

On the Hilbert-Huang Transform Data Processing System Development

By Semion Kizhner, Thomas P. Flatley, Dr. Norden E. Huang,
Karin Blank, Evette Conwell
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt Road, Greenbelt MD, 20771
301-286-7029
tom.flatley@nasa.gov

Darrell Smith
Orbital Sciences Corporation

Abstract— One of the main heritage tools used in scientific and engineering data spectrum analysis is the Fourier Integral Transform and its high performance digital equivalent – the Fast Fourier Transform (FFT). The Fourier view of nonlinear mechanics that had existed for a long time, and the associated FFT (fairly recent development), carry strong a-priori assumptions about the source data, such as linearity and of being stationary. Natural phenomena measurements are essentially nonlinear and nonstationary. A very recent development at the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC), known as the Hilbert-Huang Transform (HHT) proposes a novel approach to the solution for the nonlinear class of spectrum analysis problems. Using the Empirical Mode Decomposition (EMD) followed by the Hilbert Transform of the empirical decomposition data (HT) [1], [2], [3], [15], the HHT allows spectrum analysis of nonlinear and nonstationary data by using an engineering a-posteriori data processing, based on the EMD algorithm. This results in a non-constrained decomposition of a source real value data vector into a finite set of Intrinsic Mode Functions (IMF) that can be further analyzed for spectrum interpretation by the classical Hilbert Transform. This paper describes phase one of the development of a new engineering tool, the HHT Data Processing System (HHTDPS). The HHTDPS allows applying the HHT to a data vector in a fashion similar to the heritage FFT. It is a generic, low cost, high performance personal computer (PC) based system that implements the HHT computational algorithms in a user friendly, file driven environment. This paper also presents a quantitative analysis for a complex waveform data sample, a summary of technology commercialization efforts and the lessons learned from this new technology development.

TABLE OF CONTENTS

INTRODUCTION

- 1.0 THEORETICAL FOUNDATION
 - 2.0 HHTDPS PROPOSAL OVERVIEW
 - 3.0 CARDINAL POINTS
 - 4.0 DEVELOPMENT PLAN
 - 5.0 TOP LEVEL ARCHITECTURE
 - 6.0 EMPIRICAL MODE DECOMPOSITION MODULE ARCHITECTURE
 - 7.0 INSTANTANEOUS FREQUENCY AND HILBERT SPECTRUM MODULE ARCHITECTURE
 - 8.0 GRAPHICAL USER INTERFACE ARCHITECTURE
 - 9.0 SYSTEM OPERATIONS AND TEST RESULTS
 - 10.0 TECHNOLOGY COMMERCIALIZATION
 - 11.0 LESSONS LEARNED FROM PROPOSAL DEVELOPMENT
- CONCLUSIONS
REFERENCES
BIOGRAPHY
ACKNOWLEDGEMENTS

INTRODUCTION

Oscillatory phenomena are omnipresent and the desire to learn about the *oscillatory behavior* of a *signal* is both natural and practically useful. Signal characteristics, the study of a signal by its decomposition into simpler components, the heritage signal processing method for linear systems and data, and the novel method that is applicable to non-linear and non-stationary data are presented in this introduction.

The *characteristic parameters* of an oscillatory signal are period T , oscillation fundamental frequency f_0 , amplitude a , phase $\theta(t)$, phase shift C within $\theta(t)$, polarization, as well as

signal average value, energy and power. These are described in detail in Section 1.5.

When dealing with a signal that is not susceptible to a brute force study of its oscillatory behavior, it is tempting to decompose the signal into simpler oscillatory behavior components (*analysis*) and in such a way that allows straightforward reconstruction of the source signal from the decomposition components (*synthesis*). The signal synthesis from a subset of components yields the source signal *approximation* and always related to an approximation *residue* affects. The decomposition, in turn, can use some set of basic functions (*basis*) for component representation, such as polynomials or trigonometric sine and cosine functions, or periodic exponential functions of a complex variable. Any *complete set* of functions can be used to represent arbitrary functions. There are also multitudes of ways to decompose *anything*, be to a simple signal with a constant amplitude or a complex non-linear function. For example, a signal with a constant amplitude in time $x(t)=\frac{3}{4}$ can be represented as a sum of squared trigonometric cosine functions of time (*components*), namely

$$\frac{3}{4} = x(t) = \cos^2(t-120) + \cos^2(t) + \cos^2(t+120), \text{ for all } t.$$

However, there is very little knowledge about this signal oscillatory behavior (or absence of it) available from such *arbitrary decomposition* into non-linear cosine-based components.

The *heritage spectrum* analysis method is based on the Fourier theoretics, a linear decomposition that is especially convenient when a signal originates within a *linear and stationary oscillatory process*. The Fourier series for periodic signals with a finite period $T>0$ or the Fourier Integral Transform (also called the Fourier Transform) for spectrum analysis of non-periodic linear and stationary functions is implemented for signals that satisfy the *Dirichlet criteria* which is described in Section 1.5. A signal with period $T=0$ is a signal with a constant amplitude in time. It has no nonzero frequency oscillatory components and is also fully covered by the Fourier series. This *linear decomposition* into sine and cosine components, employed by Fourier series and the Fourier Transform, is of obvious advantage since it provides a direct solution for a linear and stationary signal that satisfies the Dirichlet criteria, where it is proven to be applicable. For example, the Fourier Series for signal $x(t) = \frac{3}{4}$ that satisfies these conditions for period $T=0$ (number of discontinuities is 0, number of extrema points is 0 and the integral over period T is finite and equals 0), returns the constant component $\frac{3}{4}$ only which contains all knowledge about the signal oscillatory behavior, namely its absence. The exclusive heritage use of the Fourier series functions, the trigonometric functions of sine and cosine, also has the following three reasons [16]. Given that we want a *time invariant* representation of signals, since there is usually no natural origin of time, leads to trigonometric

functions that are the eigenfunctions of time translation. *Linear systems* also have the same eigenfunctions – the complex exponentials that are equivalent to the real trigonometric functions. The third good reason for the Fourier functions is that the synthesis of the band limited physical signal from equally spaced samples taken at a rate of at least twice higher than the signal's highest frequency is simple to understand as a consequence of the Nyquist sampling theorem.

Most natural phenomena are non-linear and non-stationary, and direct application of the Fourier spectrum analysis may lead to undesirable affects and unrelated physical interpretation. There is presently *no engineering tool* for systematic spectrum analysis and synthesis of nonlinear and nonstationary data. Nonlinear and nonstationary processes are *complex processes that evolve in time*, such as speech or music and their properties are statistically non-invariant over time. A theoretical attempt in this endeavor was proposed by Dennis Gabor in 1946 in his "Theory of Communication" as depicted in the book on Gabor analysis and *Gabor frames* [14]. However, an effective computational method using Gabor frames is yet to be developed.

Dr. Norden E. Huang recently proposed a novel *Empirical Mode Decomposition* (EMD) computational method for non-linear and non-stationary signals. The EMD output basis functions, the *Intrinsic Mode Functions* (IMFs), are derived from the data and are susceptible to the Hilbert Transform for spectrum analysis, the *Hilbert-Huang Transform* (HHT) [1], [2], [3]. The Implementation of the Digital Hilbert Transform is using the FFT. This paper describes the development of a novel engineering tool, the *HHT Data Processing System* that implements the HHT and allows a user to make use of HHT similar to the FFT for spectrum analysis of nonlinear and nonstationary data. The theoretical foundations of the HHT and the heritage FFT methods used in HHTDPS are discussed below.

1.0 THEORETICAL FOUNDATION

The theoretical foundation of HHTDPS comprises those of the HHT and the heritage Fourier Transform (used in evaluating the Hilbert Transform), as well as the discrete equivalents of these transforms, the Discrete Hilbert Transform (DHT) and the FFT. The FFT is used to implement the DHT.

The theoretical foundations of the HHT method are related to nonlinear and non-stationary data analysis and synthesis and also comprise the relation between the heritage Fourier Transform and the Hilbert Transform [1], [2]. Because nonlinear and non-stationary systems are *complex systems* we will give a brief overview of the nature of complex systems that facilitates insights into the HHT method.

1.1 On Complex Systems

The HHT method and its implementation, the HHTDPS, are applicable to non-linear and non-stationary systems and data that originate in such systems. These are complex systems and we discuss in brief the contemporary approach in studying them via empirical engineering and reduction as well as the view of their nature.

In *empirical engineering*, where observation is fundamental, the *complexities* of a phenomenon can be partially understood by repeated observations of a whole system under a wide range of circumstances; *all the variables are retained* and as many as possible are examined. By contrast in the *reductionist approach*, traditionally found in physics and engineering, experiments are specifically designed to simplify the study of a natural phenomenon by the elimination of all but a few variables and linear model explanation.

1.1.1 HHT and Complexity

Most phenomena are complex. A swirling vortex in a turbulent ocean cannot be expressed in terms of individual water molecules. All this waits for a model that will reveal the complex and time-dependent interaction of known and unknown variables. *For nonlinear systems the whole is greater than the sum of the parts*, and they can only be understood by examining “global” behaviors in addition to a detailed examination of the individual features of which they are comprised. The continued study of complexity as promulgated by the HHT and HHTDPS is based on the premise that even if it does not provide exact solutions, so is the case with other methodologies, and we are aware that no matter how many details are uncovered there will always be unknowns beyond the sum of current knowledge. *This paper describes a novel empirical approach of the HHT to the solution of spectrum analysis problems for nonlinear and nonstationary data that originate in complex systems. However, the HHT is also based on established theories of complex systems and signal processing.*

1.1.2 Complexity Ingredients

For complexity to emerge, two ingredients are necessary. The first, and foremost, is an *irreversible medium* [8] in which things can happen: this medium, for example, is *time*. The reason for stating the apparently obvious is that the laws of motion on the microscopic level do not distinguish one direction of time from another. Yet we know from the tendency of snowballs to melt that a preferred direction of time is singled out at the macroscopic level. *Time is the irreversible medium of our problem's complexity.* The second essential ingredient is *nonlinearity* [8]. *Nonlinear systems do not obey the simple rules of addition.* In general, nonlinearity of time-dependent processes produces complex and frequently unexpected results. Irreversibility and nonlinearity characterize phenomena in every field of science and engineering. Examples in medicine include the palpitations of a heart and the human brain, which in

particular, is an extreme example of complexity achieved by biological evolution. Analysis of complexity affords a holistic perspective and with it insights into many difficult concepts, such as life, consciousness, and intelligence, that have consistently eluded science. *It is our hope that the HHT and HHTDPS-based study and empirical spectrum analysis and synthesis of nonlinear and nonstationary phenomena are steps towards progress in this domain.*

1.1.3 The Language of Complexity

Sizing up the degree of complexity of a given problem is the mission of mathematical complexity theory. Because complex problems lie beyond the scope of pen, paper, and analytical mathematics and because of their immense power, *computers* provide the only means for solving them. To tackle complexity with a computer involves a combination of subtlety and brute force. Both are characteristics of the HHT and HHTDPS.

