水波分析最適化本基函數及其分析法

摘　要

基於試驗水槽多尺度現象之波、流場數據其共關頻振(coherence)之分析結果，我們鑑得相關於試驗水波應用之最適化本基函數(basis function)及其相應之分析方式。這些結果延伸並明確具體顯現先前兩篇研究(相關於最佳 Riesz 子波及調適化連續子波轉換)與此處共關頻振研究之關係。也因為這些數據分析結果之取得，使我們體會到這些原本都是從極其獨立思惟下著手的三個課題是有其一脈相承、緊密相扣、也始料未及的因果關係。基此，本文表述了此處子波共關頻分析之優越表現，並整理修飾先前文章，使其更清晰明確闡述此一一貫學理。簡而言之，此處涵蓋了：水波訊號於離散轉換下之最小熵值 Riesz 子波、相應於此子波之連續子波函基(function basis) 、暨相關於這些轉換本質特性的探討說明，如子波係數之多餘性(redundancy)與相對之 tight 或近乎 tight 的子波架框(wavelet frame)特性、子波係數樣木空間之大小、時頻窗(time-frequency windows)之可調性及其物理含義等。

The Search of the Most Appropriate Analyzing Function Basis for Laboratory Water Waves

Yueon-Ron Lee
Institute of Harbor and Marine Technology

July 19, 1998

Contents

List of Figures

List of Tables

Abstract

In two earlier studies we have basically identified the most appropriate function basis for uses in the studies of water wave related signals, specifically, those from wave tank experiments [17, 18]. However, the two papers are basically of numerical experiments and no conclusive evidences from practical applications were given. In this report, based on the results of the studies on the coherence features in the multi-scale wave and current fields in wave tank experiments, we provide definite evidences of the superiority of the identified function basis and its related analyzing scheme. The present report covers three main subject topics concerning: best Riesz basis for our signals, continuous wavelet transform using adapted time-frequency windows, and the coherences in wave and current fields – as are indicated by three individual chapter titles. The major links among the three topics are: the proximity between signal and the minimum entropy Riesz wavelet that is associated with discrete wavelet transform; the choice of a corresponding continuous wavelet transform basis that is closely related to the best Riesz basis; and the devising of a refined algorithm for analysis that help clarify the links and also reveal the factors that lead to the success in coherence investigations. We discuss a few key issues: different intrinsic properties of frame wavelet schemes and the continuous transform approaches; the extreme redundant basis functions associated with a continuous transform versus the tight or relatively tight bases associated with wavelet frames; the sizes of sample space of transform coefficients; effects of linear filtering; and the flexibility in adapting the time-frequency windows of a continuous wavelet transform and its implication of better physics. Due to further understanding of the intimate connections among the three subject topics, the contents of the two earlier papers are mostly incorporated, but ambiguities and abstract parts are clarified and restated. Finally, the identification of coherent features for various water wave related signals solidifies the serviceability of the function basis and the associated analyzing scheme.

Chapter 1

Introduction

In our minds, there is simply neither a full-fledged analyzing function basis nor a all-purpose numerical scheme for all sorts of signals or applications; the usefulness of a particular methodology is highly case dependent. To our knowledge, the appropriate and more realistic descriptions for water wave signals, even for those that look quite regular at first place, should base on modulation concept (cf. [4, 14, 15, 19, 20, 22, 23, 26, 30, 32, 35] as well as an example given in a later section on coherence). This is most obvious or particularly true for waves with short wavelength.

From the somewhat traditional spectral perspective to the more recent wavelet viewpoint, we have: Fourier transform, short time Fourier transform (STFT), Hilbert transform (or analytical signal approach), various time-frequency transforms associated with different distributions (such as Wigner Distribution, Page distribution, Choi-Williams distribution, and etc. [8]), as well as the discrete wavelet transform (DWT) and continuous wavelet transform (CWT) (sometimes CWT is referred as integral wavelet transform, but we prefer to use continuous wavelet transform as will be explained in the text). In this introduction chapter we will basically focus on the comparison among Fourier transform, short time Fourier transform, and DWT and CWT. A closely related topic on Hilbert transform in the current interest was studied in a previous report [16].

1.1 Non-stationary Effects

It is well known that Fourier transform is suitable for characterizing stationary signals and not quite satisfactory for studying transient local phenomena. The reasons can be illustrated by the following properties of the transform.

- Functions cannot be both time- and band-limited. If a function is limited (finitely

supported) in one domain, then the independent variable of its corresponding function in the other domain stretches the entire real line ($\mathbf{R}$). In real world situations, however, signals are almost always limited in time and space; meanwhile, hardware's capability is generally band-limited. This simply implies that there is not going to be a function basis that perfectly match theory to practice. A slight variation of Fourier transform is the STFT, which is just the Fourier transform of the windowed signal (signal capped or multiplied with a window). In STFT this property is indicated by the Balian-Low theorem, which basically states that if the window function $g(t)$ of a Gabor type frame

$$g_{m,n}(t) = e^{-2\pi imt} g(t-n), \tag{1.1}$$

where $m, n \in \mathbf{Z}$, is well localized in time, then the associated Fourier transform window can not be well localized in frequency. The point here sounds a bit abstract, but, in reality, it is the trait of the following intuitive points.

- The Gibbs phenomenon states that, if there is a jump in signal, then the overshoots, occurring at both sides of the discontinuity when the inverse Fourier transform is implemented, can never disappear and remain at constant. This is equivalent to say that it takes quite a many spectral components to make up a sharp transient feature and that a local variation affects a broad range of the spectrum just as the Fourier transform of the delta function (more precisely, delta distribution) covers the whole frequency axis.

- Fourier basis functions are periodic and extend bi-infinitely; thus, signals studied are better periodic and sampled infinitely. The unavoidable side effects for not fulfilling this are many: frequency leakage, smoothing errors, edge effects due to data truncation; aliasing due to under-sampling or non-periodicity (figure 1.1 is actually a case of under-sampling, where a linear chirp is two times under sampled, i.e., sampled at a rate twice slower than Nyquist rate); and spectral variance due to finite resolution.

1.2 Windowed transforms

Most of the syndromes listed in last section can be referred as (or is related to) non-stationary effects. Both short-time Fourier transform (STFT) and wavelet transform (WT) try to remedy Fourier basis's deficiencies in characterizing transient phenomena by analyzing the set of localized signals (such as formed by varying n in equation 1.1), and they both yield local spectral information (or more precisely, local scale information, if the term "frequency", "Hz", or "spectrum" is strictly reserved for sinusoidal functions). However, due to

the Balian-Low theorem mentioned above, the waveform associated with STFT can never be truly local in time since in reality the frequency domain of discrete Fourier transform is always band-limited by obeying the Nyquist law. Whereas, wavelets can be of exactly local, at least, they must have suitable decay and without zero frequency component.

