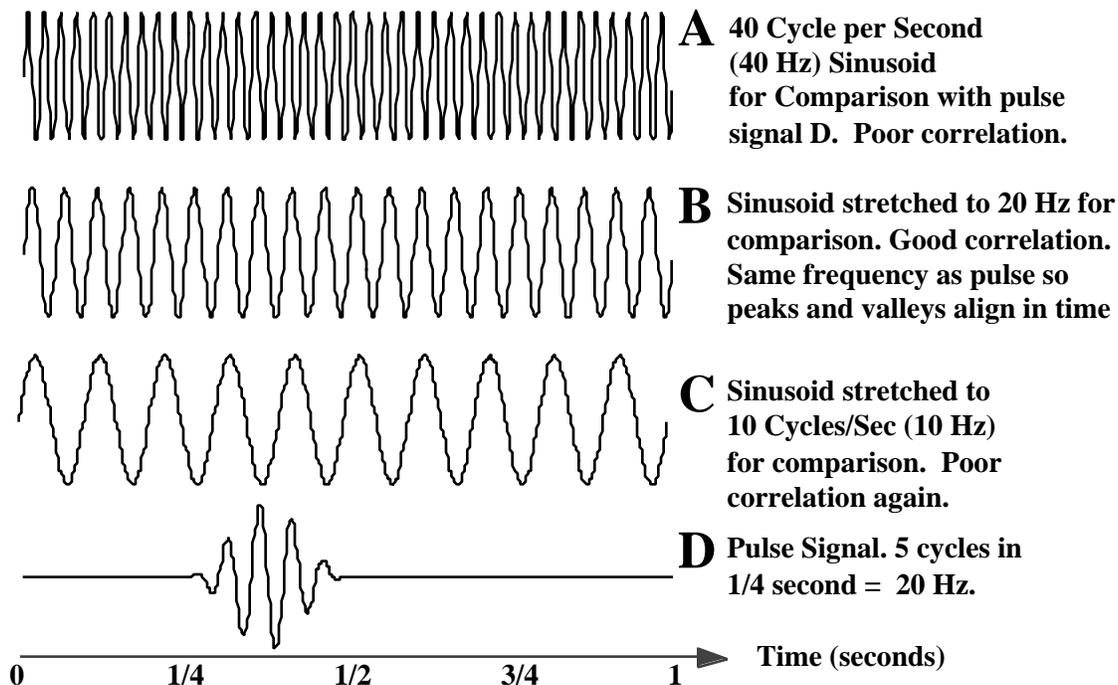


Figure 1: FFT-type comparison of Pulse Signal with several stretched sinusoids.

We can observe that the pulse has 5 cycles in 1/4 of a second. This means that it has a frequency of 20 cycles in one second or “20 Hz”. The comparison sinusoid, A, has twice the frequency or 40 Hz. Even in the area where the signal is non-zero (the pulse) the comparison is not very good.

By lowering the frequency of A from 40 to 20 Hz (waveform B) we are effectively “stretching” the sinusoid (A) by 2 so it has only 20 cycles in 1 second. We compare point-by-point again over the 1-second interval with the pulse (D) This next correlation gives us another value indicating how much of this lower frequency sinusoid (now the same frequency as our pulse) is contained in our signal. This time the correlation of the pulse with the comparison sinusoid is very good. The peaks and valleys of B and the pulse portion of D align (or can be easily shifted to align) and thus we have a large correlation value.

Figure 1 above shows us one more comparison of our original sinusoid (A) stretched by 4 and trimmed so it has only 10 cycles in the 1 second interval (C). This comparison with D is poor again. We could continue stretching and trimming until the sinusoid becomes a straight line having zero frequency or “DC” (named for the zero frequency of Direct Current) but all these comparisons will be increasingly poor.

An actual DFT (or functionally equivalent FFT) compares many “stretched” sinusoids (“analysis signals”) to the pulse rather than just the 3 shown here. The best correlation is found when the sinusoid frequency best matches that of the pulse. Figure 2 shows the first part of an actual FFT of our pulse signal D. The locations of our sample comparison sinusoids A, B, and C are indicated. Notice that the FFT tells us correctly that the pulse has primarily a frequency of 20 Hz, but does NOT tell us where the pulse is located in time!

