SIGNAL PROCESSING & SIMULATION NEWSLETTER

Note: This is not a particularly interesting topic for anyone other than those who are involved in simulation. So if you have difficulty with this issue, you may safely stop reading after page 5 without feeling guilty.

Hilbert Transform, Analytic Signal and the Complex Envelope

In Digital Signal Processing we often need to look at relationships between real and imaginary parts of a complex signal. These relationships are generally described by Hilbert transforms. Hilbert transform not only helps us relate the I and Q components but it is also used to create a special class of causal signals called analytic which are especially important in simulation. The analytic signals help us to represent bandpass signals as complex signals which have specially attractive properties for signal processing.

Hilbert Transform is not a particularly complex concept and can be much better understood if we take an intuitive approach first before delving into its formula which is related to convolution and is hard to grasp. The following diagram that is often seen in text books describing modulation gives us a clue as to what a Hilbert Transform does.

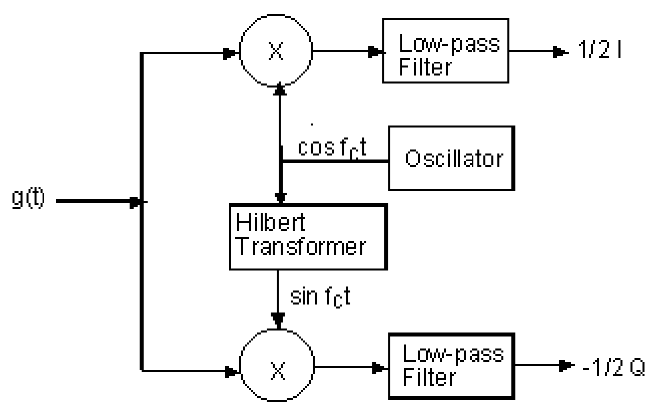

Figure 1 - Role of Hilbert Transform in modulation

The role of Hilbert transform as we can guess here is to take the carrier which is a cosine wave and create a sine wave out of it. So let's take a closer look at a cosine wave to see how this is done by the Hilbert transformer. Figure 2a shows the amplitude and the phase spectrum of a cosine wave. Now recall that the Fourier Series is written as

$$f(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega t}$$

where

$$C_n = A_n + jB_n \text{ and } C_{-n} = A_n - jB_n$$

and A_n and B_n are the spectral amplitudes of cosine and sine waves. Now take a look at the phase spectrum. The phase spectrum is computed by

$$\varphi = \tan^{-1}\frac{B_n}{A_n}$$

Cosine wave has no sine spectral content, so B_n is zero. The phase calculated is 90° for both positive and negative frequency from above formula. The wave has two spectral components each of magnitude 1/2A, both positive and lying in the real plane. (the real plane is described as that passing vertically (R-V plane) and the Imaginary plane as one horizontally (R-I plane) through the Imaginary axis)

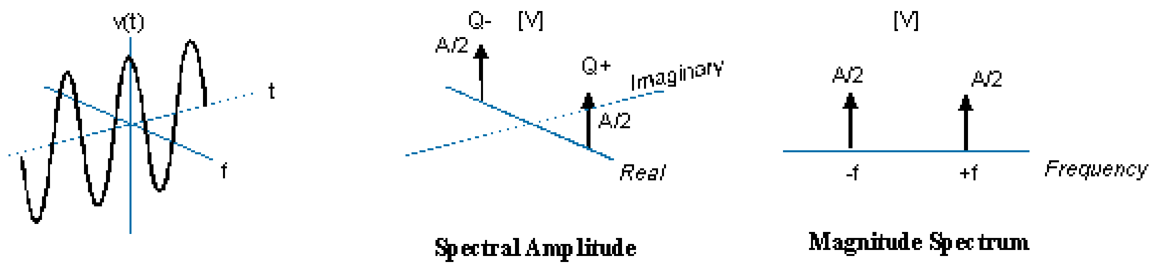

Figure 2a - Cosine Wave Properties

Figure 2b shows the same two spectrums for a sine wave. The sine wave phase is not symmetric because the amplitude spectrum is not symmetric. The quantity A_n is zero and B_n has either a positive or negative value. The phase is +90° for the positive frequency and -90° for the negative frequency.

