Signal Analysis from Wave Modulation Perspective

Part I: Characterizing the Analytic Signal Procedure

Yueon-Ron Lee
Institute of Harbor and Marine Technology

July 15, 1997

Contents

List of Figures

List of Tables

ABSTRACT

Repeatability problem of spectra for wind wave signals in wind blowing water tanks is stressed, and the deficiency of the impulse response model in serving as a black-box mechanism for comparing two spectra is also discussed. A few numerical aspects in data processing are used to highlight possible differences associated with different schemes originating from the same analytical formulation. Paradoxes regarding the analytical signal are investigated and their causes are studied and illustrated. The interplays of amplitude and frequency modulations are emphasized in characterizing the behaviors of instantaneous frequency. It is identified that local singular behavior of instantaneous frequency is associated with local irregularity of the amplitude function. We regard the Gabor's analytic signal approach as being a one-and-a-half-dimensional characterizing procedure. And it possesses most of the characters associated with Fourier analysis, such as phase noise and local transient phenomena, as well as many additional features associated with discrete numerical computations. This indicates possible difficulties when the method of analytic signal is applied to broad band processes.

Chapter 1

Introduction

In making comparison of two multi-component signals acquired before and after some external factors being introduced or measured with certain evolution being present, or in studying effects of an influential factor on the constituents of signals, a black-box mechanism of taking direct quotient of the spectral coefficients of the two signals has been most often seen. In term of Fourier analysis, the algorithm involved is basically that of a simple finite impulse response model and the whole concept can be described with the process of Fourier convolution and deconvolution. The method is conceptually easy, but it is hardly physically sound and almost always numerically error prone. Besides, there is an issue of spectral repeatability. Both numerical simulation and physically acquired data are used to illustrate this repeatability problem. The causes of the problem are attributed to phase noise and transient effects imbedded in multi-component signals. The former factor may be due to improper timing during data acquisition or to some inherent uncontrollable limitations such as those associated with random processes (e.g., those related to measurements in wind wave field or in turbulent flow field). The latter originates from sharp local and transient variations of signal. In fact, in addition to these signal-induced natures, there are still other complications that are associated with different numerical processes, such as the discrete nature of a finite resolution scheme and the possible differences in time domain or frequency domain processing. Therefore, a different perspective is always informative and possibly helpful.

1.1 The Heisenberg Uncertainty Principle

One of the well known characters for the two corresponding function spaces of the Fourier transform pair is that a function cannot be both time- and band-limited. That is to say, if a function is limited or, more formally stated, finitely supported in one domain, then the corresponding function in the other domain stretches infinitely on the real line **R**. This concept is formally stated by the Heisenberg uncertainty principle which mathematically specifies a lower limit on the area of the time-frequency window for any Fourier transform pair.

In discrete time-frequency analysis, this character is also indicated by the Balian-Low theorem, which basically states that if there exists a Gabor type frame $g_{m,n} = e^{-2\pi imt}g(t-n)$, where $m, n \in \mathbf{Z}$ and t is an independent variable of a function g, then either $g(t)$ or $G(\omega)$'s second moment of inertia goes unbounded, where $G(\omega)$ is the Fourier transform of $g(t)$ [4, 2, 13]. In all real world situations, signals studied are always limited in time (or in space) and hardware or instruments used to acquire the signals are certainly band-limited. Moreover, there always exist intrinsic differences between discrete numerical schemes and continuous analytical methods. Therefore, practical conditions never fulfill the demands of mathematical requirements, and signal analysis is truly not going to be ideal. The points stated sound a bit abstract and are analytically oriented; in the followings we shall list several issues which are quite practical and we encounter them in all cases when analyzing data using Fourier transform, and, they are more intuitive and easier to comprehend.

Fourier basis functions are periodic and extend bi-infinitely; therefore, in order to take full advantage of Fourier analysis the signals studied are better periodic and sampled infinitely. The various side effects for not fulfilling this requirement include: frequency leakage, smoothing errors, and edge effects due to data truncation and segmentation; aliasing due to under-sampling or non-periodicity; spectral variance due to finite resolution; artifact due to windowing [17, 1]. Besides, efforts to reduce a certain side effect will certainly introduce other hindrances; there are always tradeoffs.

1.2 Transient Effects and Phase Noise

The statement of the periodicity requirement as well as the existence of a lowest limit for the area of the time-frequency uncertainty window basically point out that Fourier transform is suited for characterizing stationary signals and not quite satisfactory for studying transient local phenomena. Still, there is one good example which clearly manifests this property.

The Gibbs phenomenon states that, if there is an abrupt jump or discontinuity in the signal then the overshoots, which occur at both sides of the discontinuity when performing the inverse Fourier transform from its Fourier coefficients, can never disappear and their amplitude remains at constant no matter how many spectral components are taken into account (except that the position of the overshoots approach that of the jump). In plain terms, this phenomenon illustrates the following two points. First, it takes quite a many spectral components to describe a sharp local transient feature; second, a local variation affects a board range of the spectrum just as the Fourier transform of the delta distribution covers the whole frequency axis (the delta distribution is sometimes called delta function, but it is really not a function in strict mathematical definition since it is not defined in the origin).

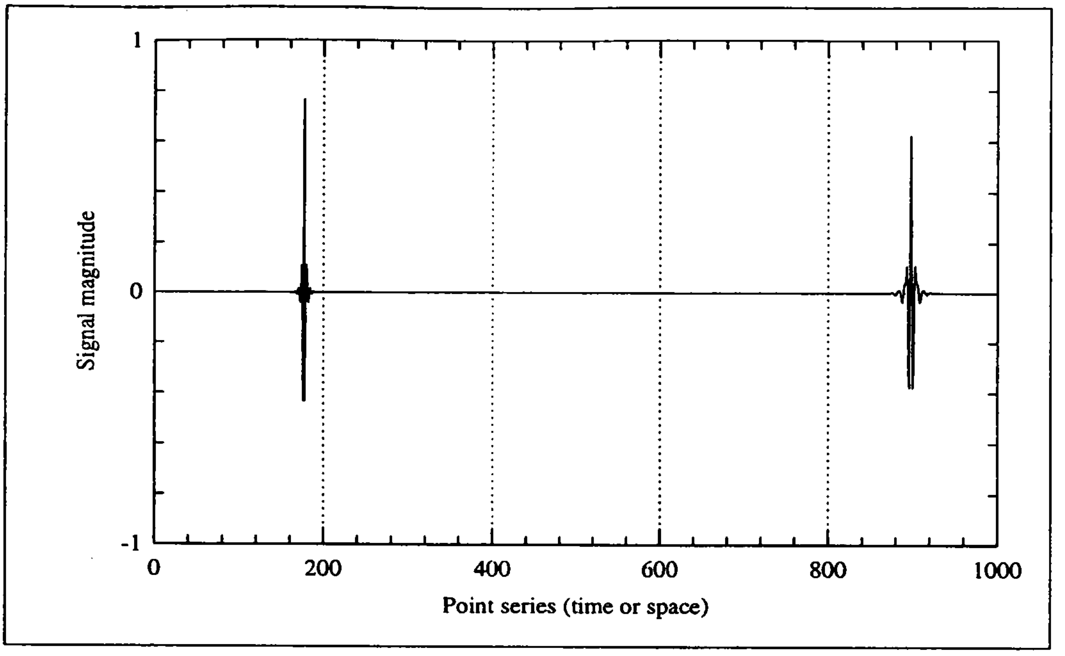
A good illustration of the phenomenon, though somewhat exaggerated, is given by Figure 1.1. In this figure two Lemarié wavelets in the two least (smallest) scales are used to represent the short local transient pulses within the signal. The left pulse corresponds to the result of the inverse wavelet transform from a unit wavelet coefficient at e600 within a 1024-point series. The right pulse corresponds to that of a unit wavelet coefficient at e470. The choice of the least two scales is to emphasize the effect of transient locality — which renders a very board distribution of power spectrum. The choice of positions 600 and 470 is somewhat intentional and somewhat arbitrary, just to give an idea about the symptom related to the positions of occurrences of the pulses, i.e., the phases — which are associated with the wiggling of the spectrum. For "intention" we mean that the greater the separation distance the more severe the wiggling. For "arbitrariness" we mean that in practical situation we generally do not have control over the timings or positions of the occurrences of local features if complexity or randomness exists. In fact these statements are the consequence of the following Fourier transform pairs:

$$f(t) \Longleftrightarrow F(\omega) \quad ; \quad g(t) \Longleftrightarrow G(\omega), \tag{1.1}$$

$$f(t-\tau) \Longleftrightarrow e^{-i\omega\tau} F(\omega), \tag{1.2}$$

$$|f(t)+g(t)|^2 \Longleftrightarrow |F(\omega)+G(\omega)|^2. \tag{1.3}$$

In the above equations the double arrow sign means that, except there may exist difference of multiplication factors, the roles of t and ω may be inter-changed — a very nice property of Fourier analysis, since these as well as other inter-changeable relations facilitate the convenience and powerfulness of the analysis. The wiggling of the spectral curve can be explained by the following two simple explanations. Firstly, a basic and useful property is that "a shift in one domain corresponds to a modulation in the other domain (one of the duality properties of the Fourier transform)". It is also noted that, even the modulus of the modulated spectral coefficients for the single shifted pulse is the same as that of the spectral coefficients of the original pulse, the modulus of the spectral coefficients of the combined pulses is certainly different from the direct summation of those of the two single pulses. Secondary, a complementary character to the first, even though Fourier transform is a linear operation on component signals or on the constituents of a signal, power spectrum does



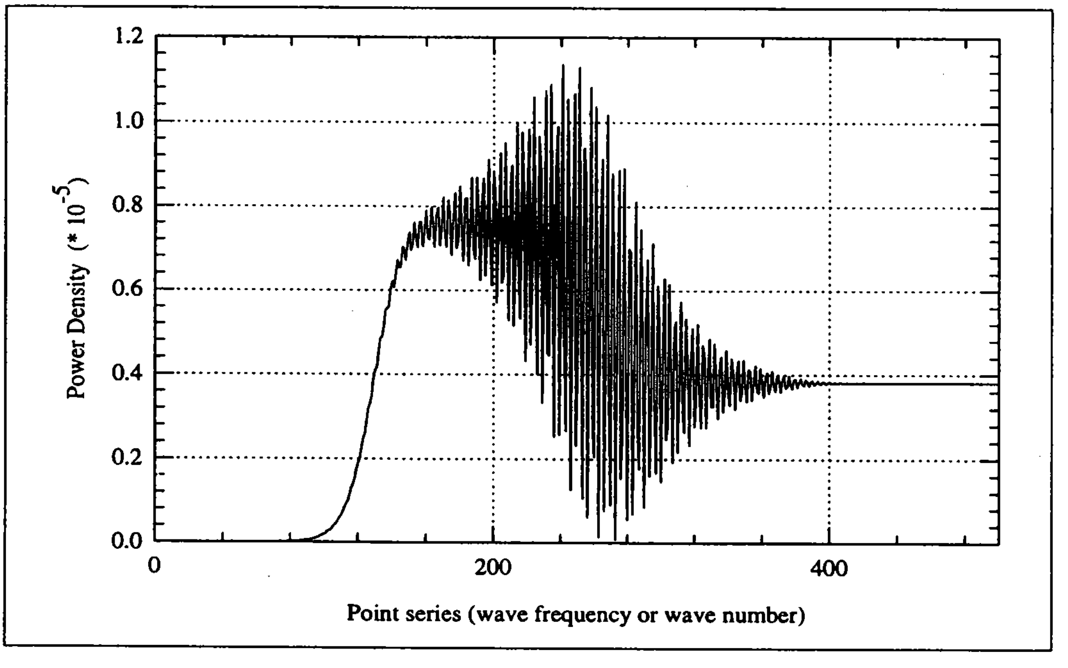


Figure 1.1: The phenomena of ambiguity and phase noise arising from the local transient features of signal are illustrated using two separated wavelets. Here two Lemarié wavelets located at two neighboring scales (the least two scales within a 1024-point series) are shown in the top figure. The power spectrum is shown in the bottom figure. The pulse at left corresponds to the inverse wavelet transform for unit wavelet coefficient e600; the right corresponds to e470.

not possess property of linear operation. That is to say, the power of the combined pulses does not equal to the sum of individual powers for the two pulses. In addition, there is an artifact which is mainly caused by our desire to attach physical meaning to the complex results of the Fourier analysis, i.e., from double-sided spectrum to single-sided spectrum when dealing with real and imaginary parts which distribute at both positive and negative frequencies. The artifact is most obvious when dealing with two-dimensional spectrum, where the symmetry of power spectrum can hardly be practically explained.

1.3 Spectral Repeatability for Wind Wave Signals

Even with the acquaintance of the previous phenomena one might not really grasp to what degree the symptoms may affect the conclusiveness of interpretations of spectra. Measurements of the water wave signals in a wind wave tank tell the story. But first let present results from a different statistical approach based on zero up-crossing method. Comparison helps illuminating different characters associated with different perspectives.

Table 1.3 shows the statistics for three sets of data from such a conventional method. The experimental data is acquired in a wind-blowing oval tank with dimension of 20.0m(L), 31cm(W), and 45cm(H). The tank is filled to a water depth of 24 cm. Detail setup of the tank can be found in Poon, et al (1992) [16]. The wind was kept blowing under the same condition for the cases shown in the table. The signal was sampled at a rate of 40 samples per second for a duration of 240 seconds. Channel 1 shows data of the aqueous flow measured at different depths from the water surface using laser Doppler anemometer (LDA) . Channel 2 shows the water surface displacement measured at nearly the same location as the LDA measurement point (1 cm separation). Statistics at channel 2 can therefore be regarded as the results from repeated measurements.

The various statistic values for all individual runs have indicated that the wave field has reached a stationary condition from the zero up-crossing point of view. However, when viewed from the spectral perspective, the idea of stationary is hardly substantiated from these same data sets. The top sub-figure of figure 1.2 shows the power spectra of the repeated measurements of water surface displacements. The bottom sub-figure shows spectra of the LDA aqueous flows measured at several different depths (3, 4, 5cm, respectively, from the still water level).

As is indicated from the top sub-figure, even though the zero up-crossing statistics has shown conditions to be in good stationary, repeatability of spectra is rather poor. Moreover, when compare the top and the bottom sub-figures it is seen that the spectral shape of the water surface displacement matches that of the aqueous flow for each individual case,

Table 1.1: The zero up-crossing statistics for three different measurements under the same wind condition. Channel 1 is for LDA aqueous flows measured at different depths from the still water surface. Channel 2 is for surface displacements measured at nearly the same location. Statistics for channel 2 can basically be regarded as results from repeated measurements. Comparisons of data at channel 2 for different cases indicate that the wave field is in good stationary. This is certainly not true when viewed from spectral perspective, as shown in subsequent figures.

