A NEW VIEW OF NONLINEAR WATER WAVES: The Hilbert Spectrum[1]

Norden E. Huang[1], *Zheng Shen*[2], *and Steven R. Long*[3]

[1]Division of Engineering Science, California Institute of Technology, Pasadena, California 91125. On leave from Laboratory for Hydrospheric Processes, Oceans and Ice Branch, Code 971, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771

[2]Division of Engineering Science, California Institute of Technology, Pasadena, California 91125 and Department of Civil Engineering, University of California at Irvine, Irvine, California 92697

[3]Laboratory for Hydrospheric Processes, Observational Science Branch, Code 972, NASA GSFC, Wallops Flight Facility, Wallops Island, Virginia 23337;
e-mail: long@osb.wff.nasa.gov

ABSTRACT

We survey the newly developed Hilbert spectral analysis method and its applications to Stokes waves, nonlinear wave evolution processes, the spectral form of the random wave field, and turbulence. Our emphasis is on the inadequacy of presently available methods in nonlinear and nonstationary data analysis. Hilbert spectral analysis is here proposed as an alternative. This new method provides not only a more precise definition of particular events in time-frequency space than wavelet analysis, but also more physically meaningful interpretations of the underlying dynamic processes.

INTRODUCTION

Historically, there are two views of nonlinear mechanics: the Fourier and the Poincaré. The traditional Fourier view is an outcome of perturbation analysis in

---

[1]The US government has the right to retain a non-exclusive, royalty-free license in and to any copyright covering this paper.

which a nonlinear equation is reduced to a system of linear ones. The final solution becomes the sum of these linear equations. In most mechanics problems, the linearized equations are second order; therefore, the solutions are trigonometric functions, and the sum of the solutions of this linear system constitutes the Fourier expansion of the "true" solution. This is thus the Fourier view: The system has a fundamental oscillation (the first-order solution) and bounded harmonics (all the higher-order solutions). Although this approach might be mathematically sound, and seems to be logical, the limitations of this view become increasingly clear on closer examination: First, the perturbation approach is limited to only small nonlinearity; when the nonlinear terms become finite, the perturbation approach then fails; Second, and more importantly, the solution obtained makes little physical sense. It is easily seen that the properties of a nonlinear equation should be different from a collection of linear ones; therefore, the two sets of solutions from the original equation and the perturbed ones should have different physical and mathematical properties. Realizing this limitation, recent investigators of nonlinear mechanics adopted a different view, that of Poincaré.

Poincaré's system provides a discrete description. It defines the mapping of the phase space onto itself. In many cases, Poincaré mapping enables a graphical presentation of the dynamics. Typically, the full nonlinear solution is computed numerically. Then the dynamics are viewed through the intersections of the trajectory and a plane cutting through the path in the phase space. The intersections of the path and the plane are examined to reveal the dynamical characteristics. This approach also has limitations, for it relies heavily on the periodicity of the processes. The motion between the Poincaré cuts could also be just as important for the dynamics. Both the Fourier and Poincaré views have existed for a long time. Only recently has an alternative view for mechanics, the Hilbert view, been proposed.

The Hilbert view is based on a new method, called empirical mode decomposition (EMD) and Hilbert spectral analysis as described by Huang (1996) and Huang et al (1996, 1998a). It has found many immediate applications in a variety of problems covering geophysical (Huang et al 1996, 1998a) and biomedical engineering (Huang et al 1998b). In this review, the new method will be summarized, and fluid mechanics examples of nonlinear water waves and turbulence data will be used to illustrate the use of this method to interpret the dynamics of these phenomena.

As the new method became available only recently, it is necessary to give a summary of it and describe some recent improvements to it here. Huang et al (1998a) clearly point out that a faithful representation of the nonlinear and nonstationary data requires an approach that differs from Fourier or Fourier-based wavelet analysis. The new method developed by Huang et al (1998a) seems to fit this need. This method uses two steps to analyze the data. The

first step is to decompose the data according to their intrinsic characteristic scales into a number of intrinsic mode function (IMF) components by using the empirical mode decomposition method. In this way, the data are expanded in a basis derived from the data itself. The second step is to apply the Hilbert transform to the IMF components and construct the time-frequency-energy distribution, designated as the Hilbert spectrum. In this form, the time localities of events will be preserved, for frequency and energy defined by the Hilbert transform have intrinsic physical meaning at any point. We will introduce the whole process by starting from the Hilbert transform.

THE HILBERT TRANSFORM

For an arbitrary time series, $X(t)$, we can always have its Hilbert transform, $Y(t)$, as

$$Y(t) = \frac{1}{\pi} P \int \frac{X(t')}{t - t'} \, dt', \tag{1}$$

where P indicates the Cauchy principal value. This transform exists for all functions of class L^p (see, for example, Titchmarsh 1948). With this definition, $X(t)$ and $Y(t)$ form a complex conjugate pair, so we can have an analytic signal, $Z(t)$, as

$$Z(t) = X(t) + Y(t) = a(t)e^{i\theta(t)}, \tag{2}$$

in which

$$a(t) = [X^2(t) + Y^2(t)]^{\frac{1}{2}};$$
$$\theta(t) = \arctan \frac{Y(t)}{X(t)}. \tag{3}$$

Theoretically, there are an infinite number of ways to define the imaginary part, but the Hilbert transform provides a unique way for the result to be an analytic function. A brief tutorial on the Hilbert transform, with emphasis on its physical interpretation, can be found in Bendat & Piersol (1986). Essentially, Equation (1) defines the Hilbert transform as the convolution of $X(t)$ with $1/t$; therefore, it emphasizes the local properties of $X(t)$. In Equation (2), the polar coordinate expression further clarifies the local nature of this representation: it is the best local fit of an amplitude- and phase-varying trigonometric function to $X(t)$. Even with the Hilbert transform, there is still considerable controversy in defining the instantaneous frequency as

$$\omega(t) = \frac{d\theta(t)}{dt}. \tag{4}$$

Detailed discussions and justifications are given by Huang et al (1998a). With this definition of instantaneous frequency, its value changes from point to point

in time. Two simple examples in Figure 1 (see color figure at end of volume) illustrate this approach. Figure 1a gives the familiar sine wave changing from one frequency to another. These data are certainly nonstationary, a characteristic that repeatedly demonstrates the power of wavelet analysis. The wavelet spectrum in color and the Hilbert analysis representation as a thin line through the wavelet spectrum are shown in Figure 1b. Their projections on the frequency-energy plane are shown in Figure 1c. The comparison is clear: The Hilbert representation gives a much sharper resolution in frequency and a more precise location in time. The second example is the common exponentially damped oscillation. The data, wavelet and Hilbert representations, and their projections are given in Figures 1d–f, respectively. Again, it can be seen that the Hilbert representation gives a superior resolution in time and frequency. Based on these comparisons, we can conclude that wavelet analysis indeed improves the time resolution compared with the Fourier method. Wavelet analysis gives a uniform frequency resolution, but as can be seen, the resolution is also uniformly poor.

Convenient and powerful as the Hilbert transform seems, by itself it is not usable for general random data, as discussed by Huang et al (1998a). In the past, applications of the Hilbert transform have been limited to narrow band data; otherwise, the results are only approximately correct (Long et al 1993b). Even under such restrictions, the Hilbert transform has been used by Huang et al (1992) and Huang et al (1993) to examine the local properties of ocean waves with detail that no other method has ever achieved. Later, it was also used by Huang (1995) to study nonlinear wave evolution. For general application, however, it is now obvious that the data will have to be decomposed first, as proposed by Huang et al (1998a).

Independently, the Hilbert transform has also been applied to study vibration problems for damage identification (Feldman 1991, 1994a,b, Feldman & Braun 1995, Braun & Feldman 1997, and Feldman 1997). In all these studies, the signals were limited to "monocomponent" signals, i.e. without riding waves. Furthermore, the signals have to be symmetrical with respect to the zero mean. Thus, the method is limited to simple, free vibrations. Although Prime & Shevitz (1996) and Feldman (1997) have used it to identify some of the nonlinear characteristics through the frequency modulation in a nonlinear structure, the limitation of the data renders the method of little practical application in both identifying and locating the damage. The real value of the Hilbert transform had to wait to be demonstrated until Huang et al (1998a) introduced the empirical mode decomposition (EMD) method, which is based on the characteristic scale separation. The EMD method was developed to first operate on the data being processed and to then prepare it for the Hilbert transform. Therefore, we will discuss the time scale problem next, since this concept is central to this new approach.
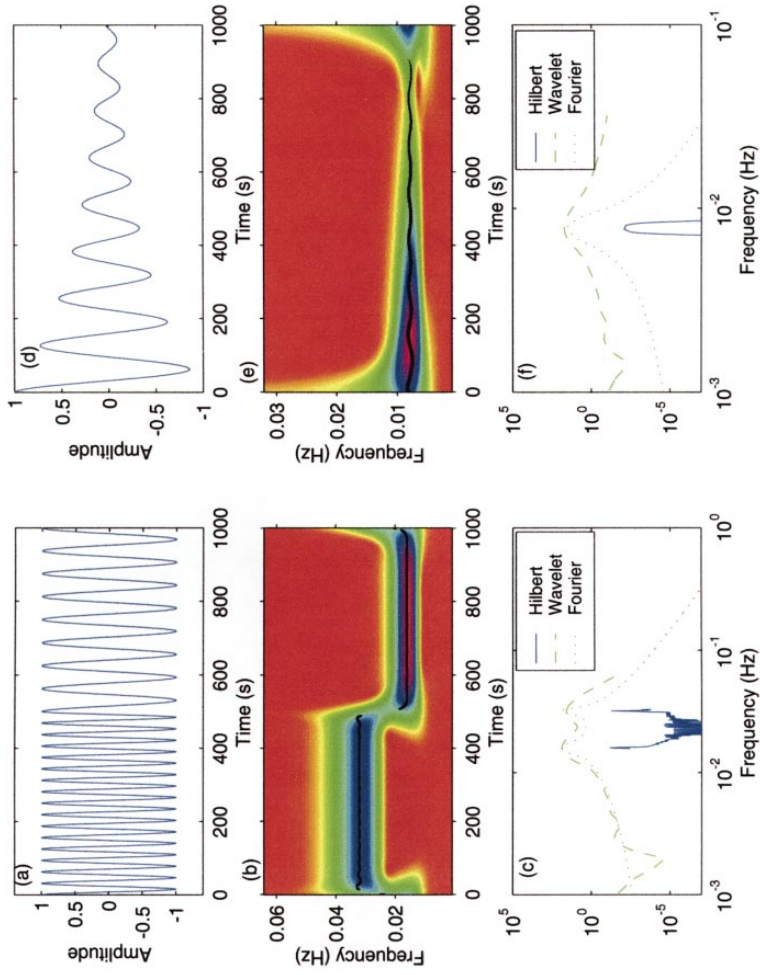
Figure 1 Comparisons of Wavelet and Hilbert representation for simple symmetric data. The Hilbert transform can be applied to these types of data to give better time-frequency resolution without difficulty.

THE CHARACTERISTIC SCALES

According to Drazin (1992), the first method of time series analysis is inspection by eye. This approach is, of course, subjective. But a trained eye can detect many trends and patterns of the data that are hard to quantify. Even to the untrained eye, there are certain properties of the data that are easy to pick up. Let us take, for example, the stationarity, the periodicity, the overall trend, and various scales defined by the time lapses between specific types of points. Valuable as these insights are, inspection by the eye alone is too subjective to be of any serious use. Of the various quantities the eye can pick up, the time scale is one that can be quantified most easily.

In interpretation of any physical data, the most important parameters are the time scale and the energy distribution with respect to it. There is no difficulty in defining the local energy density, but up until now, no clear definition of the local time scale has ever been given. In Fourier analysis, the time scales are defined as the periods of the continuous and constant-amplitude trigonometric components. As discussed in Huang et al (1998a), such a definition gives only a global averaged meaning to the energy and time scales. As such, these scales are totally divorced from the reality of time variations of either the amplitude or the frequency.

Statistical definitions of the time scale have been made by Rice (1944, 1945), who computed the expected numbers of zero-crossings, and the extrema for any data under linear, stationary, and normal distributed assumptions. Mathematically, the time scales are defined for any data, $x(t)$, as follows: The locations of t for

$$x(t) = 0 \tag{5}$$

are defined as the points of zero-crossings. The time spacing between successive zero-crossing is the zero-crossing time scale. The locations of t for

$$\dot{X}(t) = 0 \tag{6}$$

are defined as the points of the extrema. The time spacing between successive extrema is the extrema time scale.

Under the linear, stationary, and normal distribution assumptions, the expected number of zero-crossings and the expected number of extrema can be computed from Rice's formulae. But those definitions offer only a global measure, which cannot be applied to real nonlinear and nonstationary data. Because of the limitations set forth in Rice's assumptions, his results have also created a paradox: in many data, the number of expected extrema computed from his formula becomes unbounded. If most data are linear and stationary, then why can we not apply the formula to them? This is because the Fourier power spectra usually have asymptotic power law forms. For example, if the spectrum has

a -3 power law, then m_2 is unbounded. For a white noise or a delta function, the spectrum is white and then even the zero-crossing is undefined. Take ocean wave data as an example. The asymptotic form of the frequency spectrum has a power law form with the power between -4 and -5 (see, for example, Phillips 1958, Toba 1973, Phillips 1977, Kitaigorodskii 1983, Banner 1990, Belcher & Vassilicos 1987). Then, according to Rice's formula, the expected number of extrema is unbounded. Yet we can certainly count the extrema without any difficulty. This dilemma, however, has not yet caused most investigators to question the formulae and the assumptions involved, but it has led them to reject any formula that involves moments higher than the 4th. Such an approach has limited the statistical measure of time scales to the computation of the zero-crossings only. Hence the statistics of the zero-crossings are too crude to be of any real use.

The spacing of the extrema certainly offers a better measure of time scale, because this approach can measure wide-band data with multiple riding waves. It certainly agrees with our intuition of the time variations of the data. Refined as the extrema criterion is, it is not always precise enough. If one examines the data more closely, one will find that even the spacing of the extrema can miss some subtle time-scale variations, because there are weak oscillations that can cause a local change in curvature but not create a local extremum, a phenomenon known as hidden scales. To account for this type of weak signal, we introduce still another type of time scale based on the variation of curvature. Mathematically, this is equivalent to finding extreme values of

$$\frac{\ddot{X}}{(1 + \dot{X}^2)^{\frac{3}{2}}}. \tag{6a}$$

Dynamically, the curvature is equivalent to the measure of the weighted acceleration. Any change of sign of the curvature indicates a change of the sign of the force. As such, the curvature variation indeed has strong dynamic significance. As we see, if the extrema statistics have already encountered difficulties in models, the extreme values of the curvature would involve the 8th moment of the spectrum from the data. Trying to compute it under the linear and stationary assumption is impossible. Fortunately, this difficulty is but a mathematical artifact, a consequence of the linear and stationary assumptions invoked. We certainly can compute the curvature and its extrema, and then count them. Consequently, the failure of Rice's formulae is another indication of what we believe to be the falsehood of the commonly invoked assumptions of linearity and stationarity.

We now have three methods of measuring the time scales: the time between successive zero-crossings, the time between successive extrema, and the time between successive curvature extrema. In each case, the time span is a local

measure of the time variation of events. In the case of extrema and curvature spans, the local time scale counts all the waves, whether they cross the zero line or not. The aim is to define a local time scale of oscillation that will change from one extreme (by the restoring force through the zero point) to the other extreme of the opposite sign. This is the characteristic time scale. It is local, and it represents only one mode of oscillation. So we regard it as the intrinsic scale of the oscillation.