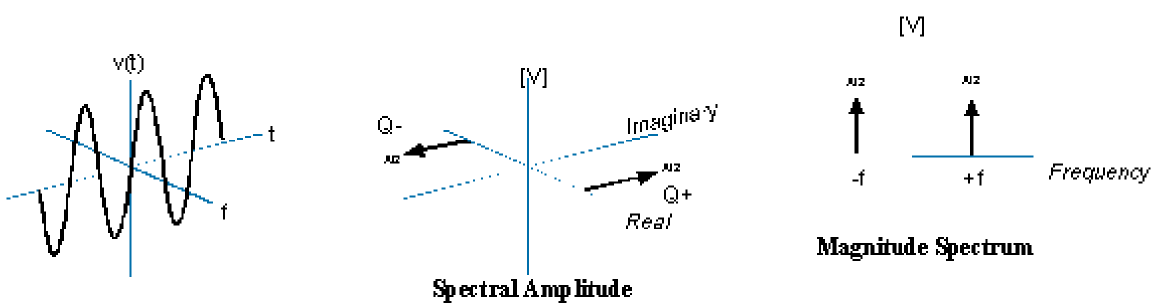

Figure 2b - Sine Wave Properties

Now we wish to convert the cosine wave to a sine wave. There are two ways of doing that, one in time domain and the other in frequency domain.

Hilbert Transform in Frequency Domain

Now compare Figure 2a and 2b, in particular the spectral amplitudes. The cosine spectral amplitudes are both positive and lie in the real plane. The sine wave has spectral components that lie in the Imaginary plane and are of opposite sign.

To turn cosine into sine, as shown in Figure 3 below, we need to rotate the negative frequency component of the cosine by +90° and the positive frequency component by -90°. We will need to rotate the +Q phasor by -90° or in other words multiply it by -j. We also need to rotate the -Q phasor by +90° or multiply it by j.

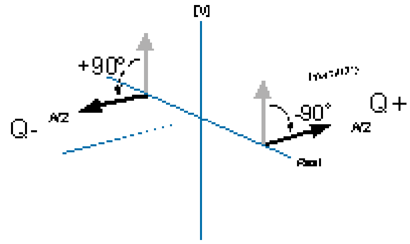

Figure 3 - Rotating phasors to create a sine wave out of a cosine

We can describe this transformation process called the Hilbert Transform as follows:
All negative frequencies of a signal get a +90° phase shift and all positive frequencies get a -90° phase shift.

If we put a cosine wave through this transformer, we get a sine wave. This phase rotation process is true for all signals put through the Hilbert transform and not just the cosine.

For any signal g(t), its Hilbert Transform has the following property

$$\hat{G}(f) = \begin{matrix} -j & \text{for } f > 0 \\ j & \text{for } f < 0 \end{matrix}$$

(Putting a little hat over the capital letter representing the time domain signal is the typical way a Hilbert Transform is written.)

A sine wave through a Hilbert Transformer will come out as a negative cosine. A negative cosine will come out a negative sine wave and one more transformation will return it to the original cosine wave, each time its phase being changed by 90°.

$$\cos wt \rightarrow \sin wt \rightarrow -\cos wt \rightarrow -\sin wt \rightarrow \cos wt$$

For this reason Hilbert transform is also called a "quadrature filter". We can draw this filter as shown below in Figure 4.

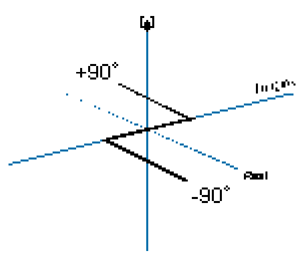

Figure 4 - Hilbert Transform shifts the phase of positive frequencies by -90° and negative frequencies by +90°.

So here are two things we can say about the Hilbert Transform.

1. It is a peculiar sort of filter that changes the phase of the spectral components depending on the sign of their frequency.
2. It only effects the phase of the signal. It has no effect on the amplitude at all.