Case : f0w6030.dat (f1 p3 c1 s9)
Date : 01/05/96
Time : 02:19:05.48
Specifics :

Sampling frequency : 40 Hz
Sampling time length : 240 Sec

Ch	#_W	H.1	H.2	H.3	H1/10	H1/3	H1/2	H.ave	H.rms	T1/10	T1/3	T1/2	T.ave	T.rms
1.	554	26.74	23.67	23.04	20.20	17.38	15.92	12.03	12.93	.44	.44	.44	.43	.44
2.	572	2.90	2.69	2.61	2.33	2.04	1.89	1.44	1.54	.44	.44	.43	.42	.42

Case : f0w6040.dat (f1 p3 c1 s9)
Date : 01/05/96
Time : 02:14:05.76
Specifics :

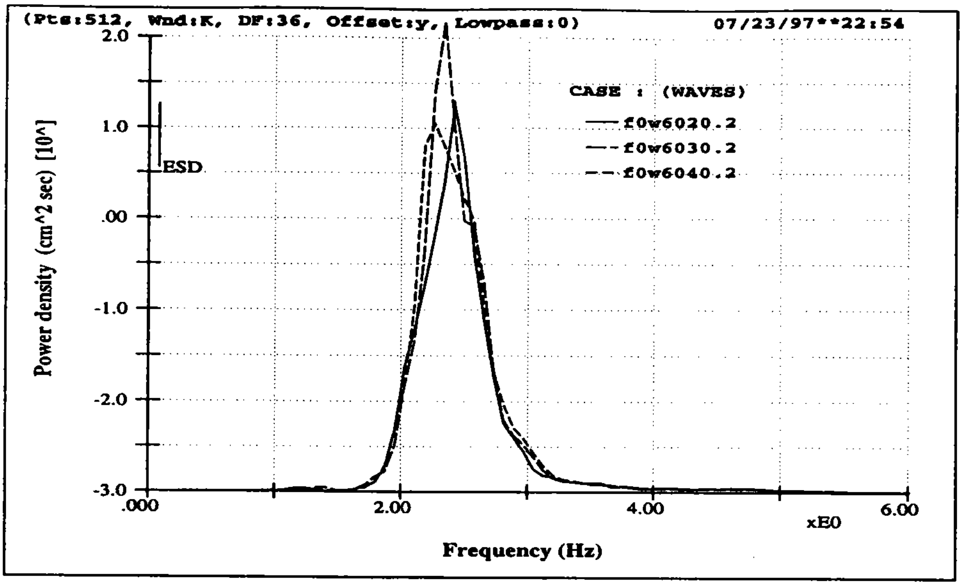
Sampling frequency : 40 Hz
Sampling time length : 240 Sec

Ch	#_W	H.1	H.2	H.3	H1/10	H1/3	H1/2	H.ave	H.rms	T1/10	T1/3	T1/2	T.ave	T.rms
1.	546	22.81	20.65	20.52	17.04	14.48	13.32	10.35	11.01	.45	.44	.44	.44	.45
2.	563	2.98	2.88	2.72	2.35	2.04	1.89	1.46	1.55	.44	.43	.43	.43	.43

Case : f0w6050.dat (f1 p3 c1 s9)
Date : 01/05/96
Time : 02:00:30.72
Specifics :

Sampling frequency : 40 Hz
Sampling time length : 240 Sec

Ch	#_W	H.1	H.2	H.3	H1/10	H1/3	H1/2	H.ave	H.rms	T1/10	T1/3	T1/2	T.ave	T.rms
1.	546	17.62	16.49	16.30	14.02	11.94	10.97	8.22	8.88	.45	.44	.44	.44	.45
2.	562	2.83	2.74	2.72	2.36	2.06	1.91	1.45	1.55	.44	.44	.44	.43	.43



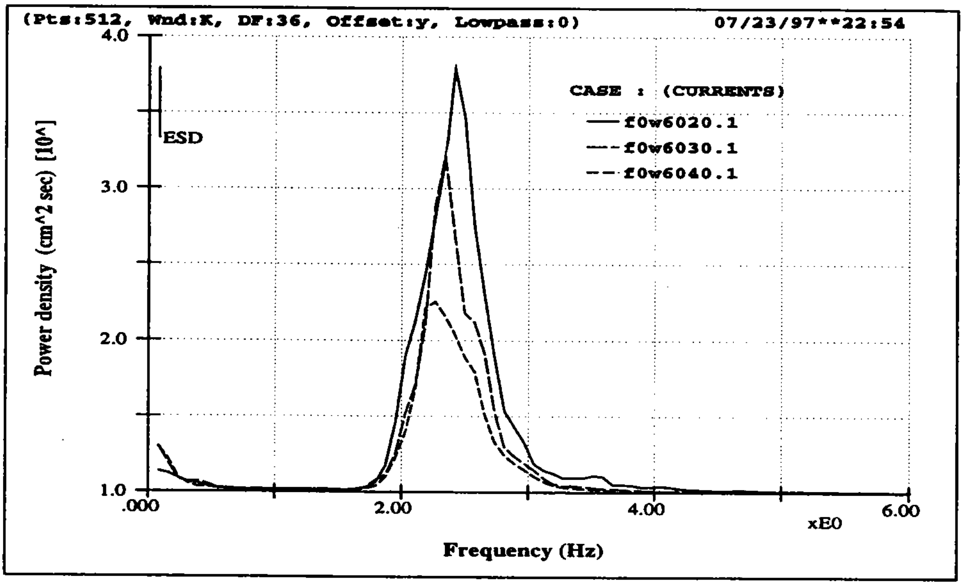


Figure 1.2: Spectra for the surface displacement and aqueous flow measurements. The top sub-figure shows power spectra of the repeated measurements of water surface displacement. The bottom sub-figure shows power spectra of the LDA aqueous flows at different measurement depths. Compared with table 1.3, it is seen that the repeatability of power spectra is rather poor even though the zero up-crossing statistics has indicated the existence of a good stationary condition. This also implies the likelihood of poor performances of spectral coherence and signal deconvolution for two signals acquired at different batches, either under the same wind condition or not.

i.e, there exists high coherence between the displacement and aqueous flow measurements within the same case run. The points here give lucid illustration of the effects of phase noise on spectral outcomes, since the displacement and aqueous flow measurements are acquired at almost the same location (with 1 cm horizontal distance) there is no phase noise between the two spectra for a single case run, on the contrary, we just don't have any control over the phases for different case runs.

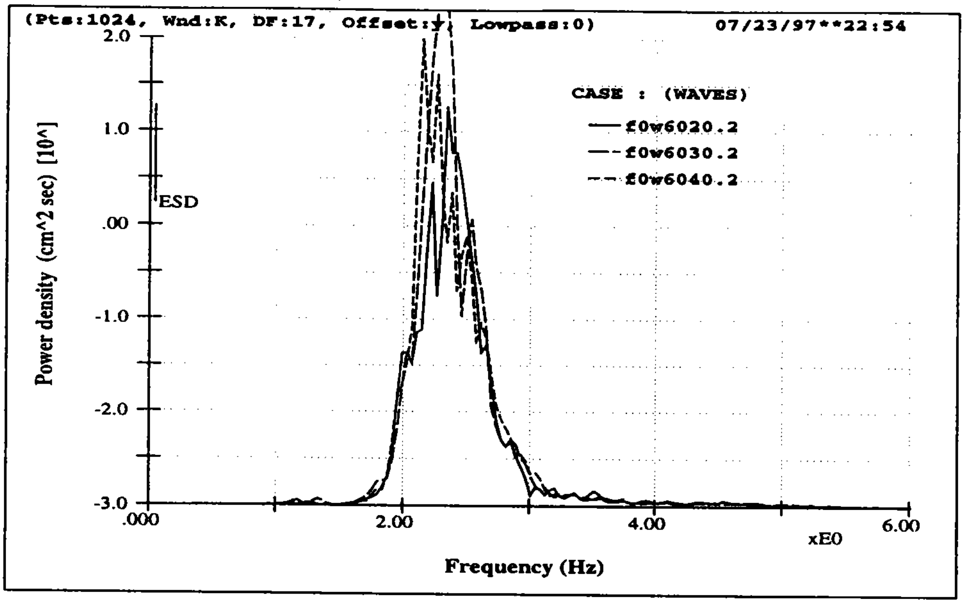
In calculating these power spectra standard treatments of FFT are employed. Related processing parameters are shown in the figures. For all cases the total length of the data is multi-segmented with 50% overlapping and Blackman window is applied to each section. Figure 1.3 shows the results using segment length of 512 points (with an approximate degrees of freedom of 36). Power spectra using a different segment length of 1024 points (with an approximate degrees of freedom of 17) are shown in figure 1.3. While the spectral resolution of the former figure is inferior, the repeatability of power spectra in the latter is significantly worsen. These behaviors also somewhat manifest the statements given in previous sections.

A more direct explanation for this repeatability problem can be simply given as the rapid diminishing as well as the irregularity of the auto-correlation coefficient function (or distribution) as shown in Figure 1.4. In the figure the auto-correlation coefficient functions of two wave gauges located at up-wind and down-wind positions are shown. Their correlation levels are quite low, and the physical interpretation for it is that wind waves lose their identities extremely fast and they reflect typical phenomena related to transient or pulse natures.

1.4 Natural Frequency of the Wind Wave Tank

Well controlled experiments can often yield informative results as well as fine scale characters that are unexpected and generally not quite obvious. Example shown in this section might sound miscellaneous but it severs to illustrate several intrinsic aspects related to experiments. These aspects include: the response characters of the instruments, the accuracy of the measurements, awareness of the degree of control in the experiments. In addition, this example also illustrates different inherent properties between numerical results calculated from auto-correlation coefficients and those from power spectral estimates, even though theoretically the Fourier transform of the auto-correlation function is the power spectrum.

When the wind is blown in the oval tank energy might be picked up by the natural frequency of the tank through some gradual process or some kind of rapid excitation. Our experimental data shows that, if the oval tank is not blocked there does exist an energy



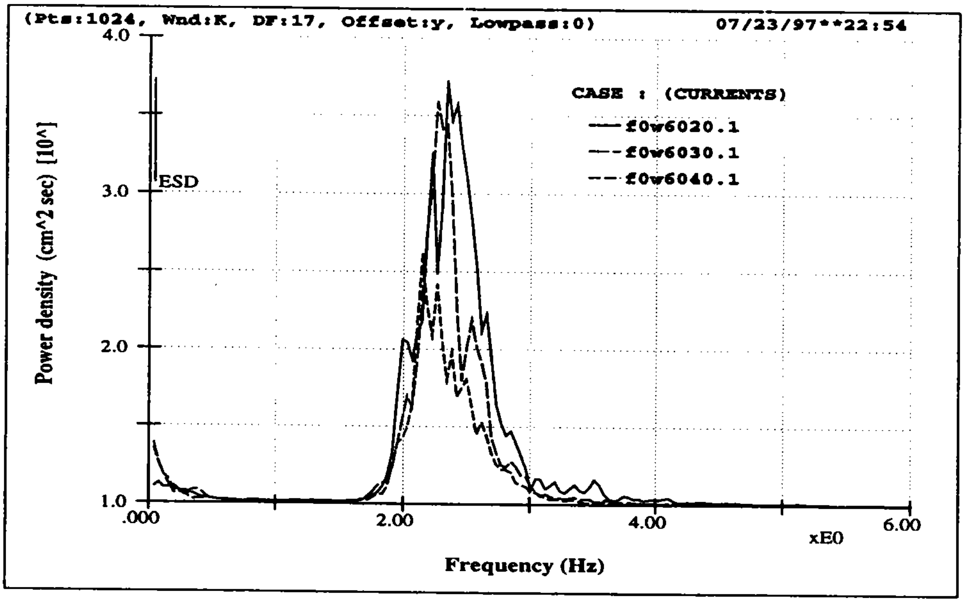
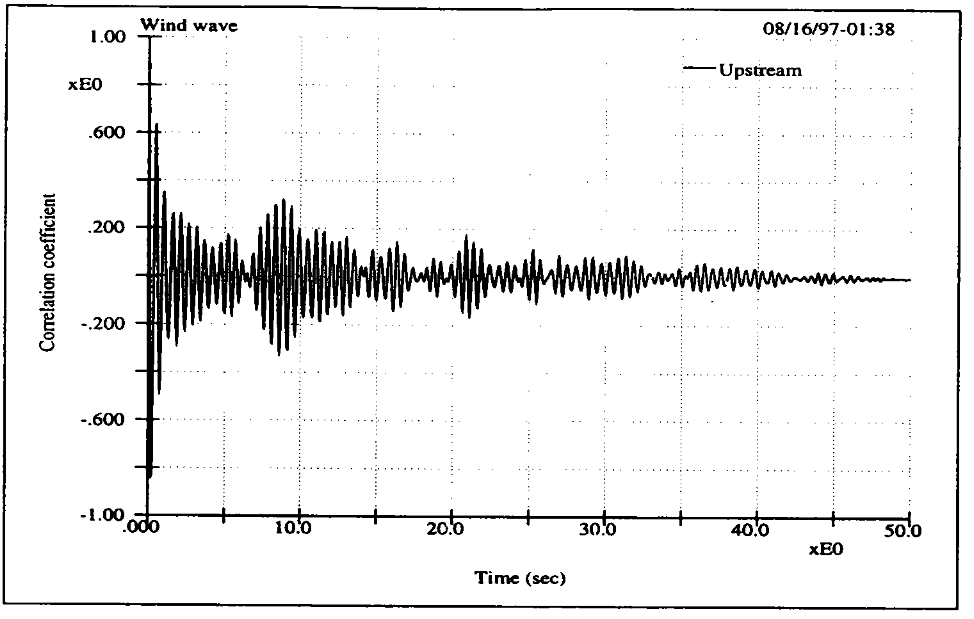


Figure 1.3: Spectra for the same data as in the previous figure but with different FFT parameters. Here a 1024-point segmentation is used. The degree of freedom is approximately halved. The resolution is improved whilst the standard deviation increases.



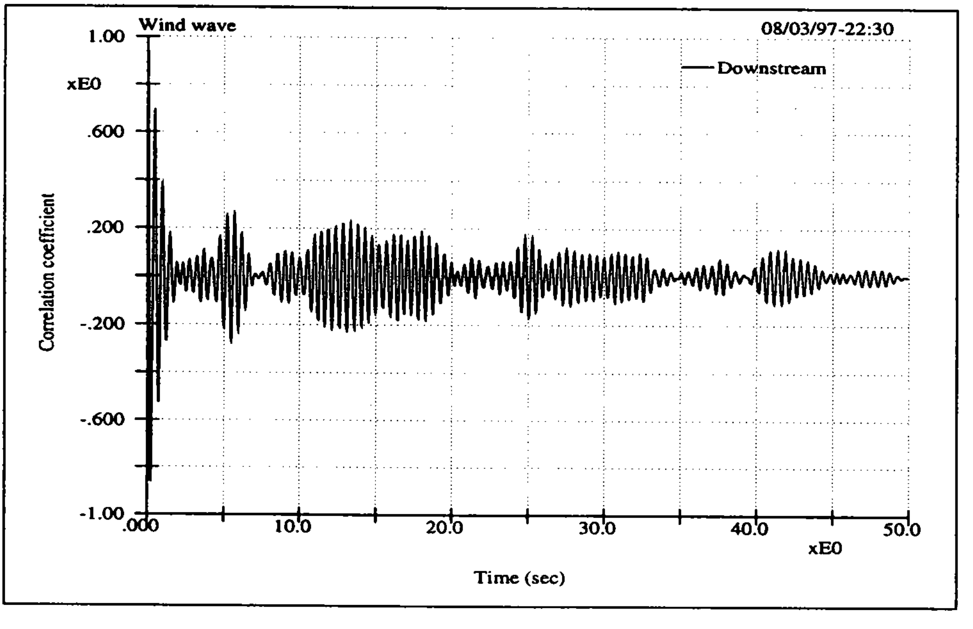


Figure 1.4: auto-correlation coefficients of wind wave signals (up- and down-stream) in the oval tank. The correlation level is low and diminishes rapidly. A reasonable argument is that these wind waves lose their identities extremely fast

pickup from the relatively short wind waves by the relatively low natural frequency which is about 30 times less than typical peak frequency of the wind waves in the tank, as will be illustrated in the followings. And in all cases there is no sign of rapid excitation.