Let outline a few properties of these two transforms:

- Both STFT and WT are windowed transforms. In STFT there exist two quite distinctive operations. The first operation is applying a suitable time-window to the signal; the second operation is performing the Fourier transform for the capped signal. The corresponding inverse transform (or reconstruction process) of STFT is naturally associated with a frequency-window and involves two similar distinctive operations too. However, in WT these two distinctive steps are not clearly observable — rather than using the very distinctive "window (either time- or frequency-window)" and "Fourier basis function (i.e., sine and cosine function)", the "window" and the "basis function" are synthesized in an inseparable specific form called "wavelet". In fact, one can clearly solidify this notion by comparing the Gabor type frame (equation 1.1) with the Morlet wavelet (equation 3.7 when the window function $g(t)$ of equation 1.1 in assumed to be a Gaussian bell and the second term in equation 3.7 is practically neglected due to its smallness. The intended purpose of either the combined operation or synthesized operation is completely the same: to provide a mechanism (or kernels) for projecting a signal into modulated or oscillating wave constituents.

- The time-frequency windows in STFT keep rigid for different scales since the window function $g(t)$ in equation 1.1 does not depend on m, i.e., their widths (usually referring to time) and heights (usually referring to frequency) do not change for all frequencies. In WT the windows are adaptive to different scales, but the sizes of different windows still keep fixed, i.e., each window's height and width are inversely proportional and the product remains constant (either for DWT or CWT). The concept of fixed size windows is illustrated by the phase planes shown in Figures 1.1 and 1.2. The windows show the results of the discrete wavelet packet transforms of a chirp signal using different bases associated with a wavelet packet. In the figures, since orthonormal bases are used, all time-frequency windows do not overlap. As to continuous wavelet transform various time-frequency windows severely tangle with each others, and we generally do not show the actual sizes and shapes of various windows — rather, each window is represented by a point (or a small area depicting the time-frequency resolution) having coordinates corresponding to its centroids in the time and frequency domains.

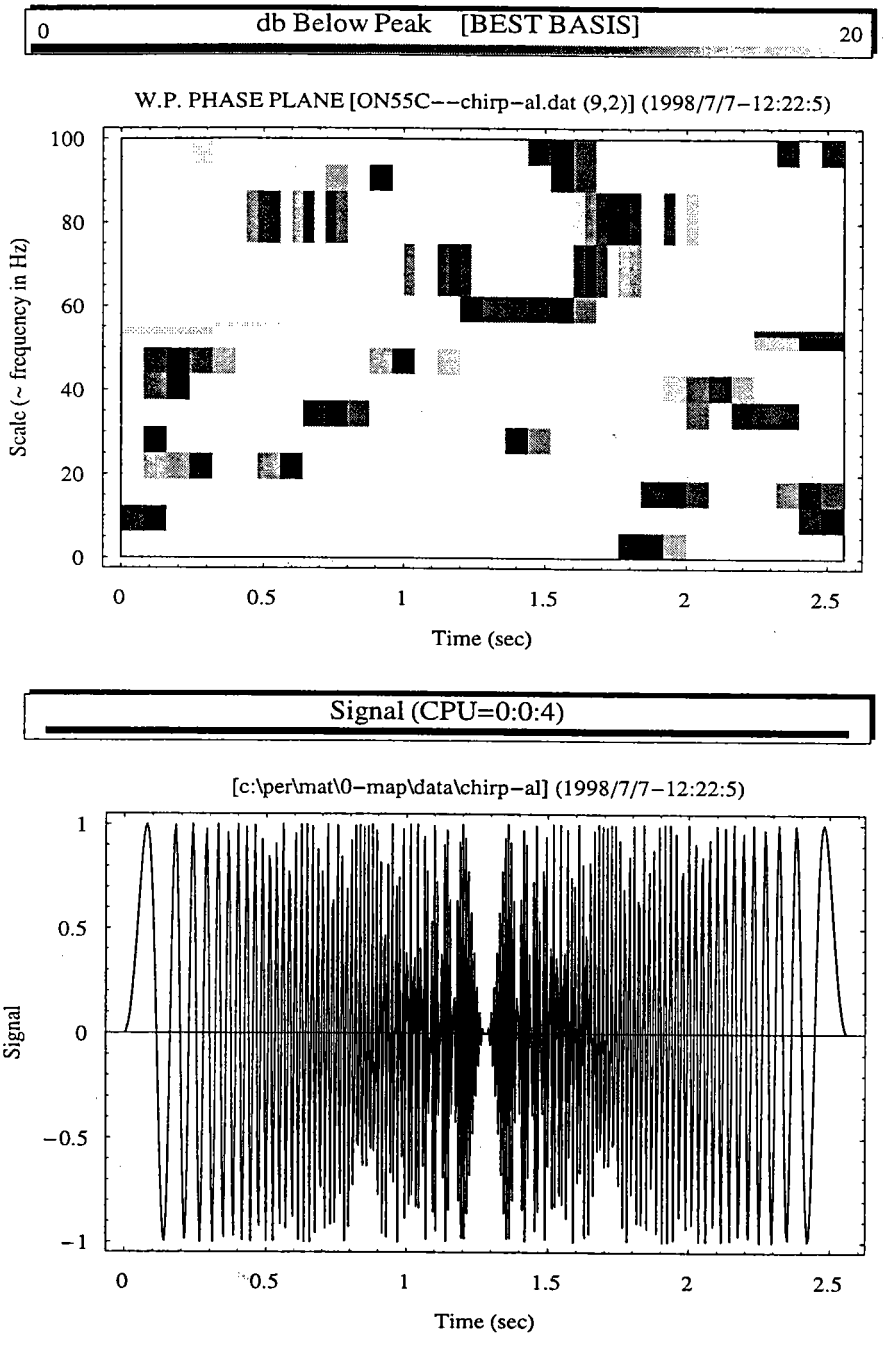

Figure 1.1: Phase plane of a wavelet packet best basis time-frequency windows (Top) using a linear chirp signal with aliasing (Bottom). Here wavelet packets associated with coiflet of 30 convolution weights is used. The original signal, if not under-sampled, should have linear instantaneous frequency distribution form 0 to 100 Hz. Note the non-symmetric effects due to nonlinear phase filtering and the scattering of windows due to composite frequency bands that form the wavelet.