Hilbert transform in Time Domain

Now look at the signal in time domain. Given a signal g(t), Hilbert Transform of this signal is defined as

$$\hat{g}(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{g(\tau)}{t-\tau}\,d\tau \qquad (1)$$

Another way to write this definition is to recognize that Hilbert Transform is also the convolution of function 1/ pt with the signal g(t). So we can write the above equation as

$$\hat{g}(t) = \frac{1}{\pi t} * g(t) \qquad (2)$$

Achieving a Hilbert Transform in time domain means convolving the signal with the function 1/ pt. Why the function 1/ pt, what is its significance? Let's look at the Fourier Transform of this function. What does that tell us? Given in Eq 3, the transform looks a lot like the Hilbert transform we talked about before.

$$F\left(\frac{1}{\pi t}\right) = -j\,\mathrm{sgn}(f) \qquad (3)$$

The term sgn in Eq 3 above , called signum is simpler than it seems. Here is the way we could have written it which would have been more understandable.

$$F\left(\frac{1}{\pi t}\right) = -j\,\mathrm{sgn}(f) = \begin{array}{ll} -j & \text{for } f > 0 \\ +j & \text{for } f < 0 \end{array} \qquad (4)$$

In Figure 5 we show the signum function and its decomposition into two familiar functions.

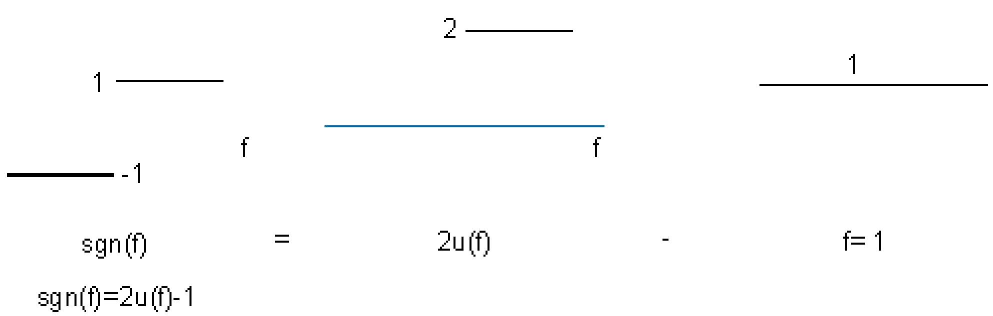

Figure 5 - Signum Function decomposed into a unit function and a constant

For shortcut, writing sgn is useful but it is better if it is understood as a sum of the above two much simpler functions. (We will use this relationship later.)

$$\mathrm{sgn}(f) = 2u(f) - 1 \qquad (5)$$

We see in 6 figure that although 1/pt is a real function, is has a Fourier transform that lies strictly in the imaginary plane. Do you recall what this means in terms of Fourier Series coefficients? What does it tell us about a function if it has no real components in its Fourier transform? It says that this function can be represented completely by a sum of sine waves. It has no cosine component at all.

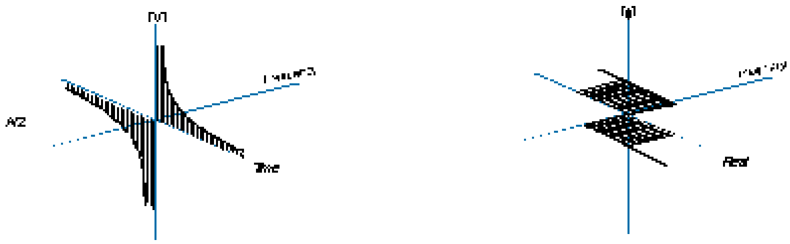

Figure 6 Function f(t) = 1/πt, and its Fourier transform

In Figure 7, we see a function composed of a sum of 50 sine waves. We see the similarity of this function with that of 1/pt. Now you can see that although the function 1/pt looks nothing at all a sinusoid, we can still approximate it with a sum of sinusoids.

The function f(t) = 1/pt gives us a spectrum that explains the Hilbert Transform in time domain, albeit this way of looking at the Hilbert Transform is indeed very hard to grasp.

We limit our discussion of Hilbert transform to Frequency domain due to this difficulty.

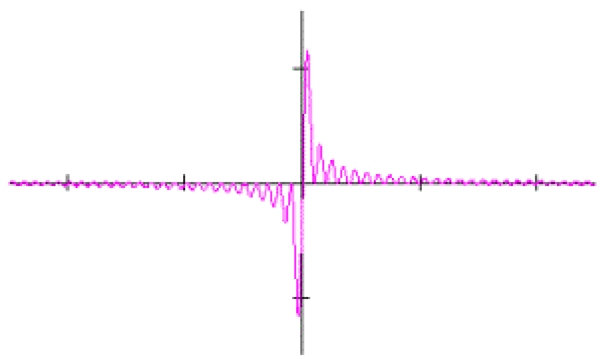

Figure 7 - Approximating function f(t) = 1/p t with a sum of 50 sine wave

We can add the following to our list of observations about the Hilbert Transform.