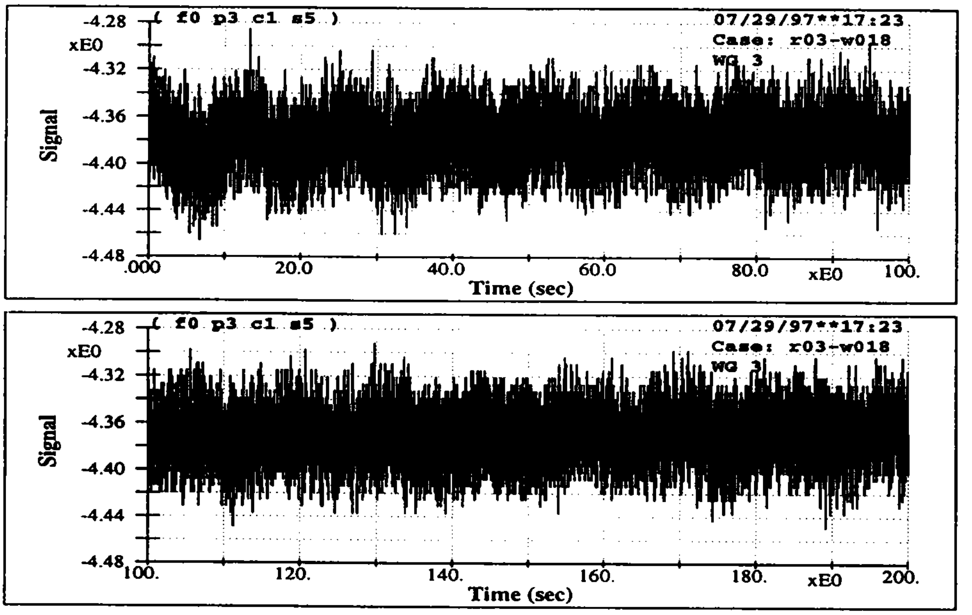


Figure 1.5: A noisy wave form where the wave of the natural frequency of the tank imbeds within.

Figure 1.5 shows a noisy wave form (which is the worst for the wave gauges we used) where the wave of natural frequency is actually imbedded. The data was acquired after the short wind waves dies away rapidly when the wind was stopped. As is seen the signal is relatively weak and almost completely submerged in the noise from the instruments. However, when the signal is smoothed and auto-correlation coefficients are calculated the period for the natural frequency is easily identify as 13.3 sec, which is in good agreement with the time needed to travel the tank using the shallow water wave limit according to the formula of $C_g = C = \sqrt{gh}$, where C_g is the group speed, C is the celerity of the wave, g is the gravitational acceleration, and h is the water depth. In the figure the tapering of the curve is associated with zero padding when calculating the coefficients.

This natural frequency can also be seen using convolution filtering. Figure1.7 shows the low-passed signal, where minimum degree of filtering using a Blackman type filter with a cutoff frequency of 3 points per cycle is adopted. As can be seen the amplitude of the wave is approximately equal to 0.15mm. The figure also indicates that most of the noise comes from relatively high frequencies and has little influence on frequencies where water waves

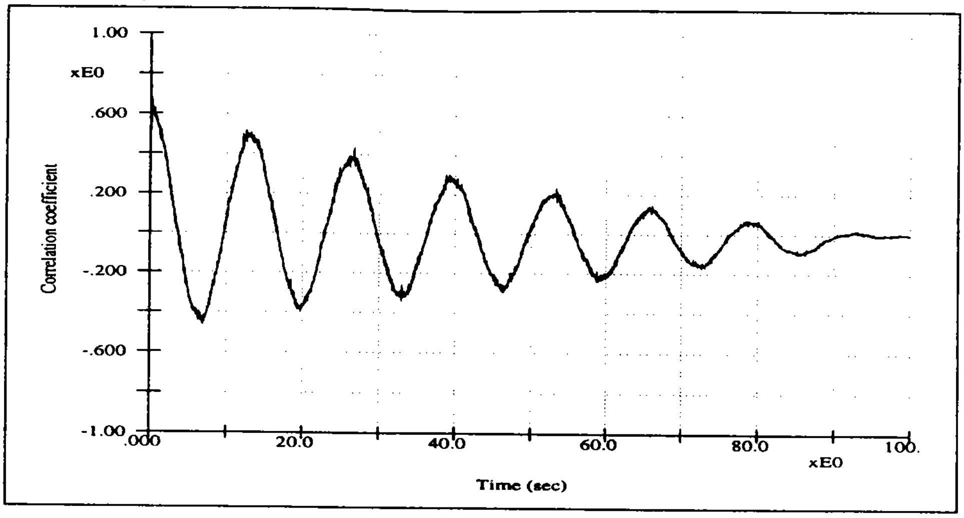


Figure 1.6: The auto-correlation coefficient function of the noisy wave form which shows the natural frequency of the oval tank.

mainly concern (about less than 10 Hz).

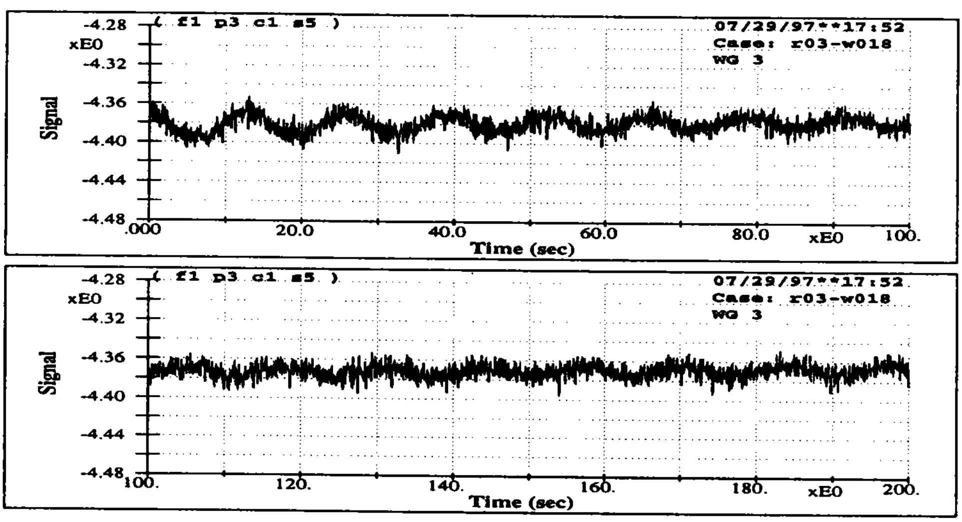


Figure 1.7: The natural frequency can also be seen from low-passing the signal. Here Blackman filter with minimum degree of filtering is used. Compared with figure 1.5 one see that most of the noise is associated with relatively high frequency, in reference to those of water waves.

Since power spectrum is the Fourier transform of the auto-correlation function, let look at the frequency domain results. Figure 1.8 shows two spectral curves using two different segmentation lengths. It may not be as easy as the auto-correlation curves in identifying the natural frequency mainly because the peak is dispersed and possibly not located near the resolution point. A few practical limitations associated with numerical aspects of discrete Fourier transform such as segmentation length, frequency leakage, edge effects, windowing, and spectral resolution, again play the important roles; while the auto-correlation is implemented in the time domain and have different features in these regards.

Some additional points for this section are given in the followings.

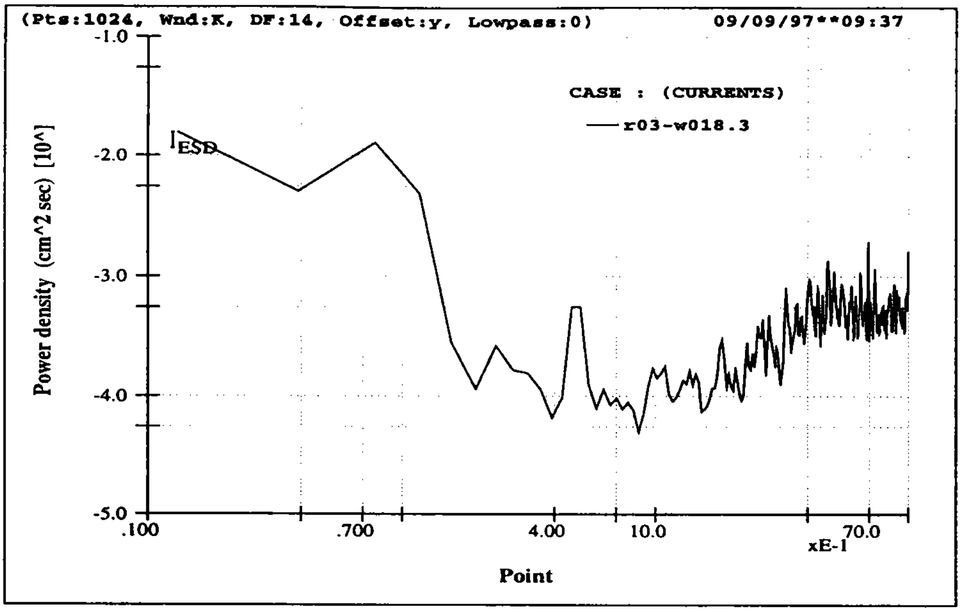
- There are intricate details in the numerical processes that render differences between the two seemingly identical implementations of a single formulation, that is to say, we have an additional concern that different implementations of a single formulation may yield results that should be intuitively the same. And this is in addition to those differences between discrete and continuous transforms.

- The noise level is generally higher than that of measurement accuracy. In our experiments the wave gauges were calibrated using a vernier with 0.05mm accuracy. Response characters of instruments, including resolution and time responses of acquisition hardware and pre-filtering instruments are verified using a WaveTek signal generator. No deterioration of signals can be identified for the instrument setups.

- Noise are of high frequency and can mostly be filtered out.

- In the case shown here the end of the oval tank was not blocked by wave absorber; if the tank was blocked no such frequency could be detected, it is therefore anticipated that there is no adverse effects of resonant excitation from the tank.

- Wave reflection coefficients are also estimated by separating incident and reflected waves using the wave separation technique proposed by Goda and Suzuki [6]. Several cases in which mechanically generated waves of longer wave periods are analyzed. The estimations from these cases should yield conservative values when compared to wind wave cases. Table 1.4 shows numerical results of an estimation, in which the energy reflection is about one percent. It is anticipated that for wind wave conditions the reflection should be milder. In our view we regard this energy reflection level as possibly having the same degree of uncertainty associated with the numerical calculation, since the processing method basically uses FFT tactics.



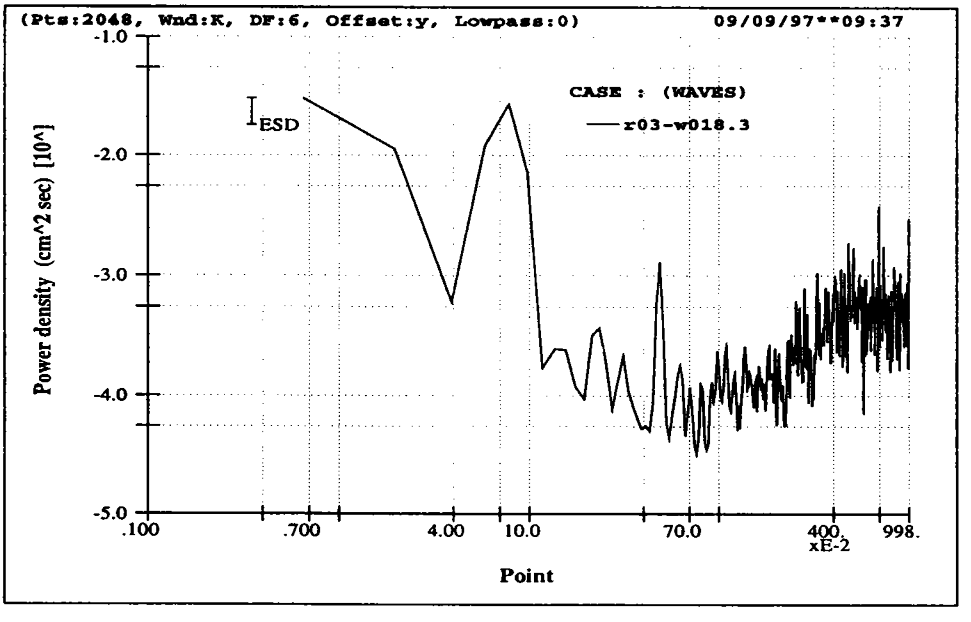


Figure 1.8: Power spectral curves for signal shown in figure 1.5. Since power spectrum is the Fourier transform of the auto-correlation function of the signal, the comparison with the auto-correlation curve indicates the existence of intricate differences between the two numerical implementations for a single formulation.

Table 1.2: Energy reflection level for a mechanically generated wave in the oval tank based on Goda and Suzuki's method [6] for separation of incident and reflected waves.

Case : Kr_Oval.dat (slp3cl)
Date : 04/29/95
Time : 13:50:27.84
Specifics : Oval tank Kr

Sampling frequency : 200 Hz
Sampling time length : 52 Sec

Ch	#_W	H.1	H.2	H.3	H1/10	H1/3	H1/2	H.ave	H.rms	T1/10	T1/3	T1/2	T.ave	T.rms
1.	110	31.95	30.77	30.61	30.61	29.09	27.69	23.31	23.99	.49	.48	.48	.46	.47
2.	106	31.10	30.60	28.11	28.84	26.41	25.82	24.05	24.17	.48	.48	.48	.48	.48
3.	112	3.47	3.40	3.35	3.27	3.01	2.88	2.48	2.54	.47	.47	.48	.45	.46
4.	106	3.44	3.43	3.38	3.38	3.27	3.21	2.96	2.97	.47	.48	.48	.48	.48
5.	108	3.42	3.17	3.13	3.19	2.99	2.90	2.65	2.67	.49	.49	.49	.48	.48

Energy reflection level (%): 0.8424

Program parameters etc. ::

Kr	:	.0918	r.ch1	:	4	r.ch2	:	3
			depth	:	24.0000	del.L	:	9.0000
L.min	:	20.0000	T.min	:	.3581	M.max	:	57
L.max	:	180.0000	T.max	:	1.2983	M.min	:	16
smo.flag	:	1	f.cut	:	67	aR.sq	:	57.6631E 3
fft.point	:	4096	b.pnt	:	1	aI.sq	:	68.4510E 5

Reconstructed incident and reflected waves :

Ch	#_w	H.1	H.2	H.3	H1/10	H1/3	H1/2	H.ave	H.rms	T1/10	T1/3	T1/2	T.ave	T.rms
4	42	3.79	3.55	2.85	3.26	2.28	1.85	1.45	1.59	.48	.48	.48	.48	.48
4	42	.34	.31	.26	.28	.21	.16	.12	.14	.49	.48	.48	.47	.47

1.5 Spectral Comparison and Signal Deconvolution

Let suppose we have an input signal and an output signal. The input signal is, say, an original signal without being influenced by external factors, or a signal measured before it is affected by some sort of structure. And the output signal is, say, the altered signal due to the introduction of those external factors or structures. One simple intuitive thinking regarding the identification of the effects of these influential factors or structures on the constituent components of the signal is to compare the input and output spectra. In this section we will explain why this approach is generally problematic. In fact the idea of this spectral comparison is just a concept manifesting a deconvolution process and can be illustrated with an extremely simple blackbox mechanism as shown in figure 1.9. The basic reasoning comes from the convolution duality property of Fourier transform,

$$h(t) \Longleftrightarrow H(\omega) \tag{1.4}$$

$$h(t) \star f(t) \Longleftrightarrow H(\omega)F(\omega). \tag{1.5}$$

Here again the double arrow sign means that the role of t an ω can be inter-changed. Refer-