Zero-crossing is a very crude measure of the data. Unless the data are truly narrow band, there might be many extrema between two consecutive zero-crossings. Our eyes are much more sensitive to the variations of the spacing between extrema, and these variations offer a more detailed measure of the given phenomena. Yet the time lapses between extrema have been problematic. The fourth moments for many phenomena are not convergent, so the expected number of extrema is impossible to compute, even though it might be easily counted. This paradox is easily resolved by considering a bold concept: The Fourier power-law spectra of most data are artificial. Most of the high-frequency components are from the spurious harmonics from either nonlinearity (singular points, such as corners, and cusps in the data train), or nonstationarity. Following Huang et al (1998a), the time scale between extrema is the key and therefore will be used as the time scale in the decomposition.

THE EMPIRICAL MODE DECOMPOSITION METHOD: THE SIFTING PROCESS

As discussed by Huang et al (1996, 1998a), the empirical mode decomposition method is necessary to deal with both nonstationary and nonlinear data. Unlike almost all the previous methods, this new method is intuitive, direct, a posteriori, and adaptive, with the basis of the decomposition being derived from the data. The decomposition is developed from the simple assumption that any data consist of different simple intrinsic modes of oscillations. Each mode may or may not be linear, and will have the same number of extrema and zero-crossings. Furthermore, the oscillation will also be symmetric with respect to the "local mean." The term local mean is an oxymoron. Any mean will involve a time scale to define it. Here, however, the mean is defined through the envelopes without resorting to any time scale. (Its precise definition is given below.) With this definition of local mean, modes of different time scales can be separated by their characteristic scales, defined as the time lapses between the successive extrema. At any given time, there might be many different coexisting modes of oscillation, each superimposed on the others. The final complicated data results. Once separated, each mode should be independent of the others; they have no multiple extrema between successive zero-crossings. Thus each is designated

as an intrinsic mode function (IMF) by the following definitions: (*a*) in the whole data set, the number of extrema and the number of zero-crossings must either equal or differ at most by one; and (*b*) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

An IMF represents a simple oscillatory mode as a counterpart to the simple harmonic function, but it is much more general. With the definition, one can decompose any function as follows: (*a*) Identify all the local extrema, then connect all the local maxima by a cubic spline line as the upper envelope; (*b*) Repeat the procedure for the local minima to produce the lower envelope. The upper and lower envelopes should cover all the data between them. Their mean is designated as m_1, and the difference between the data and m_1 is the first component, h_1, i.e.

$$X(t) - m_1 = h_1. \tag{7}$$

The procedure is illustrated in Huang et al (1998a).

Ideally, h_1 should be an IMF, for the construction of h_1 described above should have required it to satisfy all the requirements of an IMF. Yet, even if the fitting is perfect, a gentle hump on a slope can be amplified to become a local extremum in changing the local zero from a rectangular to a curvilinear coordinate system. After the first round of sifting, the hump may become a local maximum. New extrema generated in this way recover the proper modes lost in the initial examination. In fact, the sifting process can recover signals representing low-amplitude riding waves with repeated siftings.

The sifting process serves two purposes: to eliminate riding waves, and to make the wave profiles more symmetric. While the first condition is absolutely necessary for separating the intrinsic modes and for defining a meaningful instantaneous frequency, the second condition is also necessary in case the neighboring wave amplitudes have too large a disparity. Toward these ends, the sifting process has to be repeated as many times as is required to reduce the extracted signal to an IMF. In the subsequent sifting process steps, h_1 is treated as the data; then

$$h_1 - m_{11} = h_{11}. \tag{8}$$

After repeated sifting, i.e. up to k times, h_{1k} becomes an IMF, that is

$$h_{1(k-1)} - m_{1k} = h_{1k}; \tag{9}$$

then it is designated as

$$c_1 = h_{1k}, \tag{10}$$

the first IMF component from the data.

As described above, the process is indeed like sifting: to separate the finest local mode from the data first, based only on the characteristic time scale. To guarantee that the IMF components retain enough physical sense of both amplitude and frequency modulations, the number of times the sifting process repeats has to be limited. Too many sifting cycles could reduce all components to a constant-amplitude signal with frequency modulation only. Then, the components would lose all their physical significance. A simple criterion for stoppage is when the number of extrema equals the number of zero-crossings. The original Cauchy-like convergence criterion introduced by Huang et al (1998a) should be used with great care, because the deviations between successive siftings are controlled primarily by the appearance of new extrema from their previously hidden state. Such problems can be resolved now with curvature sifting. Therefore, the criterion for stoppage can be simplified, as proposed here.

Overall, c_1 should contain the finest scale or the shortest period component of the signal. We can separate c_1 from the rest of the data by

$$X(t) - c_1 = r_1. \tag{11}$$

Since the residue r_1 still contains longer-period components, it is treated as the new data and subjected to the same sifting process as described above. This procedure can be repeated for all the subsequent r_j's, and the result is

$$r_1 - c_2 = r_2,$$
$$\cdots \tag{12}$$
$$r_{n-1} - c_n = r_n$$

The sifting process can be stopped by any of the following predetermined criteria: either when the component c_n or the residue r_n becomes so small that it is less than the predetermined value of substantial consequence, or when the residue r_n becomes a monotonic function from which no more IMF can be extracted. Even for data with zero mean, the final residue still can be different from zero. If the data have a trend, the final residue should be that trend. By summing up Equations (11) and (12), we finally obtain

$$X(t) = \sum_{j=1}^{n} c_j + r_n. \tag{13}$$

Thus, one can achieve a decomposition of the data into n-empirical modes, and a residue r_n, which can be either the mean trend or a constant. As discussed here, to apply the EMD method, a mean or zero reference is not required; EMD needs only the locations of the local extrema. The zero references for each component will be generated by the sifting process. Without the need of the

zero reference, EMD avoids the troublesome step of removing the mean values for the large DC term in data with non-zero mean, an unexpected benefit.

To illustrate the sifting process, we will use a set of length-of-the-day data covering the period from 1978 to 1988. The data are given in Figure 2*a*. Clearly, the data are quite complicated, with many local extrema but no zero-crossings, because the time series represents all positive numbers. Although the mean can be treated as a zero reference, defining it is hard, for the whole process is transient. This example illustrates the advantage of adopting the successive extrema for defining the time scale; it also illustrates the difficulties of dealing with nonstationary data: Even a meaningful mean is impossible to define, but for the EMD method, this difficulty is eliminated. Figure 2*b* summarizes all the IMFs obtained from this repeated sifting process. Figure 2*b* illustrates a total of 7 components plus a residue term. In comparison to the traditional Fourier expansion, one can immediately see the efficiency of EMD.

The components of EMD are usually physical, for the characteristic scales are physical. In Figure 2*b*, we can see the yearly cycle clearly as the fifth component C_5. The first two are semi-monthly and monthly tidal modulation of the rotation speed of the earth. Different from the Fourier analysis, each component still retains both frequency and amplitude modulations. For example, the amplitude of the annual fluctuation is slightly larger at 1982, which happens to be an unusually strong El Niño year.

To demonstrate the completeness of the decomposition, the IMF components can be added back one by one to form the original data. Figure 3*a* shows the data in a dotted line and the residue term in a solid line. By itself, the residue is not an impressive running mean of the data. It should not be, for the last IMF is not a mean; it is only the residue after all the oscillatory terms have been separated from the signal. In this sense, it could be a trend. By adding the longest period IMF component, the sum gives a sense of a much better running mean, as in Figure 3*b*. The third component gives the annual cycle. The sum immediately shows the fluctuation of the length of the day by year in Figure 3*d*. By successively adding all the components, we eventually get Figure 3*h*, which is indistinguishable from the original data. The numerical difference between the sum of the IMFs and the original data is given in Figure 3*i*. The magnitude is only of the order of 10^{-15}. It is the round-off error for the computer. Thus, the completeness of the expansion is proven numerically.

To use the unique definition of instantaneous frequency, we have to reduce an arbitrary data set into IMF components from which an instantaneous frequency value can be assigned to each IMF component. Consequently, for complicated data, we can have more than one instantaneous frequency at a time locally. After decomposing the data into IMFs, and after operating on these with the Hilbert transform, we can then present the result, which we call the Hilbert spectrum.
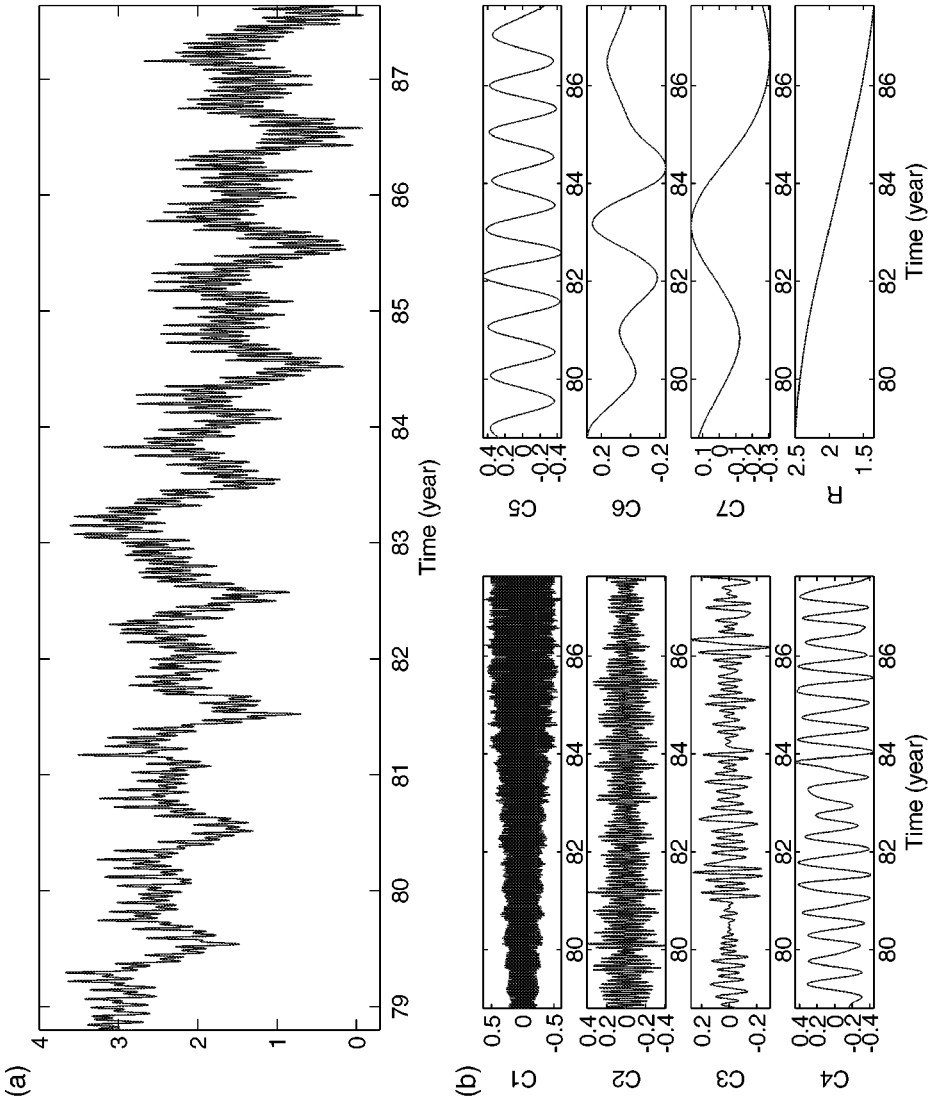
Figure 2 Demonstration of the empirical mode decomposition method from length-of-the-day deviation data; the unit is microseconds.
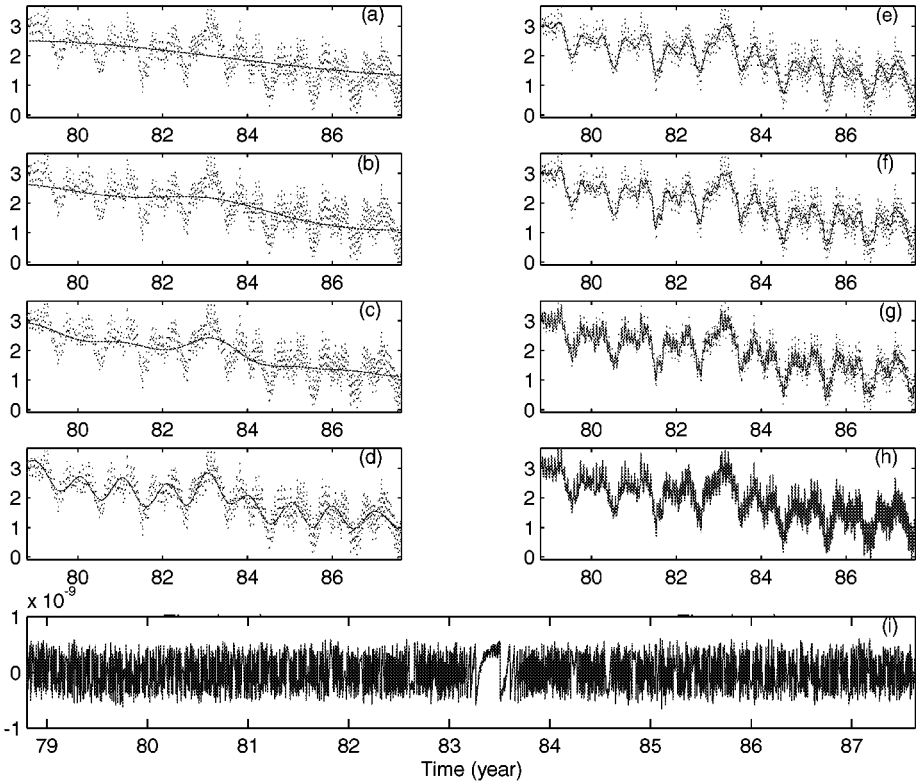
Figure 3   Reconstruction of the original data from the IMF components. The difference between the reconstructed and the original data is only 10^{-9} microsecond.

Intermittency Test

The sifting process described above seems straightforward. Yet straightforward application of the sifting method may run into difficulties when the data contain intermittency, which will cause mode mixing. We discuss this phenomenon in more detail below.

Let us consider the data given in Figure 4*a*, where there is a train of large-amplitude sine waves with another train of small-amplitude sine waves occurring intermittently. With application of the straightforward sifting, we will obtain the components as shown in Figure 4*e–h*, in which the first two IMF components contain seriously mixed modes, that is, modes of very different periods. Take the component in 4*e* as an example. The small wave train is clearly identified. Wherever the small waves are identified, the underlying
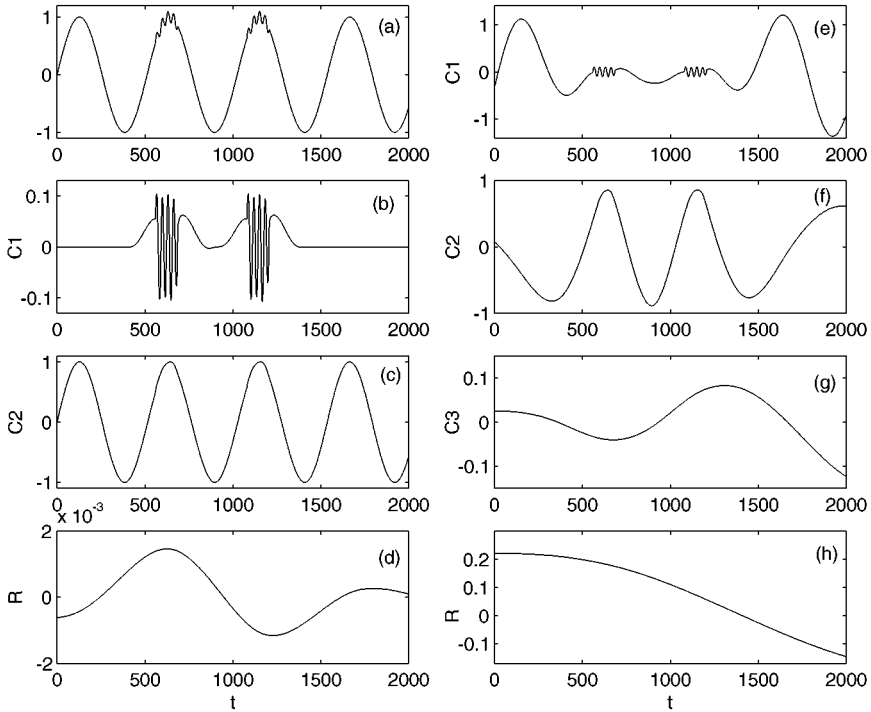
Figure 4 Effect of intermittency criterion in the EMD sifting process invoked to eliminate mode mixing. Sequence *b–d* with intermittency check; sequence *e–h* without intermittency check.

large waves will not be included in this IMF component. On the other hand, wherever there is no small-wave component, the large waves are retained as part of the component. As a result, there is a great disparity in the periods of the first IMF component. This is mode mixing; it is caused by intermittency occurring in part of the signal.

