The Hilbert Transform and Empirical Mode Decomposition as Tools for Data Analysis

Susan Tolwinski
First-Year RTG Project
University of Arizona Program in Applied Mathematics
Advisor: Professor Flaschka
Spring 2007

Abstract

In this paper, I introduce the Hilbert transform, and explain its usefulness in the context of signal processing. I also outline a recent development in signal processing methods called Empirical Mode Decomposition (EMD) which makes use of the Hilbert transform. Finally, I illustrate EMD by using the method to analyze temperature data from western Massachusetts.

1 Real Signals and the Hilbert Transform

1.1 Definition of the Hilbert Transform from Contour Integration

The Hilbert Transform and its inverse relate the real and imaginary parts of a complex function defined on the real line. The relationship given by this operation is easily derived by the application of Cauchy’s Integral Theorem to a function \( f(z) \) which is analytic in the upper half-plane, and which decays to zero at infinity. For a point \( z^* \) inside the contour depicted in Fig. (1), Cauchy’s theorem tells us that

\[
\int_\Gamma \frac{f(z)}{z - z^*} dz = \frac{1}{2\pi i} \int \frac{f(z)}{z - z^*} dz
\]

Figure 1: Contour for integral 2.
Writing $z^*$ as the sum of its real and imaginary parts, and the integral as the sum of integrals along the semi-circle of radius $R$ and the real interval $[-R, R]$, Cauchy’s Theorem becomes

$$f(z^*) = \frac{1}{2\pi i} \int_{-R}^{R} \frac{f(z)}{x - x^*} \, dx + \frac{1}{2\pi i} \int_{C_R} \frac{f(z)}{x - x^* + iy} \, dz$$  \hspace{1cm} (2)$$

As $R \to \infty$, the second term drops out (as $F$ vanishes by assumption). Now, if we replace $z^*$ with its conjugate in Eqn. (2), the integral yields zero since the pole created at $\bar{z}^*$ is not contained within the contour. Thus we can add such a term to our integrand with impunity, so that

$$f(z^*) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{f(z)}{x - x^* - iy} \, dx + \frac{f(z)}{x - x^* + iy}$$  \hspace{1cm} (3)$$

Finally, we rewrite $f(z)$ as the sum of its real and imaginary components, $f(x, y) = u(x, y) + iv(x, y)$, and take $y^* = 0$, since the aim is to relate $u(x, y)$ and $v(x, y)$ on the real line. For $u(x, 0)$, $v(x, 0)$ integrable and differentiable on $\mathbb{R}$,

$$f(z^*) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{u(x, y) + iv(x, y)}{x - x^* - iy} \, dx + \frac{u(x, y) + iv(x, y)}{x - x^* + iy}$$  \hspace{1cm} (4)$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{(u(x, y) + iv(x, y))(x - x^*)}{(x - x^*)^2 + y^2} \, dx$$  \hspace{1cm} (5)$$

$$\Rightarrow f(x^*) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{u(x, 0) + iv(x, 0)}{x - x^*} \, dx$$  \hspace{1cm} (6)$$

$$u(x^*, 0) + iv(x^*, 0) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{v(x, 0)}{x - x^*} \, dx - \frac{i}{\pi} P.V. \int_{-\infty}^{\infty} \frac{u(x, 0)}{x - x^*} \, dx$$  \hspace{1cm} (7)$$

Identifying real and imaginary parts, this gives

$$\Rightarrow u(x^*, 0) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{v(x, 0)}{x - x^*} \, dx$$  \hspace{1cm} (8)$$

$$v(x^*, 0) = -\frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{u(x, 0)}{x - x^*} \, dx$$  \hspace{1cm} (9)$$

This last result shows that on the real line, one need only specify the real component of an analytic function to uniquely determine its imaginary part, and vice versa. Note that because of the singularity at $x^*$, the integral is defined in the sense of principal values. Hence the identity is well-defined for functions which are integrable and differentiable.