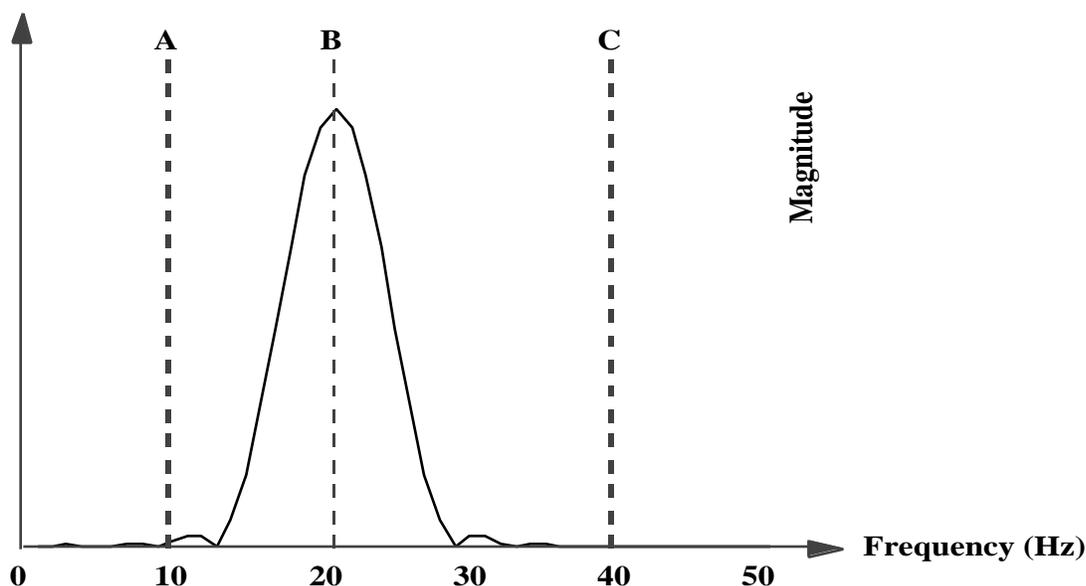


Figure 2: Actual FFT plot of above pulse signal with the three above sinusoids indicated

THE CONTINUOUS WAVELET TRANSFORM (CWT)

Wavelets are exciting because they too are comparisons, but instead of correlating with various stretched, infinite length unchanging sinusoids, they use smaller or shorter waveforms (“wave - lets”).that can start and stop where we wish.

By stretching and shifting the wavelet numerous times we get numerous correlations. If our signal has some interesting events embedded, we will get the best correlation when the stretched wavelet is similar in frequency to the event and is shifted to line up with it in time. Knowing the amounts of stretching and shifting we can determine both location and frequency.

Figure 3 demonstrates the process. Instead of sinusoids for our comparisons, we will use wavelets. Waveform A shows a Daubechies 20 (Db20) wavelet about 1/8 second long that starts at the beginning ($t = 0$) and effectively ends well before 1/4 second. The zero values are extended to the full 1 second. The point-by-point comparison with our pulse signal D will be very poor and we will obtain a very small correlation value.