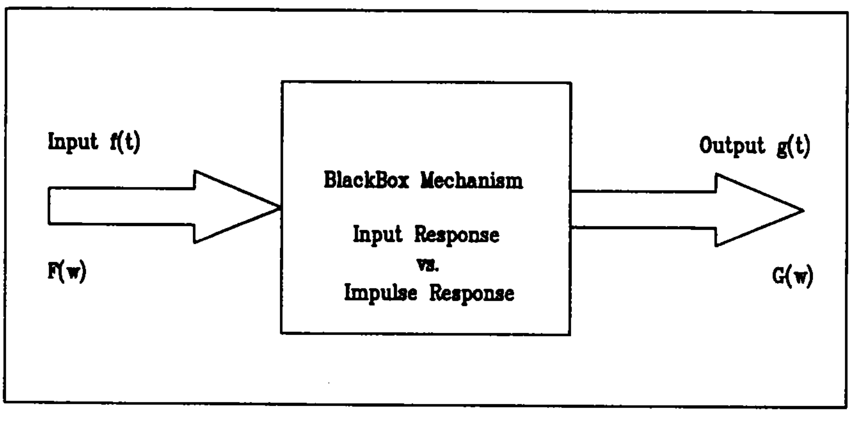


Figure 1.9: The simple blackbox here is to illustrate the following problem: Does direct quotient of spectral coefficients of two spectra physically significant? If direct division of the two spectra is taken then this blackbox implies that the output is the convolution of the input signal and a certain impulse response function, or alternatively speaking, the blackbox mechanism is the deconvolution between the output signal and the input signal. The concept is intuitively simple, but it is generally of little use due to the fact that the process is extremely error prone as explain in the text. The figure also indicates the inherent problems concerning making direct detail comparison of two spectra when there exists possible randomness.

ring to figure 1.9, if one assumes that

$$H(\omega)F(\omega) = G(\omega), \tag{1.6}$$

it seems, therefore, quite straightforward to say that individual effect on each frequency component is simply the division of two spectra,

$$H(\omega) = \frac{G(\omega)}{F(\omega)}. \tag{1.7}$$

And the response function $h(t)$ is simply as

$$h(t) = \mathcal{F}^{-1}\left[\frac{G(\omega)}{F(\omega)}\right]. \tag{1.8}$$

The problems here are manifolds, both practically and analytically. They all arise from the occurrence of $F(\omega)$ in the denominator. Practically, as are indicated in the earlier sections, due to finite resolution and discrete nature of data processing, the spectral periodogram estimate is only the expectation value of the power spectrum of a continuous distribution and this estimator suffers severely from having large variance, with efforts to improve it being rather inefficient [17]. Analytically, if only there exists a single frequency resolution point where there is no power content then the inverse transform is simply non-existent. Moreover, in many physical situations the power spectra are often quite narrowed banded, or stated otherwise, peaked only around a small region within the whole range of frequency axis, therefore, for frequency components with little energy the division of two spectral coefficients is extremely error prone. This symptom is generally referred as amplitude equalization in radar imaging terminology. The amplitude equalization results in a high-pass window function in the frequency domain. In the presence of additive noise, this window further amplifies the noise [18]. Being aware of this fact, one should avoid direct source deconvolution; nevertheless, it was not uncommon to see that similar algorithm of taking direct division of two power spectra is used to make judgements when dealing with somewhat complicate problems where phase noisy and transient effects are significant.

When viewed from the impulse response point of view, the above mentioned direct deconvolution has a physical interpretation in the time domain as being the process of a finite impulse response model (FIR) (equation 1.8, [1]). In this regard there is one related model called infinite impulse response model (IIR). For these two methods it seems that up until now we can say little about one method being superior to the other. The finite impulse response (FIR) model has the advantage of being with finite support in the response function, i.e., the number of convolution weights is finite. The disadvantage is that the response function basically states that not only previous input will have influence on the output signal but also that the parts of signal that will only happen (or created) in the future will have influence on signal right now or even far a long time ago. These implications of FIR are simply unrealistic and questionable. For infinite impulse response model the advantage is that it is somewhat more reasonable because only one side of the response function exists, which means that only previous parts of the signal will leave traces on its following signal contents. However, there is severe disadvantages too. The IIR is a non-linear model and there is no simple formulation exists for deriving response functions, and there seems to be no fully automatic procedure for computer implementation. Moreover, they need fundamental assumptions which are generally of personal preference, in some sense this means that the choice of the infinite impulse response function is somewhat subjective and suffering from no unique numerical algorithm.

1.6 Motivations and the Modulation Perspective

The wordings of these several sections up until now seem not to favor the adoption of power spectra in making detail comparison of two signals. This is by no means to say that power spectral tactics is not powerful. It just implies that different approaches have different realms of strength as well as their own frailties. The overall summary of the weak point of traditional Fourier analysis is its deficiency in characterizing local or transient phenomena. In fact, Fourier analysis should not be blamed for this deficiency for we know at the very beginning that all spectral components are certainly not varying in spatial or temporal domain (since all are sinusoids with constant amplitudes and frequencies). If spectral approach is forced to cope with this situation it will mimic the signal with many additional components and severely increases the tangling of spectral components; and practically, the finite resolution in discrete Fourier transform will let the spectral results suffer harshly. Therefore, for studying signals with many local and transient properties one's obligation is to find additional stratagems.

The intention of this report is to view signals, especially for water wave related signals, in somewhat different ways, mainly from modulation perspective. We will only concern what is termed as the "analytic signal procedure", which in more plain terms is studying the modulations of waves. There seems to be no clear definitions for the terms of modulations. Modulations imply vague ideas of temporal or spatial variations, such as amplitude modulation, phase modulation, and frequency modulation. Related concepts are instantaneous or local amplitude and frequency. In fact, the word "modulation" tells the basic difference between this study and general spectral analysis. In spectral analysis the two corresponding domains are time and frequency, respectively, i.e., each domain has only one single variable and the correspondence is one to one (we mean the number of coefficients in both domains), therefore they are all uniquely defined. From modulation perspective if the signal is in time domain then the corresponding domain has variables both in time and frequency; but the correspondence is not truly from one to two dimensions under the analytic signal procedure (not all frequencies exist at a particular time as is the case of general time-frequency analysis [1]). Concerns regarding the natures of this mapping naturally arise. And, these are the primary tasks that will be worked on. Related questions like: Is there unified approach within the analytic signal method? Are the results unambiguous and how are they interpreted? Are they physically explainable and practically meaningful? And, what can be said when comparing this method with spectral analysis?

Before the end of this chapter we give several words on our own limited understanding of the broad aspects of signal analysis. It is proper for us to be aware that there exist many different perspectives in the studies of signal and within individual perspective there may also exist different regimens. Acknowledging this, we could wish to be somewhat more objective and be giving our judgements more cautiously and unbiasedly. So, we make it clear that the present study only occupies a small territory in the field of time-frequency analysis; and time-frequency analysis is further encompassed within a board category of time-scale analysis. Furthermore, the present study will mainly use a procedure devised by Gabor in 1946, which is generally referred as "the analytic signal approach". To the author's knowledge this method was only scarcely used in our field of water wave related signals and, in some sense or other, the interpretations of the results [12, 15] were not without doubts. These skeptics are associated with the present interests as well as the desires to have a general idea of the usefulness of the method. Both modeled numerical example and data from experiments [2] will be used to illustrate various aspects in these regards.

[1] Occasionally one may band-pass filter the signal. This will give several frequencies at a particular time, such as examples given later.

[2] Applications to experimental data or wave characterizations from this perspective will be mainly given in the second part of this report.

Chapter 2

The Analytic Signal and the Instantaneous Frequency

The usefulness of a particular method of data analysis is highly signal dependent. Not only specific features exist for each individual method, but also tradeoffs and complications are all over to be found either among different methods or even among different schemes within a single category. The scope discussed here covers only a small spot in the field of signal analysis. In fact the term "modulation" just reflects a very vague idea of "changing", but it seems intuitively beneficial if we can characterize what are changing using simple concepts of local frequency, local phase, and local amplitude, especially if it is possible to view a signal as a single modulated wave form with a single time-varying amplitude and a single time-varying frequency. However, the plain facts of its fate are just not that simple and straightforward; that is why we have to bear with a variety of approaches and to open our eyes to broader outlooks. It is hoped that every single step helps in moving toward next advance.

2.1 The Complex Signal

Measurable signals are certainly real. However, it is often advantageous to associate a real signal $s_r(t)$ with a sensible imaginary counterpart $s_i(t)$. The real and imaginary parts form a complex signal $z(t)$. A complex function allows us to define its modulus (or amplitude) function $a(t)$ and phase function $\phi(t)$ of a complex exponential. The derivative of the phase yields the natural definition of instantaneous frequency $\omega_i(t)$ (this is viewed from the time and temporal frequency domains, if it is viewed form the space and spatial frequency domains the equivalent term is local wavenumber). The simple mathematical description is

$$z(t) = s_r(t) + s_i(t) = a(t)e^{i\phi(t)}, \tag{2.1}$$

with

$$\omega_i(t) = \phi'(t). \tag{2.2}$$

The main concern here is what is the sensible imaginary part and how to define it and this concern is sure to influence the exploitation on instantaneous frequency. Up until now it is generally regarded as an open question regarding the proper definition of the complex signal [3]. Here we should also point out that in the realm of signal analysis most researchers still regard instantaneous frequency as merely a primitive concept rather than a question of mathematical definition. The issues are, at best, whether a particular definition can match our intuitive thinking; whether their results can provide adequate explanations for the physics that might be of our own logical reasoning only; and whether the intuitive assumptions induce additional concerns which might be counterintuitive and possibly bring us to new discoveries.

There are basically two methods to define the complex signal. One is the method of quadrature model and the other is the method of analytic signal. Any real signal $s(t)$ can be expressed as

$$s(t) = a(t)\cos\phi(t). \tag{2.3}$$

For the quadrature model method the complex counterpart, $z(t)$, of the signal should intuitively be

$$z(t) = a(t)e^{i\phi(t)}. \tag{2.4}$$

Problems arise concerning the uniqueness of the representation of equation 2.3. And in fact there are an infinite number of ways to devise the complex form. In 1946 Gabor proposed a definition for the complex signal [5] which avoided the uniqueness problem and his method is generally called the analytic signal procedure. He introduced the complex signal as

$$\begin{aligned} s(t) &\iff S(\omega), \\ z(t) &= 2\frac{1}{\sqrt{2\pi}}\int_0^\infty S(\omega)e^{i\omega t}d\omega, \end{aligned} \tag{2.5}$$

where $S(\omega)$ is the spectrum of the real signal $s(t)$. The factor 2 in the equation is introduced so that the real part of the complex signal is equal to the original signal. As is clear from this definition, the complex signal $z(t)$ is the inverse Fourier transform of a single-sided spectrum with no components in negative frequency but with the same positive spectral components as those of $S(\omega)$; in other words, the Fourier transform of the complex signal defined in this way will have only positive frequency components. This somewhat matches our fondness that no negative frequency should ever occurs, even though at this time we do not have a real signal. In the following section we will give somewhat detail descriptions on how the analytic signal method is implemented, so that one might have as clear as possible understanding of the natures of this approach, especially when it turned out that the situations (or the computational results) were far complicate and hard to be answered from the author's prior understanding in a preliminary stage or during its first encounters.

2.2 The Analytic Signal Procedure

The form of the analytic signal in terms of the real signal is [3]

$$z(t) = s(t) + i\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}\frac{s(\tau)}{t-\tau}d\tau. \tag{2.6}$$

The symbol $\mathcal{P}$ in the equation means that the integration is carried out based on the rule of Cauchy principal value. In the following we will go through the process of how the analytic signal method is treated and, when possible, give its physical implication from experiences acquired. By doing so we do not mean to duplicate the work of the previous authors. But rather, only when one is aware of the details of manipulation and actually implements them that one is more equipped with forthright understanding of the intrinsic nature of the approach. And this is also helpful in gaining clues to the explanations of any unexpected outcomes and in clearing skeptical issues that might be bothering — whether they are associated with numerical aspects or they are of analytical natures, or they may be just caused by bugs in the computer program. In fact, the process described below had been somewhat gone through in the early stage of programming. But at that time more thoughts were put on derivation aspects, and little attention was put on their implications and practical considerations. It was when many hard-to-explain results were found that the author came back to have a thorough reexamination of the processes and put more thought on what the details might imply and implemented more numerical simulations. It was finally conceived that there was no bug [1] as can be found by the numerous checks in the most early program written long ago; but rather, some of the outcomes as well as a few "paradoxes" known in literatures are explainable since then, at least to the scope of our knowledge.

[1] In fact, if not being fully considerate is considered as a bug, then, practically, there is something wrong in the program. An example of this is the need of base band conversion as will be given in the second part of the report.

The imaginary part of equation 2.6 is the Hilbert transform of the signal, $\mathcal{H}[s(t)]$,

$$\mathcal{H}[s(t)] = \widehat{s(t)} = \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}\frac{s(\tau)}{t-\tau}d\tau, \tag{2.7}$$

The principal value operation means that the integration is implemented in way of

$$\mathcal{P}\int = \lim_{\epsilon\to 0}\left(\int_{-\infty}^{t-\epsilon_1} + \int_{t+\epsilon_2}^{\infty}\right), \tag{2.8}$$

with $\epsilon_1 = \epsilon_2$. Let

$$f(t) = \frac{1}{t}, \tag{2.9}$$

and using the duality relation of eq. 1.4, the Hilbert transform is simply the convolution of $s(t)$ and $f(t)$,

$$\widehat{s(t)} = \frac{1}{\pi}(s\star f)(t). \tag{2.10}$$

It's Fourier transform is

$$\mathcal{F}[\widehat{s(t)}] = \widetilde{S}(\omega) = \frac{1}{\pi}S(\omega)F(\omega). \tag{2.11}$$

After separating $S(\omega)$ and $F(\omega)$ the Cauchy principal value operation is associated with $f(t)$, whose Fourier transform is

$$\begin{aligned}\mathcal{F}[f(t)] = F(\omega) = \mathcal{P}\int_{-\infty}^{\infty}\frac{e^{-i\omega t}}{t}dt = \\ \mathcal{P}\int_{-\infty}^{\infty}\frac{\cos(\omega t)}{t}dt - i\int_{-\infty}^{\infty}\frac{\sin(\omega t)}{t}dt.\end{aligned} \tag{2.12}$$

The integrant in the real part of this equation is antisymmetry and the Cauchy principal value integration is zero. The latter integral does not need the principal value sign because $\sin(\omega x)/x$ is finite for all x in particular at $x = 0$. With a symmetrical integrant, since

$$\int_0^{\infty}\frac{\sin\omega x}{x}dx = \mathrm{sgn}(\omega)\int_0^{\infty}\frac{\sin u}{u}du \tag{2.13}$$

one basically know that $F(\omega)$ does not dependent on the variation of ω. But this integration is no trivial task and the known way is to invoke the complex integral calculus and using the residue theorem for integration, as can be found in the well written textbook by Greenberg [7]. The final result is a simple relation, which only depends on the sign of ω:

$$F(\omega) = \begin{cases} -i\pi\,\mathrm{sgn}(\omega) & \omega \neq 0 \\ 0 & \omega = 0. \end{cases} \tag{2.14}$$

Therefore, the Fourier transform of the analytic signal $\mathcal{A}[s(t)]$ is

$$\mathcal{F}[\mathcal{A}[s(t)]] = S(\omega) + i\mathcal{F}[\mathcal{H}[s(t)]](\omega) = \begin{cases} 2S(\omega) & \omega > 0 \\ 0 & \omega \leq 0. \end{cases} \tag{2.15}$$

The above equation yields

$$\widetilde{S}(\omega) = \begin{cases} -iS(\omega) & \omega > 0 \\ iS(\omega) & \omega \leq 0. \end{cases} \tag{2.16}$$

Making use of this relation the Hilbert transform is efficiently calculated by a simple word in ASYST language (which is equivalent to a subroutine in some programming languages). Table 2.2 lists the word. Detail manipulation of analytic signal approach is given here not for mere analytical interest; but rather to disclose the intrinsic nature of this method [2]. The hope is that the arguments given later will be more precise as well as more objective. The algorithm stated above and the program piece tell us that the basic tactic is composed of a few processes that manipulate the contents of the results of FFT on the original signal. This probably inscribes our doubts that properties for the FFT, as are discussed in the previous chapter, are going to be enlisted in this approach. Examples and explanations for both acquired data and modeled signals will be given.