3. The signal and its Hilbert Transform are orthogonal. This is because by rotating the signal 90° we have now made it orthogonal to the original signal, that being the definition of orthogonality.
4. The signal and its Hilbert Transform have identical energy because phase shift do not change the energy of the signal only amplitude changes can do that.

Analytic Signal

Hilbert Transform has other interesting properties. One of these comes in handy in the formulation of an Analytic signal. Analytic signals are used in Double and Single side-band processing (about SSB and DSB later) as well as in creating the I and Q components of a real signal.

An analytic signal is defined as follows.

$$g_+(t) = g(t) + j\,\hat{g}(t) \qquad (6)$$

An analytic signal is a complex signal created by taking a signal and then adding in quadrature its Hilbert Transform. It is also called the pre-envelope of the real signal.

So what is the analytic signal of a cosine?

Substitute cos wt for g(t) in Eq 6, knowing that its Hilbert transform is a sine, we get

$$g_+(t) = \cos(wt) + j\sin(wt) = e^{jwt}$$

The analytic function of a cosine is the now familiar phasor or the complex exponential, e^{jwt}.

What is the analytic signal of a sine?

Now substitute sin wt for g(t) in Eq 6, knowing that its Hilbert transform is a -cos, we get once again a complex exponential.

$$g_+(t) = \sin(wt) - j\cos(wt) = e^{-jwt}$$

Do you remember what the spectrum of a complex exponential looks like? To remind you, I repeat here the figure from Tutorial 6.

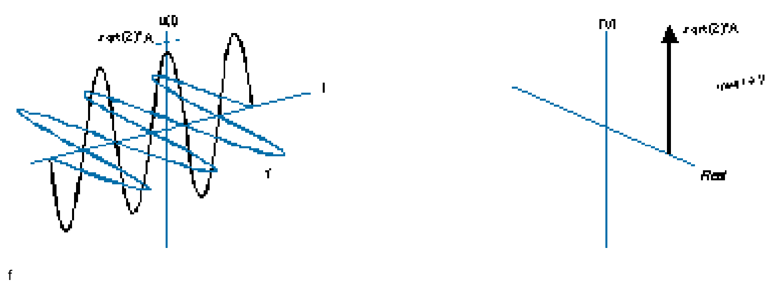

Figure 8 - Fourier transform of a complex exponential

We can see from the figure above, that whereas the spectrum of a sine and cosine spans both the negative and positive frequencies, the spectrum of the analytic signal, in this case the complex exponential, is in fact present only in the positive domain. This is true for both sine and cosine and in fact for all real signals.

Restating the results: the Analytic signal for both and sine and cosine is the complex exponential. Even though both sine and cosine have a two sided spectrum as we see in figures above, the complex exponential which is the analytic signal of a sinusoid has a one-sided spectrum.

We can generalize from this: An analytic signal (composed of a real signal and its Hilbert transform) has a spectrum that exists only in the positive frequency domain.

Let's take at a look at the analytic signal again.

$$g_+(t) = g(t) + j\hat{g}(t) \tag{7}$$

The conjugate of this signal is also a useful quantity.

$$g_-(t) = g(t) - j\hat{g}(t) \tag{8}$$

This signal has components only in the negative frequencies and can be used to separate out the lower side-bands.

Now back to the analytic signal. Let's extend our understanding by taking Fourier Transform of both sides of Eq 7. We get

$$G_+(f) = G(f) + j(-j\,\mathrm{sgn}(f)G(f)) \tag{9}$$

The first term is the Fourier transform of the signal g(t), and the second term is the inverse Hilbert Transform. We can rewrite by use of property $\operatorname{sgn}(f) = 2u(f) - 1$ Eq 9 as

$$G_+(f) = 2G(f)u(f) \tag{10}$$

One more simplification gives us

$$G_{+(f)} = \begin{cases} 2G(f) & \text{for } f > 0 \\ G(0) & \text{for } f = 0 \\ 0 & \text{for } f < 0 \end{cases} \tag{11}$$

This is a very important result and is applicable to both lowpass and modulated signals. For modulated or bandpass signals, its net effect is to translate the signal down to baseband, double the spectral magnitudes and then chop-off all negative components.