To overcome the mode mixing, a criterion based on the period length is introduced to separate the waves of different periods into different modes. The criterion is set as the upper limit of the period that can be included in any given IMF component. With this criterion introduced, the result is shown in Figure 4*b–d*. Clearly, the intermittent small wave was separated from the large waves.

Any additional criterion introduced in the sifting process implies an intervention with a subjective condition. Such intervention could cause severe bias in the final result; therefore, the introduction of any additional condition should be justified with clear and strong arguments. As a rule, data should be processed first without any added conditions. The intermittency criterion should

be introduced only when the sifted results clearly show the problem of mode
mixing. Mode mixing cannot be justified from physical grounds, for any oscil-
lator cannot have great disparity in its periods.

THE HILBERT SPECTRUM

Having obtained the intrinsic mode function (IMF) components, one will have
no difficulty in applying the Hilbert transform to each of these IMF components
and computing the instantaneous frequency according to Equation (4). After
performing the Hilbert transform to each IMF component, the original data can
be expressed as the real part (RP) in the following form:

$$X(t) = RP \sum_{j=1}^{n} a_j(t) e^{i \int \omega_j(t)dt}. \tag{14}$$

Here we have left out the residue r_n on purpose, for it is either a monotonic
function or a constant. Although the Hilbert transform can treat the monotonic
trend as part of a longer oscillation, the energy involved in the residual trend
could be overpowering. In consideration of the uncertainty of the longer trend,
and in the interest of information contained in the other low-energy but clearly
oscillatory components, the final non-IMF component should be left out. It
could be included, however, if physical considerations justify its inclusion.

Equation (14) gives both amplitude and frequency of each component as
functions of time. The same data, if expanded in the Fourier representation,
would be

$$X(t) = RP \sum_{j=1}^{\infty} a_j e^{i\omega_j t}, \tag{15}$$

with both a_j and ω_j constants. The contrast between Equations (14) and (15)
is clear: the IMF represents a generalized Fourier expansion. The variable
amplitude and the instantaneous frequency have not only greatly improved the
efficiency of the expansion, but also enabled the expansion to accommodate
nonstationary data. With the IMF expansion, the amplitude and the frequency
modulations are also clearly separated. Thus, we have broken through the re-
striction of the constant amplitude and fixed frequency imposed by the Fourier
expansion, and arrived at a variable amplitude and frequency representation.
Now let us illustrate the difference between the two expressions graphically
in the frequency-energy-time space. Figure 5a (see color figure at end of
volume) represents a set of highly transient data. In the Fourier expansion,
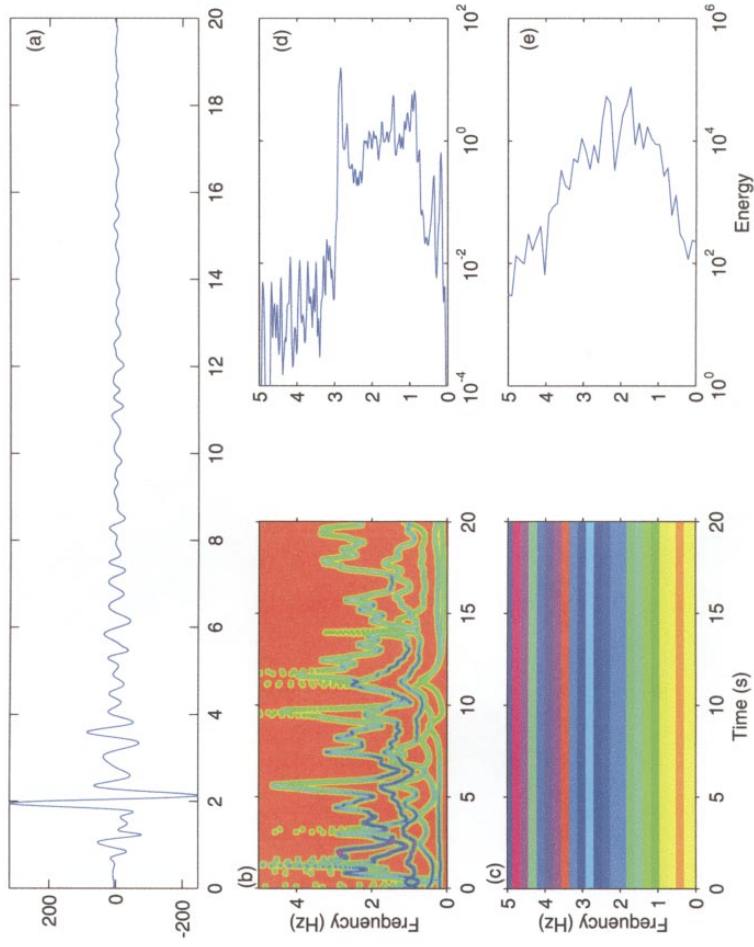the frequency and amplitude are not time dependent; therefore, all Fourier

Figure 5 The comparison of Hilbert and Fourier spactral analysis for highly transient data. The Fourier method cannot reveal of the variation of the signal with time.
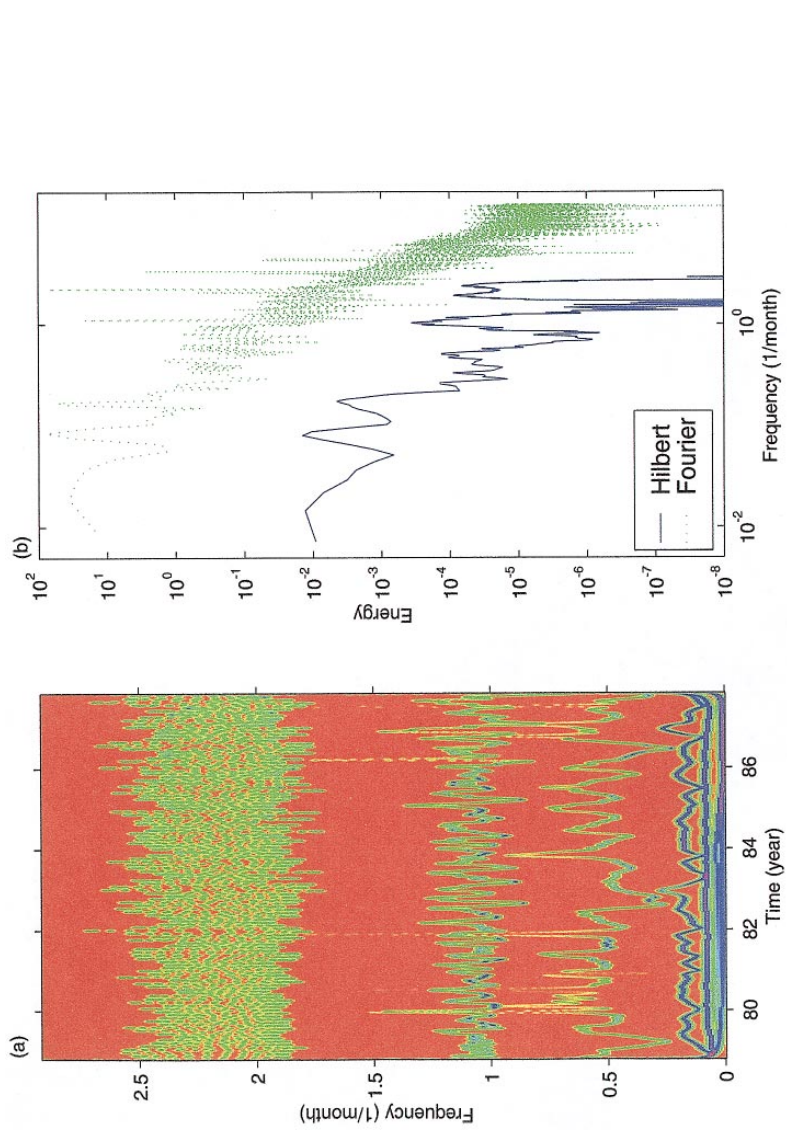
Figure 6 The Hilbert and marginal spectra for the length of the day data. It gives detailed time variation of the frequency with time.

components are represented by rectangular blocks with thickness of $d\omega$, as in Figure 5c. Consequently, the only information is the projection of the blocks on the frequency-energy plane. This is why the Fourier spectra can be meaningful only for stationary data. The same data, if expanded in terms of IMF, will produce the result as given in Figure 5b. As the frequency of each component is a function of time, it is a curve on the time-frequency plane. Furthermore, because the amplitude (or the energy) of each component is also a function of time, the final energy representation is a curve in the three-dimensional space of frequency-energy-time. This frequency-time distribution of the amplitude is designated as the Hilbert amplitude spectrum $H(\omega, t)$, or simply the Hilbert spectrum. If amplitude squared is preferred to represent energy density, then the squared values of amplitude can be substituted to produce the Hilbert energy spectrum just as well.

Various forms of the Hilbert spectra presentations can be made: color-coded maps and contour maps all with or without smoothing. The Hilbert spectrum in the color map format for the length-of-the-day data is given in Figure 6a (see color figure at end). The Hilbert spectrum appears only in the skeleton (or line) form with emphasis on the frequency variations of each IMF, while the wavelet analysis result usually gives a smoothed energy contour map with a rich distribution of higher harmonics. The skeleton presentation is more desirable, because it gives more quantitative results. Bacry et al (1991) have tried to extract the wavelet skeleton as the local maximum of the wavelet coefficient, but even that approach is encumbered by the harmonics. If more qualitative results are desired, a fuzzy view can also be derived from the skeleton presentation by using two-dimensional smoothing. We will discuss the significance of the difference between the Hilbert and wavelet presentations further on in this review.

With the Hilbert spectrum defined, we can also define the marginal spectrum, $h(\omega)$, as

$$h(\omega) = \int_0^T H(\omega, t)dt. \tag{16}$$

The marginal spectrum offers a measure of the total amplitude (or energy) contribution from each frequency value. It represents the cumulated amplitude over the entire data span in a probabilistic sense. In Figures 5d and 6b, the solid lines give the corresponding marginal spectrum of the Hilbert spectrum given in Figures 5b and Figure 6a, respectively. The lack of harmonics is clearly demonstrated. Furthermore, Figure 5d showed a much richer energy content in the low-frequency range than the corresponding Fourier spectrum in Figure 5e. This is usually the case, for the constant amplitude and frequency Fourier representation would never be able to depict the true energy content. It should

be pointed out that the marginal spectra should not be used for any nonstationary data, for the marginal spectra are the projections rather than the substance of the real frequency-energy-time distribution.

As pointed out by Huang et al (1996), the frequency in either $h(\omega, t)$ or $h(\omega)$ has a totally different meaning from the Fourier spectral analysis. In the Fourier representation, the existence of energy at a frequency ω means that a component of a sine or a cosine wave persisted through the time span of the data. Here, the existence of energy at the frequency ω means only that in the whole time span of the data, there is a higher likelihood for such a wave to have appeared locally. In fact, the Hilbert spectrum is a weighted non-normalized joint amplitude-frequency-time distribution. The weight assigned to each time-frequency cell is the local amplitude. Consequently, the frequency in the marginal spectrum indicates only that the likelihood of an oscillation with such a frequency exists. The exact time of that oscillation is given in the full Hilbert spectrum.

Having defined the Hilbert spectrum, we thus have a real frequency-energy-time representation of the data that is quantitative. With it, Huang et al (1998a) have defined the degree of stationarity (DS) as

$$DS(\omega) = \frac{1}{T} \int_0^T \left(1 - \frac{H(\omega, t)}{h(\omega)/T} \right)^2 dt. \tag{17}$$

This definition of degree of stationarity is very similar to the intermittency used in the wavelet analysis proposed by Farge (1992). A degree of statistical stationarity is also defined by Huang et al (1998a). The instantaneous energy, IE, can also be defined as

$$IE(t) = \int_\omega H^2(\omega, t) d\omega. \tag{18}$$

VALIDATION AND CALIBRATION OF THE HILBERT SPECTRUM

Through empirical mode decomposition and the associated Hilbert spectral analysis, we obtained the probabilistic Hilbert spectrum representation of the nonlinear and nonstationary data. Now we will validate the approach and the results, and calibrate its fidelity against the best existing method, wavelet analysis.

Let us first consider the following mathematical model:

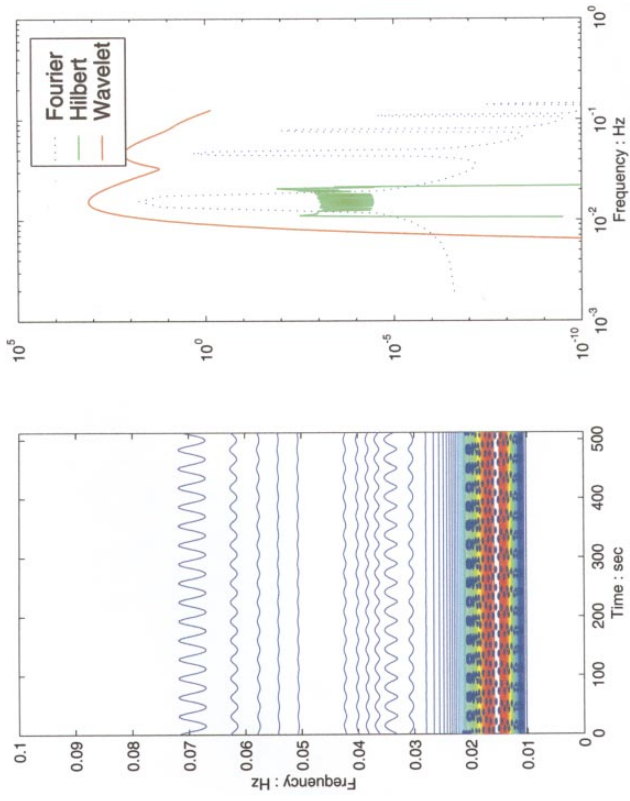$$X(t) = \cos(\omega t + \varepsilon \sin 2\omega t). \tag{19}$$

Figure 7 Comparison of Hilbert and Wavelet spectra for the intra-wave modulation case. The Hilbert view avoids the harmonics, and it gives the real intra-wave frequency modulations.

According to the classic wave theory, this expression is a clear case of intra-wave frequency modulation. The frequency, Ω, at any time is simply

$$\Omega = \omega(1 + 2\varepsilon \cos 2\omega t). \tag{20}$$

Yet from Equation (19), it is easy to show that

$$X(t) = \left(1 - \frac{\varepsilon}{2}\right) \cos \omega t + \frac{1}{2} \cos 3\omega t + \ldots, \tag{21}$$

which is similar to the second-order approximation of the Duffing equation through perturbation analysis. The Hilbert and wavelet spectra for these are given in Figure 7a (see color figure). Here we have two views for the same mathematical expression. Both representations can be used to construct the original curve, but they convey very different physical meanings. Clearly, the one based on classical wave theory is the more physical one, for it is how we define the function. From this example, we can see that there are two types of frequency modulations, inter-wave and intra-wave. The first type is familiar to us; the frequency of the oscillation gradually changes as do the waves in a dispersive system. Technically, in dispersive waves, the frequency also changes within one wave, but that was not emphasized either for convenience, or for lack of a more precise frequency definition. The second type is less familiar, but it is also a common phenomenon: if the frequency changes from time to time within a wave, its profile can no longer be described by simple trigonometric functions. Therefore, any wave profile deformation from the simple sinusoidal form implies intra-wave frequency modulation. In the past, such phenomena were treated as harmonic distortions. The purpose of the harmonics is not to represent the true frequency distribution, but rather to represent the waveform. The marginal spectra of the Hilbert and wavelet spectra together with the Fourier spectrum are shown in Figure 7b. Here the Hilbert spectrum clearly depicts the modulation of the frequency as shown by Equation (19). Wavelet analysis again gives a poor frequency resolution. Thus, we show in detail that most waveform deformations are better viewed as intra-wave frequency modulations. This is the core of the Hilbert view: it is more physical.