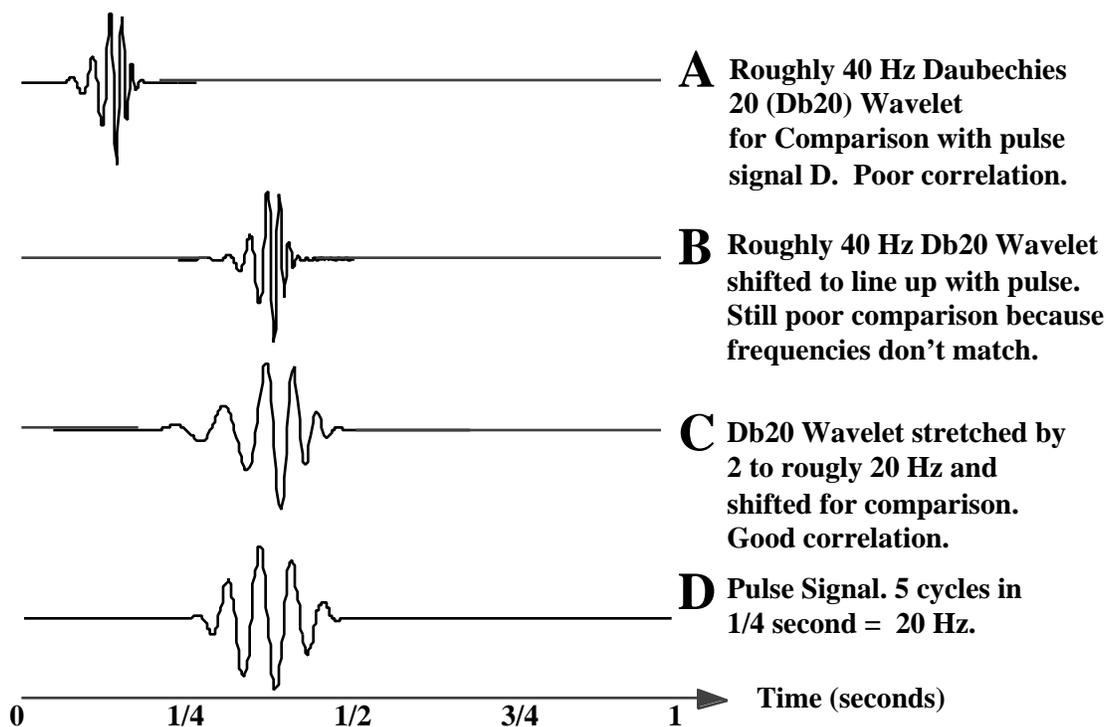


Figure 3: CWT-type comparison of Pulse Signal with several stretched and shifted wavelets.

In the previous FFT/DFT discussion we proceeded directly to stretching. In the Wavelet Transforms we shift the wavelet slightly to the right and perform another comparison with this new waveform to get another correlation value. We continue to shift until the Db20 wavelet is in the position shown in B. We get a little better comparison than A, but still very poor because B and D are different frequencies.

After we have shifted the wavelet all the way to the end of the 1 second time interval we start over with a slightly stretched wavelet at the beginning and repeatedly shift to the right to obtain another full set of these correlation values. C shows the Db20 wavelet stretched to where the frequency is roughly the same as the pulse (D) and shifted to the right until the peaks and valleys line up fairly well. At this particular shifting and stretching we should obtain a very good comparison and large correlation value. Further shifting to the right, however, even at this same stretching will yield increasingly poor correlations.

In the CWT we thus have one correlation value for every shift of every stretched wavelet. To show the data for all these stretches and shifts, we use a 3-D display with the stretching (roughly inverse of frequency) as the vertical axis, the shifting in time as the horizontal axis, and brightness (or color) to indicate the strength of the correlation. Figure 4 shows a Continuous Wavelet Transform (CWT) display for this particular pulse signal (D). Note the strong correlation of the 3 larger peaks and valleys of the pulse with the Db20 wavelet, the strongest being where all the peaks and valleys best align.

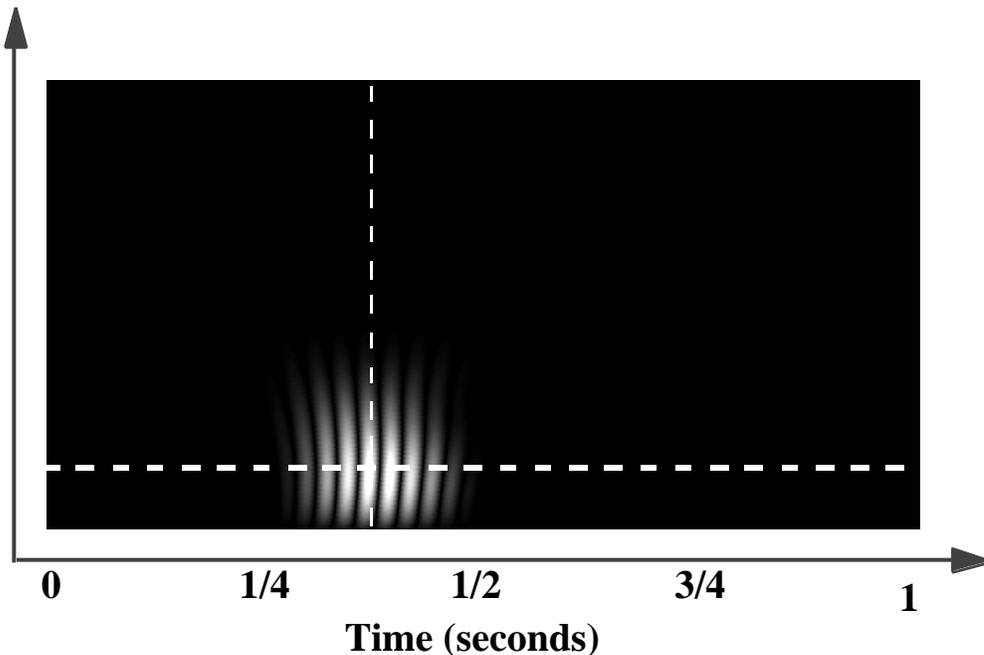


Figure 4: Actual CWT display indicating the time and frequency of the Pulse Signal.