1.2 Theoretical Foundations of Signal Spectrum Analysis and Signal Synthesis

Like with many other cases in human endeavor we know some and desire to know more. For example, in trigonometry we are concerned with the problem of “solving a triangle” – finding all the sides and angles of a triangle, when some of these are known. In spectrum analysis we are concerned with a given a set of samples (*measurements*) of a complex phenomenon and are interested in deriving knowledge of its oscillatory behavior from these measurements. How much do we have to know about the data, and what is the minimal amount of information we must know to derive the spectrum? To approach this question we consider a few following basic issues related to oscillatory behavior.

1.2.1 Signal Oscillatory Behavior

Intuitively a signal oscillatory behavior comprises:

- a point of utmost instability or a local maxima extrema, or a waveform crest point followed by
- a zero-crossing point followed by
- a point of utmost stability or a local minima extrema. or a waveform throw or point of equilibrium
- a repetition of the above which determines the time scale of an oscillatory behavior.

Is it possible to derive an oscillatory behavior's *time scale* from a complex waveform using the *input physical signal maxima and minima extrema points*? In a linear system, all data measurements are used. In a complex nonlinear set of data measurements, more general information (extrema points) may be used, in addition to the whole data set, in order to derive the nonlinear and nonstationary features. Papers [1], [2], [3] propose to use the *upper and lower extrema envelopes* for this purpose in a process of successive *median* determination (sifting process) until the absolute value of the difference between the total number of extrema points and zero-crossings is less or equal to 1. This

process results in an *oscillatory behavior's time scale* or Intrinsic Mode Function (IMF).

The Fourier approach of finding a large number of harmonics has been replaced in HHT by a multi-step sifting process of finding an oscillatory behavior time scale basis function, an IMF. An IMF satisfies the oscillatory behavior criteria and is close to being a *symmetric function*. The symmetry will then contribute to the *IMF bases orthogonality*. It is intuitively obvious that such a process will separate the finest time scale from a complex waveform first. An example of such a separation of a complex signal $x(t)$ into its time scale components $x_1(t)$ (constant 1) and $x_2(t)$ (a sine function) is as follows:

$$\begin{aligned} x_1(t) &= 1 \\ x_2(t) &= \sin(t) \\ x(t) &= x_1(t) + x_2(t) = 1 + \sin(t), \\ \text{where } 0 &\leq t \leq 10 \text{ and time increment step is } 0.1. \end{aligned}$$

The first sift of the signal $s(t)$, where $s(t)=x(t)$ yields median constant $(2+0)/2=1$. Subtracting this median from signal $s(t)$ yields the pure sinusoid that is the first IMF₁ with period $T=1$ and fundamental frequency $f_0=1/T=1$. Subtracting this IMF₁ from physical signal $x(t)$ yields the constant residue $x_1(t)=1$ with period $T=0$; Whilst the full separation of time scales was achieved – from finest scale of $\sin(t)$ to the widest time scale of residue $x_1(t)=1$. It can be remarked that an *IMF is an eigenfunctions of the sifting process* - if we input to sifting an IMF the result is just the input IMF with scaling factor 1.

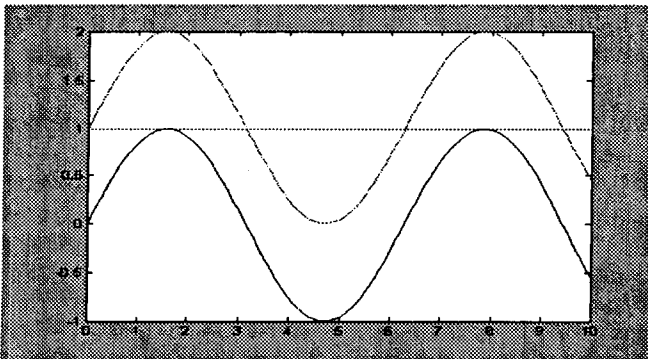


Figure 1. Separation of Oscillatory Scales by Sifting, cyan signal is $x_1(t)=1$, blue signal is $x_2 = \sin(t)$, green signal is $x(t) = \sin(t) + 1$

1.3 Frequency Spectrum

Frequency spectrum determination and interpretation are performed in a few heritage ways such as Fourier, Wavelets and Hilbert. We will consider in this paper the Fourier and Hilbert spectrum analysis and physical signal synthesis, and their relationship.

1.3.1 Fourier Spectrum Analysis

Spectrum analysis of a waveform function is difficult to perform in time domain. However, spectrum analysis is

straightforward in frequency domain and Fourier spectrum analysis provides an efficient transform from time to frequency domain, as well as the methodology for time signal synthesis from the frequency domain back to the time domain.

The Fourier spectrum analysis became a practical engineering tool with the advent of its computer-based implementation as a Discrete (also called Digital) Fourier Transform (DFT). The Fourier spectrum analysis allows one to describe a time domain waveform in both frequency-amplitude and frequency-phase domain in very clear terms. Because of this we give a short overview of the Fourier spectrum analysis.

Before asking the questions of what the Fourier Transform or Hilbert Transform are, it might be well to ask how they came about. Why were these mathematical tools developed?

1.3.2 On Waveforms and Fourier Spectrum

A waveform is a general function of time $x(t)$ that seemingly oscillates between some local maxima and minima extrema points or from a point of extreme instability (maxima) to a point of an equilibrium (minima) with a zero-crossing in-between in time. A waveform, if it can be described as a function, may be a periodic or a non-periodic function. A periodic waveform, in turn, may be a *basic sinusoid* function such as described by a trigonometric sine function $x(t) = \sin(t)$ or a cosine function $x(t) = \cos(t)$, or have a more complex form, but such that it still has a period T , the smallest time interval for which $x(t)=x(t+T)$ for any t . The period definition and other waveform *oscillation characteristic parameters* are elaborated below.

1.4 Oscillation Characteristic Parameters

We will describe the waveform oscillation characteristic parameters on the example of a plane waveform being a basic sinusoid and described by a counterclockwise circular motion $x(t)$ (pure oscillation or "tone") with angular argument ωt and beginning at angle C (phase shift).

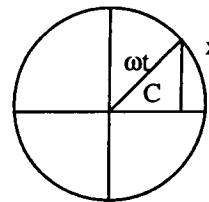


Figure 2. Circular Motion Parameters

Once the sinusoidal (circular) motion has been completely described with respect to time as time function $x(t)$, where

$$x(t) = a \cdot \sin(\omega \cdot t + C)$$

a new description of the signal with respect to frequency and phase shift can be obtained. This can be done in a

three-dimensional coordinate system with amplitude, time and frequency axes. This coordinate system comprises the amplitude-time plane (Oat) and the amplitude-frequency plane (Oaf). The sinusoid is passing through point $f_0=1/T$ on the frequency axis in a plane parallel to the amplitude-time plane. The projection of the sinusoid on the amplitude-time plane gives the time history of the sinusoid and its projection on the frequency plane takes the form of an impulse, a pulse of *instantaneous rise and fall*, with amplitude equal to the sinusoid's time plane amplitude a . Some additional information – the phase shift C is needed, however, to fix the sinusoid's position relative to the coordinate system zero time reference O [11], [12].

For a basic sinusoidal waveform the constant term of the phase offset, called phase shift C , provides the additional information needed to fix the waveform position in time in relation to coordinate frame origin O . The phase shift C for a sinusoid that was obtained from measurements rather than from a formula, where it may be given explicitly, can be determined by looking at the positive peak closest to time point of reference O , the zero time origin. Peak at zero is conventionally taken as phase shift 0 . When peak occurs after zero the sinusoid is *delayed and having negative phase, because it was expected before time 0*. When peak occurs before zero the sinusoid would have been *advanced and its phase is positive, because peak is expected at time 0 but occurs earlier*. This range $[-180, +180]$ is called *modulo 2π phase representation* as opposed to a continuous phase representation if a reference can be attached to the sinusoid.

For any electromagnetic wave the fourth (polarization) parameter that describes its electric field vector is needed to fully characterize the wave in 3-dimensional space.

The signal period T , frequency f_0 , its amplitude a , phase shift C are the sinusoid oscillation characteristic parameters for a plain waveform of one independent variable t . The basic sinusoid parameters f_0 , a , C in the frequency and phase domain can be then used to synthesize the waveform back into the time domain.

1.4.1 Angular Phase and Frequency Relationship

The basic waveform function's time-dependent variable argument ($\omega t + C$) is also called **angular phase**. It comprises two parts. The time-variable part ωt provides all information needed to determine the fundamental frequency f_0 . Namely,

$$\omega = 2\pi f_0$$

It is obvious that if we denote the angular phase as

$$\theta(t) = \omega t + C$$

then the phase derivative is the angular frequency ω

$$\theta'(t) = \omega.$$

In other words, in the simplest case the oscillation angular frequency ω can be determined as the derivate of angular phase, or phase change rate of a basic sinusoid waveform. The phase shift can then be expressed as

$$C = \theta(t) - t\theta'(t).$$

This derivation of the oscillation characteristic parameters from phase is strongly intuitive for a sinusoid because the phase is the total advancement of the waveform in time and a higher phase rate speeds up its amplitude zero-crossing points and giving a direct correlation between phase and frequency.

It is reasonable from this insight of frequency as phase derivative for the basic waveform to attempt this frequency determination approach to waveforms with more complex forms of theoretical phase or phase data obtained from measurements. For example, when phase shift is not a constant, but itself is some function of time, even the above sine function becomes a complex waveform with a very general phase function of time $\theta(t)$. Indeed, even for this general case, if $\theta(t)$ is a phase description, then consider the phase change rate at two successive time points t_1, t_2 :

$$(\theta(t_2) - \theta(t_1)) / (t_2 - t_1)$$

This average angular phase rate in time interval $\Delta t = [t_1, t_2]$, when Δt approaches 0 becomes $\theta'(t)$ or instantaneous frequency [4]-[5]. This representation of angular phase was given in the polar coordinate system. In a rectangular coordinate system (Oxy) a function $X(t)$ may be represented as a function of a complex variable:

$$X(t) = x(t) + j*y(t)$$

The amplitude and phase can correspondently be expressed as

$$a = \sqrt{x^2(t) + y^2(t)} \text{ and } \theta(t) = \arctang(y(t)/x(t)).$$

Nonsinusoidal waveforms, for example, even one generated by a circular motion with a more complex structure than a basic sinusoid for the angle phase, require a new look on angular phase and its susceptibility to yield information on motion frequency.