This simple example again illustrates that the instantaneous frequency, with intra-wave frequency modulation defined by the EMD and Hilbert spectrum, does make physical sense. In fact, such an instantaneous frequency presentation reveals more details of the system: it reveals the variation of the frequency within one period, a view never seen before.

The above examples have not only validated the EMD and the Hilbert spectrum representation, but also clarified the conditions under which spurious

harmonics are generated in Fourier-based analysis: nonlinearity and nonstationarity. In the past, this crucial problem has not been examined carefully. Typically, perturbation analysis gives a solution in series expansion. Each term is the solution of a linear equation. Although infinite series expansion is so powerful that it can approximate some transient phenomena with uniform amplitude components, their physical meaning has never been examined critically. The mathematical success has obscured our physical insights. With these idealized examples, we have established the validity and the limitations of Hilbert spectral analysis. Next, we will present some applications in both numerically computed results from low-dimensional nonlinear equations, and some data from observations.

CLASSIC NONLINEAR SYSTEMS

The advantage of studying classic nonlinear systems is their simplicity, yet they contain all the essentials of the possible nonlinear effects. All these systems have been studied extensively; therefore, although most of their dynamic characteristics are familiar, their detailed physics may not be. We cite two examples: the Duffing equation and the Rössler equation, both used by Huang et al (1998a).

The Duffing Equation

We use the classic Duffing equation to illustrate intra-wave frequency modulation. The Duffing equation is

$$\frac{d^2x}{dt^2} + x - \varepsilon x^3 = b \cos \upsilon t, \tag{22}$$

in which ε is a small parameter. This equation can be written in a slightly different form as

$$\frac{d^2x}{dt^2} + x(1 - \varepsilon x^2) = b \cos \upsilon t, \tag{23}$$

where we have factored out x and $(1 - \varepsilon x^2)$. This equation can be viewed as a nonlinear spring with a variable spring constant $(1 - \varepsilon x^2)$. If ε is zero, this is a simple oscillator with constant period. If ε is not zero, the spring constant is no longer a constant; it becomes a function of position. The period is longest when the position is near the origin, and the shortest when the position is at the maximum displacement. Thus, this nonlinear oscillator has variable frequency within one cycle of oscillation. Clearly, this is a case of intra-wave frequency modulation.
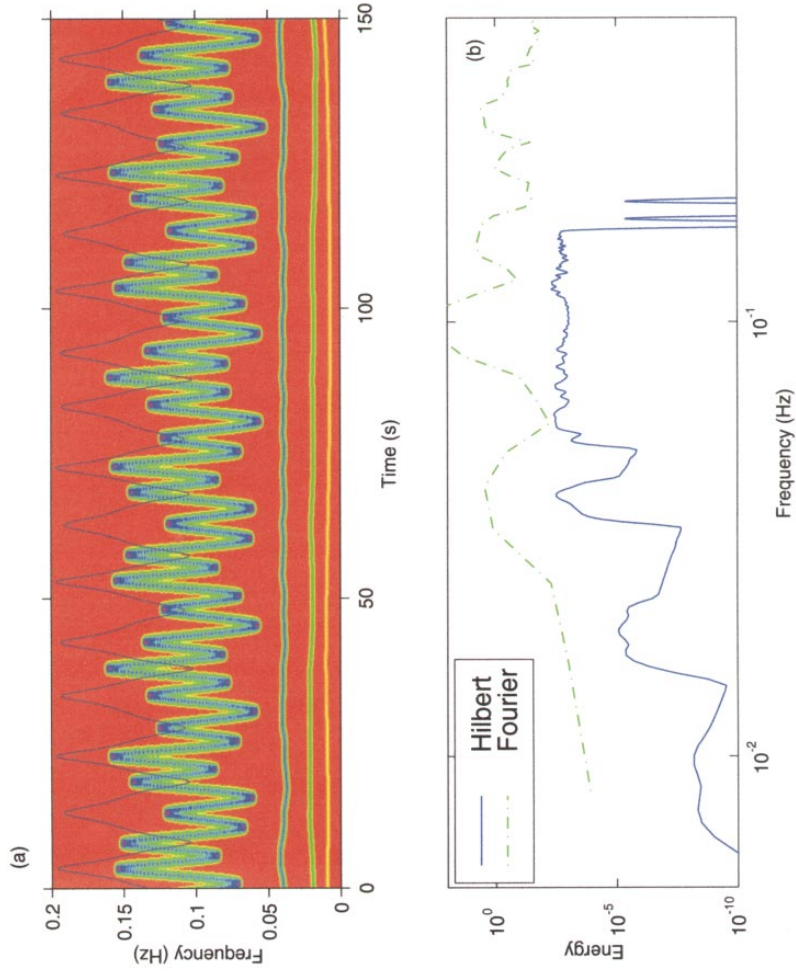
Figure 9 The Hilbert and Marginal spectra for the Duffing equation solution. The Hilbert reveals the detailed frequency modulation of the data.

Traditionally, this problem has been treated by perturbation methods. If one uses the straightforward perturbation method, one gets the secular term as

$$x(t) = \cos t + \varepsilon \left(\frac{3}{4} t \sin t + \frac{1}{32} (\cos t - \cos 3t) \right). \tag{24}$$

The homogeneous solution of Equation (22) is given in Drazin (1992) as

$$x(t) = a \cos \omega t + \frac{1}{32} \varepsilon a^2 (\cos \omega t - \cos 3\omega t) + O(\varepsilon^2 a^5) \tag{25}$$

and

$$\omega = 1 - \frac{3}{8} \varepsilon a^2 + O(\varepsilon^2 a^4) \quad \text{as } \varepsilon \to 0. \tag{26}$$

The above solutions work only for small ε. If we adopt the initial values $[x(0), x'(0)] = [1, 1]$; and $a = 1, b = 0.1, \varepsilon = 1, \upsilon = 1/25$ Hz, we cannot use the solution to evaluate the functional value of x anymore, for ε is now finite. The full solution of Equation (22) subject to the initial conditions given here has to be computed numerically. The numerical solution is given in Figure 8a. The waveform is far from sinusoidal. All the deformations indicate nonlinear effects, traditionally represented by harmonics. The full solution in the phase plane is given in Figure 8b, and we see that the locus of the solution tends to bunch into three distinguishable paths. This suggests that the motion within this time period contains a period-three oscillation in addition to the forcing function and the intrinsic oscillation time scales. These data, when subjected to the empirical mode decomposition method, yield four components and a residue, as shown in Figure 8c.

The first component is the intrinsic oscillation of the system subject to the forcing, the second component is the forcing function, and the third component is the period time scale. All of these components are physical. With these intrinsic mode functions, the Hilbert spectrum can be constructed as in Figure 9a (see color figure) together with the original data. The most interesting component is the intrinsic oscillation term. In the Hilbert spectrum, it is represented by an oscillatory line indicating the frequency change from one instant to another. A detailed comparison between the Hilbert spectrum and the data shows the uneven frequency variations from one wave to the next and is a faithful and detailed representation of the fact that the motion has a different frequency within one oscillation. The intrinsic frequency shows strong intra-wave frequency modulation, which is presented as a variable frequency oscillating between 0.06 and 0.16 Hz, with a mean around 0.11 Hz, the averaged frequency as predicted by the Hamiltonian method. The detailed variations of the intrinsic frequency indicate that it contains both inter- and intra-wave frequency modulations. The forcing function is also clearly shown with the expected frequency.
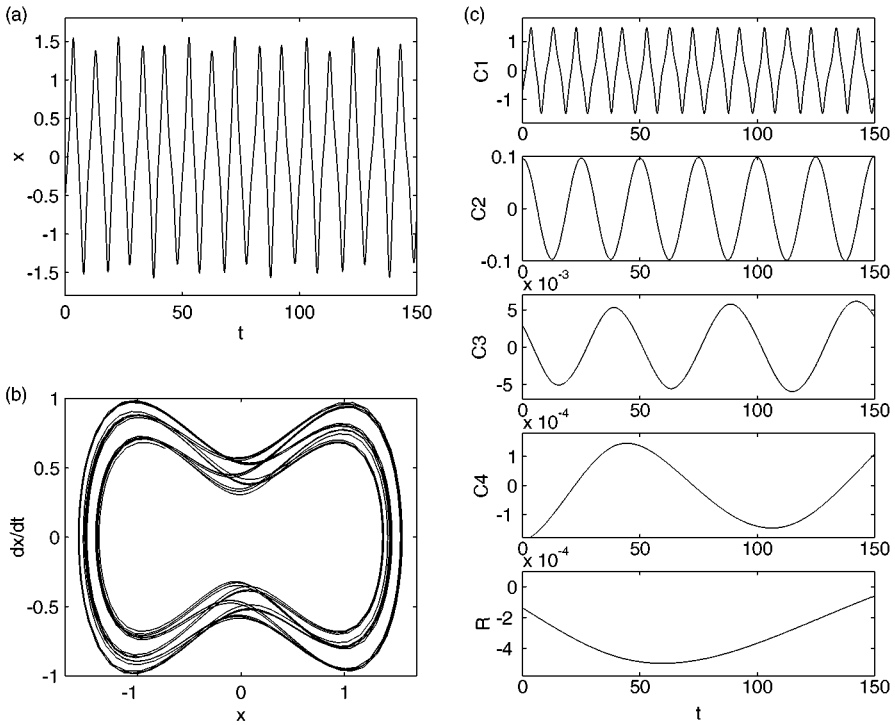
Figure 8   Data, phase diagram, and IMF components for the solution of the Duffing equation. Trajectories of motion show clear bunching at the preferred paths, indicating additional time scales.

The long-period component is the low-frequency and low-amplitude signal. It represents the slow, aperiodic wobbling of the phase depicting the period-three bunching of the paths. It is real. If we compute a longer time series, there will be still longer-period motions. As the motion is already chaotic, the path may never repeat itself.

A wavelet representation of the same data is given by Huang et al (1998a). Because the Morlet wavelet used here is Fourier based, the variation of the frequency has to be represented by harmonics. The marginal Hilbert spectra together with the Fourier spectrum of the data are given in Figure 9*b*. Here, the lack of harmonics in the Hilbert spectrum is clearly shown. This is a totally different view of the nonlinear system. It tracks the instantaneous change of the waveform by small changes of frequency rather than by using harmonics. Although both the Fourier and the Hilbert representations show the same effect of waveform deformations, the physical meaning is very different: The Hilbert representation gives a true physical interpretation of the dynamics by

indicating the instantaneous value of the frequency, and thus the Hilbert view
is much more physical.

The Rössler Equation

The nonlinear effect can also be represented by another classical example in
the Rössler equation as

$$\frac{dx}{dt} = -(y + z),$$

$$\frac{dy}{dt} = x + \frac{1}{5}y, \tag{27}$$

$$\frac{dz}{dt} = \frac{1}{5} + z(x - \mu),$$

in which μ is a constant parameter. The numerical value of x computed with
$\mu = 3.5$ is given in Figure 10a with its phase diagram in Figure 10b. This is the
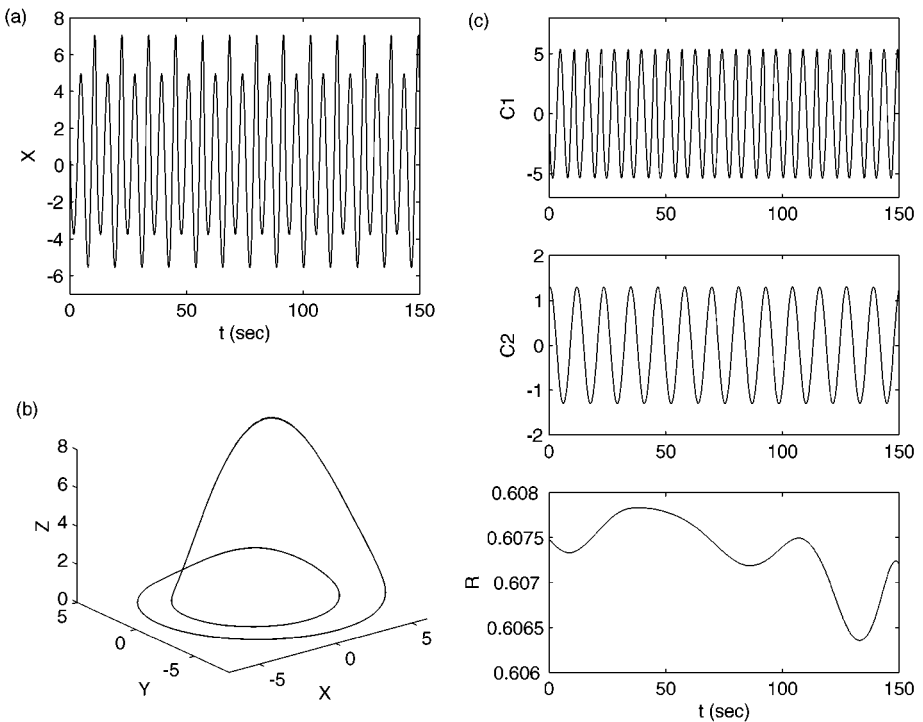
Figure 10   Data, phase diagram, and IMF components for the solution of the Rössler equation.
This is the case of period doubling, which indicates two time scales.

case of period doubling, when, starting from any point, one needs to spend twice the simple oscillation time to return to the original state. Obviously, there must be two time scales. The empirical mode decomposition method indeed gives precisely two components and a numerically insignificant residual error term in Figure 10*c*. As this is the period doubling case, we should only expect two time scales, as shown in the IMFs. The same case will require many harmonics, but the harmonics fail to give any indication of period doubling, because the harmonics represent the waveform deformation instead of the period doubling.

The Hilbert spectrum of the result is presented in Figure 11*a* (see color figure), in which the frequency of the first component fluctuates over a considerable range. But nowhere is the frequency value higher than 3 Hz. The wavelet spectrum for the same case would have harmonics for very high frequency, as shown in Huang et al (1998a). If we compute the marginal spectrum and plot them together with the Fourier spectrum in Figure 11*b*, the difference is also clear: The lack of harmonics in the representation increases the clarity of the final result. Furthermore, the main peak of the Fourier spectrum represents only a global weighted mean frequency. Its does not represent any true value during the oscillation.

With these examples, we have demonstrated that the new idea of intra-wave frequency modulation can easily depict the minute variations of the waves in the IMF component and the data. Again, with the Hilbert spectrum as a guide, the unevenness of the intra-wave frequency variation can be shown to have one-to-one correspondence with the variations of the waveform in the data.