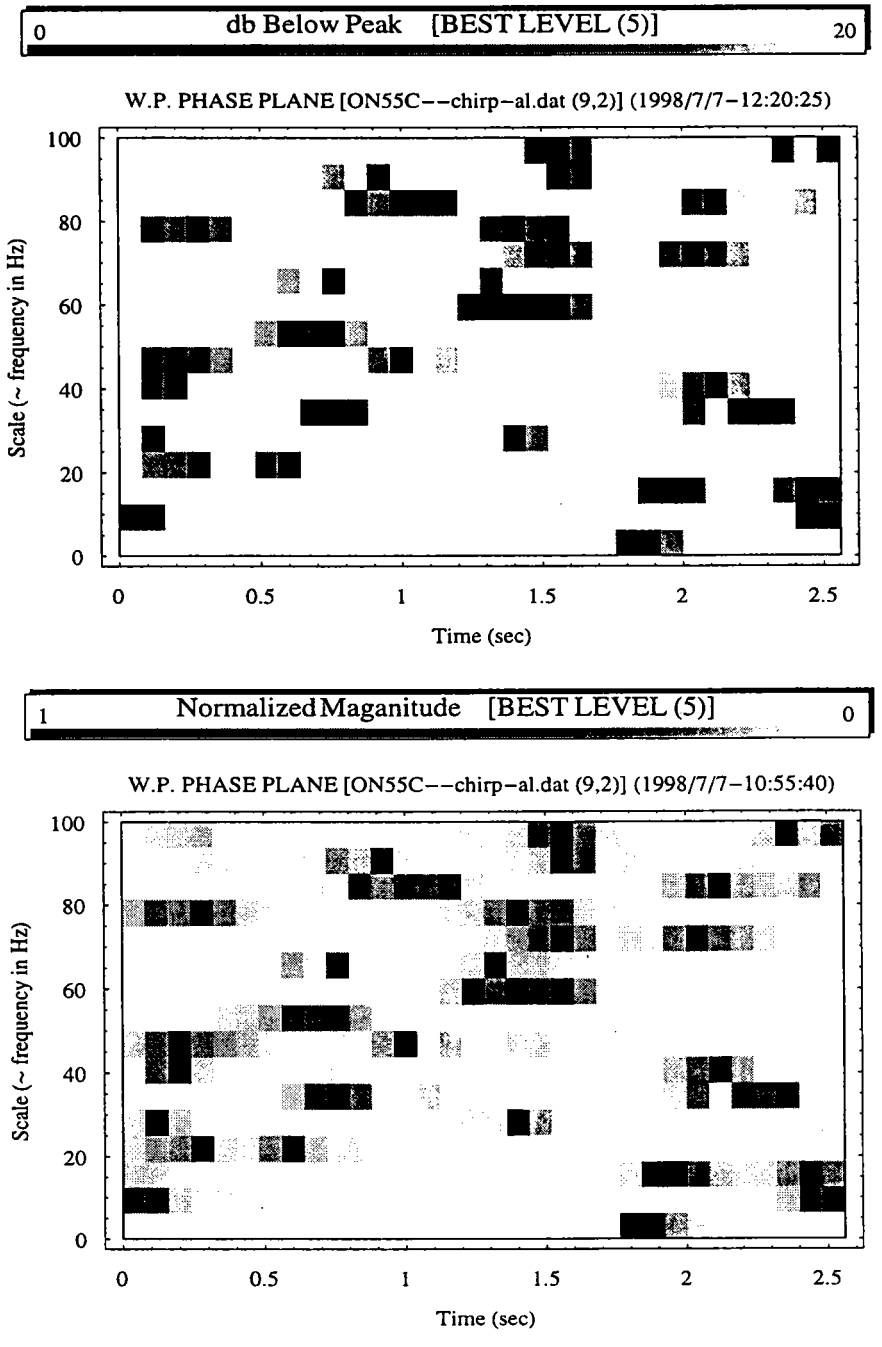

Figure 1.2: Phase planes (Top: logarithmic measure; Bottom: linear measure) of wavelet packet best level time-frequency windows using the same linear chirp and wavelet packets as in last figure. In view of the fact that a single orthonormal mother wavelet can yield many different wavelet packet representations, that there are basically infinitely many wavelet bases, as well as that we may use different graphic renderings, we are certainly in the dilemmas of choosing an appropriate basis.

- The function basis of STFT is the unique orthonormal Fourier basis comprised of sine and cosine functions; whereas, for WT, apart from the very loose constrain that the basis function (or the mother wavelet) satisfies the admissibility condition (for CWT) or stability condition (for DWT), there is virtually no restriction on the choice of basis functions. The Fourier coefficients of STFT represent local spectral information, which still have the precise meaning of "frequency". In WT a wavelet coefficient refers to a specific scale rather than a specific "frequency". And we generally suffer from its physical interpretability due to the following reasons: (1) No unique basis. That is to say, the analyzing function or mother wavelet can be designed in a plenty of ways, and the basis functions related to the mother wavelet can be either dependent or independent (orthogonal or non-orthogonal). (2) Scale does not have unit. This, together with point (1), severely hampers out ability to directly perceive its size and physical shape; and, (3) No fixed algorithm to implement wavelet transform. Many tricks and various adaptations exist: such as, the present adaptation using flexible time-frequency windows for CWT (to be studied in a later chapter), multi-voice [12] or multi-wavelet [11, 10, 33] frames, and discrete wavelet transform using different dilation factors other than the most often seen value 2 [1]. It is noted that all these various techniques may not be so disturbing in certain fields (such as fields in data transmission or signal decomposition and reconstruction) as they may in our studies focusing on physics.

- In general, the dilation lattice is in logarithmic measure for DWT (e.g., the $a_0{}^j$ in the stability condition to be mentioned) and in linear measure for discrete STFT (e.g., the $e^{-i2\pi mt}$ in the above mentioned Gabor type frame). Continuous transforms do not involve lattice. The concept of lattice is associated with the the concept of time-frequency density (which is defined as the inverse of the product of discrete steps of dilation and translation) [12]. For STFT frames, due to the existence of Nyquist frequency, the time-frequency density must not go beyond the value of generalized Nyquist density, $(2\pi)^{-1}$. For wavelet transform, however, there is no such a clear-cut limit of time-frequency density [12]. Moreover, Balian-Low theorem depicts thatthere is no good time-frequency localization for a STFT frame if constructed under a strict time-frequency lattice; on the contrary, numerous wavelet bases with good time-frequency localization have been given [6, 12, 27]. These physically imply that WT may provide better zoom-in. The existence of a lattice structure may be convenient as well as inconvenient. For water waves, since we don't expect any significant gaps in the scale contents in any reasonable physical process, besides, an interaction process should always involve scopes of time and spatial scales that are "changing"

or "evolving" at least in a relatively continuous sense, we will see that we generally do not appreciate the use of frames.