The cardinal problem is how to recover the phase of a complex waveform. The controversy is caused by different methods of phase recovery, and did not escape from the phase recovery provided by the Hilbert Transform. Phase recovery from function values is an inversion problem, and like many inversion problems, is extremely difficult to solve. The HHT theory alleviates this controversy and yields powerful spectrum results for complex waveforms [1]-[3]. Although it does not assume data linearity or stationarity, it

does use fundamental knowledge that the measurements data originated from an oscillatory phenomenon, that its spectrum exists, and that the zero-crossings are more fundamental to spectrum determination than amplitude, allowing data normalization to reduce the Bedrosian [6] and Nuttall [7] affects.

1.5 Heritage Fourier Spectrum Analysis Tools

Nonsinusoidal periodic waveforms can be composed out of simple sinusoids and described by its component sinusoids of frequencies $f_0, 2f_0, \dots, kf_0, \dots$ in the amplitude, time and frequency space, as was a single sinusoid. Thus a nonsinusoidal periodic waveform has a spectrum, although it is not obvious what this spectrum may be, without looking at its frequency domain. To derive the spectrum of a complex waveform requires spectrum analysis tools.

The two traditional tools of Fourier Spectrum analyses are

- 1) *Hardware spectrum analyzer*
- 2) *Fourier spectrum analysis* – a mathematical technique that allows describing a time domain waveform in terms of both frequency-domain amplitude and phase.

The hardware spectrum analyzer is based on the Fourier spectrum analysis performed in hardware. It has a very fine tuned oscillator and precise time interval generator – Giga Hertz frequencies and pico-seconds time interval resolution. The better the frequency analyzer the more expensive it is, with costs as high as a hundred thousands of dollars. By fine sampling of a high oscillatory signal and coarse resolution it obtains a few thousands of data samples in a fraction of a seconds and performs a FFT to yield spectrum with that resolution.

1.5.1 Fourier Spectrum Analysis

Fourier spectrum analysis (Fourier Analysis) offers the option of spectrum results in rectangular form, which consists of real and imaginary part of the complex frequency domain. This complex frequency domain is a rectangular coordinate system comprised of the Real and Imaginary axes and was briefly described above.

The Fourier spectrum analysis is comprised of

*Fourier Series and the
Fourier Integral.*

Unfortunately this classic mathematical apparatus is frustrating for all but the simplest waveforms. However, if a waveform can be sampled and digitized with a waveform digitizer, the Discrete Fourier Transform (and its computationally effective implementation – the FFT) can be used to perform the Fourier spectrum analysis for the digitized waveform. Since its general introduction in 1965, the FFT has been the commonly used algorithm for

evaluating the DFT. *The FFT algorithm can be implemented in software, firmware or hardware.* The heritage and the Hilbert Transform (used in HHT) discrete methods are based on the Fourier spectrum analysis as we briefly describe below.

1.5.2 Frequency Domain Definition by Fourier Series

The general form of the Fourier Series for a function $x(t)$ with period T is

$$x(t) = a_0 + (a_1 \cos \omega_0 t + b_1 \sin \omega_0 t) + (a_2 \cos 2\omega_0 t + b_2 \sin 2\omega_0 t) + \dots$$

where

$$a_0 = \frac{1}{T} \int_0^T x(t) dt$$

$$a_n = \frac{2}{T} \int_0^T x(t) \cos(n\omega_0 t) dt$$

$$b_n = \frac{2}{T} \int_0^T x(t) \sin(n\omega_0 t) dt, \text{ where}$$

$$\omega_0 = 2\pi f_0, \text{ where } f_0 = 1/T$$

and where f_0 is the fundamental frequency. The Fourier Series yields a discrete spectra. For example a Fourier series can easily represent a square wave with period T . Textbooks contain the Fourier series terms for many useful waveforms. Sometimes a seemingly periodic waveform can be approximated with waveforms for which the series is known. As long as a periodic waveform can be mathematically described as a function of time $x(t)$, that meets Dirichlet conditions, its Fourier series can be written. However, this may not always be easy in practice, say with pulse functions that have discontinuities. If we approximate such practical function with a continuous “ideal” to the actual function and find the Fourier series for the “ideal”, it may yield good estimates for the practical frequency spectrum.

1.5.3 Fourier Transform

The Fourier series is a useful tool for investigating the spectrum of periodic waveforms, but the world is not made up exclusively of periodic waveforms. More complex waveforms do have a spectrum as we can hear the lightning bolt’s electromagnetic spectrum on a common radio receiver. For nonperiodic functions the Fourier Integral is used for treating a nonperiodic function as a function with period T approaching infinity. The Fourier Integral can be derived from the Fourier series and yields a continuous spectrum caused by infinite T . The integral is only applicable for nonperiodic waveforms. Phase seems to approach a continuous function of frequency too. *The Fourier Integral is also called the Fourier Transform (F), as opposed to the Fourier Series that is only applicable for periodic functions.* Because the integral is taken from minus to plus infinity ($-\infty, +\infty$) “negative frequencies” appear for

non-periodic waveforms spectrum derived from the Fourier Integral that have a non-complicated interpretation in the OATF plane. *The Fourier Series (similar to Fourier Transform) can also be expressed in exponential form, which will yield a negative frequency domain.*

The Fourier Transform yield the frequency spectrum in rectangular form or as a vector of complex variable values f:

$$F(x(t)) = X(f) = \text{Re}(X(f)) + j\text{Im}(X(f))$$

or complex-valued function in rectangular form. Results can also be obtained in polar form (phasor form). The amplitude and phase of the frequency components can be derived from $X(f)$ as described below..

1.5.4 Signal Synthesis from Fourier Spectrum

The original signal's time component #k can be synthesized from the Fourier spectrum amplitude and phase for each frequency f.

Fourier spectrum frequency component amplitude $A(f)$
 $|A(f)| = \sqrt{\text{Re}^2 X(f) + \text{Im}^2 X(f)}$

Fourier spectrum frequency component phase $\theta(f)$
 $\theta(f) = \tan^{-1}(\text{Im}X(f)/\text{Re}X(f))$.

1.5.5 Digital Fourier Transform

The Digital Fourier Transform (DFT) is viewed as a discrete approximation of the Fourier integral. For a data vector comprising N data values the computational complexity of a DFT is $O(N^2)$. An algorithm that implements the DFT definition and for which the computational complexity is $C \cdot \log_2 N \cdot N$ is called a *Fast Fourier Transform or FFT*. The DFT and FFT operate on *finite sequences* (N) - sets of data with each point *discretely* and *evenly spaced* in time and obtained from analog signals by *windowing* and *sampling* at a selected sampling rate.

Given a phenomenon, if you can acquire and sample it, you can apply the FFT to it. The FFT assumes periodicity in all cases. It also assumes that the windowed data repeats with the period equal to the window length in time and beyond both sides of the physical window. The FFT gives exact results for the digital data set. However, the continuous phenomena function was windowed and sampled and these operations that are used to digitize the continuous analog function, introduce issues that require careful interpretation of the FFT result if we are to apply the result to understanding the spectrum of the original continuous function. For example, the window length and shape may affect the so-called leakage of power to adjacent frequencies. Because of these problematic issues some techniques are used for improving the FFT results, such as:

Filtering, Smoothing, Plotting axis' range selection, Signal averaging to remove additive noise, Removing the mean that often improves amplitude resolution, Use sample rates

greater than twice the highest frequency in the signal, Phase is only valid for existing components, Remove delay to reduce phase, Change windows to change leakage, Change sample rates for different resolutions, Aliasing may not be unavoidable and there is a need, if it occurred, to rationalized what components are aliases.

Similarly, improving actions will be required for the spectrum results from a Hilbert Transform.

Examples of Discrete Fourier Transform Application and Interpretation of the spectrum results

Some signals are recognized by senses - touch, sound, sight or smell. About some signals we may say that they are time histories of a phenomenon. The classification of signals may be comprised of periodic and non-periodic deterministic signals, random signals whose value at a future time cannot be predicted with certainty (noise) and transient signals. The periodic and random signals are infinite in length by definition. Transient signals, signals that have finite duration, because of their finite length are deterministic.

Periodic Signal Definition

For a periodic signal there exist *the smallest* time window of length T, such that placing a cut-out with a window T seconds long (say, 1/8 seconds) over the signal, sliding this window in either direction by an integer multiple of T gives the identical picture of the signal.

For example,

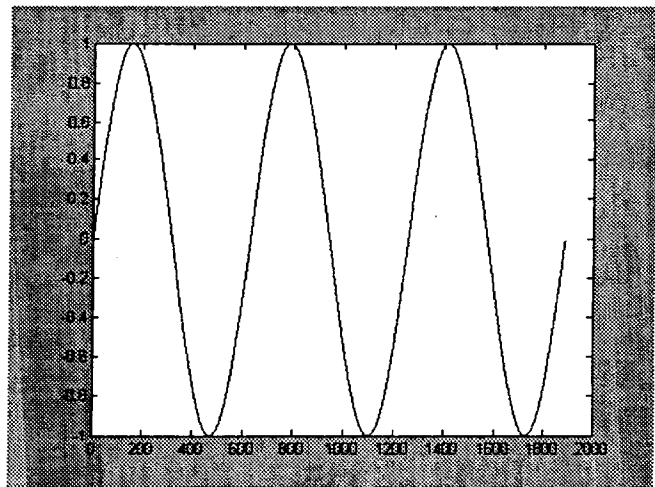


Figure 3. Periodic $\sin(t)$ signal, $T=2 \cdot \pi$ for $0 \leq t \leq 2 \cdot \pi \cdot 3$, increment step is 0.01

Any-Period-T Definition

The original picture P_t in the window was originally placed over the signal, say the left edge, coincides with time value t. However, once the window is placed at time t, sliding it on either end of this window will yield the same picture as in the original window P_t for a periodic signal with period T.

For example, two windows of the same length T placed at different origins ($t=0$ and $t=0.05$) depict different curves, but each will repeat itself when shifted to left or right by T . The following Figure 4 and Figure 5 demonstrate this example.