2.3 Local Implication of the Analytic Signal

If a signal is expressed as a complex form then it is easy to have intuitive idea about the local information on amplitude and phase. The terms "local" has very distinct implication when compared with that of spectral concept, for we know that every spectral component has constant amplitude (but with different phase) at all time. Therefore all local features of a signal are all induced by the complication of phases from spectral point of view. From the point of view of the analytical signal method amplitude and phase are all single valued at a particular time. Therefore, these local values are anticipated to be the results of some sorts of combinations from all the spectral constituents at their own particular times. This statement highlight the motives of the present study to look into the features of these combinations. What can be expected? It would be fair to state that our standing point is that one should not favor the present method simply because the method is currently used. We should discuss both its merits and its drawbacks. Therefore, we will first give a brief summary of a few known paradoxes or counterintuitive properties regarding the analytic signal. We

[2] An alternative approach implemented in the time domain is given in later section.

Table 2.1: An ASYST word (equivalent to subroutine in some other computer languages) that calculates the Hilbert transform of an array. The word takes a one dimensional input argument. As can be seen in the program, the basic tactic is related to several processes which manipulate the contents of the FFT on the input signal. Are these going to inscribe the properties related to FFT into the results of the analytic signal approach? Examples and explanations will be given.

```
\ ------------------------------------------------------------------------
\  A small program piece which finds the imaginary part of the real signal
\     based on the analytic signal method.
\  The program makes use of the final results of complex calculus based on
\     Cauchy principal value integration.
\  The length of the array will be automatically truncated to some
\     maximum allowable power of 2.
\ ------------------------------------------------------------------------
: my.hilbert
  fft   []size   n.fft.pts  :=
  dup   becomes>  t1

  dup   sub[ 1 , n.fft.pts  2  /  ]
  0     +1      z=x+iy   *
  t1    sub[ 1 , n.fft.pts  2  /  ]  :=

        sub[ n.fft.pts  2  /  1 +   ,  n.fft.pts  2 / ]
  0     -1      z=x+iy   *
  t1    sub[ n.fft.pts  2  /  1 +   ,  n.fft.pts  2 / ]  :=

  t1    ifft
  zreal
;
\ ------------------------------------------------------------------------
```

then present a few numerical experiments and try to provide explanations for some of the paradoxes. In some cases, to the author's knowledge, I have not learned of the existence of well furnished explanations for them. Before we go further let state the basic differences between quadrature signal and the analytic signal.

2.4 Frequency Contents of the Amplitude and Exponential Modulation Parts of the Complex Signal

The analytic signal is complex as is given by equation 2.6; therefore, for both the quadrature model and the analytic signal procedure the signal can always be put into polar form

$$z(t) = a(t)e^{i\phi(t)} \tag{2.17}$$

In the quadrature model $a(t)$ and $\phi(t)$ are not uniquely determined and the pair can assume arbitrary forms. In the analytic signal approach $a(t)$ and $\phi(t)$ are unique and have very specific feature. This specific feature has the physical interpretation which states that the low frequency content of $z(t)$ is relevant to amplitude modulation part $a(t)$ and the high frequency content is in its complex exponential $e^{i\phi(t)}$. Since the Fourier transform of an analytic signal does not have negative frequency components, the above feature is just the manifestation of the property of "shift versus modulation", $f(t)e^{i\omega_0(t)} \Longleftrightarrow F(\omega - \omega_o)$. If the amplitude term has a frequency higher than that of the exponential term then the shift will not be able to move the spectrum to the right hand side of the frequency axis (being aware that spectrum of a real function is double sided with conjugate symmetry). It is therefore clear that the amplitude function of the quadrature model has a higher degree of variation than that of the analytical signal method; otherwise the quadrature model signal is simply the analytic signal, except, there might be a shift in frequency. The energy of the difference (not the difference of the energy) between the analytic signal and quadrature signal gives an indication of the closeness of the two signals. Its value is twice the energy in the negative part of the spectrum of the quadrature signal. But in general the quadrature signal can be put in a form close to analytic signal so that both have similar implications.

A brief summary of this section is: without concerning the difference of units, the degree of frequency modulation of amplitude is generally milder than that of phase. This also implies that the variation of instantaneous frequency is anticipated to be more violent. And the analytic signal yields a least modulated wave amplitude envelope.

2.5 Hilbert Filters in Time and Frequency Domains

In a previous section we gave a short program piece which does the Hilbert transform of the signal and the algorithm used is based on processes of frequency domain operation. As is also clear from equation 2.10 that the transform can also be implemented in the time domain in way of convolution if we can devise an impulse response function for the equation $f(t) = 1/t$. Based on its Fourier transform, equation 2.14, as well as on the Parks-McClellan algorithm for finding a "minimax" fit to a desired frequency response [14] the finite impulse response function can be designed (the discontinuity at the origin must be avoided). The minmax method minimizes the maximum deviation of the actual response from the desired response according to three parameters: the length of the filter in time domain, the start and end frequencies of the window in frequency domain. The differences between the results calculated from the previous program piece and those using the Parks-McClellan algorithm again are tradeoff properties. There are greater edge effects using the former method whereas there will be wiggling above and below the desired frequency response windows in the latter. Figure 2.2 shows envelope curves for a signal derived from the Hilbert transforms implemented in time and frequency domains, respectively. The length of the convolution filter weights determines the relative role between the wiggling and the edge effects. A figure for a set of parameters is shown in figure 2.1. Where the number of convolution weights is chosen to be 25 and the start and end frequencies are chosen to be 1.1 and 18.9 Hz, respectively, for a data scheme with Nyquist frequency of 20 Hz. In fact, we have seen again that numerical results from the analytic signal procedure are certainly not without artifacts due to its discrete and finite resolution natures both in the time and frequency domains, just as what we have encountered when dealing with Fourier transform. The points here partly explain the doubts listed at the beginning of this chapter.

2.6 Incompatible Concepts of "Instantaneous" and "Frequency"

To the author's own intuitive thought, the concept associated with "instantaneous" is just not compatible with that associated with "frequency". For "frequency" we mean that it is the intrinsic nature of sine and cosine functions; and sine and cosine functions simply resent to be "instantaneous" — at least there should exist a few cycles for the "frequency" to be meaningful. So, what are the resulting phenomena that are both intuitive and counterintuitive, and, what kind of message can the instantaneous frequency deliver for us? Besides, the complex form of $a(t)e^{i\phi(t)t}$ may give us strong and straightforward indication that the instantaneous frequency should be relatively independent of the amplitude since the instantaneous frequency is the derivative of phase. Moreover, what we generally heard of are more on the instantaneous frequency and it seems that much less attention was pay to the amplitude function $a(t)$. Is the situation really so, or, to what degree the inter-connection between the amplitude and phase functions should be emphasized? Subjects given in the next chapter will focus on these. Where we will first state some of the paradoxes, both known or from points based on our own experience and understanding. Phenomena will be illustrated using simple examples. Characterizations of their properties and explanations for their behaviors will be attempted.

Before ending this chapter, let restate that the proper description of changing frequency is not a settled question. Therefore there might exist alternative definitions for the instantaneous frequency [3]. In the present study the complex signal is restricted to that of Gabor's analytic signal procedure. And the instantaneous frequency is defined to be the derivative of the phase. Under this definition the average frequency is given by the time average of the instantaneous frequency weighted by the local energy density. This feature is consistent with the concept of geometrical moments that we are familiar with.

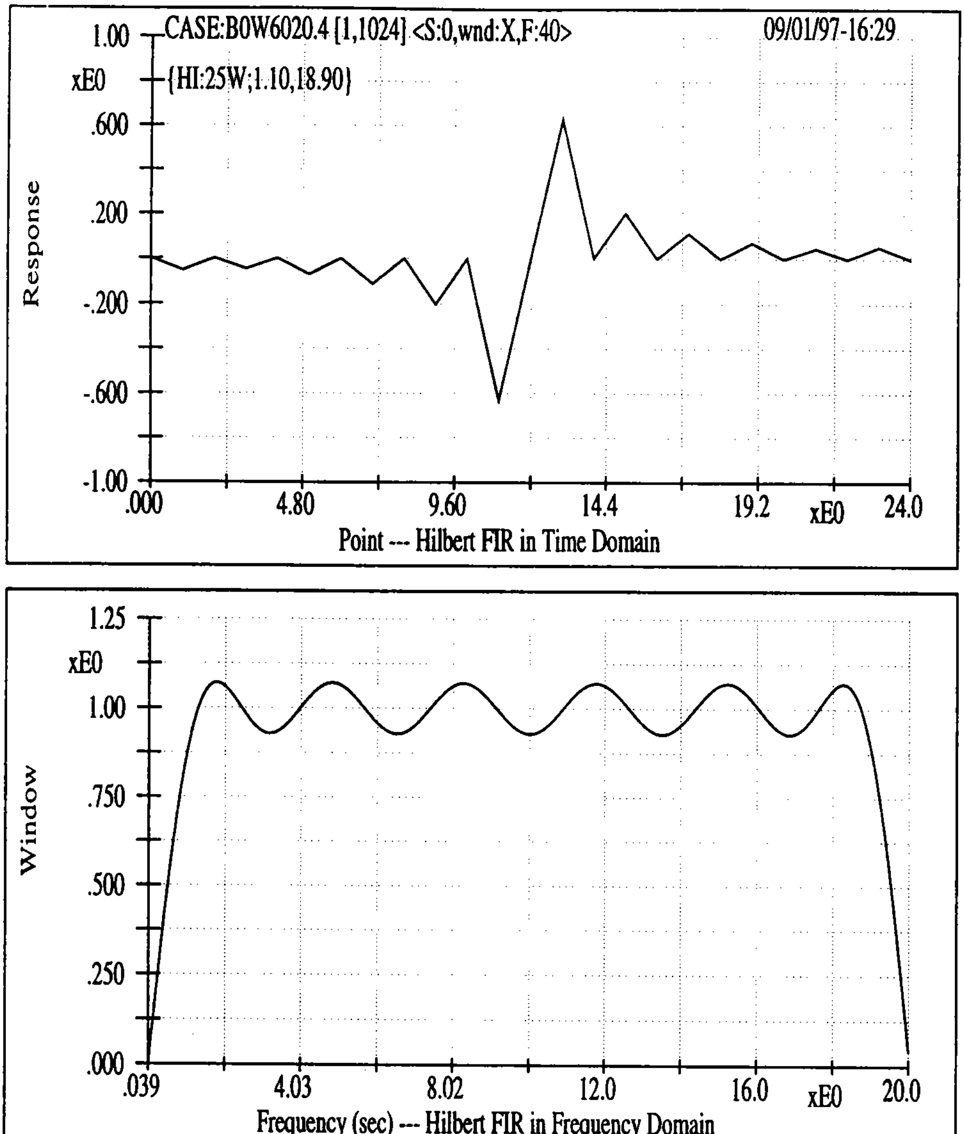


Figure 2.1: The Hilbert transform filter pair based on Parks-McClellan minimax algorithm. The length of the convolution filter is 25-point and the start and end frequencies are 1.1 and 18.9 Hz, respectively, for a data scheme of Nyquist frequency of 20 Hz. There are tradeoffs between this algorithm and that shown in table- 2.2. The present method results in wiggling of the response window function in frequency domain, which in effect causes up and down ripples of the envelope of the amplitude function; the latter has more edge effects.

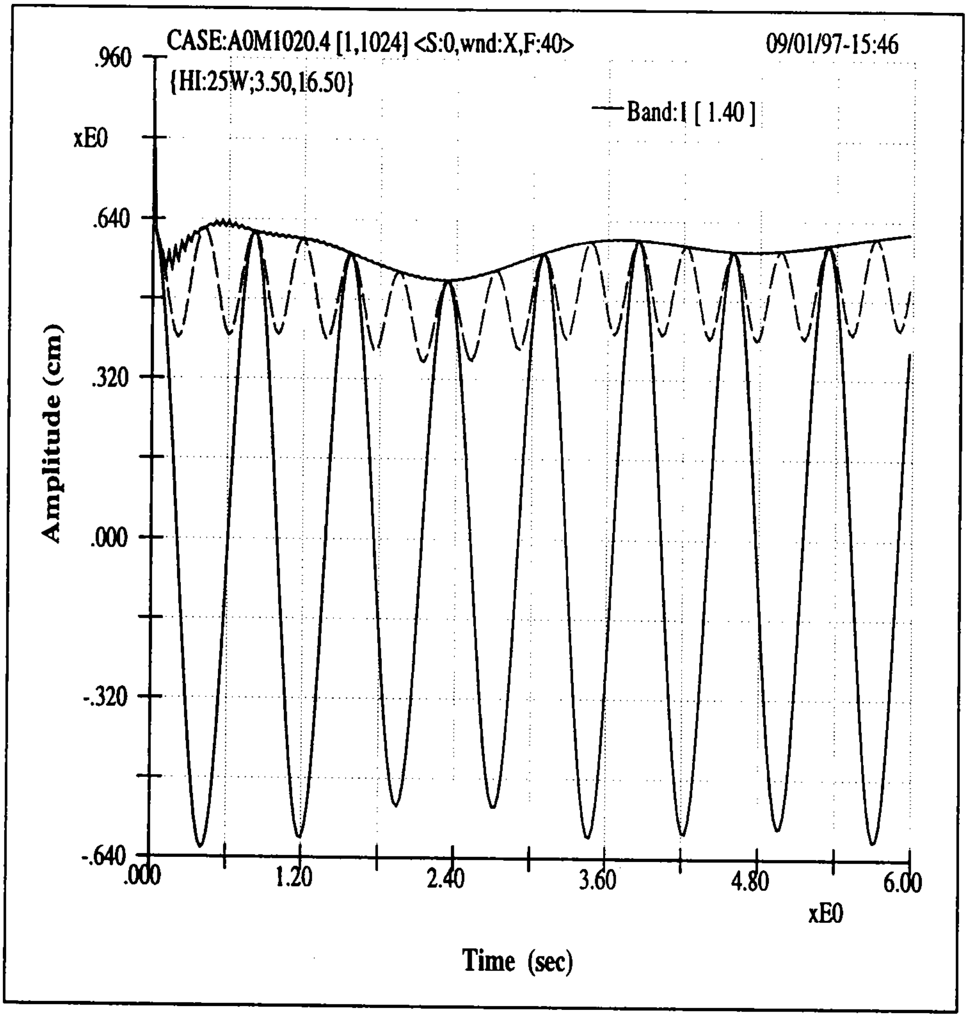


Figure 2.2: The envelope curves derived from the Hilbert transforms implemented in time (solid line) and frequency (dashed) domains, respectively.

Chapter 3

The Interplays Between Frequency and Amplitude Modulations and Their Characterizations

Paradoxes can sometimes be the many faces of an object, as well as be the same look of different issues. In the first situation different perspectives give different interpretations. In the second situation a certain feature can be associated with events of different natures. It is therefore true that they provide us more opportunities to characterize signals. However, one is also likely to be trapped in dilemma and riddled with flaws. In our view we regard that these paradoxes are relevant to the incompatible concepts between the "instantaneous" and the "frequency", since for "frequency" we mean there should be at least a few cycles for it to be meaningful, and this situation is simply not "instantaneous".