Both CWT and DWT implement a process of integral wavelet transform over the real line **R** in a continuous sense. Maybe, we can say that the difference is due to the use of different integration symbols: $\sum$ and $\int$. Digitally sampled signals are certainly discrete, but this "discrete" is not relevant to the methodology of CWT or DWT. The main difference, from application point of view, is that there is no practical interest of reconstruction (or inverse transform) for CWT due to the redundant or non-orthogonal nature of its wavlet coefficients. Both methods are capable of decomposing either functions defined over the real line or signals sampled discretely. In reality, applying CWT to sampled data is implemented in discrete manner; vis-à-vis, doing DWT for an unlimited ladder, such as that of the standard multiresolution analysis of Meyer and Mallat [24], is able to describe any function in infinite detail, i.e., over the whole real line. In fact some of the formulas in DWT are quite helpful in the explanation of our present interests by providing concepts for generalization in the limit sense. The concept of unlimited ladder of DWT is illustrated by examples of the blow-ups of a segment of wavelet curve as shown in Figure 3.1, in which we also illustrate the possible bizarre behaviors of certain wavelets.

We note that we restrict our scope to $L^2(\mathbf{R})$ Banach space for all the contents in the present report, i.e., the Hilbert space, since some of the statements may not apply to other function spaces or (sub)classes [12, 27]. Nevertheless, most of the intricacies that are different for different spaces are only of analytic interest (e.g., on the existence of multi resolution analysis (MRA), on the regularity and differentiability of wavelets and the associated scaling functions). For practical applications it is far enough to restrict to the Hilbert space, i.e., a space of functions with finite energy contents.

1.3 The motivation

The most intuitive footing to use localized transforms in our applications can be stated as: if we perceive our signal as composed of waves which are limited in life span, i.e., constituent waves are evolving with time and in space, then it is natural to adopt wavelet as our analyzing function; furthermore, in addition to this modulation nature, if we also acknowledge that intrinsic instability due only to nonliner surface boundary condition is everywhere to be found for even regular water waves, then it is still quite possible that wavelet decomposition can provide better descriptions of physics for stationary signals than what can be provided by Fourier decomposition. Besides, one clear advantage of using wavelets is the possible flexibility in constructing their wave forms.

In fact the three major subject topics – the titles of the next three chapters – recognized initially only these notions and set out quite independently; the apprehensions of their links are gained somewhat by chance and, of course, by going through more readings. We hope that this report well serves the chore: explanations of reasons and outcomes concerning the search of a better analyzing function basis for our water wave related signals.

Chapter 2

The Best Riesz Wavelet for Laboratory Water Waves

In any experimental setup it is almost always true that various modeling laws (or scaling laws) can at best be partially satisfied. The situation is further complicated in a diversified multi-scale system. In the current wave tank experiments for a wind, wave, and rain coupling system significant distortions are also acknowledged since limitation in space in experimental setup imposes severe restrictions upon the developments of the interaction process. Results of earlier researches as well as our experiences indicate that the identification of the interaction features requires the deployment of optimized analyzing schemes using sensitive and appropriate basis functions . The present chapter serves to provide initial justification of the method that will be used to analyze the data acquired. More specifically speaking, we are going to identify the most appropriate basis among Riesz wavelet bases for our signals. The wavelets of this basis are the DWT counterparts for the analyzing basis functions used in CWT and have maximum similarities to our signals

In fact, if we had not found conclusive results from applying the analyzing function basis to our experimental data and also had not seen the obvious superiority of wavelet coherences in contrast to spectral coherences, the decency of the topic of this chapter would not be fully appreciated and the factors leading to the usefulness of the basis functions would not become as ostensible as they should.

2.1 Introduction

The powerfulness of wavelet analysis comes from its flexibility in devising the analyzing wavelets as well as its adaptability in forging the algorithms. However, versatility does not

come without ambiguity. For example, the power spectrum of a function is shift-invariant; whereas, wavelet spectra are highly shift-variant [12, 25]. Figure 2.1 shows such a property. Moreover, in the investigations of wavelets, It seems that the bearing of wavelets' physical implications is not in proportion to that of their analytic interests. In addition, intrinsic properties of different signals in different applications are quite different and individual objectives are also rather specific; therefore, the bonding between wavelets and signal analysis is more established in some fields than the others.

In studying the physics of certain phenomena using wavelets one of the most intriguing questions is how to choose the analyzing wavelet(s). The concern here is quite in contrast to a few studies where they are mainly numerically or analytically oriented. For example, in coding of images and acoustic signals the goals are simple: maximum compression, minimum overhead of handling, as well as highest effectiveness and lowest distortion; in his situation mathematical relevance between signal and wavelet may be materialized in higher degree than physical pertinence needs to be unfolded [9, 10, 12].

For our interests in characterizing the physics of water wave related phenomena, it seems, at first, that the aspiration is certainly not on "efficiency" or "compactness". However, with the understanding that compactness of a coding means closeness between signal component(s) and analyzing function(s) and that wave forms which are not resemble to our signals (or signal components) are unlikely to yield any significant information on physics; it is, therefore, justified for us to find the wavelets that provide the most efficient or most economical representations for out signals. This viewpoint is related to the concept of entropy.

The works here are mainly numerical experiments on measuring the "distances" between our signals and various Riesz wavelet bases given in several wavelet treatises [6, 12, 27, 29]. No attempt to make new constructions of bases or to extend the existing constructions is made (for each wavelet sub-group mentioned below different mother wavelets with different support lengths can be constructed using the same algorithm for that subgroup, but we will come to realize that this is actually not necessary when the associated two-scale scaling function or father wavelet is not changed.) Nevertheless, we have tried to include various categories of Riesz wavelets that we are aware of. These wavelets are dyadic wavelets with "mathematical sampling rate" 1 (no unit). They are of most practical interests in applications for discretely sampled signals. Furthermore we restrict our scope to laboratory water waves. The criterion used is the entropy statistics of discrete transform coefficients.