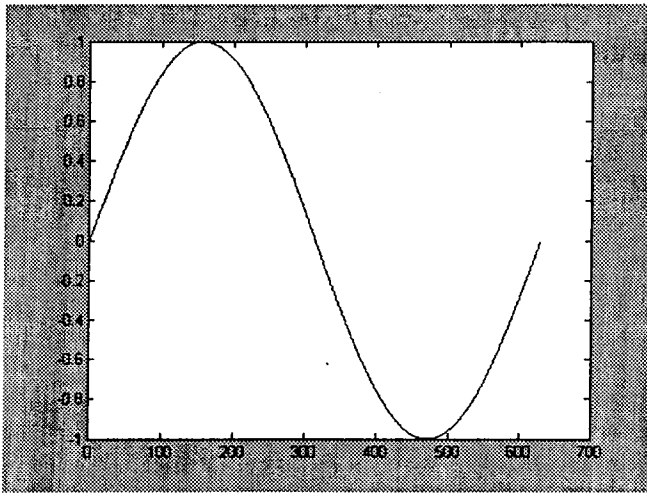


Figure 4. Window for $0.0 \leq t \leq 2\pi$

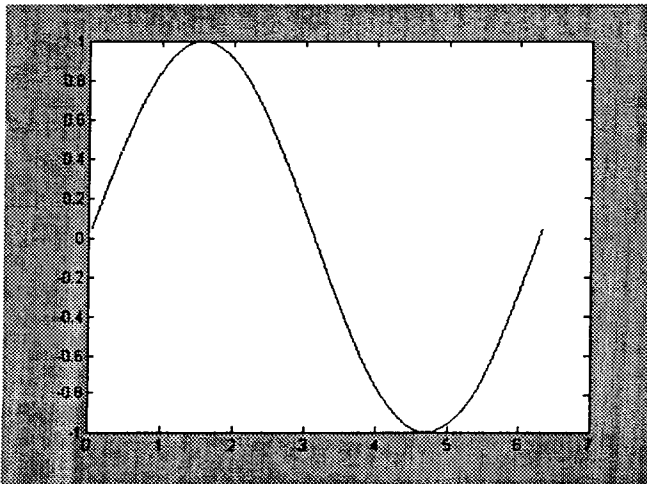


Figure 5. Window $0.0+0.05 \leq t \leq 2\pi + 0.05$

Dirichlet Criteria Definition for a Periodic Signal

A periodic signal satisfies the Dirichlet criteria if the criteria's following three conditions are satisfied in "Any Period T " or any period, namely:

- The number of *discontinuities* is *finite* in any period
- The number of *extrema points* is *finite* in any period
- The *integral* of the signal absolute value is *finite* in any period or over $[t, t+T]$ for any t in $-\infty < t < +\infty$.

Signal Average Value

For a periodic signal the average value of the signal equals the area bounded by one cycle of the signal function divided by T . For a random signal the period T is infinite. However, the average value can be obtained by integrating the signal

over $[-T, T]$, dividing it by T and making T approach infinity. This yields the mean value of the random signal.

For a transient signal, "average value" is not meaningful because if we follow the random signal definition of average, the area under a transient signal is finite and because T moves to infinity the average becomes 0 and that is not meaningful. *This is why transient signals or intermittency artifacts must be taking out in the beginning of an EMD sifting process.*

Signal Energy and Power

The characteristics of signal and power are far more useful than the average value for signal characterization.

Signal Energy

Energy is the work potential or the amount of work done and has units of joules in the "Systeme International" (SI) system of measurement units comprised of {meter, kilogram, second, ampere}. The *signal energy* for an arbitrary signal $x(t)$ in time T is defined as the integral of $x^2(t)dt$ over T , with no scaling factors. It is evident that the energy of a non-transient signal over the entire timeline is infinite and converges for a transient signal to a finite value. The *HHTDPS Hilbert Spectrum* is using IMFs energy for spectrum demonstration in different colors. The *square root of energy* may also be used for better spectrum visualization.

Signal Power

The rate of doing work is the signal power. Energy utilization may not be constant and vary from instant to instant. *Instantaneous power* is the true rate of change of energy in a system. In most cases, however, power refers to average power or signal energy over time T divided by time T . This is also called the mean squared value of the signal and the square root of it is called root mean squared value. Signal power may also be useful in the HHTDPS Hilbert Spectrum interpretation.

Harmonic Series

The fundamental frequency f_0 of a sinusoid with period T is $1/T$. A harmonic series is a set of sinusoids whose frequencies are harmonically related, meaning that frequency f of each sinusoid is an integer multiple of the fundamental frequency f_0 . Fourier developed the theory for obtaining the harmonic series that provides the best fit to a given periodic function. Finding the amplitude values of the harmonic components f giving the closest approximation to the original signal is termed Fourier analysis.

Waveform Synthesis and Gibb's Phenomenon

It is advisable to verify a spectrum analysis by synthesizing the original waveform from its constituent spectral components. For example, when the Fourier harmonic series is derived for a square wave, the exact square wave is not regained (synthesized) for two reasons. First, the total number of frequency components is not added in. Even adding 20 components shows that the deviation from the

square wave is obvious. This is due to the second reason, the *Gibb's Phenomenon* for signals with instantaneous transitions (discontinuities), such as in the square signal pulse edges or at boundary points for any signal in discrete analysis. The Gibb's phenomenon is always equal to 8.95% overshoot of the discontinuity amplitude. Adding more terms will take out more ringing at the end but the Gibb's phenomenon still occurs. There are methods of reducing the Gibb's phenomenon as much as by a factor of 7 [16].

Harmonic Series Generation and Waveform Synthesis Details

In a harmonic series generation and waveform synthesis application first we specify a period window by three relevant parameters:

- Window point of origin O
- Window length D
- Sampling resolution ∇t

We specify a window origin at O on the function argument timeline, window length D and sampling interval ∇t in seconds:

$$\{O, \nabla t, D\}$$

The number N of data points generated is determined from the window size D and sampling interval as

$$N=(D/\nabla t) + 1$$

Example 1.

A sinusoid in harmonic notation is used in this example. The fundamental frequency of this waveform function is $f_0=25$ and the number of cycles per second 25 Hz. The chosen in this example window spans from 0.0 to 0.31 seconds. As seen from the graph below the function curve has 8 maxima points. A one second long window would yield 100 data points and correspondently $100*8/32$ or 25 maxima points that, in turn, corresponds to 25 Hz fundamental frequency f_0 .

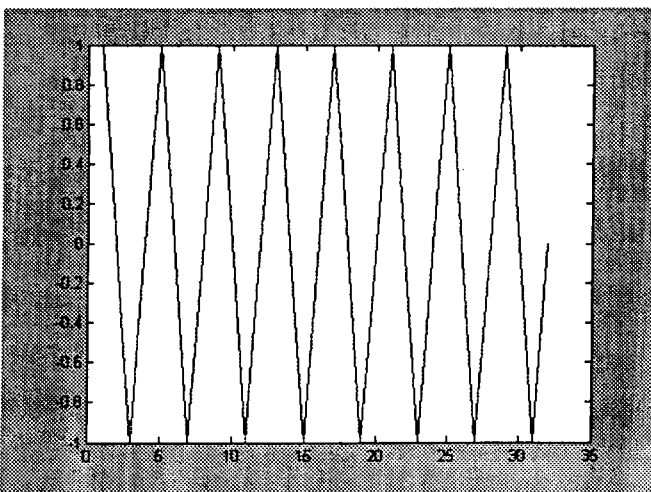


Figure 6. Window Parameters {O, ∇t , D},
 $x(t)=\cos(2*\pi*25*t)$ where $0<=t<=0.31$

Applying an FFT to the signal provides the correct frequency domain information for the properly windowed and sampled version of the waveform. The Frequency resolution interval ∇f for a sampling window containing N points at sampling interval ∇t is defined as

$$\nabla f = 1/N\nabla t$$

This means that each count in the FFT output frequency vector corresponds to ∇f .

In our example $N = (0.31/0.01 + 1) = 32$ and $\nabla t = 0.1$ or

$$\nabla f = 1/32*0.01 = 1/0.32 = 3.125 \text{ Hz per count.}$$

The Nyquist frequency is the highest frequency sinusoid that can be defined at a given sampling rate. For equally spaced samples ∇t the Nyquist frequency is $1/2\nabla t$. In our example the Nyquist frequency is $1/2*0.01=50$ Hz.

The FFT standard output arrangement (DFT) is as follows:

$$\{(The \text{ dc component count } C_0) + C_1, +C_2 \dots C_{N/2} N/2 \text{ Nyquist frequency } \dots -C_1 N-i \dots -C_1 N\}$$

To be able to synthesize a sinusoid of frequency F we need to sample it at a rate of 100 Hz (0.01 sec) and frequency resolution ∇f a window wide enough to fit N data samples – such that $N/2 * \nabla f \Rightarrow F$, namely that $N \Rightarrow 2F/\nabla f$. For our example $N \Rightarrow 2*50/3.125=32$.

For example, if $F=1.5$ GHz we would need to sample it within a spectrum analyzer at a sampling rate of $2F$ or more or equal to 3 GHz. At frequency resolution ∇f we would need $N=3G/\nabla f$ sampling points. The FFT is fast to process 2^{10} or 1024 data points. This means that we must select $N = 1024$ or deal with resolution $3 \times 10^9/2^{10} = 2^{28-10} = 2^{18}$ or frequency resolution in units of megahertz. As HHTDPS data size limitations are concerned the are similar to that of an FFT. For an FFT the preferable data size is ≤ 4096 data points. High frequency signals require short windows (a few milliseconds) and low frequency resolution, similar to the principles of a hardware spectrum analyzer.

In the FFT output vector of complex numbers the real part is arranged as depicted in the following vector. The significant amplitude frequency components occur on counts 8 from the top and count 8 from the bottom. A components' frequency for a corresponding count k can be derived from this count k and ∇f as

$$f_k = k*\nabla f$$

The frequency domain vector count 8 corresponds to frequency f_8 or

$$f_8 = 8*3.125 \text{ Hz} = 25 \text{ Hz}$$

`>>z = real(f)`
`z =`

0.0000 DC Component is zero for this function
 0.0000 count 1 of Positive frequencies down to
 0.0000 count 2 the middle point Nyquist count

.....
 -0.0000

16.0000 count 8 from top yielding in Hz the fundamental frequency $f_0 = 8 \text{ counts} * 3.125 = +25 \text{ Hz}$.
 The amplitude of the 25 Hz component is $16/32 = 0.5$ ($N=32$ is the normalization factor), where 0.5 is half of the signal sinusoid's amplitude of 1.0.
 The negative frequency has the second half of the amplitude 0.5 with its *negative sign implied*.

.....
 -0.0000 Last count $N/2-1$ of positive frequencies
 -0.0000 Nyquist frequency
 -0.0000 Last count of Negative frequencies

.....
 16.0000 count 8 from the bottom

.....
 -0.0000 count 2
 0.0000 Count 1 of Negative frequencies and going up!

Synthesizing the Waveform Using Fourier Spectrum Components

The Fourier frequency domain vector amplitudes A at counts 8 (from top and from vector bottom) are 16. The renormalized amplitude is A/N or equal to $16/32=0.5$. Whilst, we have two frequency components and $x(t)=0.5 \cos(2*\pi*25*t) + 0.5 \cos(2*\pi*-25*t) = \cos(2*\pi*25*t)$. This is because $\cos(t)$ is an even function and $\cos(t)=\cos(-t)$ for any t .