3.1 Paradoxes Regarding Instantaneous Frequency

The phenomena related to the paradoxes given here had long been bothering me in interpreting computational results of experimental wave data, but it was until I read some literatures that I formally grasped the existence of these difficulties or only partially understood problems. And they also intrigue me to refresh my own understanding and to do more works on this topic. In this section brief descriptions of these paradoxes will be given. In Addison, additional concerns which seem have not been well illustrated are also discussed. It is important for us to understand these difficulties so that our interpretations are not misleading. Numerical simulations as well as practical data will be used to explain these concerns. First let list the paradoxes mainly given in [3]:

- Instantaneous frequency may not be one of the frequency in the spectrum.
- Instantaneous frequency of a signal with discrete line spectrum (i.e., the signal is composed of only a finite number of discrete frequencies,) may be continuous and range over an infinite number of values.
- For a band-limited signal the instantaneous frequency may well go outside the band, both higher or lower.
- Although there is no negative frequency components for the analytic signal, the instantaneous frequency may be negative.
- While we start out thinking the instantaneous frequency as a local concept, why must we use the whole signal at all time, as is revealed by the definition for the Hilbert transform, to calculate the instantaneous frequency?

Additional points are:

- The last point in the paradoxes listed above indicates the dependency between amplitude modulation and frequency modulation. And if they are dependent, to what degree will one affect the other? That is to say, the point is how instantaneous frequency is affected by amplitude distribution. The question here is relevant to the intuitive thought that the instantaneous frequency should not concern too much on the amplitude function, since the instantaneous frequency is simply the derivative of phase.
- As were stated in the last chapter that there are complications in discrete Fourier transform when compared to continuous analysis, as well as differences between time domain and frequency domain operations. The question here is whether similar complications remain when dealing with the analytic signal. Considerations like these may lead to characterization of the approach.

To illustrate these paradoxes detail investigations by numerical modeling are helpful. First let us consider the analytic signal of a real signal with three sinusoidal components,

$$s(t) = ae^{i\alpha t} + be^{i\beta t} + ce^{i\gamma t}. \tag{3.1}$$

This example is basically the same as that given in Cohen [3] but with an additional third term; in addition, further considerations for cases with various combinations of parameters are given. The enhanced modeling gives additional information regarding these paradoxes.

If $a, b, c, \alpha, \beta,$ and γ are all taken to be constants, the spectrum of this signal consists of three delta functions (or distributions) at $\alpha, \beta,$ and γ,

$$S(\omega) = a\delta(\omega - \alpha) + b\delta(\omega - \beta) + c\delta(\omega - \gamma). \tag{3.2}$$

If $\alpha, \beta,$ and γ are further taken to be positive, then the signal is analytic. Solving for the phase and amplitude functions, one has

$$\phi(t) = \arctan\left(\frac{a\,\sin(\alpha\,t) + b\,\sin(\beta\,t) + c\,\sin(\gamma\,t)}{a\,\cos(\alpha\,t) + b\,\cos(\beta\,t) + c\,\cos(\gamma\,t)}\right) \tag{3.3}$$

$$\begin{gathered} A^2(t) = \\ a^2 + b^2 + c^2 + \\ 2ab\,\cos((\alpha - \beta)\,t) + 2ac\,\cos((\alpha - \gamma)\,t) + 2bc\,\cos((\beta - \gamma)\,t), \end{gathered} \tag{3.4}$$

and taking the derivative of the phase we obtain

$$\begin{aligned} \omega_i(t) \quad = \quad & \frac{1}{A^2(t)}[a^2\alpha + b^2\beta + c^2\gamma + \\ & (\alpha + \beta)ab\,\cos((\alpha - \beta)\,t) + \\ & (\beta + \gamma)bc\,\cos((\beta - \gamma)\,t) + \\ & (\alpha + \gamma)ac\,\cos((\alpha - \gamma)\,t)]. \end{aligned} \tag{3.5}$$

Figure 3.1 shows the instantaneous frequency and the amplitude modulation of the complex signal for the parameter set of $[a, b, c; \alpha, \beta, \gamma] = [-1.2, 1, -1; 1.6\times 2\pi, 3.2\times 2\pi, 3\times 2\pi]$. This figure illustrates the first three paradoxes listed above.

Figure 3.2 shows the results where there are slight modifications to the frequencies of the three sinusoids as $[a, b, c; \alpha, \beta, \gamma] = [-1.2, 1, -1, /1.592 \times 2\pi, 3.183 \times 2\pi, 3.024 \times 2\pi]$. As can be seen the positions of occurrences of the sharp variations of instantaneous frequency do not change very much, but the general shape of frequency modulation changes dramatically. It also shows that the instantaneous frequency may go negative.

Other interesting phenomena are also shown by considering a few additional parameter sets, such as those shown in figures- 3.3 and 3.4. Here all the three parameter sets have the same frequency components and the same values of modulus (absolute values). The

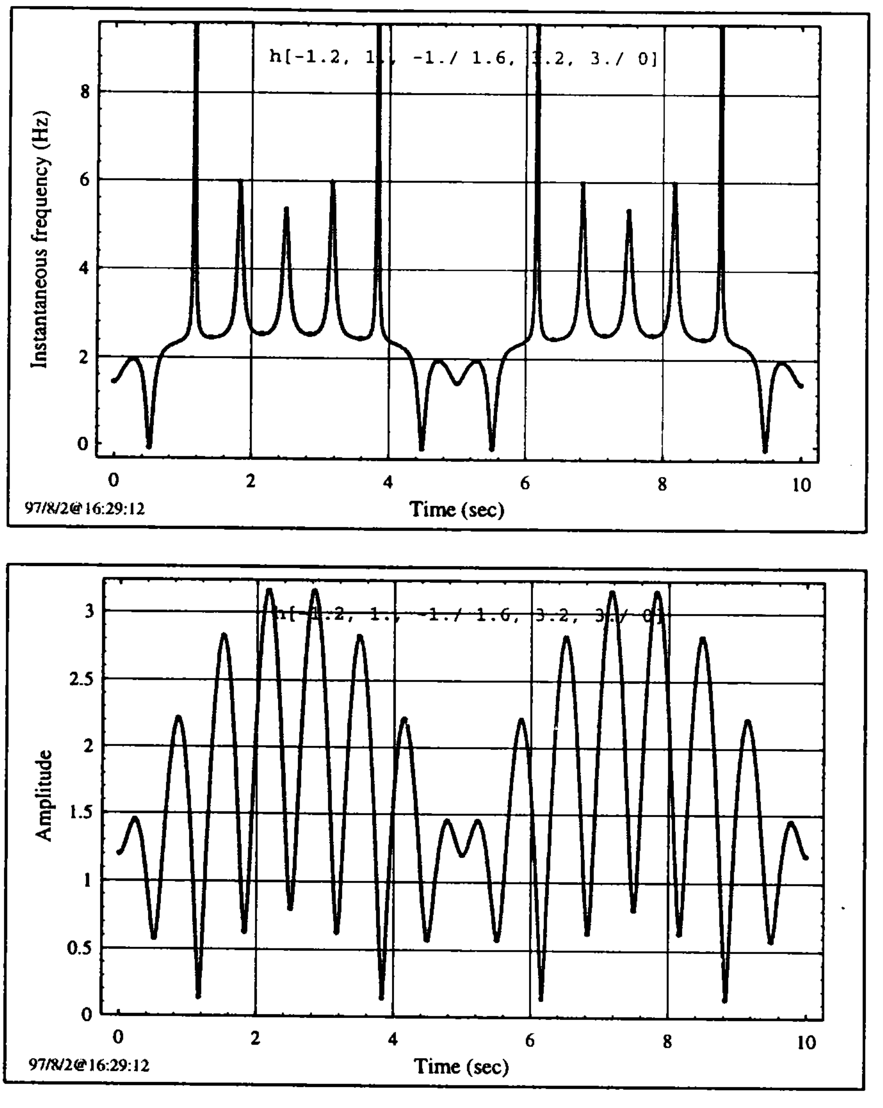


Figure 3.1: The instantaneous frequency and amplitude function for the parameter set $[a, b, c; \alpha, \beta, \gamma] = [-1.2, 1, -1, /1.6 \times 2\pi, 3.2 \times 2\pi, 3 \times 2\pi]$.

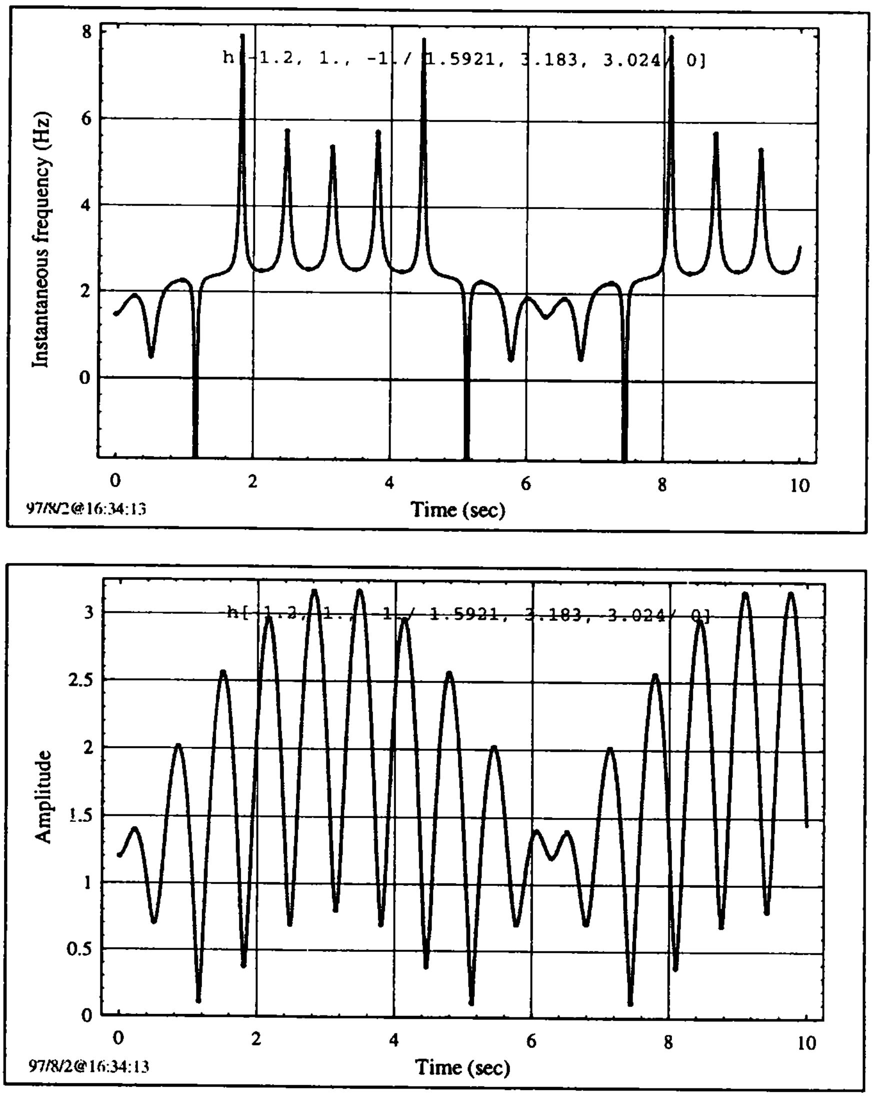


Figure 3.2: The instantaneous frequency and amplitude function corresponding to the parameter set $[a, b, c; \alpha, \beta, \gamma] = [-1.2, 1, -1, /1.592 \times 2\pi, 3.183 \times 2\pi, 3.024 \times 2\pi]$.

only differences are on their relative phases. The relative phases in these examples are given as the negative signs associated with the amplitude, which is equivalent to a phase shift of π radian. Figure 3.3 shows the modulations corresponding to two parameter sets as $[a, b, c; \alpha, \beta, \gamma] = [-2, 1.5, 2, /2.5 \times 2\pi, 2.75 \times 2\pi, 3 \times 2\pi]$ and $[a, b, c; \alpha, \beta, \gamma] = [-2, -1.5, 2, /2.5 \times 2\pi, 2.75 \times 2\pi, 3 \times 2\pi]$. Figure 3.4 shows those for $[a, b, c; \alpha, \beta, \gamma] = [-2, -1.5, -2, /2.5 \times 2\pi, 2.75 \times 2\pi, 3 \times 2\pi]$. As can be seen both the amplitude and the instantaneous frequency in the two figures have totally different behaviors. It is also shown that there are points near which the instantaneous frequency may approach infinity, both in positive and negative directions.

3.2 Effects of Support Length and Differentiability of the Amplitude Function on the Instantaneous Frequency

The terms "support length" and "differentiability" (and many others as mentioned below) concern more on continuous functions distributed over the whole real line and are analytically oriented; but they serve well in fulfilling the purpose of the present chapter for explaining features from discrete numerical computations, in which these terms are basically of no significance.

The previous examples clearly illustrated the first four paradoxes. But there seem to be few clues for the other points. Let us consider a different case using a chirp signal

$$z(t) = a(t)e^{i(\alpha t + \beta t^2)}. \tag{3.6}$$

For this case, if $a(t)$ is finitely supported in time axis then $z(t)$ may not be qualified as analytic signal, since the spectrum of $a(t)$ stretches the whole frequency axis and the modulation term may not be able to shift the negative frequencies fully toward the positive part. For $z(t)$ to be analytic $a(t)$ is better to be band limited. For clearing this point let use a bell shape normal distribution $a(t) = e^{-\sigma t^2}$ as an example. Since Fourier transform of $a(t)$, $\mathcal{F}[a(t)](\omega) = A(\omega)$, has the same Gaussian distribution as $a(t)$, the amplitude function is infinitely supported in both the time and frequency domains, therefore, the complex signal is a quadrature model signal rather than an analytical signal. However, it is reasonable to take an envelope function as close as to the Gaussian function in the time domain such that it is only finitely supported in the frequency axis (one can just does it in reverse way by assuming a spectrum very close to the infinitely supported Gaussian bell spectrum). We therefore see that a slight change can change the status of analytic signal from being existent to being non-existent. And this point clearly explains why the "local instantaneous frequency" is