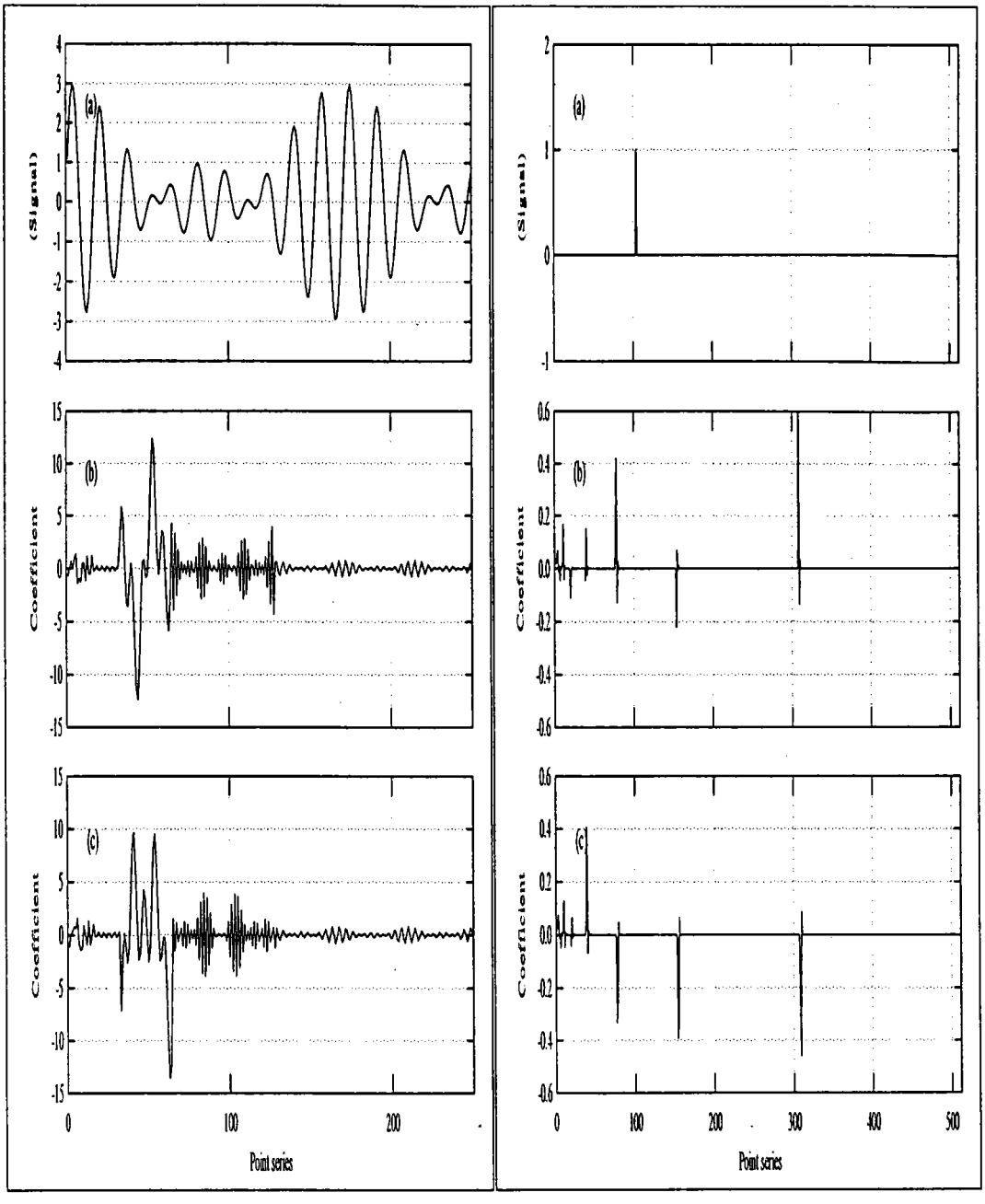

Figure 2.1: The shift non-invariant property of two wavelet transforms. Each top sub-figure shows individual signal. The middle shows the wavelet coefficients. The bottom shows the wavelet coefficients for the shifted signal (Right figure: 20 points to the left (using BO22D); Left figure: 3 points to the left (using ON33A)). It is also noted that, even though Fourier power spectrum is shift-invariant, Fourier spectral coefficients (without the second power) is still shift-variant. This property is linked to the poor performance of coherences associated with orthonormal cases as to be explained in a later chapter.

2.2 Wavelet bases tested

The Riesz wavelet bases used here can basically be divided into four categories: orthonormal (ON), semi-orthogonal (SO), bi-orthogonal (BO), and orthonormal wavelet packets bases. Within the first orthonormal category there are several different groups: Daubechies wavelets (both most asymmetric and least asymmetric), Coiflets, Meyer wavelet, and Battle-Lemarié wavelets.

No effort is made to give detail accounts of all the wavelets used here; only main properties will be briefed. But first, let us state the related notations and conventions. Let a function or a signal be denoted by $f(t)$; the two-scale scaling function of a Riesz basis be $\phi(t)$; the associate mother wavelet be $\psi(t)$ and its Riesz wavelets be $\psi_{j,k}(t) = \sqrt{2^j}\psi(2^j t - k)$, where $j, k \in \mathbf{Z}$, i.e., the space V_j (formed by $\psi_{j,k}$, $k \in \mathbf{Z}$ for a given j) in the multiresolution ladder are nested in $\cdots \subset V_{-1} \subset V_0 \subset V_1 \cdots$, and the finest and the coarsest scale space, say, for a 1024-point signal, are V_{10} and V_0, respectively; the number of filter coefficients or the number of convolution weights be N if the associated wavelet is finitely supported (support length equals $N - 1$); the dual wavelet and dual scaling function, if exist, be $\widetilde{\psi}(t)$ and $\widetilde{\phi}(t)$; the inner product be $\langle \cdot, \cdot \rangle$; and the Kronecker delta be $\delta_{j,k}$, $j, k \in \mathbf{Z}$, which is equal to 0 for $j \neq k$ and 1 for $j = k$.

Up until now, all practical wavelets of discrete transform are associated with the theory of multiresolution analysis (MRA) [24]. And when talking about Riesz wavelets there always exist dual wavelets except orthonormal wavelets which are self-dual. Any discrete wavelet transform involves two convolution operations: one yields detail information; another yields smooth information [29]. Convolutions can either be implemented in direct way in time domain for compactly supported wavelets or in indirect way in frequency domain. We list the basic properties (restricted to real-valued wavelets) and give the symbols of representation for each category and subgroup as follows.