This synthesis of the frequency components yields exactly the original signal function $x(t)= \cos(2*\pi*25*t)$, as is guaranteed by the Fourier Digital Transform definition.

Example 2.

Let us derive the Fourier spectrum for a constant function, $x(t)=C=1$ for $0 \leq t \leq .31$ in increments of .01.

$f=\text{fft}(x)$

$f=$
 32 %dc component with amplitude $32/N=32/32=1.0=C$.
 0
 0
 ...
 0

The signal synthesis yields $x(t) = a_0=1=C$, as expected.

Signal Synthesis

$x(t) = a_0 + \cos(2*\pi*t) =$
 $dc/N + \cos(2*\pi*t) = 32/32 + \cos(2*\pi*t)$ or
 the exact input signal $1 + \cos(2*\pi*t)$.

Another two illustrative examples would be a square pulse and a constant time signal analysis and synthesis.

1.6 Hilbert Transform Spectrum

The Hilbert Transform is known for a long time. However, it is not easily interpretable when used for spectrum analysis. To arrive at a meaningful Hilbert Transform spectrum interpretation we will first examine the more familiar Fourier spectrum analysis interpretation.

1.6.1 Hilbert Transform Interpretation

In 1930 Denis Gabor defined a generalization of the Euler's formula in the form of a *complex analytic signal*, namely:

$$\Psi(t) = v(t) + j\gamma(t), \text{ where}$$

$$\gamma(t) = H[(v(t))] = -1/\pi \int_{\Omega} v(t)/(\eta-t) d\eta, t \in \Omega$$

$\Psi(t)$ is unique and it is an analytic signal by definition.

The Hilbert Transform is just another integral transform with kernel equal to $1/(\pi(t-\eta))$. It is a Cauchy Principal (P) integral and is denoted by symbolic notation letter P in front of the integral

$$-1/\pi P \int_{\Omega} v(t)/(\eta-t) d\eta$$

The Cauchy Principal integral is an integration technique that was proposed by Cauchy and solves the singularity problem at point $t=\eta$. Point η is surrounded by a small interval $[\eta-\epsilon, \eta+\epsilon]$. The integral is taken from the left to $\eta-\epsilon$ and from the right to $\eta+\epsilon$. In the vicinity of point η the integral is taken along a positive semi-circle of radius ϵ around point η . The integration is completed by allowing ϵ to approach 0.

Gabor used the analytic signal to introduce the Hilbert Transform (H) to signal processing for one-dimensional waveform $v(t)$. When the Hilbert transform is applied to a general waveform function $v(t)$, we obtain another function of time $\gamma(t)$.

The complex variable conjugate pair of functions $\{v(t), \gamma(t)\}$ then comprises the above analytical signal $\Psi(t)$ that can then also be expressed as

$$\Psi(t) = a(t)e^{j\theta(t)}, \text{ in which}$$

$$a(t) = [v^2(t) + \gamma^2(t)],$$

$$\theta(t) = \arctan((\gamma(t)/v(t)).$$

This analytic signal form is similar to the Fourier frequency component expression in terms of amplitude and phase, which were used in above examples for component synthesis. In the Hilbert Transform the amplitude and phase

are time variables while in Fourier analysis they are fixed for each f:

$$\omega(t) = d\theta(t)/dt = \theta'(t)$$

There is considerable controversy in using the Hilbert Transform for spectrum analysis [5], [6], [7]. For example, [5] lists five paradoxes associated with the Hilbert Transform spectrum analysis.

To avoid these paradoxes, the concept of a Intrinsic Mode Function (IMF) was developed. IMFs are signals with well-defined instantaneous frequency. Symmetrical by nature, with riding waves eliminated, these data sets can be processed by the Hilbert transform to obtain meaningful results. For a signal to be considered an IMF, it must fulfill the following conditions: “(1) in the whole dataset, the number of extrema and the number of zeros crossings must either equal or differ at most by one; and (2) at any points, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero” [1]. To obtain IMFs from a typical dataset, the EMD method was developed by Huang and others to filter and break down the data. When the Hilbert Transform is applied to a normalized IMF symmetric function, it yields powerful spectrum results and meaningful interpretation of physical oscillatory phenomena. Following is a brief description of the Hilbert Transform and its application IMF spectrum analysis. The IMF represents a generalized Fourier expansion.

1.6.2 Hilbert Spectrum Definition and Interpretation

In 1743 Leonard Euler introduced the identity

$$\exp(jz) = \cos(z) + j\sin(z)$$

This identity was used 100 years later by Fourier in the Fourier Transform. 150 years later Charles P. Steinmetz would use this identity to describe the complex notation of basic harmonic waveforms to electrical engineering

$$\exp(j\omega t) = \cos(\omega t) + j\sin(\omega t) \quad (\omega=2\pi f),$$

These are the bases for the Fourier Transform. Gabor’s generalization of the Euler formula and analytic signal

$$\Psi(t) = v(t) + i\gamma(t)$$

are the bases for the Hilbert Transform.

While the Fourier Transform of $v(t)$ yields the spectrum of $v(t)$ if the signal is linear and stationary and satisfies the Dirichlet conditions, the Hilbert Transform is applicable to a wider class of phenomena, such as non-linear and non-stationary functions. However, the Hilbert Transformation of a function of time $v(t)$ is another function of time of different shape $\gamma(t)$.

Can the Hilbert Transform $\gamma(t)$ yield information about the frequency spectrum of $v(t)$?

It can be shown [13] that

$$H[\cos(\omega t)] = +\sin(\omega t)$$

$$H[\sin(\omega t)] = -\cos(\omega t)$$

From the first of the above two transforms and by the definition of the analytic signal the function

$$\exp(j\omega t) = \cos(\omega t) + j\sin(\omega t)$$

is an analytic signal. This allows us to derive the Hilbert spectrum for the function $\cos(\omega t)$ in the following example.

1.6.2.1 Hilbert Spectrum for a simple sinusoid

Let $v(t) = \cos(\omega t)$. Then

$$\gamma(t) = H[\cos(\omega t)] = +\sin(\omega t).$$

$$\Psi(t) = v(t) + i\gamma(t) = \cos(\omega t) + j\sin(\omega t)$$

$$\theta(t) = \arctang(\sin(\omega t)/\cos(\omega t)) = \arctang(\tan(\omega t)) = \omega t$$

$$\theta'(t) = (\omega t)' = \omega$$

or the same Hilbert frequency spectrum as the heritage Fourier Series for a sinusoid waveform, as was expected. This example’s graphical results are depicted in the following section 9.0.

1.6.2.2 Hilbert Spectrum for a Complex Waveform

Along these lines, under what conditions could

$$d(\arctang(\gamma(t))/v(t))/dt$$

be interpreted as instantaneous frequency for a more complex than basic sinusoid waveform? This question is considered in detail in [1]. Furthermore [1], pp. 928 the definition of an interpretable Hilbert spectrum is derived as follows.

Having obtained the intrinsic mode functions (components), we apply the Hilbert transform to each IMF, and compute for each IMF the instantaneous frequency function as the derivative of the Hilbert transform phase function $\theta(t)$. Then we synthesize the IMF component k from its Hilbert Transform amplitude and phase functions as

$$a_k(t)e^{j\theta_k(t)}$$

Since $\omega(t) = d\theta(t)/dt$

we can rewrite the above expression as

$$a_k(t)e^{j\int \omega_k(t) dt}$$

The input signal $X(t)$ can then be synthesized as

$$X(t) = \sum a_k(t)e^{j\omega_k(t)t}$$

for all IMF components $k=1, \dots, n$.

This equation as stated in [1] pp. 928 enables to represent the amplitude and the instantaneous frequency as functions of time in a 3-D plot, in which the amplitude can be contoured over the frequency-time plane as the z-axis surface or by using a color scale in the 2-D frequency-time plane to represent amplitude. This plot of a superposition of all colored $\omega_k(t)$ in the instantaneous frequency-time plane $\{\omega, t\}$ is called the Hilbert spectrum $H(\omega, t)$.

1.6.3 Marginal Spectrum Definition and Interpretation

Based on Hilbert spectrum $H(\omega, t)$ we define marginal spectrum as

$$h(\omega) = \int H(\omega, t) dt$$

The marginal spectrum [1] offers a measure of total amplitude (or energy) contribution from each frequency value over the entire data span time interval $[0, T_{end}]$.

Example 3.

Hilbert Spectrum and Hilbert Marginal Spectrum

The length of day measurements file `Lodt78.dat` [1] was processed by the Empirical Mode Decomposition module producing the file of 11 Intrinsic Mode Functions stored in file `Lodt_p.dat`. The resulting IMFs were then processed by the HHTDPS' Hilbert Transform module with Hilbert Spectrum and Hilbert Marginal Spectrum options. The corresponding results are demonstrated below in Figure 7 and Figure 8.

The Hilbert Spectrum and the Hilbert Marginal Spectrum are described in more detail in Section 9.2.

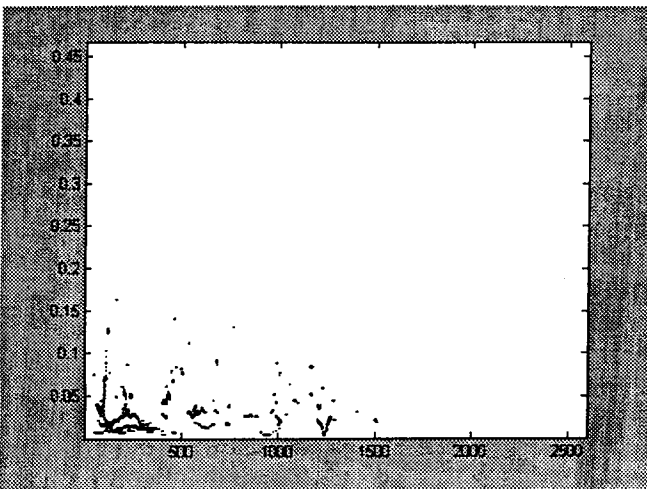


Figure 7. `Lodt78_p.dat` Hilbert Spectrum for parameters $\{nyy=200, t0=1, t1=2688\}$

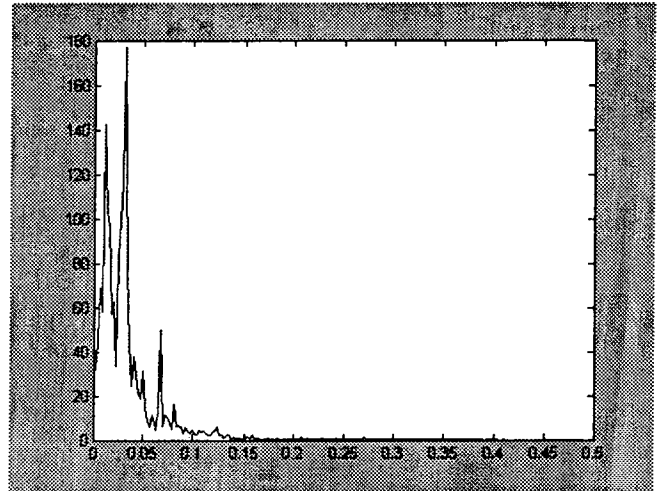


Figure 8. Hilbert Transform Marginal Spectrum

1.6.4 Hilbert-Huang Transform Synthesis

The waveform synthesis in the HHT is a direct summation of physical signal's IMFs generated by EMD. Presently there is no other way to synthesize the physical signal waveform from its Hilbert Spectrum.