We take a small leap of abstraction, and in the spirit of the preceding derivation, define the Hilbert transform of a real function of a real variable, $\phi(x)$:

**Definition** Let $\phi(x)$ be integrable and differentiable on $R$. Then the **Hilbert transform of $\phi(x)$**, denoted $H\{\phi(x)\}$, is given by

$$H\{\phi(x)\} = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{\phi(x)}{\xi - x} \, d\xi$$  \hspace{1cm} (10)$$

Comparing this definition to eqn.(8), we see that if $\phi$ is the imaginary part of some analytic function on the real line, then $H\{\phi(x)\}$ gives the real component of the same function on the real line. It seems natural to suspect a form like eqn.(9) will give the inverse Hilbert transform; we will see later on that this is indeed the case.
1.2 Properties of the Hilbert Transform and Inverse

The form of definition (10) suggests an alternate view of the Hilbert transform of \( \phi(x) \): as a convolution of \( \phi(x) \) with \( -\frac{1}{\pi x} \). Using the convolution theorem on this definition provides some intuition as to the action of the Hilbert transform on a function in the frequency domain. Recall that Fourier transforms are useful for computing convolutions, since the transform of the convolution is equal to the product of the transforms. Let \( F\{\cdot\} \) denote the Fourier Transform operation. In this paper, we adopt the convention of engineers and define

\[
F\{f(t)\} = \hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \\
F^{-1}\{\hat{f}(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t} d\omega
\]

Then

\[
F\{H\{\phi(t)\}\} = F\left\{\frac{1}{\pi} \int_{-\infty}^{\infty} \phi(\xi) \frac{d\xi}{\xi - t}\right\} \\
= F\left\{\phi(t) * \left(-\frac{1}{\pi t}\right)\right\} \\
= \hat{\phi}(\omega)F\left\{-\frac{1}{\pi t}\right\}
\]

But what is the Fourier transform of \( -\frac{1}{\pi t} \)? It is worth memorizing that \( F\{-\frac{1}{\pi t}\} = i\text{sgn}(\omega) \). To see this, consider the inverse transform of \( \text{sgn}(\omega) \). This function can be written as \( 2H(\omega) - 1 \), with \( H \) the Heaviside distribution. Then we can approximate the inverse transform by the sequence of distributional actions

\[
F^{-1}\{2H(\omega) - 1\} = \frac{2}{2\pi} \int_{0}^{\infty} e^{-\epsilon\omega} e^{i\omega t} d\omega - \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} d\omega \\
= \frac{1}{\pi} \frac{e^{i\epsilon(\omega + it)}}{\omega - i(t + i\epsilon)} \bigg|_{0}^{\infty} - \delta(t) \\
= \frac{i}{\pi} \left(\frac{1}{t + i\epsilon} - \delta(t)\right)
\]

Now, taking the limit of this expression as \( \epsilon \to 0 \), we should obtain the expression whose Fourier transform is the signum function. The first term gives

\[
\lim_{\epsilon \to 0} \frac{1}{\pi} \frac{1}{t + i\epsilon} = \frac{i}{\pi} \left(\frac{1}{t} - i\pi\delta(t)\right)
\]

so that

\[
F^{-1}\{2H(\omega) - 1\} = \frac{i}{\pi} \frac{1}{x + i0} \equiv \frac{i}{\pi} \left(\frac{1}{t} - i\pi\delta(t)\right) - \delta(t) \\
= \frac{i}{\pi t}
\]

This confirms that taking the Hilbert transform of a function in \( t \) is equivalent to multiplying the function’s transform by \( i \) times its sign at every point in the frequency domain. In other words, the Hilbert transform rotates the transform of a signal by \( \frac{\pi}{2} \) in frequency space either clockwise or counter-clockwise, depending on the sign of \( \omega \).
A related fact: like the Fourier transform, the Hilbert transform preserves the "energy" of a function. Define a function’s energy by

\[ E_f = \int_{-\infty}^{\infty} |f(t)|^2 dt \]  

(22)

Since the Hilbert transform just multiplies the Fourier transform by \(-i \text{sgn}(\omega)\), it clearly leaves the energy of the transform unchanged. By Parseval’s theorem, we know that \( E_\hat{f} = E_f \), and so \( E_{H(f)} = E_f \). Some other useful properties of the Hilbert transform are easily verified:

- The Hilbert transform and its inverse are linear operations.
- An extension of the property described above is that \( F \{ H^n f(t) \} = (i \text{sgn}(\omega))^n \hat{f}(\omega) \)
- Taking multiple Hilbert transformations of a function reveals the following identities: \( H^2 = -I \Rightarrow H^4 = I \Rightarrow H^2 = H^{-1} \)