Comparison with the Hamiltonian Solution

From the above examples, one can see that Hilbert spectral analysis offers more detailed information than the Hamiltonian system, in which the averaged frequency is defined as

$$\frac{\partial J}{\partial t} = 0, \quad \text{and} \quad \omega = \frac{\partial \theta}{\partial t} = \frac{\partial H(J)}{\partial J}, \tag{28}$$

with J as the averaged action density, defined by

$$J = \frac{1}{2\pi} \iint dp \, dq, \tag{29}$$

in which θ is the angular variable and H is the total Hamiltonian in terms of the action density variable. In these canonical expressions, the most important parameter is the averaged period or frequency, based on which the Poincaré section and the modern topological view of the dynamic system are built. With such a view, the shape of the phase plane is not as important as the time needed to trace a full cycle. As long as they are closed curves, they are topologically
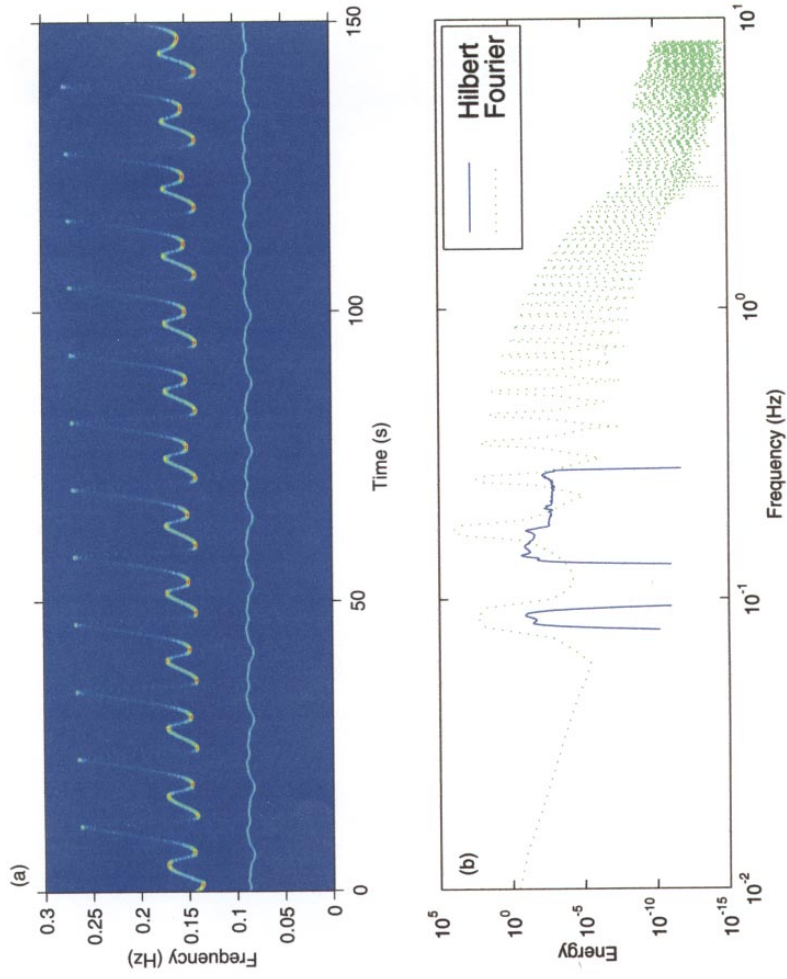
Figure 11 The Hilbert and Marginal spectra for the Rössler equation solution. The peak of the Fourier spectrum is not any real scale of the oscillation, but a weighted mean.

equivalent. Yet the different shapes of the phase curves represent different details of the oscillations. Such details can be represented only by the instantaneous frequency, as shown above. The motion described by nonlinear equations clearly requires the instantaneous frequency variation to specify its full physical details. Thus, the Hilbert view is preferred over the previous view of classical nonlinear mechanics. We will now use the new method to view nonlinear water wave problems.

WATER WAVE PROBLEMS

Now we will turn to real physical phenomena, the problem of water waves. All water waves are nonlinear (Whitham 1974); therefore, harmonic analysis has always been an inseparable part of wave phenomenon studies. Yet such an approach offers a confused view: As water waves are dispersive, waves of different frequencies will propagate at different phase velocities. In terms of harmonics, however, we are forced to separate the waves into two categories: free waves and bounded waves. Harmonic components do not obey the dispersive relationship. Consequently, for a given wave component of certain frequency and wave number, one must first determine whether it is a free or a bounded wave before discussing its propagating properties. For a random wave field, the simple, direct, and logical way to represent it is through its spectrum. The harmonics in the Fourier spectrum, however, create a real problem in spectral analysis of how to determine which component is free and which is bounded. In the traditional Fourier view, the typical spectrum has a rather narrow peak and a wideband tail. Near the peak region, the waves are mostly free waves, which propagate according to the dispersion relationship. Toward the tail, the waves are mostly harmonic in nature, but even here, free waves are still possible. Therefore, for any given frequency or wave number, the waves can either be free or bounded harmonics of other free waves. As bounded harmonics, a particular wave can be the harmonic of the free fundamental waves with $1/n$ times its frequency. Thus, the true nature of any component in a wave spectrum can never be exactly determined. This is the consequence of using Fourier analysis.

Thus we must point out here that the harmonics are a mathematical artifact. The characteristics of nonlinear water waves are represented in the deformation of the wave form. Now, we will examine the nonlinear wave problem with Hilbert spectral analysis. The central idea is to use intra-wave frequency modulation to explain wave form deformation. Using the nonlinear wave phenomena as examples, we can contrast the different views and gain new insights, which will also help us in interpreting more complicated cases in natural phenomena with high degrees of freedom. We will start with the classic Stokes wave.

The Stokes Wave

The Stokes wave is one of the first successes of a mathematical description of
natural phenomena (Stokes 1847). The classical Stokes wave profile reveals the
nonlinearity by its sharpened crests and rounded-off troughs. Longuet-Higgins
(1963), Huang et al (1983, 1984, 1990a,b) and Shen et al (1994) have modeled
the asymmetric form, which gives the skewness in the water surface eleva-
tion distribution. In those attempts, harmonics were used as a mathematical
tool. As they are modeling the waveform deformation, that approach is per-
fectly legitimate. All the harmonics, however, can model only the up-and-
down asymmetry. Front and back asymmetry is known to exist and has been
modeled by Longuet-Higgins (1982). We will see what the Hilbert view can
reveal.

A short section of the wave record used by Huang et al (1998a) is reproduced
in Figure 12. Applying the EMD to these data produced eight components.
The most important one is the second component, which accounts for almost
all the energy of the data. Its Hilbert spectrum is given in Figure 12*a*, while
the corresponding Morlet wavelet spectrum is given in Figure 12*b*, all with the
wave profile superimposed on them. Clearly, the waves are nonlinear, for the
Hilbert spectrum shows instantaneous frequency modulation, the hallmark for
nonlinear effects. The fluctuation of the frequency is not exactly symmetric
with respect to the wave profile, but exhibits a slight phase shift toward the
wave front, as shown in the detailed plot of the Hilbert spectrum. The highest-
frequency part of the wave was always seen to be aligned with the wave front,
indicating that this part of the wave has a higher instantaneous frequency, or
a sharper change in its phase. The lowest-frequency part of the wave always
aligned with the wave back and the trough. This is in general agreement with
the sharpened crest and rounded-off trough profile, but the description is even
more exact: The Hilbert spectrum also pointed out the front-back asymmetry
as modeled by Longuet-Higgins (1982). In that study, Longuet-Higgins also
invoked a shifted phase for the harmonics. The front-back asymmetry can also
be seen from the corresponding wavelet spectrum shown in Figure 12*b*, in which
the harmonics are concentrated slightly in the front of the wave crest. Such a
detail could never have been detected with standard Fourier spectral analysis.
This also indicates that the traditional Stokes wave model does not give a true
description of the water waves.

The marginal spectra from both the Hilbert and the wavelet spectra are shown
in Figure 12*c*, together with the Fourier spectrum. This comparison again shows
the difference between the Hilbert and Fourier views: In the Hilbert view, there
are no harmonics, but there are many sub-harmonics. This, too, serves as an
indicator of nonlinearity. The frequencies of the waves are modeled by fluctua-
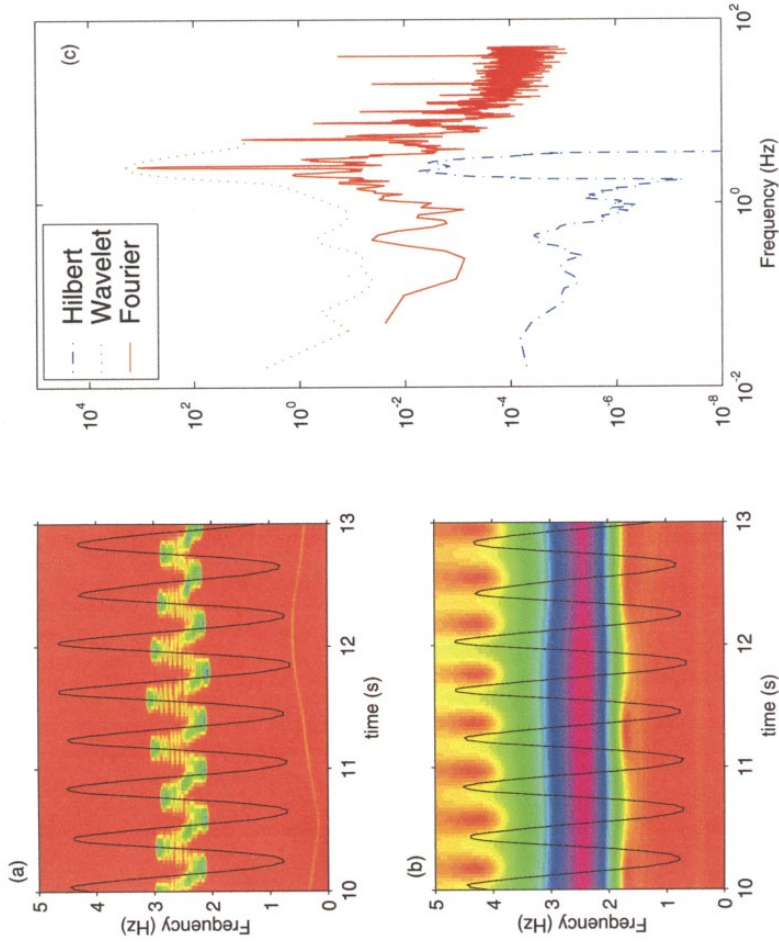tions in frequency from time to time in the Hilbert view, while the same change

Figure 12 The selected Hilbert, Wavelet and their marginal spectra for the laboratory data of the Stokes waves. Both Wavelet and Hilbert spectral analysis reveal the front-and-back-asymmetry. The Hilbert spectra offers better frequency resolution.

is modeled as superimposed constant frequency and amplitude sinusoidal components in the Fourier view. These Fourier components are a mathematically correct decomposition for the data, but they do not make physical sense, because pure sinusoidal waves are not a solution for any equation governing water surface motions. The wavelet spectrum, while correctly depicting the front-back asymmetry, gives very poor frequency resolution. By comparison, the Hilbert view gives an overall superior representation of the phenomena for both time variations and frequency resolution. Now let us move to the wave evolution problem.

Wave Evolution: Fusion

In nonlinear wave dynamics, there is an intriguing problem regarding wave evolution. In numerous controlled experiments (Lake et al 1977; Lake & Yuen 1978; Melville 1982; Su et al 1982; Chereskin & Mollo-Christensen 1985; and Huang et al 1996) the main frequency is seen to shift to a lower frequency, a downshift. This innocent phenomenon presents a serious conflict with wave theory.

According to the kinematics of wave trains, the movement of a constant phase is given by

$$\theta(x, t) = constant \ t \tag{30}$$

in which $\theta(x, t)$ is a slowly varying phase function of position and time. The wave number k and frequency n can be defined as

$$k = \nabla\theta, \quad n = -\frac{\partial\theta}{\partial t}. \tag{31}$$

Both the wave number and frequency are also assumed to be slowly varying functions of time and position. From Equation (31), one can immediately obtain the kinematic conservation equation of the waves as

$$\frac{\partial k}{\partial t} + \nabla n = 0. \tag{32}$$

From Equation (32), the frequency of a stationary wave train should be constant. The laboratory setting is precisely the stationary case, yet the frequency downshift has been observed routinely.

Theoretical study of the downshift problem has developed almost in parallel with the experimental side. The first attempt to model the downshift was based on the nonlinear Schrödinger (NLS) equation (Yuen & Ferguson 1978a,b). The results produced by this approach, however, predict only cyclic recurrence of downshift and upshift. Later, the problem was studied with the modified nonlinear Schrödinger (MNLS) equation, derived by Dysthe (1979) with the added

slow drift term and used by Lo & Mei (1985); the wideband case of the nonlinear Schrödinger equation derived by Trulsen & Dysthe (1996); the Zakharov integral equation used by Caponi et al (1982); and the exact hydrodynamic equations used by Dold & Peregrine (1986). All the studies found recurrence in two-dimensional wave trains. Yet, by adding a simulated wave-breaking term, Trulsen & Dysthe (1990) were able to obtain permanent downshift. By adding wind and eddy viscosity, Hara & Mei (1991) also found downshift. With these results, most investigators believed that the dissipation mechanism must be important. Such a conclusion was further supported by studies by Hara & Mei (1994); Poitevin & Kharif (1991, 1992); Skandrani et al (1996); Uchiyama & Kawahara (1994); and Kato & Oikawa (1995), all with some type of damping.

Recently, Trulsen & Dysthe (1997) extended the investigation to three-dimensional cases, and found that a downshift is possible by allowing oblique sideband perturbations. A quantitative analysis of past results showed that all are within the possible range of three-dimensional perturbations. With one exception (Huang et al 1996), in all those studies, theoretical or experimental, the results were obtained through Fourier analysis. The downshift was defined as the shift of the peak frequency of the spectrum. As shown by Huang et al (1996), the Fourier spectrum is a very poor way to analyze either the downshift phenomenon or the shift of the peak frequency of the spectrum; indeed, it is an inadequate way to quantify the downshift. Yet, Trulsen (1998) still tried to explain the downshift (crest pairing) as a consequence of beating in linear dispersion among different Fourier modes. Any such beating or modulation will have to be reversible, but the downshift in water wave evolution is irreversible. There are many unsolved difficulties in the present Fourier view of the downshift problem. This will be illustrated through an examination of the experimental evidence.

The laboratory data were collected by Huang et al (1996) in the NASA Air-Sea Interaction Research Facility located at Goddard Space Flight Center's Wallops Flight Facility, Wallops Island, Virginia. The wind-wave tank is 91.5 cm wide, 122 cm high, and 1830 cm long, with an operational water depth of 75 cm. For a complete description of the facility and its capabilities, see Long (1992). For this example, waves were generated by a programmable wave maker set at 2.5 Hz. Wave data were collected at eight stations along the tank covering the fetch from 3 to 15 meters. The raw data of the wave elevations and the corresponding Fourier spectra can be found in Huang et al (1996). They further showed that the Fourier spectra offered a very poor indicator for frequency downshift. If we adopt the definition of the peak frequency as the measure of frequency downshift, the only station that reported a downshift is station 8. If one used peak frequency as a measure, then what spectral resolution should one use? Take the data from stations 7 and 8 of the laboratory experiment as an example. Huang et al (1996) identified the downshift based on the spectral peak method to be between these two stations. Let us re-examine this approach
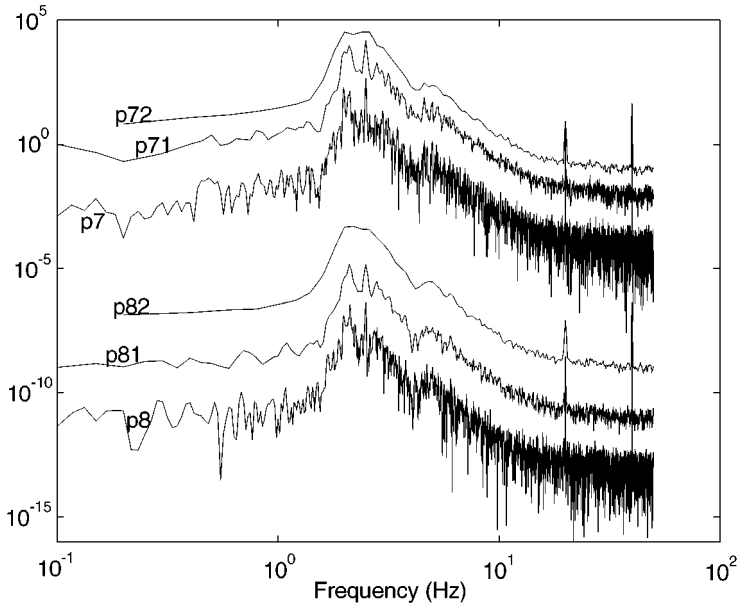
Figure 13   Fourier spectra for two stations with different frequency resolutions, to demonstrate that downshift of the spectral peak depends on frequency resolution.

carefully. The spectra of these two stations with various frequency resolutions are given in Figure 13. Starting with the full data length of 6000 points, the spectra are marked p7 and p8 in Figure 13. The spectra with 3000 points are marked with p71 and p81; the spectra with only 1000 points are marked with p72 and p82. The spectral pair, p7 and p8, does not show the peak downshift. The pair p71 and p81 show a tie. Only the pair with the lowest resolution showed downshift. Yet counting the real numbers of the waves through their total phase angle variations has shown downshift long before the waves reach station 8, as shown in Huang et al (1998a). The downshift started long before the peak frequency changed. This raises a question about the definition of downshift. Is the peak frequency change a good measure of downshift? Or is Fourier spectral analysis a good tool for studying downshift? The answer to both these questions is no.