1.7 Empirical Mode Decomposition Algorithm

The Empirical Mode Decomposition method is, in simplest terms, a filter that sifts through data and breaks it down into simpler components, the IMFs. IMFs retain the features of their parent data set and can uncover oscillatory trends that are not easily visible in the original. This break down is not haphazard; summation of the IMFs will return the parent dataset. This allows the IMFs themselves to be used for data analysis and manipulation without use of the Hilbert Transform.

The EMD process is relatively straightforward. A set of $x(t)$ values, the original dataset, is inputted for evaluation. From here begins the sift process. The data is then examined for local maximum and minimum extrema; those values are set aside. Using the extrema as knots, an upper envelope, with the maximum values, and a lower envelope, with the minimum values, is created via cubic spline interpolation. These envelopes are then linearly averaged together to create a mean envelope. This mean envelope represents the general trend of the data. This envelope is then subtracted from the parent data to create an IMF candidate. The IMF candidate signal is then tested if it passes the definition of an IMF, i.e., the number of extrema differs from the number of zero crossings by no more than one. If the IMF candidate set meets the criteria, a counter is incremented; otherwise the counter is set to zero. A check is then done on the counter – if it meets a user-defined value the IMF candidate is returned as an IMF. If not, the sifting process is executed again on the candidate, regardless if it passes the criteria or not.

Data returned as an IMF is stored and then subtracted from the parent data. The difference is used as a parent dataset for

a new sifting process. This is repeated until a dataset contains only two extrema. This data set is considered the "residual" and is returned with the IMFs.

2.0 HHTDPS PROPOSAL OVERVIEW

The proposal was to develop a generic, low cost, high performance PC based computer system that implements the HHT computational algorithms in a user friendly, file driven environment. This Huang-Hilbert Transform Data Processing System (HHTDPS) will then be tested using data provided by a number of appropriate application teams, and the data analysts from the respective provider organizations will evaluate the results.

The implementation will port and optimize Dr. Huang's existing preliminary C-language and Matlab code to the PC system, and baseline its functional performance. We will then run a series of simulations and benchmarks to measure system speed and identify processing bottlenecks. This information will be used to target specific functions for hardware acceleration. Computationally intensive routines will be implemented as Field Programmable Gate Array (FPGA) and Digital Signal Processing (DSP) hardware functions, and integrated into the baseline system to produce the final high performance system. Finally, we will test the new system on a few commercially viable applications and deliver a complete HHT tool and its documentation.

The proposed PC / FPGA / DSP hybrid system leverages state-of-the-art commercial data processing hardware, and the significant experience of the Code 565 engineering team with this technology. This project also benefits from a significant cost savings through the sharing of advanced design tools that already exist within the Code 565 infrastructure.

The proposed system has numerous potential NASA and commercial applications. The HHTDPS could be used to support the development and analysis of adaptive optics, finite element models, stray light models, orbit and attitude models, thermal models, and image and science data processing for space and earth science missions. There are also prospective space flight applications of HHT such as satellite sensor performance degradation analysis and compensation, as well as large structures deformation analyses as those of the Space Station and compound mirror observatories. Commercially, the medical industry could benefit from electronic data analysis of cardio-vascular and neurological systems functions. Critical security systems and forensic investigative tools using advanced bio-pattern analyses (voice and eye "fingerprints", sound and video decomposition) applications would benefit law enforcement. Seismology data analyses would have many applications, including oil and mineral exploration, natural disasters (earthquake) data analyses for developing better models by NASA, National Geodetic Survey (NGS) and Federal Emergency Management Agency (FEMA), spacecraft

components environmental vibration testing data analyses (aerospace industry and NASA) and industrial tool wear detection (commercial manufacturing). In addition to these current high volume commercial applications, this new technology could enable new applications such as the development of nano-tubes for use in medical and commercial fluid mechanics applications.

2.1 Goal and Objectives

The goal of the proposal was to develop a novel engineering tool for spectrum analysis of nonlinear and nonstationary data. In addition, it must have a user friendly, operational quality HHT data processing system that supports a wide range of applications, and ready for commercialization. The objectives are as follows: Port and optimize existing HHT software to a PC platform. Simulate and benchmark the HHT algorithms and identify bottlenecks. Implement the bottlenecks in Reconfigurable Computing (RC) hardware and deliver a low cost, high performance HHT implementation.

In the development process it was expected to demonstrate:

- 1) That the published results [1], [2], [3] can be reproduced by a computer code
- 2) That the code works and yielding similar results as given in [1] on the same input
- 3) Determine the HHT method capabilities and limitations by testing against different data sources at different data rates, volumes and magnitudes.

3.0 HHTDPS Implementation Cardinal Points

There were a few cardinal points that affected the HHTDPS Phase 1 development. These points stem from the intention to use HHT on-board a spacecraft whilst the necessity of having all HHTDPS components very compact and computationally fast, and considerations typical to the development of a project on schedule and within a tight budget. These points are described below.

3.1 Implementation in C-Language

The heritage Matlab functions are used for spectrum analysis. However, Matlab is a *large* general purpose *system* and because it is an interpreted language it will always be slower than compiled code. It would be difficult to fly a Matlab program in a "flight box" configuration since it would require the overhead and additional complexity of having to run Matlab to interpret the code. Therefore there is a need to convert basic spectrum analysis Matlab scripts to C-Language a cost of losing the mathematically inclined Matlab programming environment. This conversion to compiled code will also permit the integration of the spectrum analysis functionality into the HHTDPS, as well as remove the dependence for users of the system to have Matlab.

In addition, EMD's code must be constructed from publicly available patents' descriptions and published papers and developed in *plain C* rather than C++. This makes it more

susceptible to future implementation in hardware, since many systems currently do not fully support C++.

3.2 Algorithmic Definitions

In addition to language definition, it was necessary to provide precise definitions of many of the elements required to create the EMD. Although it is easy at a top level to use terms such as “extrema”, “envelope”, and “intermittency” and make perfect sense, a different level of detail is required to implement these items in a programming language. Particularly in C, which contains little beyond basic mathematical constructs, does this become evident as even the most intuitive of concepts must be broken down to a most basic level.

3.2.1 Extrema Points Definitions

The extrema points’ definitions must be systematic throughout the HHTDPS code. Although this may seem simple to define as a maxima extrema point is defined as an internal data set point x_i , such that $y_{i-1} < y_i > y_{i+1}$ and minima extrema point is defined by $y_{i-1} > y_i < y_{i+1}$ this will not account for two of the following situations and may cause serious problems in creation of the envelope.

The first is that an extrema may actually resemble a plateau, that is, there may be two or more points equal to each other prior to changing direction. If these are not identified, the envelope will not be correctly created which can cause severe problems in the generation of the IMFs.

The second situation comes directly from with the methods used to deal with the first. That is, one could try to compensate by defining $y_{i-1} < y_i \geq y_{i+1}$ as the maxima and minima extrema point is defined by $y_{i-1} > y_i \leq y_{i+1}$. However, this can cause detection of “false extrema”, i.e. plateaus that do not change direction after the level section. Therefore a better definition for finding an extrema is:

Maxima:

$$y_{i-1} < y_i > y_j$$

where $j > i$ and no point between j and i contains a value of $y > y_i$

Minima:

$$y_{i-1} > y_i < y_j$$

where $j > i$ and no point between j and i contain a value of $y < y_i$

Although it may seem trivial to define this to such an extent, it is something that can be easily overlooked if considering a classical definition of extrema

3.2.2 Upper and Lower Envelopes Definitions

An upper envelope, or upper awning, is created by connecting the maxima extrema points of a data set and a lower envelope is created by connecting the minima extrema points. For the connection, a piecewise cubic spline is used which provides a smooth curve between extrema points that is continuous in the second derivative. Although a cubic

spline may not most accurately reflect the trends of the data such as a piecewise hermite spline, the IMFs created with it prove to be more symmetrical than with other methods.

3.2.3 Prediction

A problem with the creation of the envelopes is that they must cover the entire dataset, however it is rare, if ever, that extremas will be found precisely at the ends of the dataset. For this we need a pair of maxima and minima extrema points that lies to the right and to the left of the end points of the data set. These four extrema points must be predicted. The prediction of the four points, like any prediction, is difficult to do. This is attested to by multitudes of prediction algorithms and not a single one has been accepted as a rule. Envelopes’ extensions to cover data set window by prediction of a set of points and find one maxima and minima outside envelopes is followed by truncation of the mean data set to bring it back into the signal envelope. The HHT processes are very sensitive to end points’ selection. This is why three prediction options are available in the HHTDPS: *pattern, linear, and mean*.

Pattern prediction statistically analyzes the extremas at the borders of the dataset and estimates the distances and amplitudes of the predicted extrema based on those results. Linear prediction originally involved using the linear predictive coding algorithm; however this was found to be processor intensive and could occasionally cause data degradation. It was replaced with an algorithm that simply copies the extrema points closest to the data borders. Finally, mean prediction uses the average distances and heights of the extrema points to generate predictions. All these algorithms have their respective advantages and disadvantages.

Problems with prediction can result in an extrema point at a far distance from the last extrema point which may cause waves in the cubic spline that connects these two points. This becomes problematic even in the sifting for the first few IMFs and may lead to data corruption or degradation.

3.3 Intermittency Definition, Detection and Removal

Intermittent or “riding waves” affect the outcome of the IMF generation. In general, riding waves are defined as a transient signals that are interrupting the predominant pattern of the wave. Particular to intermittency detection, intermittency is defined as any set of a maximum and minimum extrema such that the distance between the adjacent extrema points is less than the designated threshold, η . In EMD during the process of removing intermittency from an IMF, a search is made in the signal for all maximum and minimum extrema which the distance between any two extrema is greater than or equal to η .