From this last property, the definition of the inverse Hilbert transform is apparent:

**Definition** The inverse Hilbert transform of \( \psi(x) \), denoted \( H^{-1}\{\psi(x)\} \), is given by

\[ H^{-1}\{\psi(x)\} = -\frac{1}{\pi} \text{P.V.} \int_{-\infty}^{\infty} \frac{\psi(\xi)}{\xi-x} d\xi \]  

(23)

As suggested by the result in eqn.(9), we now see that the inverse transform of \( \psi(x) \) indeed gives the imaginary part on \( \mathbb{R} \) of an analytic function with real part \( \psi(x) \) on \( \mathbb{R} \).

### 1.3 Real and Analytic Signals

It is necessary to introduce some terminology from signal processing:

- A **signal** is any time-varying quantity carrying information we would like to analyze.

- To engineers, an **analytic signal** is one with only positive frequency components. For such signals, it is possible (we shall soon see) to give a representation of the complex form

\[ \Phi(t) = A(t) e^{i\theta(t)} \]  

(24)

where the original signal is the projection of the complex signal onto the real axis.

The engineering notion analyticity is closely related to the mathematician’s notion of what it means for a function to be analytic. For a function with a strictly positive frequency spectrum,

\[ f(t) = \frac{1}{2\pi} \int_0^{\infty} \hat{f}(\omega) e^{it\omega} d\omega \]  

(25)

If we wish to extend this function on \( R \) to the entire complex plane, extend \( t \to z = (t + is) \in \mathbb{C} \). Then the function on the complex plane may be expressed

\[ f(z) = \int_0^{\infty} \hat{f}(\omega) e^{i(t+is)\omega} d\omega \]  

(26)

\[ = \int_0^{\infty} \hat{f}(\omega) e^{-s\omega} e^{it\omega} d\omega \]  

(27)

Thus, it is plausible that the integral will return an analytic function so long as \( s > 0 \). In other words, a function with a positive frequency spectrum will have an analytic extension in the upper half-plane. So an engineer’s analytic function corresponds to an extended function which is analytic to the mathematician for \( \{ \text{Im}(z) > 0 \} \).
• Once we have such an analytic representation of the form \( A(t)e^{i\theta(t)} \), we can talk about the instantaneous phase \( \theta(t) \) of the signal, and its time-varying amplitude \( A(t) \).

• The **instantaneous frequency** is the rate of change of the phase, \( \theta'(t) \). We can compute the expected value of the frequency in the time domain by

\[
< \omega > = \int_{-\infty}^{\infty} \omega(t)|\Phi(t)|^2 dt
\]

\[
= \int_{-\infty}^{\infty} \omega(t)A^2(t)dt
\]

On the other hand, we may also calculate it in frequency space:

\[
< \omega > = \int_{-\infty}^{\infty} \omega|\hat{\Phi}(\omega)|^2 d\omega
\]

\[
= \int_{-\infty}^{\infty} \omega \hat{\Phi}(\omega)\hat{\Phi}^*(\omega)d\omega
\]

\[
= \int_{-\infty}^{\infty} F^{-1}\{\omega \hat{\Phi}(\omega)\} \Phi^*(t)dt \quad \text{(by Parseval’s Theorem)}
\]

\[
= -i \int_{-\infty}^{\infty} \Phi^*(t)\Phi'(t)dt
\]

\[
= \int_{-\infty}^{\infty} \left( \theta'(t) - i \frac{A'(t)}{A(t)} \right) |A(t)|^2 dt
\]

Now, since \( A(t) \) is real and \( L^2(R) \), the second integrand forms a perfect derivative which vanishes when integrated over the whole real line. Thus from the analyses carried out in both time and frequency space, we may make the identification \( \omega(t) = \theta'(t) \).