Another complication arises for the spectral peak measure. Granted the peak criterion, the downshift occurs at station 8. Yet Huang et al (1996) have counted the waves in the laboratory data through the total phase changes. They found that waves started to disappear at station 5. The number of waves missed increases with the distance from the wave maker. Then what will be the state of the waves at stations 5, 6, and 7? There should be no downshift based on spectrum peak,

but clearly the total number of waves has decreased. This decrease, however, occurs only at certain very local regions. Such local change makes the data nonstationary, the condition Fourier analysis is ill-equipped to deal with. As shown above in the Rössler equation, the Fourier spectral peak represents only the global mean frequency. It is not sensitive to the local change of frequency as in the local and discrete downshift phenomena. Therefore, we conclude that the spectral peak is not a good measure for downshift.

Huang et al (1996) opted for the use of Hilbert analysis, with which variation of frequency can be defined much more precisely and locally. Phase variation can be presented in two ways: first, the total phase value changes with respect to the reference station, e.g. station 1. This revealed that the decrease of the total phase values was an integer multiple of 2π. Secondly, the relative phase variation can be presented in a joint distribution with the elevation. The selected results are shown in Figure 14*a–d*. The discrete location of the phase variation is

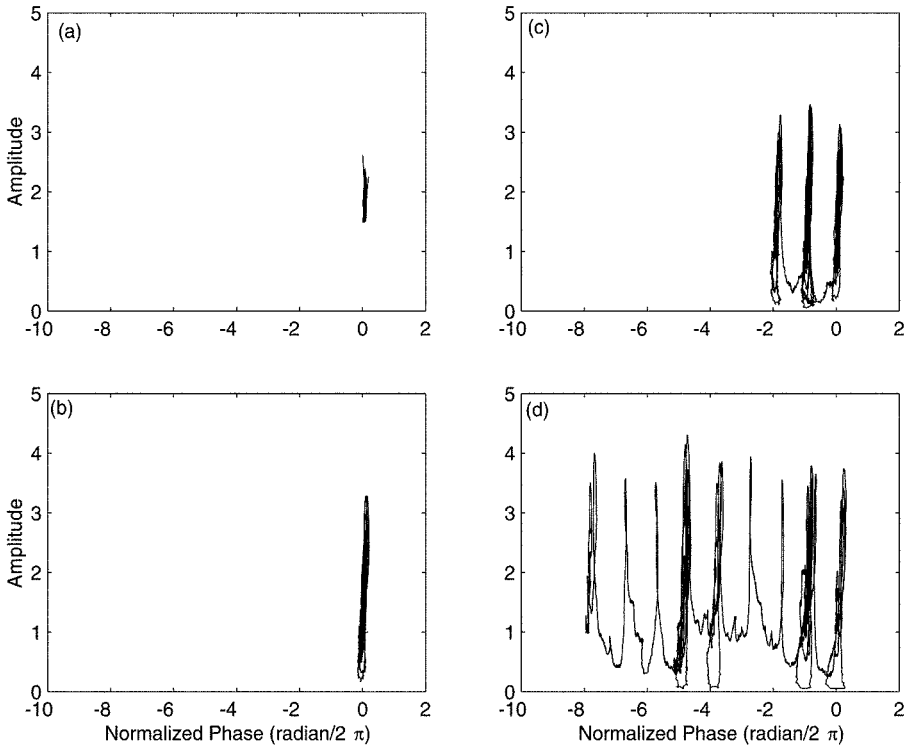*Figure 14*   Joint distribution of phase and amplitude and waves at different stations all relative to the first station. Phase variation is discrete.
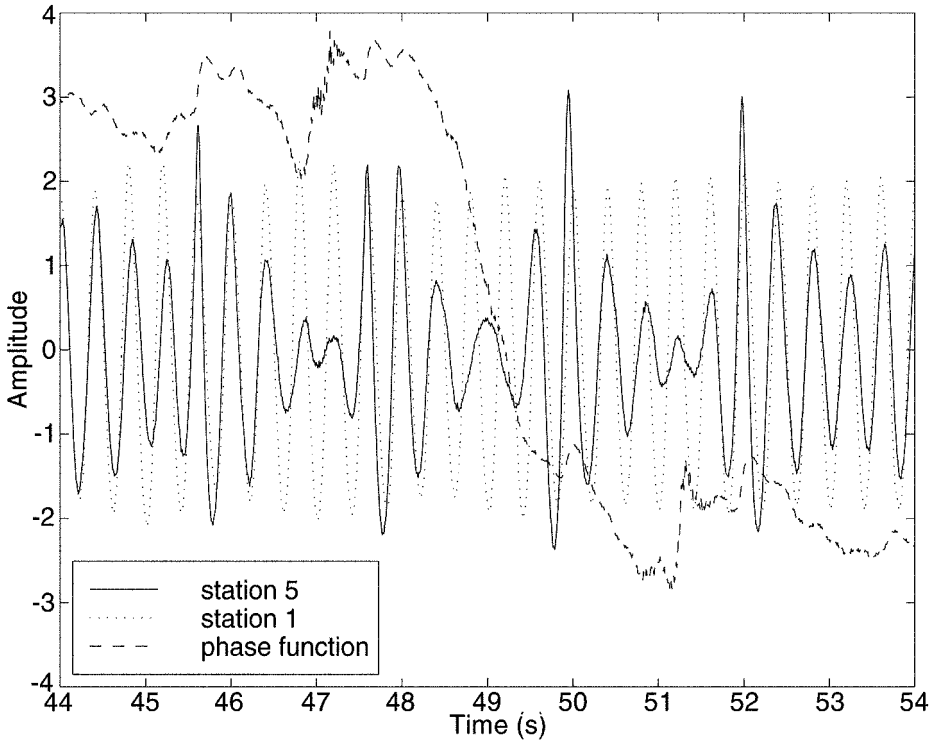
Figure 15   Fusion of two waves into one also coincides with a phase jump. Magnitude of phase jump at 49 seconds is exactly 2π.

a sharp deviation from the traditional picture of slowly varying phase, frequency, and amplitude. It was further shown that the wave evolution process is similar to that of fusion, in which two waves fused into one locally and discretely at the point where the phase jump occurred. A detailed example given in Huang et al (1998a) is reproduced here in Figure 15, in which fusion is vividly illustrated near the 49-second location on the time axis.

To summarize the findings on nonlinear wave evolution from the experimental study by Huang et al (1996), we have to emphasize that the wave frequency downshift in the evolution indeed seems to be not a slowly varying process, but rather a sudden jump. This presents a difficulty to theoretical analysis too, for all of the theoretical models, such as NLS, MNLS, and others, are based on the slowly varying phase, frequency, and amplitude. The process observed is local, and the variation noted is discrete. Waves are lost in the process in which n waves are fused into $n-1$ waves. This is the phenomenon of the "missing

crest" as observed by Lake et al (1977), "crest paring" observed by Ramamon-jiarisoa & Mollo-Christensen (1979), and "fusion" observed by Huang et al (1996). To reconcile this experimental observation with theoretical models is a critical subject for future wave studies.

We must break with the earlier paradigm of wave analysis, and emphasize again that Fourier analysis is not a good method for studying waves. The reasons are many: water waves are nonlinear; therefore, we should not expect to use a linear expansion and be able to represent it. With the Fourier expansion, the harmonics have only a mathematical significance, but no physical meaning. Furthermore, as the wave evolution is local, Fourier expansion simply cannot represent this nonstationary process. As shown in the Rössler equation above, the only way the Fourier method can represent a local frequency change is through harmonics. But such a representation is no longer local. With these concerns, we can examine random wave problems next.

Random Ocean Wind Waves

The logical measure of a random wave field is its various statistical measures (Huang et al 1990a) and spectra (Huang et al 1990b). Traditionally, the spectral representations are all Fourier based. Huang (1995) and Huang et al (1996, 1998a) proposed the Hilbert spectral analysis, which offers a new view of the wave spectra. The data used by them were collected at a coastal tidal station at the rate of 1 Hz. With Hilbert spectral analysis, a section of the result is presented in Figure 15, in which the corresponding wavelet spectrum is presented in a colored contour and the wave profile is also shown. By comparison, the sharpness of the Hilbert spectrum is again evident; it can track minute variations of the energy and frequency. In this comparison, the Hilbert spectrum indeed gives a quantitative indication of energy and frequency variation with time. The frequency variation is especially interesting, for this is the first time that any method has revealed how fast frequency can change with time in a wave train.

When marginal spectra are computed from both the wavelet and Hilbert spectra, the results reveal properties similar to those seen in the Fourier spectrum. Therefore, all three of the spectra are plotted in Figure 16*b* (see color figure). The comparison is clear: The leakage of wavelet analysis causes such smoothing that the result ceases to have any quantitative value. It seems that the wavelet spectrum indeed resolves the nonstationarity to a certain degree, but it has the poorest uniform frequency resolution.

Another interesting point is that the Hilbert spectrum contains almost no energy beyond 0.25 Hz, while the Fourier spectrum has a power law tail to the limit of the Nyquist limit. The form of the spectral tail has been studied extensively (see, for example, Huang et al 1990b). In reality, energy in this high-frequency range is heavily contaminated by harmonics (Huang et al 1981). As
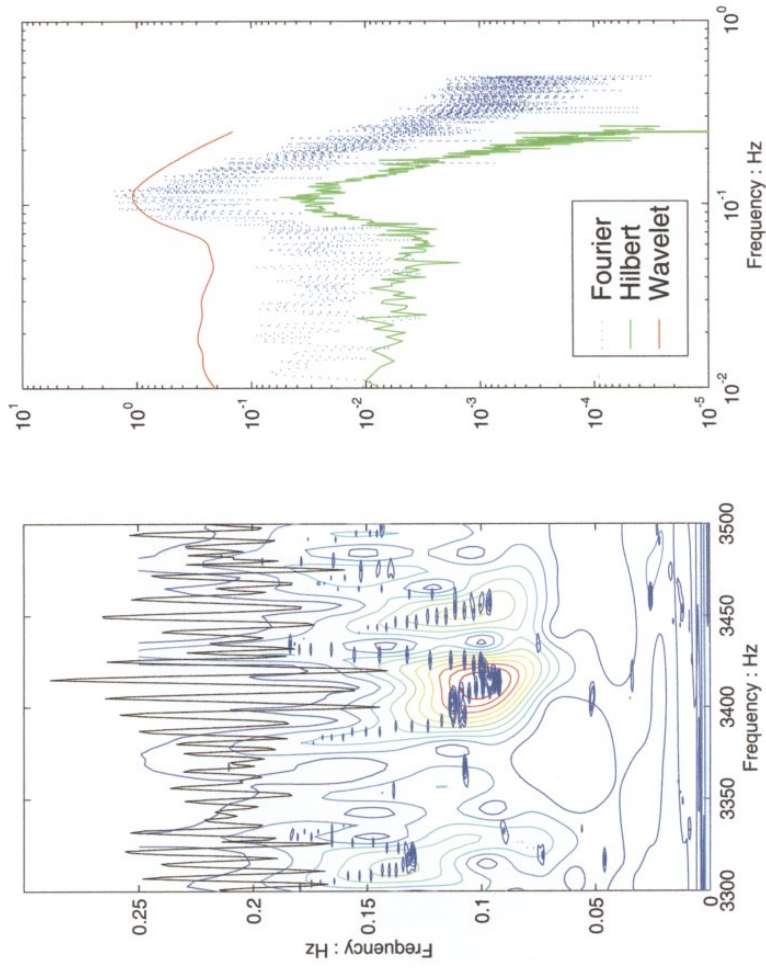
Figure 16 A selected section of the Hilbert and Wavelet spectra together with the wave profile. The Hilbert spectrum can track the frequency variation closely. The marginal spectra show that the Wavelet spectrum has very poor frequency resolution.

a result, the functional form could be a mathematical artifact. Once the true energy content of the free wave is known, the dilemma of determining the expected number of extrema from wave spectrum moments could be resolved easily. As discussed above, the expected number of extrema is proportional to the ratio of the fourth to the second moment of the spectrum. For most of the Fourier spectra, the power law form of the spectral tail makes the fourth moment unlimited. The Hilbert spectrum gives a practical cut-off frequency, an elusive limit that for some time has bothered many investigators working on statistical representations of the wave field.

Other than the cut-off limit, the experimental verification of the form of the equilibrium range originally proposed by Phillips (1958) should now also be re-examined in light of the new Hilbert view. As the Hilbert spectrum is truly based on the local scale without contamination of the harmonics, it should be used in testing the theoretical result based on the local scale dynamics. Whether the spectral form should be slightly modified from that used in the more recent proposals by Phillips (1985); Toba (1973); Phillips (1977); Kitaigorodskii (1983); Banner (1990); and Belcher & Vassilicos (1987) is a problem that needs to be resolved with more studies.

In addition to the spectral form, the wave train properties can also be studied by Hilbert transform as shown by Huang et al (1996), in which they found a strong indication of the discrete characteristics of the wave field. Such an observation was supported indirectly by the studies of Shen & Mei (1993). This is another problem needing further study.

TURBULENCE DATA

Turbulent flow is both nonstationary and nonlinear. Several heuristic models are presently competing to represent the flow. A first hypothesis assumes the turbulence fields consist of superimposed waves; then turbulence is the result of exchanging energy among the waves of comparable wave numbers. A second hypothesis assumes the turbulence fields consist of localized vortices that are locally coherent. These vortices are highly nonlinear, yet they could also maintain their identities without significant interactions among them. A third hypothesis assumes the turbulence fields consist of superimposed wave packets like eddies. This is the wave model with intermittence. And finally, a fourth hypothesis assumes the turbulence is characterized as pure noise. These views are consistent with the more rigorous mechanical approach presented in Sagdeev et al (1988), where the flow fields are divided into weak and strong turbulence. In weak turbulence, flow can be represented by weak interactions among waves of comparable wave numbers, while in strong turbulence, the nonlinearity of the waves is strong even if the interactions among them are weak. To identify

the basic building block is an elusive goal (Pullin & Saffman 1998). In fact, most phenomenological studies of turbulence are confined to probabilistic and spectral properties as reviewed by Farge (1992) and Nelkin (1994).

Traditionally, most turbulence data have been analyzed by Fourier-based methods. As pointed out by Farge (1992), a great risk of uncritical use of the analysis is to misinterpret the functions used in the analysis as characteristic of the phenomena. Turbulence is nonlinear and nonstationary; Farge (1992) hence rightfully argued that a localized expansion should be preferred over the space-filling trigonometric functions used in Fourier analysis. Consequently, she proposed wavelet analysis as the solution (see, for example, Meneveau 1991). Farge's objection is the same as that of Huang et al (1998a) on the application of Fourier spectral analysis to nonlinear and nonstationary data. Based on Huang et al (1998a) and the argument presented above, Hilbert spectral analysis should be a better tool. Hilbert spectral analysis has already been applied to a section of wind data measured with a Pitot tube over the water surface during the initial stage of wind-wave generation. But the data rate is too low to be useful in investigation of turbulence. We present some recent results using the Hilbert spectrum approach on the universal equilibrium subrange of turbulence.