Complex Envelope

We can now define a new quantity based on the analytic signal, called the Complex Envelope. The Complex Envelope is defined as

$$g_+(t) = \tilde{g}(t)\, e^{j2\pi f_c t}$$

The Complex Envelope of signal g(t)

The part $\tilde{g}(t)$ is called the Complex Envelope of the signal g(t).

Let's rewrite it and take its Fourier Transform.

$$\tilde{g}(t) = g_+(t)e^{-j2\pi f_c t}$$

We now see clearly that the Complex Envelope is just the frequency shifted version of the analytic signal. Recognizing that multiplication with the complex exponential in time domain results in frequency shift in the Frequency domain, using the Fourier Transform results for the analytic signal above, we get

$$\tilde{G}_{(f)} = \begin{cases} 2G(f - f_c) & \text{for } f > 0 \\ G(0) & \text{for } f = 0 \\ 0 & \text{for } f < 0 \end{cases} \tag{12}$$

So here is what we have been trying to get at all this time. This result says that the Fourier Transform of the analytic signal is just the one-sided spectrum. The carrier signal drops out entirely and the spectrum is no longer symmetrical.

This property is very valuable in simulation. We no longer have to do simulation at carrier frequencies but only at the highest frequency of the baseband signal. The process applies equally to other transformation such as filters etc. which are also down shifted. It even works when non-linearities are present in the channel and result in additional frequencies.

There are other uses of complex representation which we will discuss as we explore these topics however its main use is in simulation.

Example

Let’s do an example. Here is a real baseband signal.
(I have left out the factor 2p for purposes of simplification)

$$s(t) = 4\cos 2t - 6\sin 3t$$

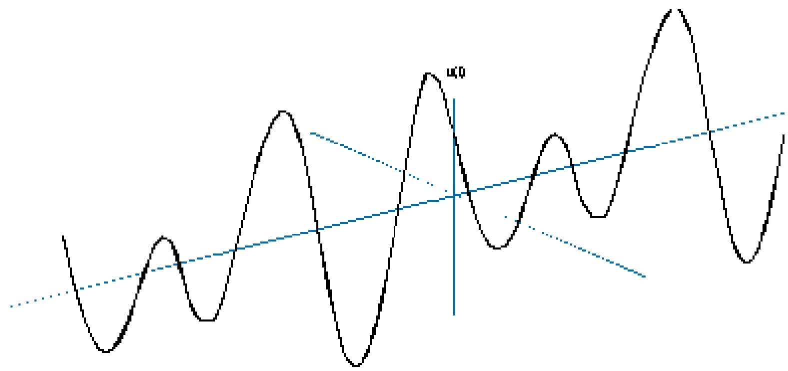

Figure 9 - A Baseband Signal

The spectrum of this signal is shown below, both its individual spectral amplitudes and its magnitude spectrum. The magnitude spectrum shows one spectral component of magnitude 2 at f = 2 and -2 and an another one of magnitude 3 at f = 3 and -3.

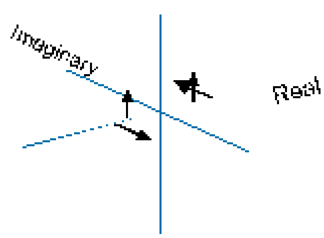

Figure 10a - Spectral amplitudes

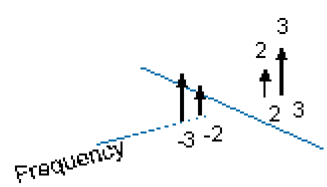

Figure 10b- The Magnitude Spectrum

Now let’s multiply it with a carrier signal of cos(100t) to modulate it and to create a bandpass signal,

$$\hat{g}(t) = s(t)\cos 100t$$

$$\hat{g}(t) = 4\cos 2t\cos 100t - 6\sin 3t\cos 100t$$

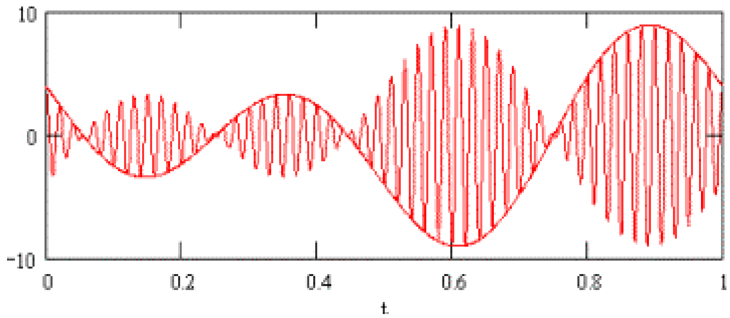

Figure 11 - The modulated signal and its envelope

The envelope of the modulated signal is the information signal

Let's take the Hilbert Transform of this signal. But before we do that we need to simplify the above so we only have sinusoids and not their products. This step will make it easy to compute the Hilbert Transform. By using these trigonometric relationships,