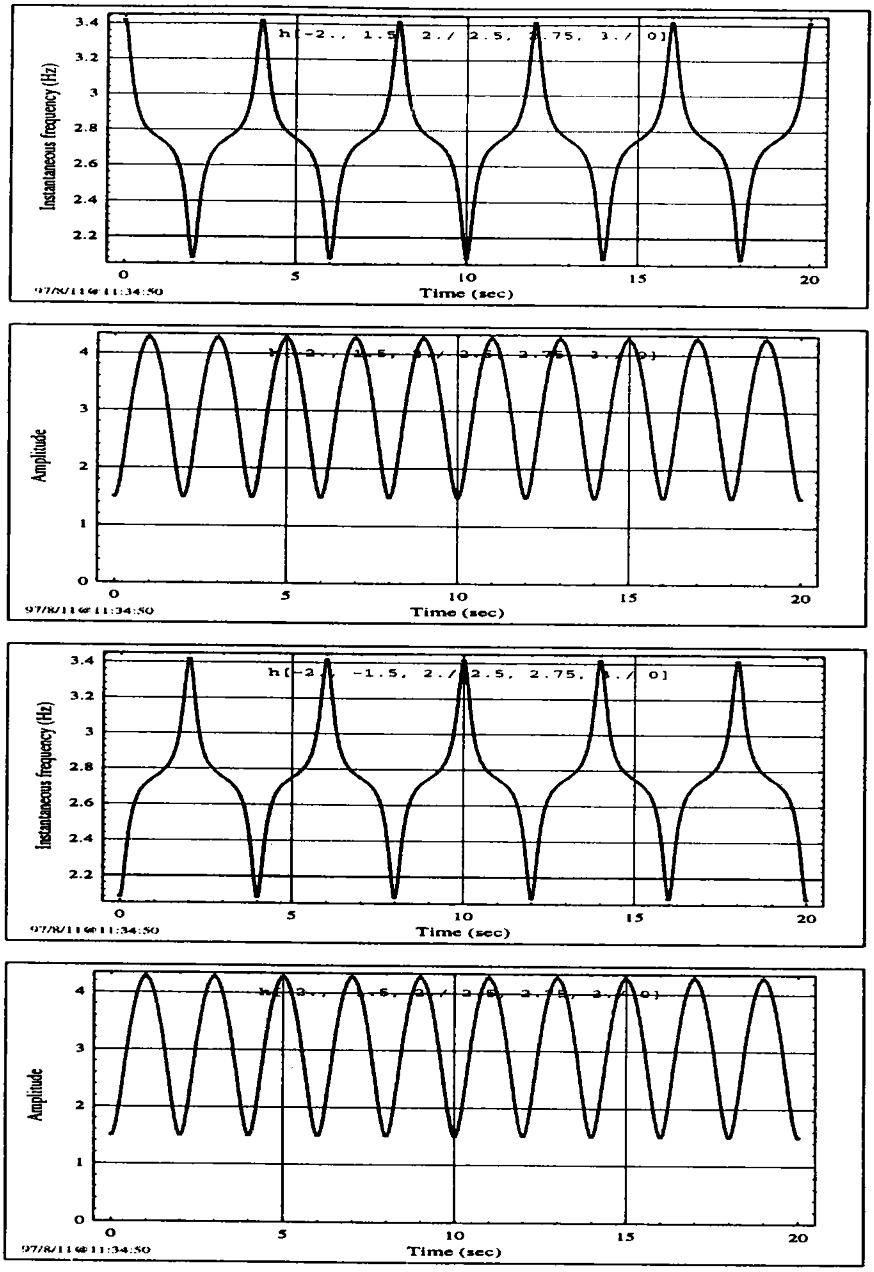
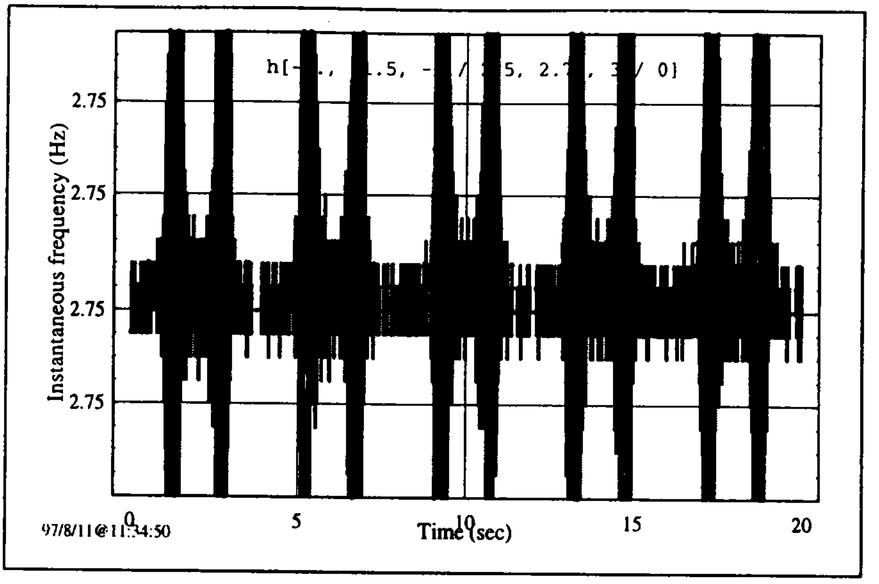


Figure 3.3: The frequency and amplitude modulations for two different sets of parameters, $[a, b, c; \alpha, \beta, \gamma] = [-2, 1.5, 2, /2.5 \times 2\pi, 2.75 \times 2\pi, 3 \times 2\pi]$ and $[-2, -1.5, 2, /2.5 \times 2\pi, 2.75 \times 2\pi, 3 \times 2\pi]$. Relatively, there is a constant phase shift in one of the sinusoidal components. The figure is to be compared with the following one.



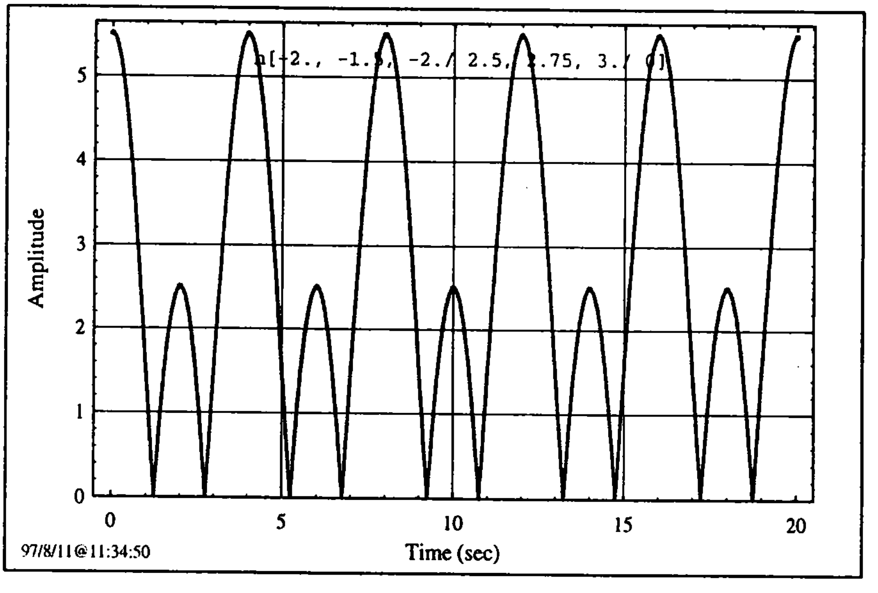


Figure 3.4: The frequency and amplitude modulations for another set of parameters $[a, b, c; \alpha, \beta, \gamma] = [-2, 1.5, -2, /2.5 \times 2\pi, 2.75 \times 2\pi, 3 \times 2\pi]$. Again, there is only a constant phase shift in one of the sinusoidal components. This figure together with the previous one indicate the sensitivity of the instantaneous frequency on the relative phases of the spectral components. One concludes that the analytic signal procedure also suffers severely from the phase noise effects just like the power spectra does. (The tick labels for the y axes of the instantaneous frequency are all of the same value; this is to reflect that the instantaneous frequencies are basically the same except that there exist unbounded points under machine-precision.)

certainly related to the parts of signal that might be million years ago or will only happen in several generations later. Moreover, the instantaneous frequency for the chirp is

$$\omega_i = \alpha + \beta t. \tag{3.7}$$

This instantaneous frequency is simply independent of $a(t)$; and, this independence is certainly opposite to the "dependence" shown in the previous examples. And it seems that we can somewhat arbitrary change the shape of $a(t)$ without losing this property of independence. However, one should pay attention that arbitrary change of amplitude may cause the analytic signal to become a quadrature model signal only. It is therefore concluded that there are interplays between the amplitude and the instantaneous frequency.

In fact the statements above is relevant to a complicate issue of differentiability, or regularity, of the amplitude function. There is a lore of analytical topics involved in this question as can be found in Daubechies' wavelet treatise [4]. Features of regularity, differentiability, support length, band width, and decay properties are intertwined and tradeoff properties of signals. The points here are: if $a(t)$ has poor regularity (a more formal term is Lipschitz regularity) than $A(\omega)$ spreads over a broad range of ω, and from the shift versus modulation property we know that we've got to have a very high frequency modulation to shift the whole band to positive axis of ω; if $a(t)$ is not differentiable at certain points then there is no chance for $z(t)$ to become an analytical signal. It is also noted that most of these mathematical complications seem to be of analytical significance for continuous functions only; it might be relatively rare to find their practical implications in discrete numerical applications. The situation here is similar to a case in which the values of entropy of the wavelet coefficients of a signal calculated using Meyer wavelet and Battle and Lemerié wavelet can hardly be differentiated (Lee & Wu's [10, 11]. While the Meyer wavelet is finitely supported and only partly differentiable the Battle and Lemerié wavelet is infinitely supported and differentiable for all values on the real line, they have the same visual look. From complex signal point of view, the situation is likely to analogous to the discussion of the "energy of the difference" between the analytical signal and the quadrature model signal, which can be of no significance.

Let us give another explanation regarding the localization of the instantaneous frequency. In other words the question is: if we need the signal at all time to calculate a local instantaneous frequency then what is the degree of influence on the local frequency from parts of signals that are some distance away. We shall discuss this question form two perspectives. Firstly, taking the basic equation for the Hilbert transform, $\mathcal{P}\int_{-\infty}^{\infty}\frac{s(\tau)}{t-\tau}d\tau$, we know that the imaginary part of the complex signal is derived from inversely weighting the signal in accordance with the separation distance measured from the point where local values are desired. And by analogy to Gibbs phenomenon we know that there is more wiggling near sharp variation points when inverse Fourier transform is carried out. With greater variations under greater weighting at sharp variation points it is anticipated that there are extreme behaviors at these points. The sharp variation points are generally associated with occurrences of small wave envelope amplitude[1]. Secondly, examples also show that amplitude envelope of large scale have little influence on the instantaneous frequency. Figure 3.5 shows the amplitudes and envelopes at several separated frequency bands for a wind wave signal. The individual bands are obtained by band-pass filtering the signal with center frequencies indicated in the figure. Figure 3.6 shows the windowed amplitudes and envelopes at corresponding frequency bands. The Hamming window is used in this case and is shown as one of curves in figure 3.7. Figure 3.8 shows the frequency modulations at four individual bands with and without such a data window as labeled in the figure. As are seen, they have almost identical variations except enhancements of singular features at a few locations. It is concluded that the variations of instantaneous frequency is mostly determined by sharp local variations rather than magnitude of envelope function. Furthermore, since the local variations are not symmetric with respect to the point of interest for most of the cases, therefore, the instantaneous frequencies may well go extremely large as well as go in the opposite direction, i.e., run negative, as is also partly implied in the Hilbert transform equation.

In all cases, both from numerical simulations and physical data, the locality of the singular region (or sharp variation region) is in proportion to the locality of the amplitude modulation; that is to say, rapid change of instantaneous frequency lasts a shorter time span, and milder variation of instantaneous frequency last longer. This reflects the symptom that the concept of "instantaneous" is not compatible with that of "frequency".

To end this chapter we conclude that the paradoxes and the skeptics mentioned earlier have been mostly explained, and we have also pointed out that the analytic signal procedure have many properties, such as phase noise, ambiguity effects, edge effects, filter lengths, finite resolutions, etc., that are similar to those of Fourier transform. This also implies possible limitations or difficulties when studying a broad band process using the analytic signal procedure.

[1]This feature is related to a character that roughness elements of the wind waves are more commonly found to situate at low wave displacements, supposed there is no obvious breaking, as will be studied in part II.

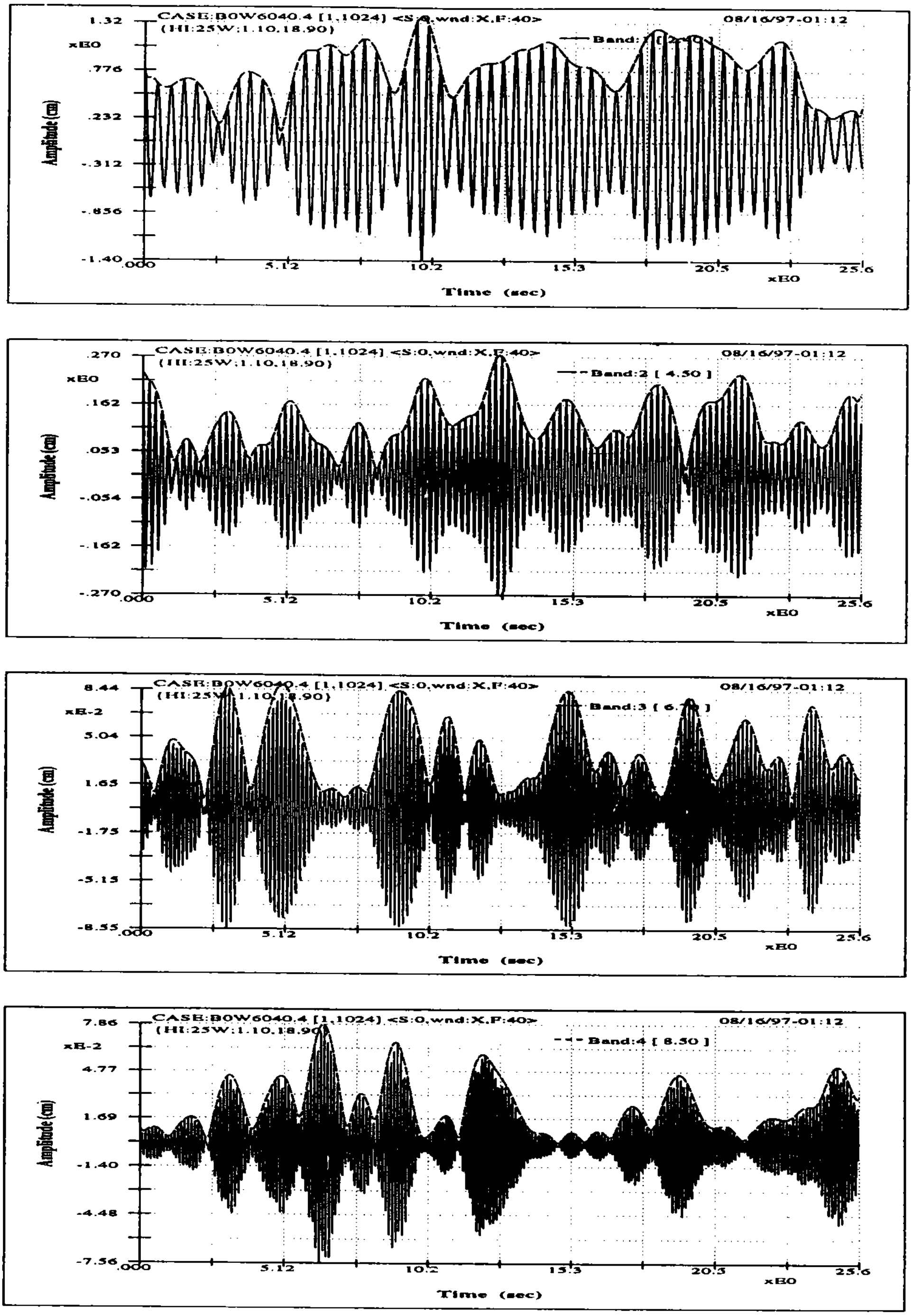


Figure 3.5: Amplitude modulations at four different bands for a wind wave signal.