According to Kolmogorov (1941), at infinite Reynolds number, all possible symmetries should be restored locally, and all turbulent flows are self-similar. At this stage, the small-scale statistics are uniquely and universally determined by the mean energy dissipation rate, ε, and a scale, I. Then, through a dimensional argument, he postulated that the energy spectrum, $E(k)$, at this range should be

$$E(k) \propto \varepsilon^{2/3} \, k^{-5/3}, \tag{33}$$

in which k is 1/I. This is the famous $-5/3$ law. Numerous observations have confirmed this formula (Frisch 1995). Yet problems still exist. On the theoretical side, this theory did not include small-scale eddy intermittence effects, which were observed first by Batchelor & Townsend (1949). During the last 50 years different models were proposed successively, but not one of them has been firmly proven to be even a good approximation up to now. Among these models, the first was proposed independently by Kolmogorov (1962) and Obukhov (1962). They conjectured that the energy transfer to small scales was a self-similar cascade with an associated multiplicative process that was approached by a lognormal distribution of the dissipation rate. Novikov & Stewart (1964) proposed another multiplicative process, which is usually called the black and white model. Mandelbrot (1974) conjectured that in the regions where this process takes place, the energy dissipations are a self-similar fractal subset, and that outside this fractal subset there is no dissipation. This model is essentially
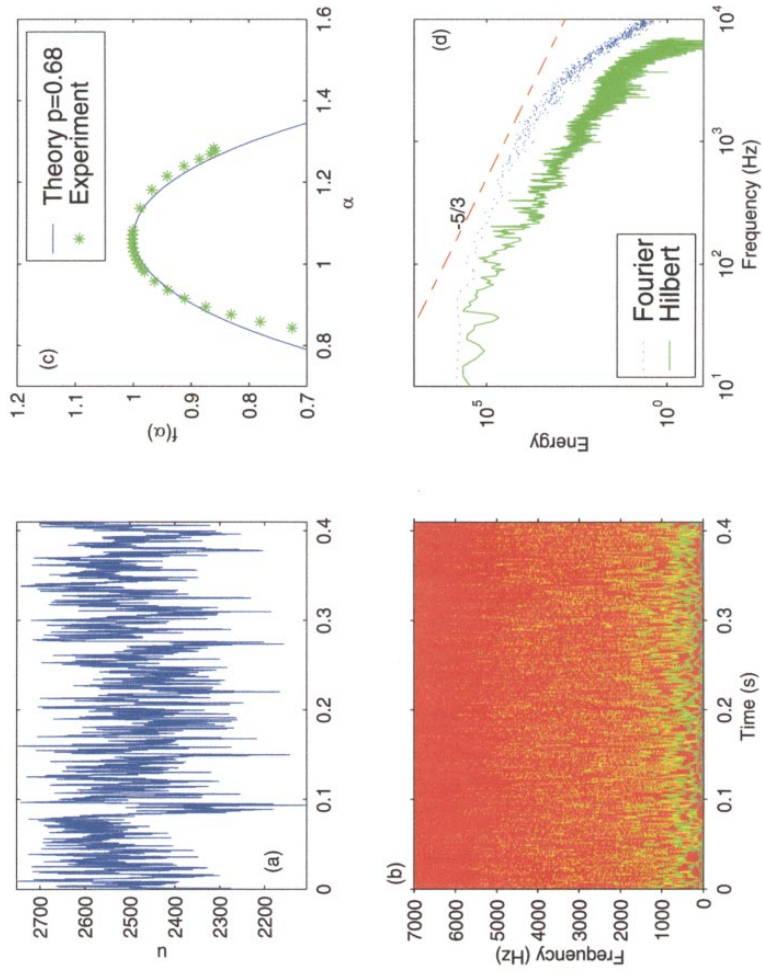
Figure 17   The wind tunnel turbulence data, with the Hilbert, singularity, and marginal spectra, respectively.

the same as that of Novikov & Stewart (1964). Parisi & Frisch (1985) modi-
fied Mandelbrot's fractal subset to a multifractal subset where the dissipation
is strong in one section and weak in another section for any given scale, rather
than absolutely null. The corresponding model resulting from this is named the
multifractal model (Meneveau & Sreenivasan 1991).

On the experimental side, all confirmations to date are based on the Fourier
spectrum. As the scale is strictly local and the flow nonlinear, and because
the Fourier spectrum is more global and linear, it would certainly be more
desirable to have a more local measure of the scales to test the theory. This
desire apparently prompted Farge (1992) to propose wavelet techniques as an
alternative. Unfortunately, wavelet analysis lacks scale resolution, and is also
linear. Consequently, wavelet analysis has not produced any definite answer
despite numerous attempts. With the introduction of EMD and Hilbert spectral
analysis, we now have a method to accurately visualize the characteristic scale
at any given location. The method can also help us resolve the hierarchical
structure of smaller and larger eddies. The data shown in Figure 17a (see color
figure) were taken in the turbulent boundary layer of a hot plate in a wind
tunnel. The mean axial speed is 13 m/s with a Reynolds number of 3.12×10^6.
Sampled at a rate of 20 kHz, the total data length is 8192 points. With the EMD
decomposition, this data set produced 11 IMF components. All the components
are of comparable magnitude, and the velocities seen are highly intermittent.
Two typical short sections of the data are plotted in detail in Figure 18a–b.
Two types of events stand out in Figure 18a: First, there are large regions
of intermittency (marked by A). Second, the intermittent region violated the
assumption of the curdling process as required in the multifractal model (marked
by B). On detailed examination, we found that events A cover over only a small
portion of the data length, but event B occurs as a rule rather than an exception.
In Figure 18(b), the cascade model seems to work well. Events A and B imply
that turbulence is most likely not a multifractal process.

With these IMFs, the Hilbert spectrum is constructed with the same frequency
resolution as the Fourier spectrum as shown in Figure 17b. The energy distribu-
tions are uniform throughout the time-frequency space, but also are intermittent.
The marginal Hilbert together with the Fourier spectra are shown in Figure 17d.
The Fourier spectrum follows a curved trend. If a $-5/3$ straight line is drawn,
it seems to fit only a small section of about one decade of around 1000 Hz
frequency range. The Hilbert spectrum shows a very broad, constant sloping
base covering about three decades. It then is seen to bend into a larger negative
slope around 5000 Hz, the frequency at which the viscous effect is becoming
dominant. Since the Hilbert spectral analysis does not admit harmonics, the
appearance of the marginal Hilbert spectrum represents motions of the physical
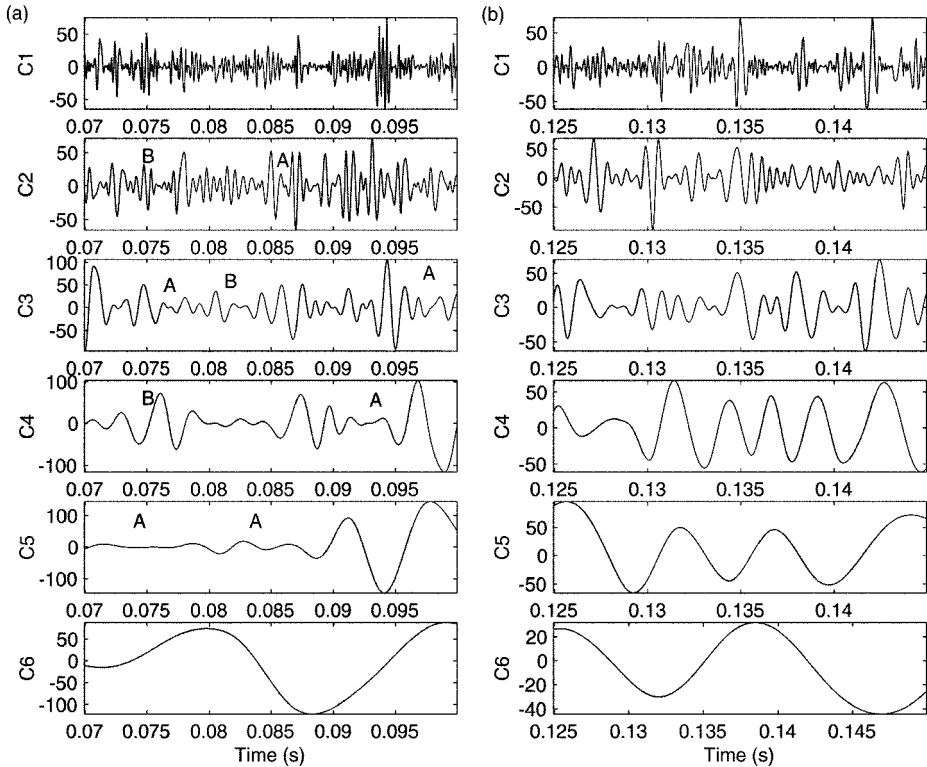scale locally. The dominant range is similar to the inertia range proposed by

Figure 18   Selected sections of the intrinsic mode function expansion of the turbulence data. Events *A* and *B* show the conflict with the fractal hypothesis.

Kolmogorov (1962). The slope of the spectrum, however, is slightly higher than the $-5/3$ power. The meaning of this difference needs to be examined.

Finally, there are an increasing number of investigators (Parisi & Frisch 1985; Saddeev et al 1988; Meneveau & Sreenivasan 1991) who have proposed to describe turbulence as a fractal process. This has prompted Aurell et al (1992) to point out that a spurious multifractal result is a possibility. In fact, when the data used here were processed in the same way as by Meneveau & Sreenivasan (1991), the resulting singularity spectrum shown in Figure 17*c* also suggested a multifractal conclusion. But detailed examination of the decomposed data in Figure 18*a* offered some contradiction to the curdling process. Thus, the multifractal state could indeed be spurious.

All these studies, however, are phenomenological. The results here are used primarily to highlight the new method for data analysis. The real dynamics need

to be studied through an approach totally different from these cited here (see
Pullin & Saffman 1998).

DISCUSSION

Through various numerical and real data we have shown that the new Hilbert
view indeed provides a clearer picture of the underlying physical processes.
Influenced by the omnipotent perturbation methods of the past for weakly non-
linear phenomena, data analysis has been dominated by Fourier-based analy-
sis. A few more remarks on Fourier analysis are necessary here. Although
the Fourier transform is valid under extremely general conditions (see, for ex-
ample, Titchmarsh 1948), to use it as a method for physical interpretation of
frequency-energy distribution was not the original intention. The Fourier ex-
pansion was originally proposed to approximate any function to any degree
of accuracy mathematically. In such an expansion, each component certainly
serves its mathematical function in the approximation, but no more. In spectral
analysis, the Fourier spectrum has indeed provided a general method for exam-
ining the global energy-frequency distributions; however, in this application,
additional physical meanings are assigned to the components, an extension
whose physical meaning has never been clearly established. For Fourier spec-
tral analysis to be meaningful, there are some crucial restrictions: the data
must be linear, and strictly periodic or stationary. Furthermore, to have good
resolution, the data have to be long. Few of the data sets, from either natural
phenomena or artificial sources, can satisfy all these strict conditions of station-
arity. Additionally, most of the natural systems are nonlinear. Almost all the
data we face will have one or more of the following problems: the total data
span is too short; the data are nonstationary; and the data represent nonlinear
processes. Facing such data, Fourier spectral analysis is of limited use. For lack
of alternatives, however, Fourier spectral analysis is still applied. As a result,
the term spectrum has become almost synonymous with the Fourier transform
of the data. The uncritical use of Fourier spectral analysis and adoption of the
stationary and linear assumptions may give misleading results. The problems
can be illustrated by the following arguments:

First, Fourier spectral analysis transfers the data from the time to the fre-
quency domain with constant amplitude and frequency trigonometric terms. In
the frequency domain, the relationship with time is totally lost. Thus Fourier
spectral analysis suffers an inherited defect for representing nonstationary data.
As the Fourier spectrum utilizes uniform harmonic components globally, it
therefore needs many additional harmonic components to simulate either the
nonstationary or the nonlinear variations of the data. To illustrate the above
point, let us consider a delta function that has a phase-locked white Fourier

spectrum. Here, many Fourier components are added to simulate the nonstationary nature of the data in the time domain, but their existence diverts energy to a much wider frequency domain. Constrained by the energy conservation principle, each component will have a relatively low energy content. The total energy is uniformly distributed over the whole time domain, which is not physical for a nonstationary process. Thus, the Fourier spectral components might make mathematical sense, but they make no physical sense.

Second, Fourier spectral analysis uses linear superposition of trigonometric functions; therefore, it needs additional harmonic components to simulate the deformations in wave profiles. Most of the deformations, as will be shown later, are the direct consequence of intra-wave frequency modulations through nonlinear effects. Thus the harmonics give a misleading energy-frequency representation for nonlinear data.

There are other variations of the Fourier-based methods, such as the spectrogram (see, for example, Oppenheim & Schafer 1989); wavelet analysis for time series (see, for example, Chan 1995, Farge 1992, and Long et al 1993a), and two-dimensional images (Spedding et al 1993); the Wigner-Ville distribution (see, for example, Claasen & Mecklenbräuker 1980 and Cohen 1995); the evolutionary spectrum (see, for example, Priestley 1965); the empirical orthogonal function expansion, also known as the principal component analysis, or singular value decomposition method; and some miscellaneous methods such as least square estimation of the trend, smoothing by moving averaging, and differencing to generate stationary data. All the above methods are designed to modify the global representation of the Fourier analysis, but they all failed in one way or the other, as discussed by Huang et al (1998a) and demonstrated here.

Finally, let us turn to the problem of nonlinearity. It has always been controversial to use the term nonlinear in association with "data". The most convincing objection to the term nonlinear data is that all data can be decomposed into Fourier series. Since the Fourier series is a linear decomposition, and each component is also the solution of a linear differential equation, then it follows that the data are the superposition of linear solutions; therefore, they should be regarded as linear. This is the typical Fourier view. With this logic, of course, all data are linear. There are various tests proposed by Priestley (1988), Bendat (1990), and Tong (1990). Unfortunately, these tests give only necessary conditions. The view advanced here, to link the intra-wave frequency modulation as an indicator for nonlinearity, also has difficulties. If one examines the classic nonlinear system as given in Equations (22) and (27), one finds that the solution forms a nonlinear equation that has a particular characteristic, intra-wave frequency modulation. Therefore, a nonlinear signal should have phase-locked harmonics, while a linear signal should have only uniformly distributed phase. Unfortunately, this condition is also only a necessary but not a sufficient one.

Many intra-wave frequency modulation cases could also be the solutions of variable coefficient linear differential equations. Equation (19) is an example, as discussed by Huang et al (1998a), and Mathieu's equation is another. Those variable coefficient linear differential equations can exhibit all kinds of non-linear behavior, including the generation of chaos. A foolproof definition will have to wait for a better understanding of nonlinear systems. For the time being, the term is used as a means of convenience as well as an attempt to describe data with different characteristics. Its use is very similar to that in Bendat (1990) and Tong (1990).

CONCLUSIONS

The empirical mode decomposition (EMD) method and the associated Hilbert spectral analysis indeed offer a powerful method for nonlinear, nonstationary data analysis. Central to the present approach is the sifting process to produce IMFs, which enables complicated data to be reduced into amplitude- and frequency-modulated form so that instantaneous frequencies can be defined. These IMFs form the basis of the decomposition and are complete and practically orthogonal. The expansion in terms of the IMF basis has the appearance of a generalized Fourier analysis with variable amplitudes and frequencies. It is the first local and adaptive method in frequency-time analysis.

A great advantage of EMD and Hilbert spectral analysis is effective use of the data. In EMD, we have used all the data in defining the longest-period component. Furthermore, we do not need a whole wave to define the local frequency, for the Hilbert transform gives the best-fit local sine or cosine form to the local data; therefore, the frequency resolution for any point is uniformly defined by the stationary-phase method or local derivative of the phase. This advantage is especially valuable in extracting low-frequency oscillations. Unlike wavelet analysis, instantaneous frequency can still be localized in time even for the longest period component without spreading energy over wide frequency and time ranges. Still another advantage of EMD and Hilbert spectral analysis is its application to transient data without zero or mean references; the trend or the DC term is automatically eliminated.