• The square of the **instantaneous bandwidth**, which measures the signal width in frequency space, is given by

\[
B^2 = \left( \frac{A'(t)}{A(t)} \right)^2 + \left( < \omega^2 > - < \omega >^2 \right)
\]

Bandwidth is defined in the frequency domain as the variance of frequency of a signal \( s(t) = A(t)e^{i\theta(t)} \). We can manipulate this expression and take inverse Fourier transforms to get the expression for bandwidth in the time domain as follows:

\[
B^2 = \int (\omega - < \omega >)^2 |\hat{s}(\omega)|^2 d\omega
\]

\[
= \int \omega^2 |\hat{s}(\omega)|^2 d\omega - < \omega >^2
\]

\[
= \int (\omega \hat{s})(\omega \hat{s})^* d\omega - < \omega >^2
\]

\[
= \int \frac{d}{dt}A(t)e^{i\theta(t)} \frac{d}{dt}A(t)e^{-i\theta(t)} dt - < \omega >^2
\]

\[
= \int \left( A'(t) + iA\theta'(t) \right) \left( A'(t) - iA(t)\theta'(t) \right) dt - < \omega >^2
\]

\[
= \int \left( \frac{A'(t)}{A(t)} \right)^2 |s(t)|^2 + \theta'^2(t)|s(t)|^2 dt - < \omega >^2
\]

\[
= \int \left( \frac{A'(t)}{A(t)} \right)^2 |s(t)|^2 dt + \left( < \omega^2 > - < \omega >^2 \right)
\]
Signals in nature are, of course, real. But it is analytic signals, with strictly positive frequency spectra, that are easy to process. How do the spectra of real signals compare? Let \( \hat{f}(\omega) \) and \( \hat{f}^\ast(\omega) \) be the Fourier transform of \( f(t) \) and its complex conjugate. We can recover \( f(t) \) from its Fourier transform by taking the inverse transform, and obtain \( f^\ast(t) \) by conjugating the resulting expression for \( f(t) \):

\[
f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t} d\omega
\]

\[\Rightarrow \quad f^\ast(t) = (f(t))^* = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{-i\omega t} d\omega\] (43)

On the other hand, we can also find \( f^\ast(t) \) from its own Fourier transform:

\[
f^\ast(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}^\ast(\omega)e^{i\omega t} d\omega
\]

(44)

Making the change of variables \( \omega \to -\omega \) in this last expression, we obtain

\[
f^\ast(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}^\ast(-\omega)e^{-i\omega t} d\omega
\]

(45)

Setting the two expressions for \( f^\ast(t) \) equal then gives the important result

\[
\hat{f}(\omega) = \hat{f}^\ast(-\omega)
\]

(46)

In other words, the negative frequency spectrum of a real signal is identical to the positive spectrum of its conjugate. This important finding allows us to represent any real signal in terms of just its positive frequencies:

\[
f(t) = \frac{1}{2\pi} \int_{0}^{\infty} \hat{f}(\omega)e^{i\omega t} + \hat{f}^\ast(\omega)e^{i\omega t} d\omega
\]

(47)

This construction of \( f \) as an analytic signal is useful, but requires prior knowledge of the Fourier transforms of the signal.

A further powerful construction to this end that uses the Hilbert transform is that of the analytic extension of a real signal. This is a function with identical positive spectrum to \( f \), modulo a constant factor, but with no negative frequencies. Call this function \( \Phi(t) \), and construct it in frequency space by

\[
\hat{\Phi}(\omega) = \frac{1}{2}(\hat{f}(\omega) + \text{sgn}(\omega)\hat{f}(\omega))
\]

(48)

The spectra of a function \( f \) and it’s analytic extension \( \phi \) are depicted below:

Now we can take the inverse Fourier transform to get \( \Phi(t) \). Since we’ve meddled with the symmetry of the spectrum, taking the inverse transform must now yield a complex result. Indeed,

\[
\Phi(t) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \left( \hat{f}(\omega) + \text{sgn}(\omega)\hat{f}(\omega) \right)e^{i\omega t} d\omega
\]

(49)

\[
= \frac{1}{2}f(t) + \frac{1}{2}F^{-1}\{\text{sgn}(\omega)\hat{f}(\omega)\}
\]

(50)

\[
= \frac{1}{2}f(t) + \frac{1}{2}F^{-1}\{\text{sgn}(\omega)\} * f(t) \quad \text{(Convolution Theorem)}
\]

(51)

\[
= \frac{1}{2}f(t) + \frac{i}{2\pi t} * f(t)
\]

(52)