$$\sin A \cos B = \frac{\sin(A+B) + \sin(A-B)}{2}$$

$$\cos A \cos B = \frac{\cos(A+B) + \cos(A-B)}{2}$$

we rewrite the above signal as

$$g(t) = 2\cos(2+100)t + 2\cos(2-100)t - 3\sin(3+100)t - 3\sin(3-100)t$$

Now we take the Hilbert Transform of each term and get

$$\hat{g}(t) = 2\sin(2+100)t + 2\sin(2-100)t + 3\cos(3+100)t + 3\cos(3-100)t$$

Now create the analytic signal by adding the original signal and its Hilbert Transform.

$$g_+(t) = g(t) + j\hat{g}(t)$$

$$\begin{aligned} g_+(t) = {} & 2\cos(2+100)t + 2\cos(2-100)t - 3\sin(3+100)t - 3\sin(3-100)t \\ & + j(2\sin(2+100)t + 2\sin(2-100)t + 3\cos(3+100)t + 3\cos(3-100)t) \end{aligned}$$

Let's once again rearrange the terms in the above signal

$$\begin{aligned} g_+(t) = {} & 2\cos(2+100)t + j2\sin(2+100)t \\ & + 2\cos(2-100)t + j2\sin(2-100)t \\ & - 3\sin(3+100)t + j3\cos(3+100)t \\ & - 3\sin(3-100)t + j3\cos(3-100)t) \end{aligned}$$

Recognizing that each pair of terms is the Euler's representation of a sinusoid, we can now rewrite the analytic signal as

$$g_+(t) = (4\cos 2t - 6\sin 3t)\, e^{j100t}$$

But wait a minute, isn't this the original signal and the carrier written in the complex exponential? So why all the calculations just to get the original signal back?

Now let's take the Fourier Transform of the analytic signal and the complex envelope we have computed to show the real advantage of the complex envelope representation of signals.

Spectrum of the Complex Envelope **Spectrum of the Analytic Signal**

Figure 12 - The Magnitude Spectrum of the Complex Envelope vs. The Analytic Signal

Although this was a passband signal, we see that its complex envelope spectrum is centered around zero and not the carrier frequency. Also the spectral components are double those in figure 10b and they are only on the positive side. If you think the result looks suspiciously like a one-sided Fourier transform, then you would be right.

We do all this because of something Nyquist said. He said that in order to properly reconstruct a signal, any signal, baseband or passband, needs to be sampled at least two times its highest spectral frequency. That requires that we sample at frequency of 200.

But we just showed that if we take a modulated signal and go through all this math and create an analytic signal (which by the way does not require any knowledge of the original signal) we can separate the information signal the baseband signal s(t)) from the carrier. We do this by dividing the analytic signal by the carrier. Now all we have left is the baseband signal. All processing can be done at a sampling frequency which is 6 (two times the maximum frequency of 3) instead of 200.

The point here is that this mathematical concept help us get around the signal processing requirements by Nyquist for sampling of bandpass systems.

The complex envelope is useful primarily for passband signals. In a lowpass signal the complex envelope of the signal is the signal itself. But in passband signal, the complex envelope representation allows us to easily separate out the carrier.

Take a look at the complex envelope again for this signal

$$g_+(t) = (4\cos 2t - 6\sin 3t)\, e^{j100t}$$ the analytic signal

$$g(t) = 4\cos 2t - 6\sin 3t$$ the complex envelope

We see the advantage of this form right away. The complex envelope is just the low-pass part of the analytic signal. The analytic signal low-pass signal has been multiplied by the complex exponential at the carrier frequency. The Fourier transform of this representation will lead to the signal translated back down the baseband (and doubled with no negative frequency components) making it possible to get around the Nyquist sampling requirement and reduce computational load.

mntcastle@earthlink.net