Other than the practical aspect, the most important conceptual innovation of the present study is the physical significance assigned to the instantaneous frequency for each mode of a complicated data set. By adopting the instantaneous frequency, we can clearly define both the inter- and intra-wave frequency modulations in a wave train. Such frequency modulations are totally lost in Fourier spectral analysis, and only the inter-wave frequency modulation can be vaguely depicted in wavelet analysis. Yet, both the inter-wave and the intra-wave frequency modulations are critical in interpretation of oscillatory phenomena.

The former explains the wave form deformation by nonlinear effects, which traditionally has been taken as the harmonic distortion; the latter explains the dispersive propagation of waves. Intra-wave frequency modulation offers new insight into nonlinear oscillation systems in more detail than the modern topological treatment. By adopting the instantaneous frequency, we have eliminated the need of not only higher harmonics to simulate nonlinearly deformed waves, but also spurious harmonics to simulate nonstationary data. We believe this new method can give us new physical insight in all other nonlinear and nonstationary phenomena. Instantaneous frequency can be defined only through the IMF, which is defined here based on local properties of the data rather than the global restrictions proposed before.

Hilbert spectral analysis is also a tool. Its use in exploring the full physical meanings of complicated data is only now beginning, and associated properties of the marginal spectra need to be explored.

ACKNOWLEDGMENTS

Visit the *Annual Reviews home page* at
http://www.AnnualReviews.org

Literature Cited

Aurell E, Frisch U, Lutsko J, Nergassola M. 1992. On the multifractal properties of energy dissipation derived from turbulence data. *J. Fluid Mech.* 238:467–86

Bacry E, Arnéodo A, Frisch U, Gagne Y, Hopfinger E. 1991. Wavelet analysis of fully developed turbulence data and measurement of scaling exponents. In *Proc. Turbulence 89: Organized Structures and Turbulence in Fluid Mechanics, Grenoble, Sept. 1989,* ed. M Lesieur, O Métais, pp. 203–15. Dordrecht: Kluwer

Banner ML. 1990. Equilibrium spectra of wind waves. *J. Phys. Oceanogr.* 20:966–84

Batchelor GK, Townsend AA. 1949. The nature of turbulent motion at large wave number. *Proc. R. Soc. London Ser. A* 199:238–55

Belcher SE, Vassilicos JC. 1987. Breaking waves and the equilibrium range of wind wave spectra. *J. Fluid Mech.* 342:377–401

Bendat JS. 1990. *Nonlinear System Analysis and Indentification From Random Data.* New York: Wiley-Interscience

Bendat JS, Piersol AG. 1986. *Random Data:*

Analysis and Measurement Procedures, New York: Wiley & Sons. 2nd ed.

Braun S, Feldman M. 1997. Time-frequency characteristics of nonlinear systems. *Mech. Syst. Signal Proc.* 11:611–20

Caponi EA, Saffman PG, Yuen HC. 1982. Instability and confined chaos in a nonlinear dispersive wave system. *Phys. Fluids* 25:2159–66

Chan YT. 1995. *Wavelet Basics.* Boston: Kluwer Academic

Chereskin TK, Mollo-Christensen E. 1985. Modulational development of nonlinear gravity-wave groups. *J. Fluid Mech.* 151: 365–77

Claasen TACM, Mecklenbräuker WFG. 1980. The Wigner distribution—a tool for time-frequency signal analysis. Part I: Continuous time signals 35:217–50; Part II: Discrete time signals 35:276–300; Part III: Relations with other time-frequency signal transformations. *Philips J. Res.* 35:372–89

Cohen L. 1995. *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall

Dold JW, Peregrine DH. 1986. Water-wave modulation. In *Proc. Intl. Cong. Coastal Engineering*, Taipei, ed. BL Edge, *Vol. 1*, 163–175, ASCE.

Drazin PG. 1992. *Nonlinear Systems.* Cambridge, UK: Cambridge Univ. Press

Dysthe KB. 1979. Note on a modification to the nonlinear Schrödinger equation for application to deep water waves. *Proc. R. Soc. London Ser. A* 369:105–14

Farge M. 1992. Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.* 24:395–457

Feldman M. 1991. Method for determination of vibratory system modal parameters using Hilbert transform. Application for Patent no. 098985, Israel, 28 July 1991.

Feldman M. 1994a. Nonlinear system vibration analysis using Hilbert Transform: I. Free-vibration analysis method, FREEVIB. *Mech. Syst. Signal Processing,* 8:2:119–27

Feldman M. 1994b. See Feldman 1994a. II. Forced vibration analysis method, FORCE-VIB. 8:2:309–18

Feldman M. 1997. Nonlinear free-vibration identification via the Hilbert Transform. *J. Sound Vib.* 208(3):475–89

Feldman M, Braun S. 1995. Identification of nonlinear system parameters via the instantaneous frequency: application of the Hilbert Transform and Wigner-Ville techniques, *Proc. Int. Modal Analysis. Conf. 13th* 637–42

Frisch U. 1995. *Turbulence.* Cambridge, UK: Cambridge Univ. Press

Hara T, Mei CC. 1991. Frequency downshift in narrowbanded surface waves under the influence of wind. *J. Fluid Mech.* 230:429–77

Hara T, Mei CC. 1994. Wind effect of the nonlinear evolution of slowly varying gravity capillary waves. *J. Fluid Mech.* 267:221–50

Huang NE, Long SR, Tung CC, Yuan Y, Bliven LF. 1981. A unified two-parameter wave spectral model for a general sea state. *J. Fluid Mech.* 112:203–24

Huang NE, Long SR, Tung CC, Yuan Y, Bliven LF. 1983. A non-Gaussian statistical model for surface elevation of nonlinear random wave fields. *J. Geophys. Res.* 88:7597–606

Huang NE, Long SR, Bliven LF, Tung CC. 1984. The non-Gaussian joint probability density function of slope and elevation for a nonlinear gravity wave field. *J. Geophys. Res.* 89:1961–72

Huang NE, Tung CC, Long SR. 1990a. Wind-Wave Spectrum. In *The Sea, Vol. 9*, pp. 197–238

Huang NE, Tung CC, Long SR. 1990b. The probability structure of the ocean surface. In *The Sea, Vol. 9*, pp. 335–66

Huang NE, Long SR, Tung CC, Donelan MA, Yuan Y, et al 1992. The local properties of ocean surface waves by the phase-time method. *Geophys. Res. Lett.* 19:685–88

Huang NE, Long SR, Tung CC. 1993. The local properties of transient stochastic data by the phase-time method. In *Computational Methods for Stochastic Processes*, pp. 253–79, ed. AH Cheng, CY Yang. Ashurst, UK: Computational Mechanics Publications

Huang NE. 1995. Nonlinear Evolution of Water Waves: Hilbert's View. In *Proc. Int. Symp. Experimental Chaos*, 2nd, ed. W Ditto et al 327–41. World Scientific

Huang NE. 1996. Computer implicated empirical mode decomposition method, apparatus, and article of manufacture. *U.S. Patent Pending*

Huang NE, Long SR, Shen Z. 1996. The mechanism for frequency downshift in nonlinear wave evolution. *Adv. Appl. Mech.* 32:59–111

Huang NE, Shen Z, Long SR, Wu ML, Shih HH, et al 1998a. The Empirical Mode Decomposition and Hilbert Spectrum for Nonlinear And Nonstationary Time Series Analysis. *Proc. R. Soc. London Ser. A* 454:903–95

Huang W, Shen Z, Huang NE, Fung YC. 1998b. Engineering analysis of biological variables: an example of blood pressure over 1 day. *Proc. Natl. Acad. Sci. USA.* 95:4816–21

Kato Y, Oikawa M. 1995. Wave number downshift in modulated wavetrain through a nonlinear damping effect. *J. Phys. Soc. Japan.* 64:4660–69

Kitaigorodskii SA. 1983. On the theory of the equilibrium range spectra of wind generated

gravity waves. *J. Phys. Oceanogr.* 13:816–27

Kolmogorov AN. 1941. The local structure of turbulence in incompressible viscous fluid for very large Reynolds number. *Dokl. Akad. Nauk SSSR* 30:9–13 (reprinted in *Proc. R. Soc. Lond. A* 434:9–13, 1991).

Kolmogorov AN. 1962. The refinement of previous hypothesis concerning the local structure of turbulence in incompressible viscous fluid at high Reynolds number. *J. Fluid Mech.* 13:82–85

Lake BM, Yuen HC. 1978. A new model for nonlinear gravity waves. Part I. Physical model and experimental evidence. *J. Fluid Mech.* 88:33–62

Lake BM, Yuen HC, Rungaldier H, Ferguson WE. 1977. Nonlinear deep-water waves: theory and experiments, Part 2: Evolution of a continuous wave train. *J. Fluid Mech.* 83:49–74

Lo E, Mei CC. 1985. A numerical study of water-wave modulation based on a higher-order nonlinear Schrödinger equation. *J. Fluid Mech.* 150:395–416

Long SR. 1992. NASA Wallops Flight Facility Air-Sea Interaction Research Facility, *NASA Ref. Pub. No.* 1277, 36 pp.

Long SR, Lai RJ, Huang NE, Spedding GR. 1993a. Blocking and trapping of waves in an inhomogeneous flow. *Dyn. Atmos. Oceans.* 20:79–106

Long SR, Huang NE, Tung CC, Wu ML, Lin RQ, et al 1993b. The Hilbert techniques: an alternate approach for non-steady time series analysis. *IEEE Geosci. Remote Sensing Soc. Lett.* 3:6–11

Longuet-Higgins MS. 1963. The effect of non-linearities on the statistical distributions in the theory of sea waves. *J. Fluid Mech.* 17:459–80

Longuet-Higgins MS. 1982. On the skewness of sea-surface slopes. *J. Phys. Oceanogr.* 12:1283–91

Mandelbrot B. 1974. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.* 62:331–58

Melville WK. 1982. The instability and breaking of deep-water waves. *J. Fluid Mech.* 115:165–85

Meneveau C. 1991. Analysis of turbulence in the orthonormal wavelet representation. *J. Fluid Mech.* 232:469–520

Meneveau C, Sreenivasan KR. 1987. Simple multifractal cascade model for fully enveloped turbulence. *Phys. Rev. Lett.* 59:1424–27

Meneveau C, Sreenivasan KR. 1991. The multi-fractal nature of turbulent energy dissipation. *J. Fluid Mech.* 224:429–84

Nelkin M. 1994. Universality and scaling in fully-developed turbulence. *Adv. Phys.* 43:143–81

Novikov EA, Stewart RW. 1964. The inter-mittency of turbulence and the spectrum of energy dissipation. *Izv. Akad. Nauk. SSSR, Ser. Geoffiz.* 408–13

Obukhov AM. 1962. Some specific features of atmospheric turbulence. *J. Fluid Mech.* 13:77–81

Oppenheim AV, Schafer RW. 1989. *Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall

Parisi G, Frisch U. 1985. Fully developed turbulence and intermittency. In *Proc. Int. School on Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, ed. M Ghil, R Benzi, G Parisi, pp. 71–88. Amsterdam: North-Holland

Poitevin J, Kharif C. 1991. Subharmonic transition of a nonlinear wave train on deep water. In *Mathematical and Numerical Aspects of Wave Propagation Phenomena*, ed. GC Cohen, L Halpern, P Joly. pp. 567–76. SIAM.

Poitevin J, Kharif C. 1992. Subharmonic transition of a nonlinear short wave train on deep water. In *Proc. Nonlinear Water Waves Workshop*, ed. DH Peregrine, pp. 54–63. Bristol UK: Univ. Bristol

Phillips OM. 1958. The equilibrium range in the spectrum of wind-generated waves, *J. Fluid Mech.* 4:426–34

Phillips OM. 1985. Spectral and statistical properties of the equilibrium range in wind generated gravity waves. *J. Fluid Mech.* 156:505–31

Phillips OM. 1977. *The Dynamics of the Upper Ocean.* Cambridge, UK: Cambridge Univ. Press

Priestley MB. 1965. Evolutionary spectra and non-stationary processes. *J. Roy. Statist. Soc. B* 27:204–37

Priestly MB. 1988. *Non-Linear and Non-Stationary Time Series Analysis.* London: Academic

Prime MB, Shevitz DW. 1996. Linear and nonlinear methods for detecting cracks in beams. In *Proc. Modal Analysis Conf.*, 14th: pp. 1437–43

Pullin DI, Saffman PG. 1998. Vortex dynamics in turbulence. *Annu. Rev. Fluid Mech.* 30:31–51

Ramamonjiarisoa A, Mollo-Christensen E. 1979. Modulation characteristics of sea surface waves. *J. Geophys. Res.* 84:7769–75

Rice SO. 1944. Mathematical analysis of random noise. *Bell Syst. Tech. J.* 23:282–310. Part II. Power spectrum and correlation functions. *Bell Sys. Tech. J.* 23:310–32

Rice SO. 1945. Mathematical analysis of random noise. Part III. Statistical properties of random noise currents. *Bell Sys. Tech. J.*

24:46–108. Part IV. Noise through nonlinear devices. *Bell Sys. Tech. J.* 24:109–56

Sagdeev RZ, Usikov Da, Zaslavsky GM. 1988. *Nonlinear Physics: From the Pendulum to Turbulence and Chaos.* Chur, Switz: Harwood

Shen Z, Mei L. 1993. Equilibrium spectra of water waves forced by intermittent wind turbulence. *J. Phys. Oceanogr.* 23:2019– 26

Shen Z, Wang W, Mei L. 1994. Fine structure of wind waves analysis with wavelet transform. *J. Phys. Oceanogr.* 24:1085–94

Skandrani C, Kharif C, Pointevin J. 1996. Nonlinear evolution of water surface waves: the frequency downshift phenomenon. *Contemp. Maths* 200:157–71

Spedding GR, Browand FK, Huang NE, Long SR. 1993. A 2-D complex wavelet analysis of an unsteady wind-generated surface wave field. *Dyn. Atmos. Oceans.* 20:55–77

Stokes GG. 1847. On the theory of oscillatory waves. *Trans. Camb. Pril. Soc.* 8:441–55

Su MY, Bergin M, Marler P, Myrick R. 1982. Experiments on nonlinear instabilities and evolution of steep gravity-wave trains. *J. Fluid Mech.* 124:45–72

Titchmarsh EC. 1948. *Introduction to the Theory of Fourier Integrals.* Oxford: Oxford Univ. Press

Toba Y. 1973. Local balance in the air-sea boundary processes. III. On the spectrum of wind waves. *J. Oceanogr. Soc. Japan* 29:209–20

Tong H. 1990. *Non-Linear Time Series: A Dynamical System Approach.* Oxford, UK: Clarendon

Trulsen K, Dysthe KB. 1990. Frequency downshift through self modulation and breaking. In *Water, Wave Kinematics*, ed. A Torum, OT Gudmestad, pp. 561–572, Kluwer: Dordrecht

Trulsen K, Dysthe KB. 1996. A modified nonlinear Schrödinger equation for broader bandwidth gravity waves on deep water. *Wave Motion* 24:281–89

Trulsen K, Dysthe KB. 1997. Frequency downshift in three-dimensional wave trains in a deep basin. *J. Fluid Mech.* 352:359–73

Trulsen K. 1998. Crest pairing prediction by modulation theory. *J. Geophys. Res.* 103: 3143–47

Uchiyama Y, Kawahara T. 1994. A possible mechanism for frequency down-shift in nonlinear wave modulation. *Wave Motion* 20:99–110

Whitham GB. 1974. *Linear and Nonlinear Waves.* John Wiley, New York: Wiley & Sons

Yuen HC, Ferguson WE. 1978. Fermi-Pasta-Ulam recurrence in the two-space dimensional nonlinear Schrödinger equation. *Phys. Fluids* 21:2116–18

Yuen HC, Ferguson WE. 1978. Relationship between Benjamin-Fier instability and recurrence in the nonlinear Schrödinger equation. *Phys. Fluids* 21:1275–78

CONTENTS