But recall from eqn. (10) that the convolution of a function with \( \frac{1}{\pi t} \) defines the Hilbert transform of that function. Thus we can write

\[
\Phi(t) = \frac{1}{2}(f(t) + iH\{f(t)\})
\]

(53)
We have thus derived a way use the Hilbert transform to construct an analytic signal \( \Phi(t) \) whose real part is the original signal \( f(t) \). As promised earlier, the analytic extension \( \Phi \) can be represented as a time-varying amplitude and phase, \( \Phi(t) = A(t)e^{i\theta(t)} \), by defining

\[
A(t) = \sqrt{\left(\frac{1}{2}f(t)\right)^2 + \left(\frac{1}{2}H[f(t)]\right)^2} \tag{54}
\]
\[
\theta(t) = \tan^{-1}\left(\frac{H[f(t)]}{f(t)}\right) \tag{55}
\]

From the standpoint of signal processing, we have cast our signal in a form that is now significantly easier to analyze. Constructing a representation of the signal with only positive frequencies allows us to talk meaningfully about signal averages and spreads in frequency space (instead of calculating zero and a huge spread between a frequency peak and its symmetric reflection across \( \omega = 0 \)).

The extended signal \( \Phi(t) \) of a real signal also has the interesting property that its real and imaginary components are orthogonal with respect to the inner product defined by

\[
\langle f(t), g(t) \rangle = \int_{-\infty}^{\infty} f(t)g^*(t)dt \tag{56}
\]

Since the real and imaginary parts of \( \Phi(t) \) are \( f(t) \) and \( Hf(t) \), one can prove the assertion of orthogonality by using Placherel’s relation as follows:

\[
\langle f(t), Hf(t) \rangle = \int_{-\infty}^{\infty} f(t)(Hf(t))^*dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)(-\text{sgn}(\omega)\hat{f}(\omega))^*d\omega \tag{57}
\]
\[
= \frac{i}{2\pi} \int_{-\infty}^{\infty} \text{sgn}(\omega)\hat{f}(\omega)(\hat{f}(\omega))^*d\omega \tag{58}
\]
\[
= \frac{i}{2\pi} \int_{-\infty}^{\infty} \text{sgn}(\omega)\left|\hat{f}(\omega)\right|^2 d\omega \tag{59}
\]
\[
= 0 \text{ (Since the spectrum of real signal is even)} \tag{60}
\]

2 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a cutting-edge method for the analysis of non-stationary signals. Recently developed by N. Huang [5], the method decomposes a signal into a sum of components, each with
slowly varying amplitude and phase. Once one has a signal represented in such a form, one may analyze the properties of each component using the Hilbert Transform. EMD is an adaptive method; that is, the decomposition it produces is specific to the signal being analyzed. In theory, the components it outputs should separate phenomena occurring on different time scales.

Each component of the EMD is called an Intrinsic Mode Function (IMF). As modes, we would like the IMFs to satisfy two criteria, so that they will resemble a generalized Fourier decomposition:

- An IMF may only have one zero between successive extrema.
- An IMF must have zero "local mean."

In the second criterion, the definition of what constitutes the local mean is somewhat vague. For some $\epsilon$, we can write the requirement as

$$TMF_\epsilon(t) = \frac{1}{\epsilon} \int_{t-\frac{\epsilon}{2}}^{t+\frac{\epsilon}{2}} IMF(r) dr = 0$$  \hspace{1cm} (61)

In practice, the scientist is free to define $\epsilon$ as she deems appropriate to her particular application. The choice determines to some extent how finely the IMFs will be resolved by essentially determining a frequency above which fluctuations will be averaged out. Obviously, the scientist should choose $\epsilon$ no greater than any time scale on which she believes interesting information will be contained.

The computation of IMFs which satisfy the two criteria and sum to create the signal goes as follows. A candidate for the highest-frequency IMF is determined by first fitting a cubic spline through all local maxima to create an upper envelope. A lower envelope is constructed in the same manner. Together, the envelopes form a candidate for the time-varying amplitude of the IMF. However, we want to ensure that the component function will have negligible local mean. The practitioner has another choice here in deciding the threshold above which the process will not accept a candidate IMF. If the average of the two envelopes does not fall uniformly within the threshold value of zero, we subtract the mean from the envelopes. The construction is repeated until the second criterion is satisfied. The result of this inner loop produces the first IMF.