1. Orthonormal wavelets (denoted as ONxxS, ONxxA, ONxxC, Meyer, and B&L, where x is an integer related to support length):

$$\psi = \widetilde{\psi},$$

$$\phi = \widetilde{\phi},$$

$$\langle \psi_{j,k}, \widetilde{\psi}_{\ell,m} \rangle = \delta_{j,\ell}\delta_{k,m},$$

$$f(t) = \sum_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k},$$

One MRA ladder (single set of frame bounds),

One filter pair (one smooth and one detail).

Within this category we have

- Daubechies most compactly supported wavelets (ONxxA): The wavelets in this group have maximum number of vanishing moments for given compatible support width. Or stated otherwise, they are the most compactly supported wavelets for given compatible number of vanishing moments. The famous most compactly supported continuous wavelet belongs to this group and has only four filter coefficients. These wavelets are quite asymmetry (so, the "A" in ONxxA). The vanishing moments and the number of filter coefficients are, respectively,

$$\int_{-\infty}^{\infty} t^{\ell}\psi(t)dt = 0, \qquad \ell = 0, 1, \cdots, x,$$

$$N = 2x,$$

  where x is the integer number in ONxxA. The minimum number of x is 2.

- Daubechies least asymmetric wavelets (ONxxS): For a given support width, these wavelets, in contrast to those of the ONxxA subgroup, are the most symmetric ones (so, the "S" in ONxxS). They have the same representations of vanishing moments and number of filter coefficients as those of ONxxA. But the known minimum number of x is 4.

- Coiflets (ONxxC): The Coiflets have vanishing moments for both ψ and ϕ; therefore, from Taylor expansion point of views [12], they have high compressibility for fine detail information (i.e., a great portion of the fine scale wavelet coefficients are relatively small); and henceforth, they have simple quadrature rule to calculate the fine smooth information (i.e., the calculation of the inner product of a function and the fine-scale scaling functions is more efficient). Since every discrete wavelet transform involves both smoothing and detailing operations, there may exist some advantages from these two properties for certain applications, such as those applications that do not stress lossless of signal contents [11, 34]. Now the vanishing moments and number of filter coefficients are

$$\int_{-\infty}^{\infty} t^{\ell}\psi(t)dt = 0, \qquad \ell = 0, 1, \cdots, x,$$

$$\int_{-\infty}^{\infty} \phi(t)dt = 1,$$

$$\int_{-\infty}^{\infty} t^{\ell}\phi(t)dt = 0, \qquad \ell = 1, \cdots, x,$$

$$N = 6x.$$

- Meyer wavelet (Meyer): The Meyer wavelet is one of the first constructed orthonormal wavelets and is the wavelet with most compact support in frequency domain (if without any assignment "finitely supported" refers to time domain). Therefore, due to contrast properties between the two Fourier domains stated in the introduction chapter, the wavelet is infinitely differentiable in time domain, i.e., has an infinite Lipschitz regularity C^{∞} and does not have exponential decay. The support length $N \to \infty$.
- Battle and Lemarié wavelet (B&L): The Battle and Lemarié wavelet of m^{th} order is constructed from the orthonormal scaling function derived by applying the standard orthonormalization trick to the m^{th} order cardinal B-spline N_m [2, 6]. For $m = 1$, it is exactly the Haar wavelet. The Haar wavelet is the only finitely supported wavelet in this group (also a case of BO11O=BO11D to be mentioned below) and is the most compactly supported discontinuous wavelet; all other wavelets in this group are infinitely supported. They all have exponential decay and possess C^{m-2} regularity. Figure 2.2 shows a comparison of Meyer wavelet and the Battle-Lemarié wavelet of fourth order. They look quite the same even though their constructions, or derivations, or formula involved are completely different.

2. Semi-orthogonal wavelets (SOxO and SOxD):

$$\psi \neq \widetilde{\psi},$$

$$\phi = \widetilde{\phi},$$

$$\langle \psi_{j,k}, \psi_{\ell,m} \rangle = \langle \widetilde{\psi}_{j,k}, \widetilde{\psi}_{\ell,m} \rangle = \delta_{j,\ell},$$

$$f(t) = \sum_{j,k} \langle f, \psi_{j,k} \rangle \widetilde{\psi}_{j,k} = \sum_{j,k} \langle f, \widetilde{\psi}_{j,k} \rangle \psi_{j,k},$$