The display shows that the best correlation occurs at the brightest point or at about 3/8 second. This agrees with what we already know about the pulse, D. The display also tells us how much the wavelet had to be stretched (or “scaled”) and this indicates the approximate frequency of the pulse. Thus we know not only the frequency of the pulse, but also the time of its occurrence!

We run into this simultaneous time/frequency concept in everyday life. For example, a bar of sheet music may tell the pianist to play a C-chord of three different frequencies at exactly the same time on the first beat of the measure.

For the simple example above we could have just looked at the pulse (D) to see its location and frequency. The next example is more representative of wavelets in the real world.

Figure 5 shows a signal with a very small, very short discontinuity at time 180. The Amplitude vs. Time plot of the signal is shown at the upper left but does not show the tiny “event”. The Magnitude vs. Frequency FFT plot tells what frequencies are present but does not indicate the time associated with those frequencies.

With the wavelet display, however, we can clearly see a vertical line at 180 at low scales when the wavelet has very little stretching, indicating a very high frequency. The CWT display also “finds” the

large oscillating wave at the higher scales where the wavelet has been stretched and compares well with the lower frequencies. For this short discontinuity we used a short wavelet (a Db4) for best comparison.

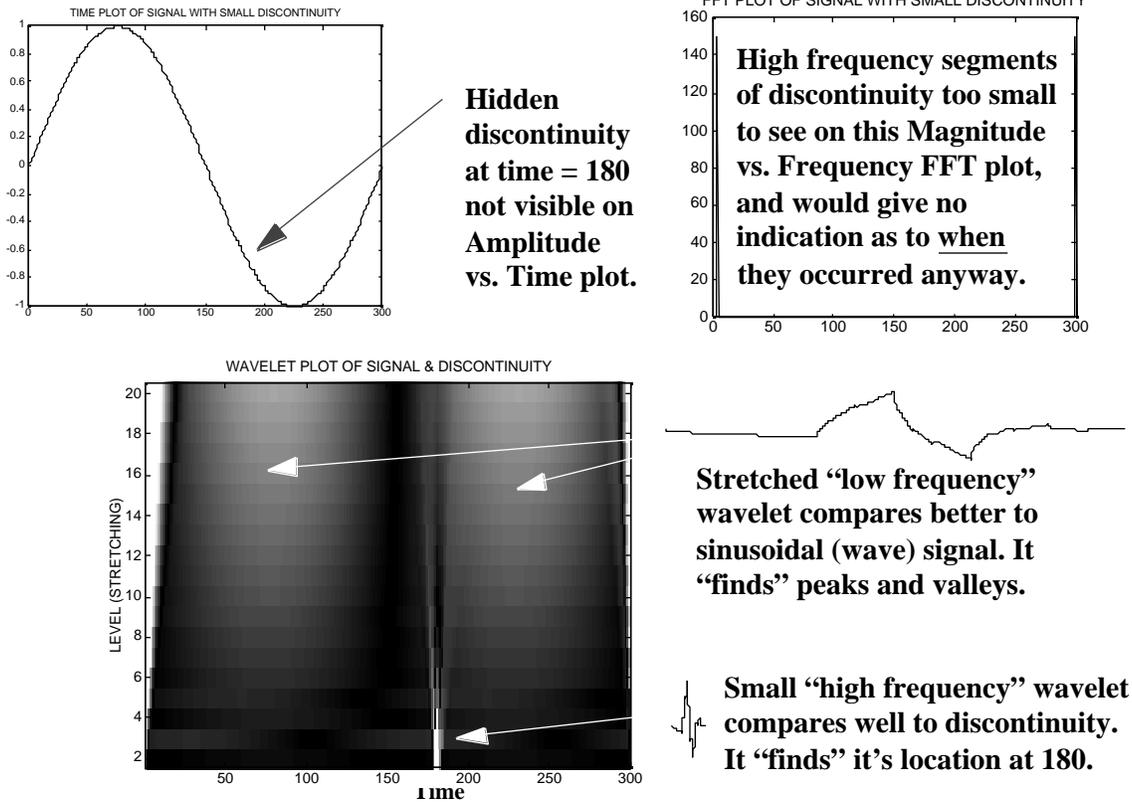


Figure 5: Comparison of Time Plot, Frequency Plot (FFT), and Wavelet Plot (CWT) of a signal with a hidden discontinuity.