Subtract the IMF from the original signal; the difference is the residue for the outer loop of the process. The construction of an IMF is now repeated for the residual signal. This EMD procedure is repeated until the residue is either constant or monotone. This final residue should reveal any trend that exists in the data.

Schematically, a code for this procedure for a signal $s(t)$ would be structured as follows:

```plaintext
Residue = s(t)
I_1(t) = Residue
i = 1, k = 1
while Residue \neq 0 or Residue is not monotone
  while I_i has non-negligible local mean
    U(t) = spline through local maxima of I_i
    L(t) = spline through local minima of I_i
    Av(t) = \frac{1}{2} (U(t) + L(T))
    I_i(t) = I_i(t) - Av(t)
    i = i + 1
  end
  IMF_k(t) = I_i(t)
  Residue = Residue - IMF_k
  k = k+1
end
```

This procedure results in the decomposition

$$s(t) = \sum_i (IMF_i) + \text{Residue}$$  \hspace{1cm} (62)
A few pictures may be worth a thousand words in this instance. Four successive iterations of the inner loop toward the construction of just one IMF for a signal are illustrated below. The top and bottom splines through the signal maxima and minima form the envelopes, and the heavy curve in between is the mean signal which gets subtracted from the signal in each case, leaving the residue shown in Fig.(3) (figures from [7]).

By construction, one expects each IMF to be locally orthogonal to the others. Each is constructed by a difference of the residual signal and the local mean of the residual signal (i.e. the residue on the $k^{th}$ step minus the final $I_i$ for a given $k$). Let $x(t)$ be the residue of the signal on the $k^{th}$ step, and let $\bar{x}(t)$ denote its local mean,

$$\bar{x}(t) = \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} x(\tau) d\tau.$$  \hspace{1cm} (63)

Then we expect the local inner product to give zero:

$$< IMF_i, \bar{x}(t) > = \int_{t-\frac{\epsilon}{2}}^{t+\frac{\epsilon}{2}} (x(\tau) - \bar{x}(t)) \bar{x}(t) d\tau$$  \hspace{1cm} (64)

$$= \bar{x}(t) \int_{t-\frac{\epsilon}{2}}^{t+\frac{\epsilon}{2}} x(\tau) d\tau - \bar{x}^2(t) \int_{t-\frac{\epsilon}{2}}^{t+\frac{\epsilon}{2}} d\tau$$  \hspace{1cm} (65)

$$= \epsilon \bar{x}^2(t) - \epsilon \bar{x}^2(t)$$  \hspace{1cm} (66)

$$= 0$$  \hspace{1cm} (67)

The $jth$ intrinsic mode function, with $j < i$, has already been subtracted off the original signal to produce $x(t)$. For $j > i$, IMF$_i$ will be subtracted from the signal before computation of IMF$_j$. By the calculation
above, then the set of Intrinsic Mode Functions into which the signal is decomposed form a locally orthogonal set. However, those calculations are not strictly true, since the IMF is not computed as a true local mean as in eqn.(63), but via envelopes instead. Thus the inner product above will likely be some small nonzero number. Include the residue as the $n+1$st IMF, and write the reconstructed signal as their sum. Denote the reconstruction of signal $s(t)$ by capital $S(t)$. Then its square is

$$S^2(t) = \sum_{j=1}^{n+1} IMF_j^2(t) + 2 \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} (IMF_i(t))(IMF_j(t))$$

(68)

The modes form a complete basis by construction (their sum reconstructs the original signal). Therefore, if they are orthogonal, the cross terms should vanish when integrated over time as in Parseval’s Inequality. The index of orthogonality provides a measure of how close the modes come to this by

$$I.O. \equiv \sum_t \left( \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} (IMF_i(t))(IMF_j(t)) \right) \frac{s^2(t)}{s^2(t)}$$

(69)

where the sum over time is over the length of the whole data series.