One MRA ladder ,

Two filter pairs ,

$$N = 3x - 1 \quad \text{for SO}x\text{D},$$

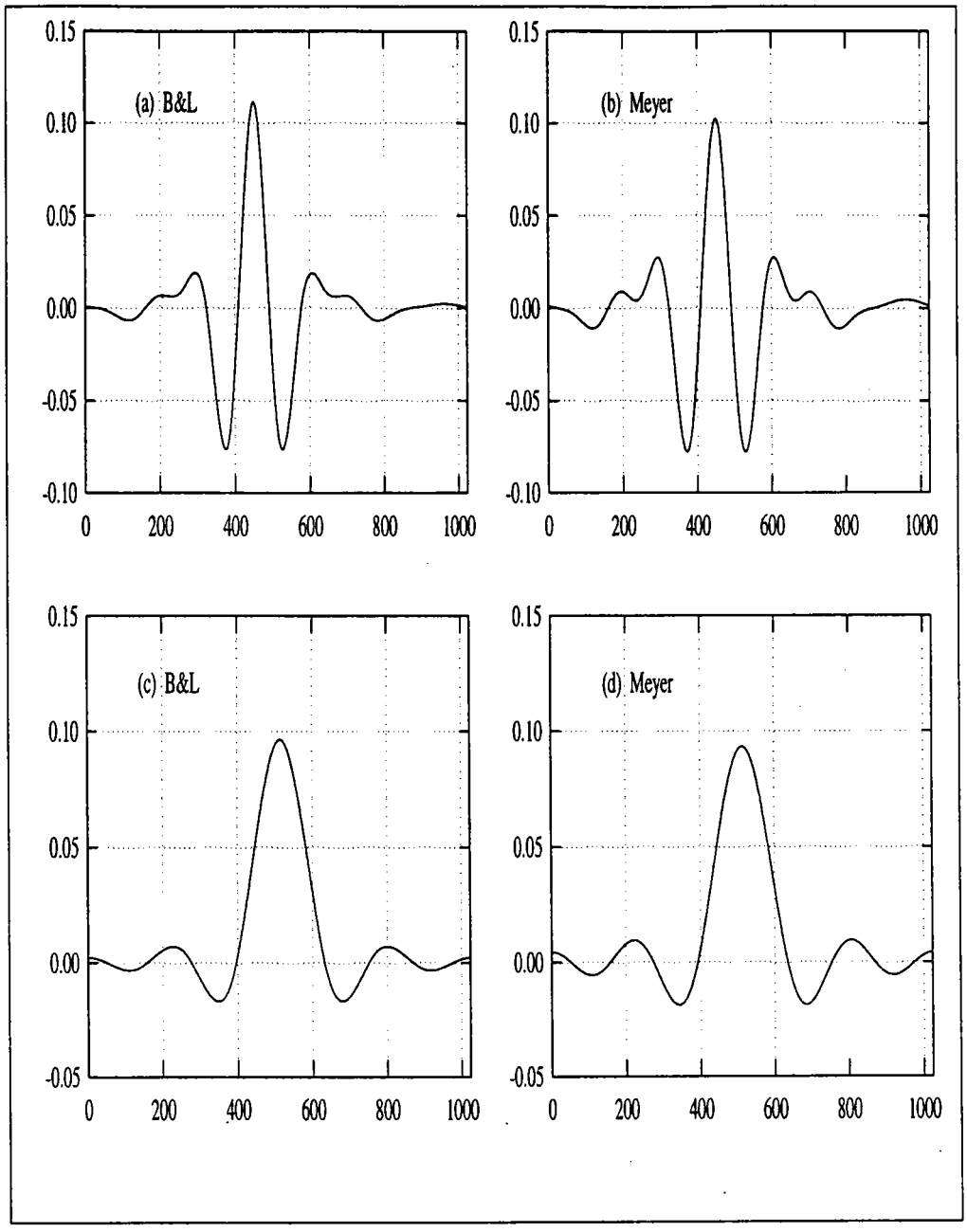

Figure 2.2: Two wavelets with similar looks but of quite different constructions and analytic properties (such as regularity, differentiability, rate of decay, support length, etc.) Many complicate aspects of discrete Riesz wavelet seem not to reflect their associations with practical concerns. The Battle and Lemarié wavelet and the Meyer wavelets are shown in(a) and (b), respectively; their corresponding scaling functions are shown in (c) and (d).

$$N \to \infty \quad \text{for SO}x\text{O}.$$

The semi-orthogonal wavelets are inter-scale, but not inner-scale, orthogonal. Their scaling functions are cardinal B-spline N_m. Although there are two distinctive (independent) filter pairs (one for decomposition one for reconstruction), there is only one MRA V_j-ladder. It was shown by Chui [6, 7] that cardinal B-spline wavelet of order higher than $m = 3$ is almost a modulated Gaussian (however, a modulated Gaussian is not a wavelet). We therefore use only B-wavelet of fourth order (i.e., Cubic B-spline wavelet where $m = 4$).

3. Bi-orthogonal wavelets (BOxyO and BOxyD):

$$\psi \neq \widetilde{\psi},$$

$$\phi \neq \widetilde{\phi},$$

$$\langle \psi_{j,k}, \widetilde{\psi}_{\ell,m} \rangle = \langle \phi_{j,k}, \widetilde{\phi}_{\ell,m} \rangle = \delta_{j,\ell}\delta_{k,m},$$

$$f(t) = \sum_{j,k} \langle f, \psi_{j,k} \rangle \widetilde{\psi}_{j,k} = \sum_{j,k} \langle f, \widetilde{\psi}_{j,k} \rangle \psi_{j,k},$$

Two MRA ladders,

One filter pair,

$$N = 2y + x - 1 \quad \text{for BO}x\text{y}\text{O and } x \text{ odd},$$

$$N = 2y + x - 2 \quad \text{for BO}x\text{y}\text{O and } x \text{ even},$$

$$N = 2y + x - 1 \quad \text{for BO}x\text{y}\text{D and } y \text{ odd},$$

$$N = 2y + x - 2 \quad \text{for BO}x\text{y}\text{D and } y \text{ even}.$$

The wavelets in this category are also constructed by Daubechies and sometimes called non-orthogonal wavelets. As is well known, no real-valued orthonormal compactly supported wavelet, except the Haar wavelet, has symmetric properties. However, from point of view of reconstructing a signal from its partially truncated wavelet coefficients, symmetry is a desired property of the filter to yield a more natural or smooth perception. The symmetry can be achieved by giving up orthogonality; if this is the case one has dual pairs for both wavelets and scaling functions. It is obvious that the semi-orthogonal case is more general than orthogonal one, and the bi-orthogonal case is even more general, as is clearly indicated by the additional freedom of dual

scaling function, reflected in the x and y in the notation BOxyO and BOxyD. Nevertheless, the wavelets in this category need only one pair of independent filters for decomposition and reconstruction even though they have two different MRA ladders with different sets of Riesz bounds. This is quite opposite to semi-orthogonal wavelets where they have one MRA ladder but with two filter pairs.

4. Wavelet packets: The wavelet coefficients derived form an orthonormal wavelet decomposition can be further decomposed by using either the same set of filter coefficients (called two-scale sequence in Chui [6]) associated with that original wavelet or by using another set of filter coefficients from different orthonormal wavelets. So, basically there can be infinitely many wavelet packet decompositions. The decompositions are of a tree-like refinement process and is called wavelet packet transform. The wavelet packet coefficients generally give better frequency resolutions with longer time supports. There are no simple formulas to describe this tree decomposition, but a schematic plot is convenient for elucidating the mechanism. Figure 2.3 shows this tree-like decomposition and gives some explanations. Due to its tree-like processing the execution time is dramatically increased. Within the length of the coefficients the number of branches and branch patterns can be chosen in any way so long as there is no overlap within any column. That is to say, any column, wide or narrow, must have one and only one contribution from all levels (rows).

   In this category we have two standards for selecting our best basis. One is just called by the same name "best basis"; another called "best level basis". For a 1024-point signal the finest level occurs at $j = \log_2 1024 = 10$ and there are 2^{10} different choices of bases. Within these 2^{10} choices the one which yields the minimum entropy is called "best basis". And if we enforce the restriction that all wavelet packets be at the same level j, then we have 10 levels (0 to 9) to choose from, and the one that yields minimum entropy is called best level basis. The indexes of the wavelet packet coefficient, including both subscript and superscript of U in the figure, i.e., its location within a specific level, determine the location (time of occurrence or support length and frequency band) of the coefficient's time-frequency window in the phase plane. Concepts regarding this wavelet packet transform can be seen in figure 1.1. The characteristics is that areas of individual windows are all equal.