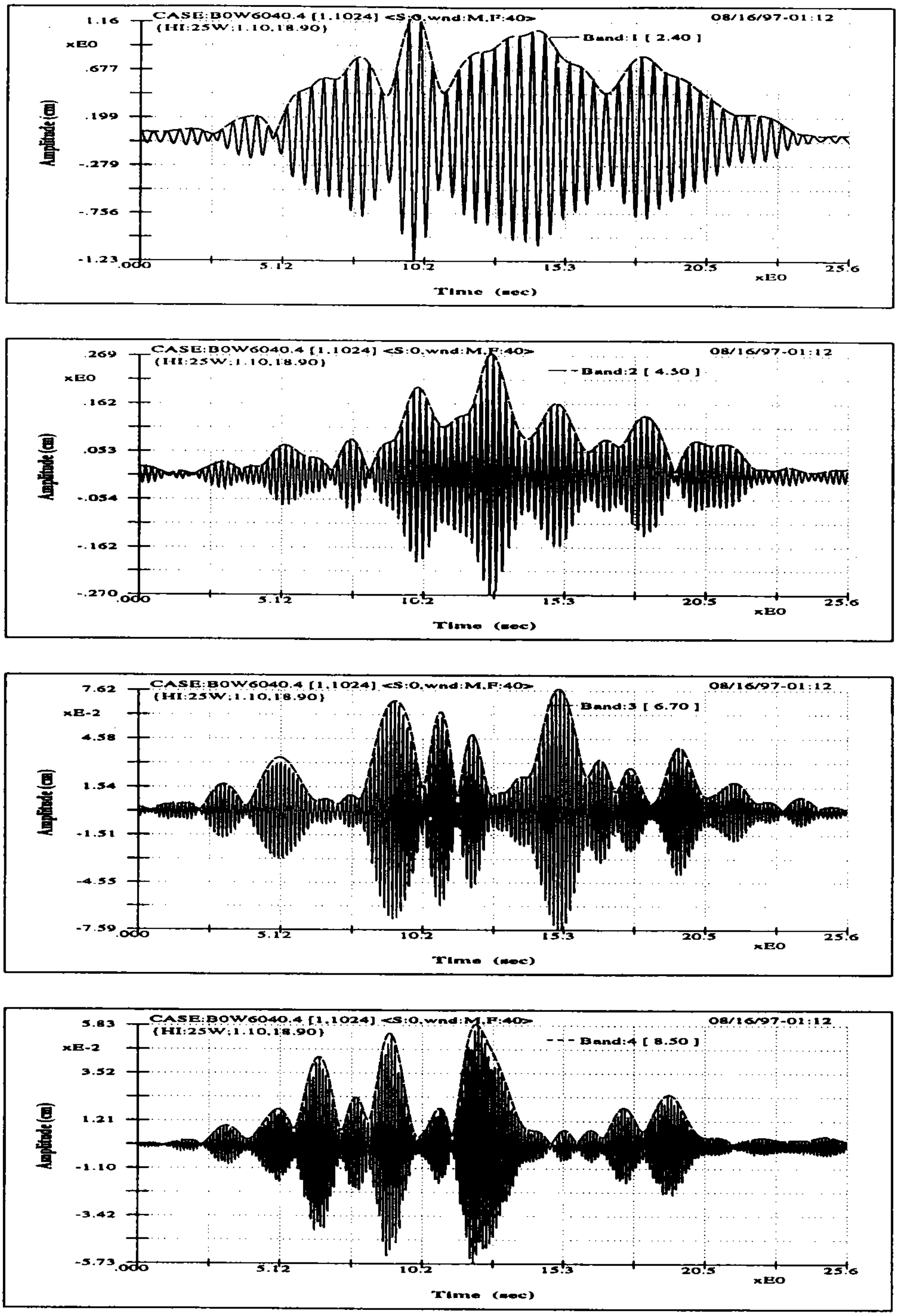


Figure 3.6: Windowed amplitude modulations for the four different bands of the same wind wave signal in figure 3.5.

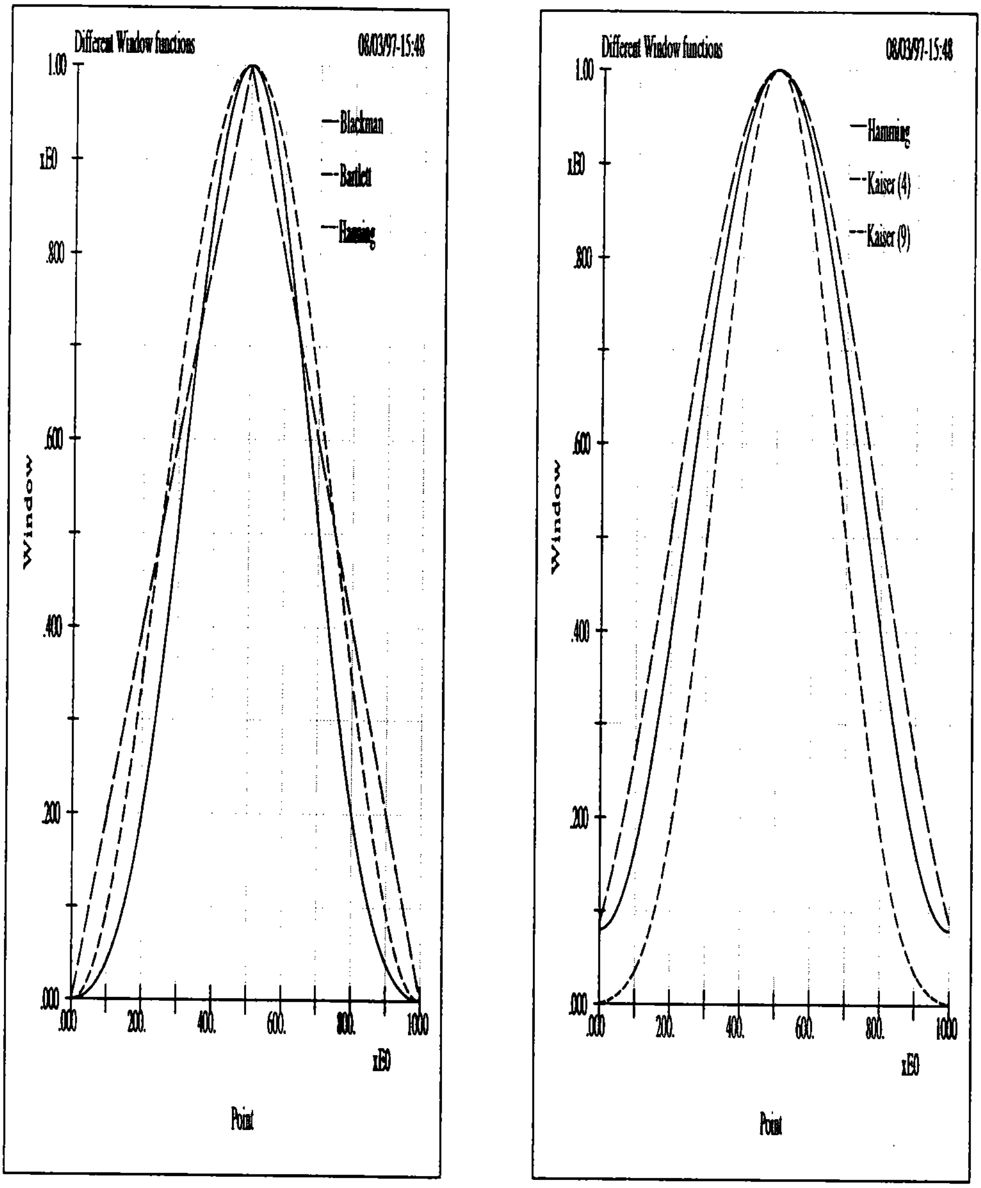
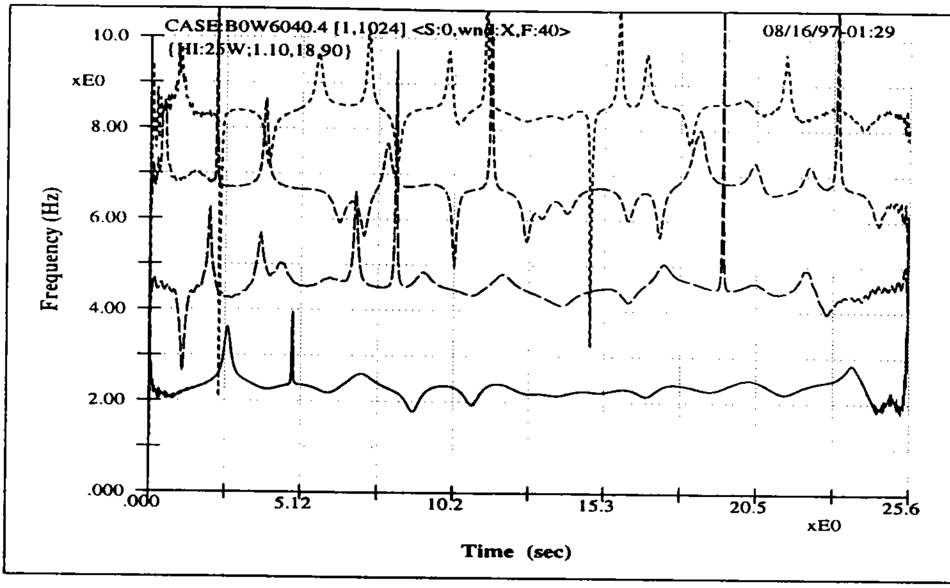


Figure 3.7: The data windows used in various numerical computations.



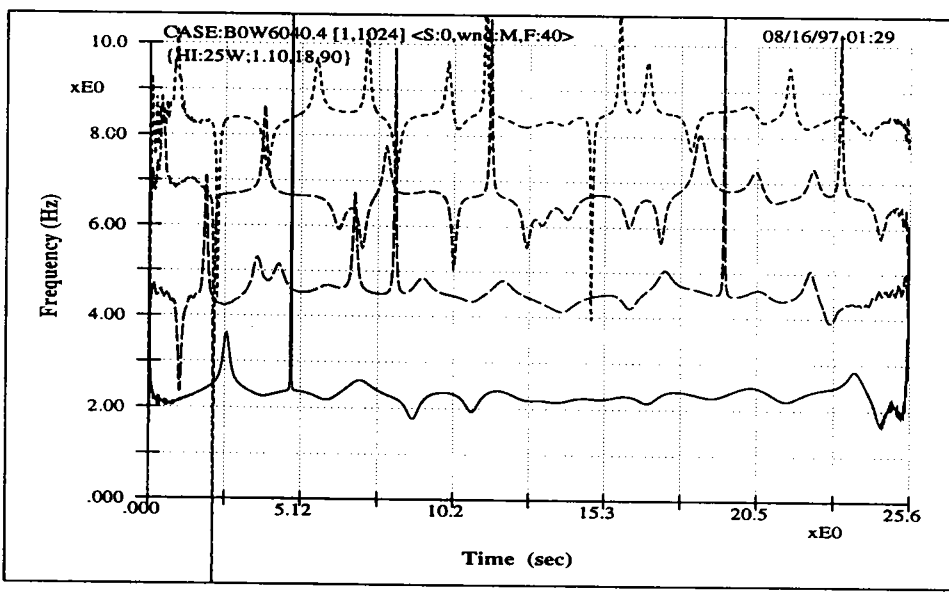


Figure 3.8: Instantaneous frequencies for the four different frequency bands with (bottom) and without (top) an amplitude window as shown in figures 3.5 and 3.6, respectively. The figure indicates that sharp variation of instantaneous frequency is mostly determined by local variation of amplitude envelope rather than its magnitude only.

Chapter 4

Conclusions

We first illustrate the problems of using a direct deconvolution process in making comparison of wind wave spectra for data acquired in an experimental water tank. The difficulties are referred as the repeatability problems of spectra and their main causes are attributed to the ambiguity effects due to transient phenomena and phase noise due to relative timing of events.

We then motivate the study of signals from modulation perspective using Gabor's analytic signal procedure. The known paradoxes as well as additional concerns or difficulties associated with this specific method are listed. Though we are not going to say that their causes are answered in satisfactory way; we do provide explanations and reasons for their existences through detail numerical modeling and thorough implementations of the procedure. A few intrinsic differences among different numerical schemes, such as those associated with discrete or continuous approach, those related to time domain or frequency domain processing, are also illustrated using practical data form experiments.

It is found that there exists profound interplay between local amplitude modulation and the instantaneous frequency, and the occurrences of singular behavior of the instantaneous frequency are associated with local irregularity of the amplitude function. That is to say, violent variations of instantaneous frequency are caused by local sharp variations of amplitude envelope. And since the occurrences of sharp variation of envelope distribution are associated with locations of small amplitude (we have not formally provided proof of this statement, but our numerical simulations as well as experimental wave data all show this tendency without exceptions), the sharp variations of instantaneous frequency always happen at points that are of little local energy. This point highlights the possible shortcomings for the analytical signal when applied to broad band processes. Besides, the analytic signal remains seriously entangled with the features of transient effects and phase noise as found in spectral analysis.

A few subject topics, mainly on the applications of the analytical signals to experimental wave data, such as characterizing the wave growth and decay time spans, frequency aliasing and base band conversion, surface roughness, bound wave system [8, 9, 19] and multiple Stokes waves, etc. will be worked on in subsequent studies.

Bibliography

[1] Bracewell, B. *The Fourier Transform And Its Applications*. McGraw-Hill Book Company, Singapore, second edition, 1986.

[2] Chui, C. K. *An Introduction to Wavelets*. Academic Press, Inc., San Diego, California, USA, 1992.

[3] Cohen, L. *Time-Frequency Analysis*. Prentice Hall PTR, Englewood Cliffs, New Jersey, USA, 1995.

[4] Daubechies, I. *Ten Lectures on Wavelets*. SIAM, Philadelphia, USA, 1992.

[5] Gabor, D. Theory of communication. *Journal of the IEE*, 93:429–457, 1978.

[6] Goda, Y, and Y., Suzuki. Estimation of incident and reflected waves in random sea experiments. In *Proc. 15th Conf. on Coastal Engineering*, pages 828–845, 1976.

[7] Greenberg, M.D. *Advanced Engineering Mathematics*. Prentice Hall PTR, Englewood Cliffs, New Jersey, USA, 1988.

[8] Lake, B.M., and H.C. Yuen. A note on some nonlinear water wave experiments and comparison of data with theory. *J. Fluid Mech.*, 83:75–81, 1977.

[9] Lake, B.M., and H.C. Yuen. A new model for nonlinear gravity waves, part 1, physical model and experimental evidence. *J. Fluid Mech.*, 88:33–62, 1978.

[10] Lee Y.R. and J. Wu. Continuous wavelet transform using a locally adapted time-frequency window. In *Proc. 18th Conf. on Coastal Engineering in Taiwan*, pages 95—106, 1996.

[11] Lee Y.R. and J. Wu. Wavelet and wavelet packet best basis for laboratory water waves. In *Proc. 18th Conf. On Coastal Engineering in Taiwan*, pages 83—94, 1996.

[12] Melville, W. K. Wave modulation and breakdown. *J. Fluid Mech.*, 128:489–506, 1983.

[13] Meyer, Y. *Wavelets and operators*. Cambridge University Press, New York, USA, 1992.

[14] Oppenhein, A.V. and R.W. Schfer. *Digital Sgnal Processing*. Prentice Hall PTR, Englewood Cliffs, New Jersey, USA, 1975.

[15] Paul, A.H., D. Xu, and J. Wu. Breaking of wind-generated waves: measurements and characteristics. *J. Fluid Mech.*, 202:177–200, 1989.

[16] Poon, Y.K., S. Tang, and J. Wu. Interaction between rain and wind waves. *J. Phys. Oceanogr.*, 22:977–987, 1992.

[17] Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flennery. *Numerical Recipes in Fortran*. Cambridge University Press, New York, USA, second edition, 1992.

[18] Soumekh, M. *Fourier Array Imaging*. Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA, 1994.

[19] Su, M.Y. Evolution of groups of gravity waves with moderate to high steepness. *Phys. Fluids*, 25:2167–2174, 1982.