We can get a feeling for how well the EMD separates time scales by running the process on a test signal with known components. Consider a superposition of two sines of different frequencies and a "chirp" signal, which oscillates with an increasing frequency in time:

$$a(t) = \cos(t^2)$$

(70)
$$b(t) = 3\cos(3t)$$

(71)
$$c(t) = \cos(t)$$

(72)
$$\text{SIGNAL}(t) = a(t) + b(t) + c(t)$$

(73)

The resulting signal and its decomposition are depicted below in Fig.(4):

![Figure 4: Test run of Empirical Mode Decomposition on a known signal.](image)

Clearly, the EMD returns component IMFs of a very close visual match to the original signal components, with nearly indistinguishable amplitudes and frequencies. In this case, the absolute value of the difference between the reconstructed and original signals is $O(10^{-15})$. 

10
2.1 The Hilbert Spectrum.

Suppose we have an Empirical Mode Decomposition of a signal \( f(t) \) into a set of component IMFs, \( \{\text{IMF}_i(t)\}_{i=1}^n \). If we compute the Hilbert transform of each IMF and construct a corresponding analytic signal for each, then we have \( n \) complex functions

\[
\begin{align*}
  z_1(t) &= \text{IMF}_1(t) + iH\{\text{IMF}_1(t)\} = A_1(t)e^{i\phi_1(t)} \\
  z_2(t) &= \text{IMF}_2(t) + iH\{\text{IMF}_2(t)\} = A_2(t)e^{i\phi_2(t)} \\
  &\quad \vdots \\
  z_n(t) &= \text{IMF}_n(t) + iH\{\text{IMF}_n(t)\} = A_n(t)e^{i\phi_n(t)}
\end{align*}
\]

The expansion of the signal in terms of these functions,

\[
f(t) = \sum_{j=1}^n A_j(t)e^{i\phi_j(t)}
\]

may be viewed as a generalized Fourier series, where the frequency and amplitude are no longer constrained to be constant. Unlike a Fourier series or transform, this representation allows a simultaneous understanding of the signal in both frequency and time. One may construct a three-dimensional plot of the amplitude at a given time and frequency (i.e., plot the amplitude on top of the time-frequency plane); the resulting distribution is called the Hilbert Spectrum of the signal, \( H(\omega,t) \). The Hilbert spectrum provides an intuitive visualization of what frequencies occurred when during the signal duration, and also shows at a glance where most of the signal energy is concentrated in time and frequency space. For the test signal of the previous section, the Hilbert spectrum is shown in the color graph below (Fig. (5)). The ordering red-orange-yellow-white indicates increasing amplitudes with black indicating zero amplitude. Each Intrinsic Mode Function is clearly identifiable as a continuous curve on the spectrum. For example, the curve of rising frequency is clearly due to the energy contained in the "chirp" component \( a(t) \), which has increasing frequency in time.

2.2 Example Application: Analysis of Temperature Data

Empirical Mode Decomposition presents a method to analyze variations in time series on time scales intrinsic to the particular data series. The growing recent concern over climate change provides an interesting setting in which to apply this decompositional method. Perhaps EMD can help provide hard answers to questions about whether apparent shifts in the weather are truly trends, or can be written off as natural cyclic climatic variations.

To illustrate what can be gleaned from these methods, I analyze daily temperature data from NOAA taken in Amherst, Massachusetts from 1988 to 2005 [8]. In what follows, all vertical axes have units of degrees Fahrenheit, unless otherwise marked.

The data series is shown in Fig. (6), with the temporal axis given by the number of days since January 1, 1988. For such a data series this long, with fluctuations occurring on time scales much shorter than the length of the entire series, we expect many IMFs in our decomposition. Indeed, the empirical mode decomposition returns 12 modes (including the residue) for this function (Fig. (14), appended).

Of course, an observer will most likely not notice the oscillations of individual temperature modes due to interference between modes; it is their superposition which gives the actual temperature. In theory, however, each mode should have the nice interpretation as setting the scale on which a resident of Amherst, Massachusetts can expect temperature fluctuations about some mean on various time scales. Consider, for example, the eighth empirical mode function as shown in Fig. (7). This mode has very regular oscillations on yearly time scales. Recall that the mode was constructed to have mean zero. Thus, the eighth mode has...
the obvious interpretation that over yearly time scales, one may expect regular fluctuations on the order of 50 degrees Farenheit about some mean. Of course, the peaks in these fluctuations line up with the summer dates in the data set, and the troughs with winter, and one should have expected to see this behavior in one of the lower-frequency modes even before computation.