This is an example of why wavelets have been referred to as a “mathematical microscope” for their ability to find interesting events of various lengths and frequencies hidden in data.

DISCRETE WAVELET TRANSFORMS

Besides acting as a “microscope” to find hidden events in our data, wavelets can also separate the data into various frequency components, as does the FFT. The FFT/DFT is used extensively to remove unwanted noise that is prevalent throughout the entire signal such as a 60 Hz hum. Unlike the FFT, however, the wavelet transform allows us to remove frequency components at specific times in the data. This allows us a powerful capability to throw out the “bad” and keep the “good” part of the data in that frequency range.

These types of transforms are called “Discrete Wavelet Transforms” (DWT). They also have easily computed inverse transforms (IDWT) that allow us to reconstruct the signal after we have identified and removed the noise or superfluous data for denoising or compression. .

“REDUNDANT” DISCRETE WAVELET TRANSFORMS (RDWT)

In one type of DWT, the Redundant Discrete Wavelet Transform or RDWT, we first compare (correlate) the Wavelet “filter” with itself. This produces a “Highpass Halfband Filter” or “superfilter”. When we compare or correlate our signal with this superfilter we extract the highest half of the frequencies. For a very simple denoising, we could just discard these high frequencies (for whatever time period we choose) and then reconstruct a denoised signal.

Multi-level RDWT’s allow us to stretch the wavelet, similar to what we did in the CWT, except that it is done by factors of 2 (twice as long, 4 times as long, etc.). This allows us stretched superfilters that can be halfband, quarter-band, eighth-band and so forth.

“CONVENTIONAL” DISCRETE WAVELET TRANSFORMS (DWT).

We stretched the wavelet in the CWT and the RDWT. In the conventional DWT, we shrink the signal instead and compare it to the unchanged wavelet. We do this by “downsampling by 2”. Every other point in the signal is discarded. We have to deal with “aliasing” (not having enough samples left to represent the high frequency components and thus producing a false signal). We must also be concerned with “shift invariance” (do we throw away the odd or the even values?--it matters!).

If we are careful, we can deal with these concerns. One amazing capability of the filters in the conventional DWT is alias cancellation where the basic wavelet and 3 similar “filters” combine to allow us to reconstruct the original signal perfectly. The stringent requirements on the wavelets to be able to do this is part of why they often look so strange (see Figure 8).

As with the RDWT, we can denoise our signal by discarding portions of the frequency spectrum -- as long as we careful not to throw away vital parts of the alias cancellation capability. Correct and careful downsampling also aids with compression of the signal. Modern JPEG compression uses wavelets. Figure 6 shows JPEG image compression. The image on the right was compressed by a ratio of 157:1 using a Biorthogonal 9/7 set of wavelets.

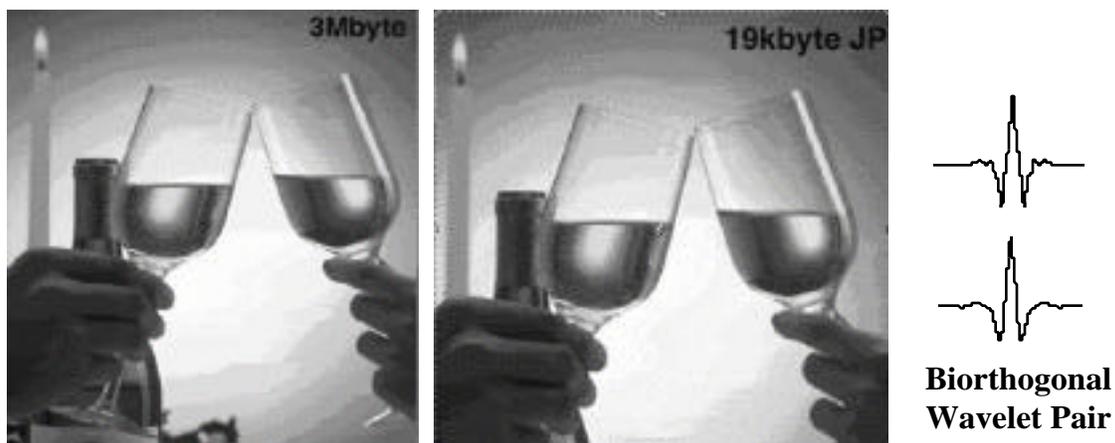


Figure 6: JPEG image compression of 157:1 using Biorthogonal 9/7 set of symmetrical Wavelets.