In order to separate the intermittency artifacts into a different IMF, those points found during the process are removed, and all values that are less than the threshold are retained, and the EMD calculation is continued. There is currently no automatic heuristics for selection of threshold

values, due to the complexity of such analysis. It is important to note that with this process all properties of the EMD will be retained, including that all IMFs generated will sum together to provide the original signal. Intermittency checking does not remove waves from the dataset, but reserves them for calculation in the next generation of an IMF.

3.6 Other Cardinal Points

There were other cardinal points which were considered in the HHTDPS Phase 1 implementation and which we will only list without further elaboration:

- IMF Normalization and Symmetry Definition (integral of the IMF is close to 0)
- IMF basis orthogonality as criteria of IMF basis quality and zero value is ideal.
- Heritage definition of orthogonality for polynomials and trigonometric functions.
- Hilbert Transform and Resulting Phase Angle
- Instantaneous Frequency Definition [4]
- HHTDPS Qualitative Spectrum Analysis
- Gibbs phenomenon and its resolution [9].
- Hardware implementation of bottleneck functions [10].

4.0 DEVELOPMENT PLAN

The development plan for the HHTDPS was stated first and then rigorously followed. The plan was for all phases of the HHTDPS. However, this paper describes only the Phase 1 activities of the development plan.

Phase 1

Develop a complete, PC-based, user-friendly software version of the HHT Data Processing System. Users supply data files on CD, floppy disk, ZIP disk or via electronic file transfer, perform their processing at the local PC, and take their processed data away with them on CD, floppy disk, ZIP disk or via electronic file transfer. Complete "user's guide" documentation is provided. This baseline system will serve as the functional benchmark during Phase 2.

Phase 2

Analyze the performance of the Phase 1 system while processing various types of input data. Identify compute intensive processes that may be accelerated via DSP and/or FPGA implementation. Generate DSP and/or FPGA modules and integrate into the Phase 1 system. Benchmark Phase 2 system performance and compare to Phase 1 performance.

Phase 3

Investigate remote user access to the HHTDPS. Generate list of issues and implementation options. Implement one or more prototype remote access functions. This phase will be completed as time allows (remote access is not a requirement).

Phase 1 Activity Plan

- Purchase PC hardware
- Obtain and analyze existing code
- List basic algorithms
- Collect user requirements, including IO data types and formats
- Develop modular system architecture (keeping Phase 2 in mind)
- Develop functional specifications for system elements
 - I/O formats
 - User I/F
 - Processing modules
- Write code for modules
- Test
- Benchmark
- Document.

5.0 TOP LEVEL ARCHITECTURE

The HHTDPS was designed as a modular system comprising the Empirical Mode Decomposition module (EMD), the Instantaneous Frequency and Hilbert Spectrum module (IFS and Data Analysis Tools) and the Graphical User Interface module (GUI). The architecture is represented in the following Figure 9.

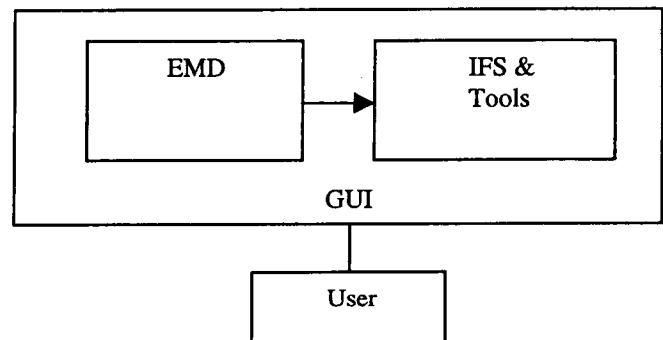


Figure 9. HHTDPS Components

6.0 EMPIRICAL MODE DECOMPOSITION MODULE ARCHITECTURE

The EMD module architecture generally follows the block diagram presented in [3] for a 1-dimensional input case. The input physical signal is loaded and stored as a physical signal array. The user can then perform the EMD process on the signal, and extract the IMFs. From there, the user has the options of obtaining additional information about the IMFs or using one of several means to convert them into the Hilbert Spectrum. All data generated by the program can be saved or exported to disk.

7.0 INSTANTANEOUS FREQUENCY AND HILBERT SPECTRUM MODULE ARCHITECTURE

The IFS module comprises the Hilbert Transform of the IMF components, phase formulation and differentiation that yields the instantaneous frequency. The instantaneous frequency for all IMFs is then further aggregated into frequency bands for a chosen band resolution that is chosen by the user. The aggregated instantaneous frequency energy is displayed as a 2-dimensional frequency-time color-coded image. This is the signal Hilbert Spectrum.

The Marginal Hilbert Spectrum Option [1] is generated by summation of energy values in the Hilbert Spectrum across the time line. This produces a typical 2-dimensional frequency-amplitude plot.

8.0 GRAPHICAL USER INTERFACE ARCHITECTURE

The Graphical User Interface is a method to integrate EMD, IFS modules and provide User with typical graphical means to run the HHTDPS functions. The GUI utilizes Visual Basic to provide a robust interface. It was designed to be intuitive to use, but does not abstract the user from control of the various functions.

9.0 SYSTEM OPERATIONS AND TEST RESULTS

The following section demonstrates results obtained using the completed system. This includes generation of IMFs using the EMD, as well as the display of the nominal and marginal Hilbert spectrum.

9.1 Empirical Mode Decomposition Results

Consider a waveform comprised of two pure sinusoids. It is expected that there will be two IMFs from EMD and each having small amplitude residuals. For k sinusoids EMD is expected to generate k IMFs for any small or large k . For example, if $k=200$ the EMD is expected to produce around 200 IMFs.

The two modes IMF01 and IMF02 are the input signal main modes, as observed from the IMFs significant amplitudes. The remaining IMFs are small residue components that can be attributed to algorithm computational traces, resulting from spline processes.

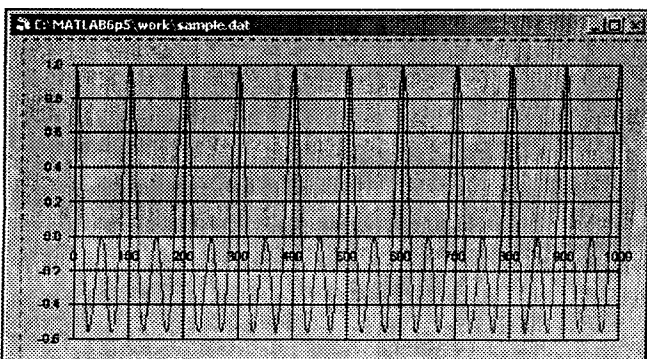


Figure 10. Input Physical Signal $y = (\cos(2\pi t) + \cos(2 \cdot 2\pi t))/2$

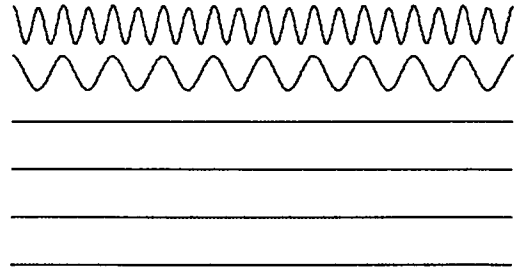


Figure 11. IMFs generated, in order from top to bottom. Last signal is the residual

The resulting IMFs must be saved and used for signal synthesis. A copy of the IMFs may be then preprocessed as required to obtain a better Hilbert Transform for spectrum analysis.

9.1.1 Decomposition Interpretation

We may begin the interpretation with the direct observation of the last decomposition component – the Residue in Figure 11. This is the data long-term trend. The next to the last decomposition components IMF05, IMF04, and IMF03 are depicting some transients resulting from the fact that the spline is only an approximation. The oscillatory behavior components are the IMF02 and IMF01, where IMF01 is the finest oscillatory scale followed by the next scale IMF02. These base functions can be further analyzed using the HHTDPS Hilbert Spectrum module as described below.

9.2 Hilbert Spectrum Analysis

The pre-conditions to get an interpretable Hilbert Spectrum are as follows within the HHTDPS are:

- 1) In HHTDPS the Hilbert Spectrum is derived from the IMFs – the output of the Empirical Mode Decomposition of the physical signal.
- 2) The IMFs, in order to be used as input to the Hilbert Transform, must be normalized in order to satisfy the Bedrosian [6] and Nutall [7] requirements. Even a sinusoid function $A \cdot \cos(t)$ must be normalized to yield a meaningful for spectrum interpretation Hilbert Transform by a factor of $1/A$. Some signal data is already normalized naturally – any signal with small amplitudes, including length of day deviation from 24 hours in microseconds or voice data as recorded on a standard PC sound card as a wave file has magnitudes between 0 and 1.

- 3) Like all discrete transforms the Hilbert Transform is an approximation and the Gibbs phenomenon end affects must be taken into consideration for the Hilbert Transform to be usable for Hilbert Spectrum derivation and interpretation. This requires preprocessing of the IMFs before using them as input to the Hilbert Transform. Namely, in order to get a “good” Hilbert Transform each IMF must have equal values

at end points, as well as the same slope. In any other case of IMF the end-affect (the Gibbs phenomena) propagates into the spectrum and may result in a broken line. Preconditioning the IMF to meet these two requirements may result in IMFs of different size and some shorter than original IMF produced by EMD. This is not a problem for the derivation of the Nominal and Marginal Hilbert Spectrum.

3) There are two types of interpretable Hilbert Spectrum data produced by the HHTDPS. The first is called "The Nominal Hilbert Spectrum" and the second type is called "The Marginal Hilbert Spectrum".

9.2.1 Nominal Hilbert Spectrum

The Nominal Hilbert Spectrum depicts IMF components instantaneous frequencies' amplitudes in time as color on a 2-D graph {Oft}.

For each IMF_i component $data_i(t)$ the ISP1 module is producing an array of instantaneous frequencies $f_i(t)$ and corresponding time amplitudes $a_i(t)$. The global maximum and minimum frequencies two values across all $f_i(t)$ are determined and the Hilbert Spectrum is derived within this range $[\max f_i(t), \min f_i(t)] = [F1, F2]$ for all i . A graph is produced of amplitudes $a(f,t)$. For each point t_j all $f_i(t_j)$ values across the IMFs vertical cut in time are examined and all k -frequency values in this cut are aggregated into a few frequency bins specified by user in the GUI window. If user selected F -bins, then $[F1, F2]$ is divided by scale $F1, F1 + F, F1 + 2 * F, \dots, F2$

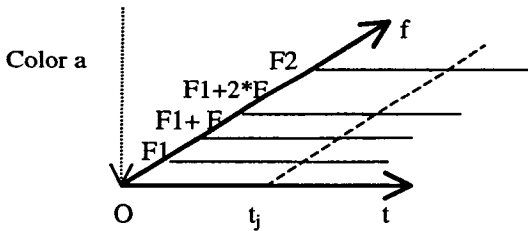


Figure 12. Hilbert Spectrum Frame {Oft, Color a}

Following is the Hilbert Spectrum for the above example with expected result being two straight lines $y=1, y=2$.