We should be able to learn new information from modes with less obvious physical interpretations. For this data set, IMF 5 is an example of one such instance (Fig. 8). The period of the oscillations given by this function is on the order of two weeks. The envelope of IMF 5 is itself almost periodic; the amplitude appears
generally to be greater during the times corresponding to winter, and about halves its size during the peaks of summer heat. We can conclude that on time scales of about a half a month, there is greater variation in the temperature in Amherst, MA during the winter than in the summer. In a full scientific study of this data, it would be interesting to investigate whether this correlates somehow with humidity or other factors.

![Figure 8: Close-up of IMF 5 compared with original series.](image)

The residue of an Empirical Mode Decomposition exhibits the overall trend in the data, if there is one. For the temperature data from Amherst, we see that all oscillatory modes are superposed on a monotone increasing base function. We can interpret the residue as a slow increase in the overall average temperature from 1988 to 2005 by about a half a degree Fahrenheit.

![Figure 9: Close-up of Residue.](image)

The error in the reconstructed signal, given by the absolute value of the pointwise difference between the original data series and the superposition of all IMFs, is of the order of machine epsilon (see Fig. (10)). The index of orthogonality for this decomposition is 0.0172, indicating a leakage of less than 2 percent of the overall signal energy during EMD. By both of these measures, the decomposition gives a very accurate representation of the data, indeed.

![Figure 10: Original Series and Reconstruction from IMFs.](image)

Not every IMF has a meaningful physical interpretation. The Huang-Hilbert spectrum of the decomposition can aid the scientist in determining which modes may be neglected. For our example of Amherst temperature data, the Huang-Hilbert spectrum makes it visually clear that the majority of the signal energy is contained in perhaps the bottom 25% of the frequencies represented in the full EMD (Fig. (11)).
Taking the Hilbert spectrum of the IMFs individually is helpful in identifying which of the modes contribute most strongly to the overall signal energy. Looking at the 7th IMF alongside its Hilbert spectrum (Fig.12), we confirm first of all that the dips in the Hilbert spectrum correspond to a decrease in frequency in the IMF, and that a broadening or shrinking of the amplitude causes a lightening or darkening of color in the spectrum. The spectrum of the 7th IMF also makes it visually clear that this mode contributes a very coherent signal to the overall data. Its spread in frequency space is narrow, and it traces out a nearly continuous curve in the joint time-frequency domain. It would appear then that the 7th IMF reveals an important piece of the dynamics of temperature in Amherst, MA.

Contrast the 7th IMF to one of the higher frequency modes, say, the 4th (Fig.(13)). It is not easily apparent from the plot of the mode in time whether the function is likely to contain meaningful information. However, a glance at its Hilbert spectrum reveals that this mode mainly describes noise, at least at the level of resolution afforded by the author’s run of the code on this data set.

A coarse-to-fine reconstruction confirms that we may neglect some of the highest frequency modes (See Fig.(15)). By the time the thirteenth down to about the 5th IMFs have been superposed, the reconstruction captures the most important features of the original data. Adding on the higher, noisier frequency modes thereafter makes the data look more complicated, but alters the overall character very little.

3 Conclusions.

We have seen that the Hilbert transform is an essential tool for conversion of signals into analyzable forms. Also, Huang’s Empirical Mode Decomposition is a powerful tool for separating phenomena that occur on different time scales from highly non-stationary signals. In application, the decomposition given by this process must be interpreted according to a thorough background knowledge of the science from which the signal arises. Finally, the Huang-Hilbert spectrum makes use of the Hilbert transform to provide a useful visual, qualitative understanding of a signal decomposed by EMD.

4 Acknowledgments.

I would like to thank Professor Flaschka for encouraging me to follow my own interests in the applications of this topic, taking the time to work right along side of me through some of its challenges, and for role modeling a genuine love for exploring new ideas. This kind of advising helped shape the kind of project that teaches how pleasurable the process of science can be.
Figure 12: 7th IMF and its Hilbert spectrum for Amherst temperature data.

Figure 13: 4th IMF and its Hilbert spectrum for Amherst temperature data.
References


Figure 14: Empirical Mode Decomposition of example data series.
Figure 15: Coarse-to-fine reconstruction of temperature signal from IMFs.