There are many types of Wavelets. Some come from mathematical expressions. Others are built from basic Wavelet Filters having as little as 2 points. The Db4 , Db20, and Biorthogonal wavelets shown

earlier are examples of this 2nd type. Figure 7 shows a 768 point approximation of a continuous Db4 wavelet with the 4 filter points (plus 2 zeros) superimposed.

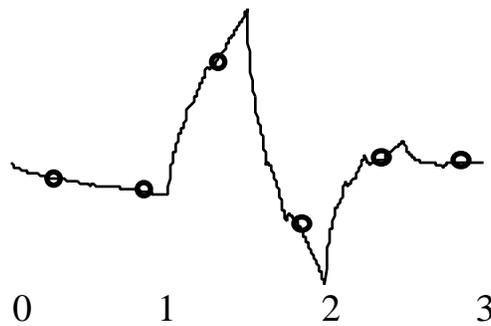


Figure 7 - Daubechies 4 wavelet with 4 original filter points and 2 end zeros

Some wavelets have symmetry (valuable in human vision perception) such as the Biorthogonal Wavelet pairs. Shannon or “Sinc” Wavelets can find events with specific frequencies (these are similar to the Sinc Function filters found in traditional DSP). Haar Wavelets (the shortest) are good for edge detection and reconstructing binary pulses. Coiflets Wavelets are good for data with self-similarities (fractals) such as financial trends. Some of the wavelet families are shown in Figure 8.

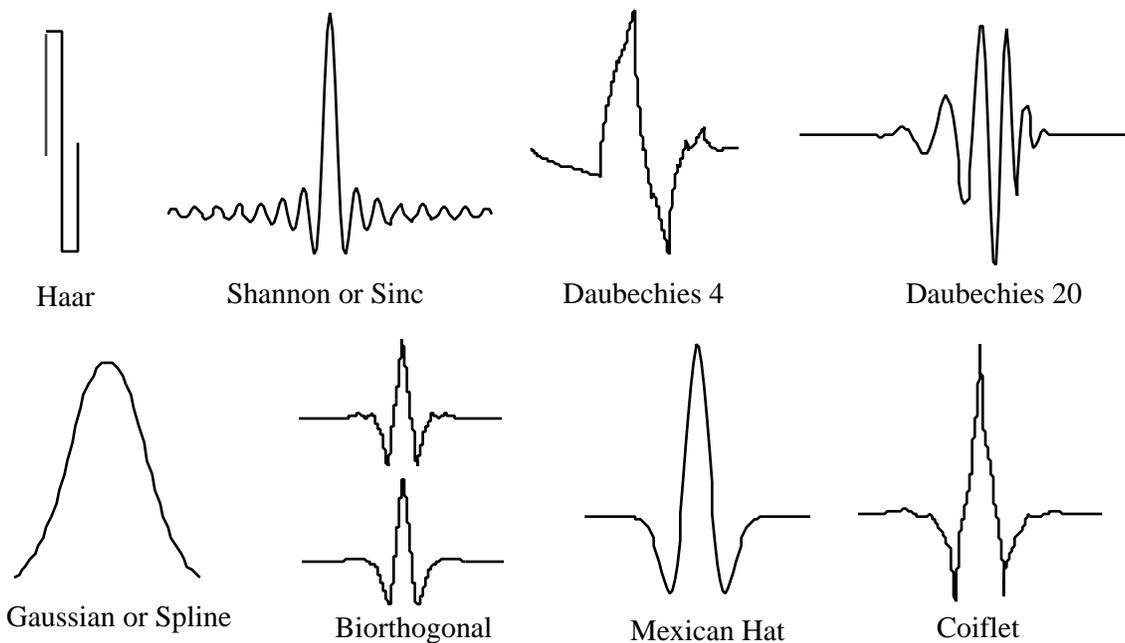


Figure 8 - Examples of types of wavelets

You can even create your own wavelets, if needed. However there is “an embarrassment of riches” in the many wavelets that are already out there and ready to go. We have already seen that with their ability to stretch and shift that wavelets are extremely adaptable. You can usually get by very nicely with choosing a less-than-perfect Wavelet. The only “wrong” choice is to avoid wavelets due to an abundant selection.

There is much more to discover than can be presented in this short overview. The time spent, however, in learning, understanding and correctly using wavelets for these “non-stationary” signals with anomalies at specific times or changing frequencies (the fascinating, real-world kind!) will be repaid handsomely.