Gibbs phenomenon affects are clearly seen at the end points and propagate into the middle of the Hilbert Spectrum data.

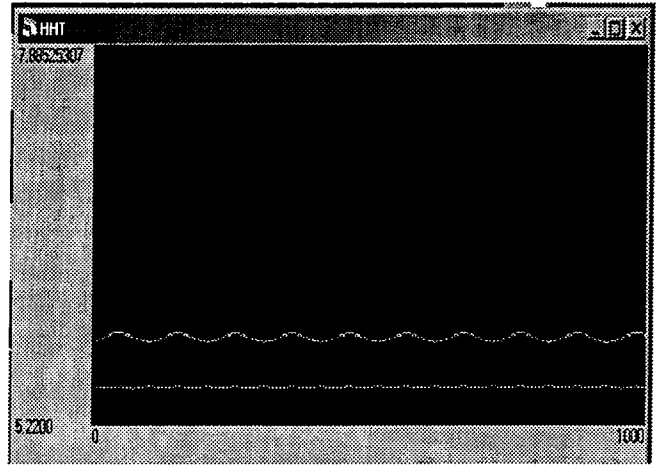
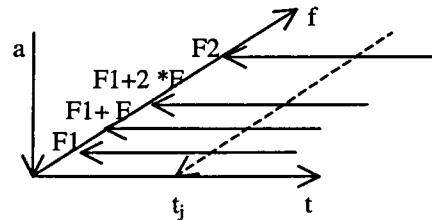


Figure 13. Nominal Hilbert Spectrum

9.2.1 Marginal Hilbert Spectrum

The Marginal Hilbert Spectrum is obtained from the Nominal Hilbert Spectrum by summing up the amplitudes along the frequency bin time lines $F1, F1+F, F1+2F, \dots, F2$ and a display in plane aOf . Adding frequency amplitudes in Figure 13 as shown in the following picture



results in a 2-D Marginal Hilbert Spectrum Figure 14

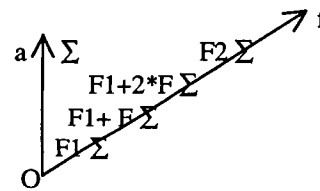


Figure 14. Marginal Hilbert Spectrum Frame {Oaf}

9.3 EMD, Hilbert Spectrum and Physical Signal Synthesis

The waveform synthesis from using Hilbert Spectrum is not feasible. However, the synthesis can be achieved directly by using the IMFs generated by the HHTDPS EMD1 module.

10.0 TECHNOLOGY COMMERCIALIZATION

The HHT method has many diverse commercial applications. The HHTDPS can be applied in a variety of fields to study phenomena such as basic nonlinear mechanics, climate cycles, solar neutrinos variations, earthquake engineering, geophysical exploration, submarine design, structural damage detection, satellite data analysis,

nonlinear wave evolution, turbulence flow, speech recognition, and biometrics, voice identification, blood pressure variations and heart arrhythmia. It also can be used as a filter for audio recordings.

This method is also used to analyze sea surface temperature data collected by NASA satellites and instruments. The National Oceanic and Atmospheric Administration uses Huang's method to analyze images from some of its Earth orbiting spacecraft. It has proven successful in connecting environmental changes to El Nino phenomena with weather changes.

NASA has patented this method. The goal of the HHT technology commercialization is to collaborate with industry and academia to get people using the method. A beta version is available for university research, while the final version of HHTDPS is available for licensing.

11.0 LESSONS LEARNED

In the words of Dijkstra testing can show presence of errors but not their absence. Testing may take longer by an order of magnitude of the time planned for testing. We have learned some lessons in new technology development and commercialization, that new technology development is expensive and requires funding. We learned about HHT details that, in-turn, helps to advance the HHT state of the art.

CONCLUSIONS

We have accomplished the goal we set out in the introduction and developed the first engineering tool based on the Empirical Mode Decomposition. We have developed a new ground system configuration for this specific project and laid the basis for its re-use in future projects, including spacecraft on-board data processing. We have developed a methodology that widens the field of digital signal processing and spectrum analysis of nonlinear and non-stationary data. We have demonstrated that the published HHT algorithms and results can be implemented independently in a computer code and have demonstrated that the code runs in reasonable time (making it a useful engineering tool) and produces the same results on the same inputs. Future work includes the development of the HHTDPS Phase 2. In Phase 2 the computationally intensive components will be implemented as Field Programmable Gate Array (FPGA) and Digital Signal Processing (DSP) hardware functions, and integrated into the baseline system comparable in performance to FFT.

REFERENCES

[1]. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analyses" by Norden E. Huang et al, Proc.R Soc. Lond. (1998) 454, 903-995

[2]. "A new view of nonlinear water waves: The Hilbert Spectrum" by Norden E. Huang et al Annu. Rev. Fluid Mech. 1999, 31 417-457

[3] Norden E. Huang Patent #6,311,130 B1
Patent Date: Oct. 30, 2001

Patent Name: Computer Implemented Empirical Mode Decomposition Method, Apparatus, and Article of Manufacture for Two-Dimensional Signals.

[4]. "Instantaneous Anything" by Leon Cohen, Proc. IEEE ICASSP-93, vol. 5, pp. 105-108, 1993

[5] "Time Frequency Analysis" by Leon Cohen, Englewood Cliffs, NJ Prentice Hall, 1995

[6] "A Product Theorem for Hilbert Transforms" by E. Bedrosian/RAND Corporation, Santa Monica CA in IEEE Proceedings, Volume 54, Oct.-Dec. 1966, pp. 458-459

[7] "On Quadrature Approximation to the Hilbert Transform of Modulated Signals" by Albert H. Nuttall/Litton Systems Inc., Waltham MA in IEEE Proceedings, Volume 54, Oct.-Dec. 1966, pp. 458-459

[8] "Frontiers of Complexity" by Peter Coveney and Roger Highfield, A Fawcett Columbine Book, 1995.

[9] "On the Gibbs phenomenon and its resolution" by David Gottlieb and Chi-Wang Shu, SIAM Rev. Vol 39, No. 4, pp. 644-668, 1997

[10] Semion Kizhner, David J. Petrick, Thomas P. Flatley, Phyllis Hestnes, Marit Jentoft-Nilsen, Karin Blank "Pre-Hardware Optimization of Spacecraft Image Processing Software Algorithms and Hardware Implementation", 2002 IEEE Aerospace Conference Proceedings Big Sky Montana, March 9-16, 2002 IEEE Catalog Number: 02TH8593C ISBN: 0-7803-7232-B

[11] "The FFT, fundamentals and concepts" by Robert W. Ramirez, Englewood Cliffs, N.J., Prentice-Hall, c1985.

[12] "Understanding Signals and Systems" by Jack Golten, The McGraw-Hill Companies, 1997

[13] "Hilbert Transform in Signal Processing" by Stefan L. Hahn 1996 Artech House Inc.

[14] "Gabor Analysis and Algorithms Theory and Applications" by Hans G. Feichtinger, Thomas Strohmer Editors, 1998 Birkhauser Boston.

[15] "On IMF Normalization" Norden E. Huang Patent Application

[16] The Art of Doing SCIENCE and Engineering", by Richard W. Hamming Gordon and Breach Science Publishers, 1997.

BIOGRAPHY

Semion Kizhner, an aerospace engineer with the National Aeronautics and Space Administration (NASA) at the Goddard Space Flight Center (GSFC), participated in the development of the Space Shuttle launched Hitchhiker carrier and several attached Shuttle payloads such as the Robot Operated Materials Processing System (ROMPS). He was responsible for establishing the Global Positioning

System (GPS) applications and test facility at GSFC and supported GPS simulations for several space projects, such as the OrbView-2, SAC-A and EO-1 spacecrafts. He is currently developing capabilities to access spacecrafts as nodes on the Internet, to accelerate generation of images derived from the GOES-8 spacecraft data [10] and implementing algorithms for high rate control loops in optical instruments. He proposed the development of the Hilbert-Huang Transform Data Processing System (HHTDPS) and has been leading the HHTDPS development team. He graduated from Johns Hopkins University with an MS degree in computer science.

Tom Flatley is an Electronics Engineer at the NASA Goddard Space Flight Center, and is currently the Chief Technologist and Acting Associate Head of the Electrical Systems Branch. From 1998-2002 he served as Associate Head of the Ground Systems Hardware Branch, and from 1993-1997 he served as head of the Flight Electrical Systems Section and Flight Component Development Group. Prior to this period he developed numerous flight and ground components and subsystems for various NASA missions, beginning in 1985. Mr. Flatley's current work includes the coordination of Re-configurable Computing technology development and Accelerated Science Data Processing activities for the GSFC Electrical Engineering Division, with the primary goal of developing RC technology and hybrid systems for flight and ground science data processing applications. He also manages several collaborations with government, industry and academic partners, and serves as liaison between technology developers and end users in the science community.

Norden E. Huang is a senior fellow at NASA, Goddard Space Flight Center. He holds a doctoral degree in Fluid Mechanics and Mathematics from the Johns Hopkins University. Dr. Norden Huang has worked on nonlinear random ocean waves spectrum determination. Recently, he developed a new method, the Hilbert Spectrum Analysis, specifically to process non-stationary and nonlinear time series. He developed the Empirical Mode Decomposition method that is the basis for the Hilbert-Huang Data Transform [1], [2], [3]. Dr. Norden Huang is a member of the National Academy of Engineering.

Karin Blank is a computer engineer with NASA, Goddard Space Flight Center. She implemented the GOES-8 color picture generation algorithms on a digital signal processor (DSP) and she is the lead software engineer for the HHTDPS development. She graduated from Worcester Polytechnic Institute with a BS degree in computer science and is currently studying for her masters in computer science at the Johns Hopkins University.

Evette Conwell is a Commercial Technology Manager, at the NASA Goddard Space Flight Center Technology

Commercialization Office. She received a BA in Mathematics from Cheyney University of Pennsylvania and an MS degree in Engineering Management from George Washington University, Washington, D.C. She had facilitated license and partnership arrangements for NASA technology transfer and commercialization with some 23 organizations in USA.

Darrell Smith is a principle software engineer with the Orbital Sciences Corporation. His areas of interest include satellite ground systems, simulators, databases, and client-server computing. He received a BS in Computer Science from the University of Maryland, College Park, and an MS in Computer Science from the Johns Hopkins University.

ACKNOWLEDGEMENTS

The authors would like to thank the NASA Goddard Space Flight Center Office of Commercial Programs and specifically Nona Cheeks, Don Vargo and Ronald W. Kaese for the stimulating discussions and encouragement to develop this new technology.