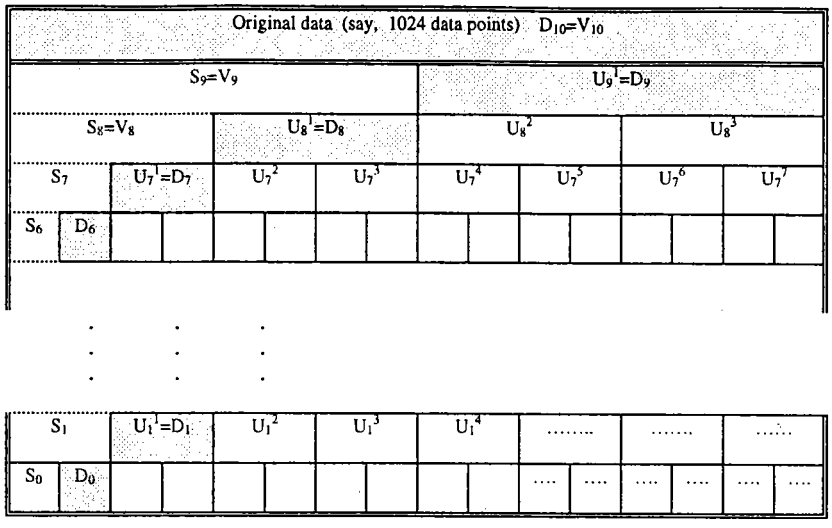

Figure 2.3: Schematic representation of the tree structures of the wavelet packet decomposition. S(=V in the text) and D stand for smooth and detail information, respectively. U with superscript larger than 1 stands for further decomposition of D by wavelet packets. All subscripts mean scale levels. All superscripts mean relative locations of the frequency bands for compatible subscript.

2.3 The entropy criterion

Entropy is a terminology of statistical physics. Since it is of statistics, it gives indication without assurance. The entropy can be viewed as a measure of the "distance" between a signal and its reconstructed one using partially truncated transform coefficients. We shall give straightforward descriptions rather than somewhat mystified notions as given in some of the readings. We therefore go through the actual procedures first and give the descriptions of its statistical meanings next. Let suppose that we have a 1024-point sampled data, then there is a set of 1024 wavelet coefficients (C=$\{c_i\}$). We take the absolute or squared value of these coefficients, sort them, and then divide the sequence into M (say, 100 or 200 or 300) divisions which are equally spaced from 0 to the maximum value of the coefficients. Then we have the statistics of occurrence for each division, and the distribution of these normalized occurrences is the probability density distribution or probability density function (denoted by pdf), say $\{p_1, p_2, \cdots, p_{M-1}, p_M\}$. The entropy is

$$H(p) = -\sum_i p_i \log p_i. \tag{2.1}$$

Where, when $p_i = 0$, it is assumed that $0 \log 0 = 0$, since in reality one can assumed that there exists an almost zero probability in that interval without affecting the total sum of probability practically, after all it is only a statistics and the modification virtually has no effect on the norm value. If absolute values of c_i are taken, $H(p)$ is called L^1-norm entropy; if squared value are taken, it is squared L^2-norm entropy. Of course another power can be used, but the squared L^2-norm, being the energy, is physically the most significant. The physical implication of this definition of entropy is described as follows: let suppose two probability distribution functions sorted in decreasing order are p and q, if p decreases faster than q, then $H(p) \leq H(q)$ [34]. Though this statement is only one-way correct (the reverse is not always true), it means that smaller entropy indicates that more energy is concentrated within a smaller number of wavelet coefficients. And hence, if only a certain number of terms is kept, the truncated error, i.e., the distance, will likely be smaller for sequence with smaller entropy.

There is another notion (sometimes referred to as the geometric notion [34]) for calculating the entropy. Again, we describe the procedures first and give the simple physical interpretation later. By setting the number of divisions to be the same as the number of coefficients and by defining probability density to be the normalized (with respect to the total power) value of the squared wavelet coefficient, that is to say, the total energy $\|C\|^2 = \sum_i |c_i|^2$ and the probability density $p_i = |c_i|^2 / \|C\|^2$. Substituting into equation 2.1 one has the entropy as

$$H(p) = \log \|C\|^2 - \frac{\sum_i |c_i|^2 \log |c_i|^2}{\|C\|^2}. \tag{2.2}$$

The notion here is simple: if one just put more weight on coefficients of small energy and less weight on coefficients of large energy (all coefficients being normalized), then the weighted energy is an indication of entropy. And since taking the log of a value is sort of a weighting operation and since the total energy is finite, small entropy therefore means that the number of significant coefficients is small, or stated otherwise, more energy is concentrated in fewer coefficients.

One equivalent indicator to entropy of a pdf is the theoretical dimension $D(p)$ and is defined as [34]

$$D(p) = e^{H(p)} = \prod_i \left(p_i^{-p_i}\right). \tag{2.3}$$

As was mentioned, entropy does not yield how conclusive the result is. But our numerical results show little ambiguity regarding the one with minimum entropy.

2.4 Results and discussions

To increase the definiteness in the comparisons, we calculate entropy based on several setups: direct coefficient entropy related to L^2-norm based on equation 2.3 (column 1 in tables 2.1 and 2.2), pdf entropy related to L^2-norm with 300 (column 2) and 200 (column 4) divisions, and pdf entropy related to L^1-norm (column 3), based on equation 2.1. Theoretical dimension for one of the setups is also given (column 5). The Tables show the results using a wind wave signal in an oval tank. It is noted that if peak frequency (or primary scale) of other signal is significantly different, then, to be consistent in comparison, the analyzed signal lengths and the sampling rates should be properly adjusted according to its peak frequency. This is because in discrete wavelet transform we need to keep track of the actual physical size of translation so as to have physical perception of the wave forms. Table 2.1 give results from all orthonormal wavelets (including B&L, Meyer, ONxxA, ONxxS, and ONxxC), semi-orthogonal wavelets (Cubic B-spline, SO3O and SO3D), as well as spectral results. Table 2.2 give results from bi-orthogonal wavelets. Many distinctive features can be derived from the tables.

- Where there are duals, the dual wavelets always give much smaller entropy than as given by their corresponding wavelets. This certainly verifies that, for our water wave signals, using

$$f(t) = \sum_{j,k} \langle f, \widetilde{\psi}_{j,k} \rangle \psi_{j,k} \tag{2.4}$$

  provides much better efficiency in decomposition and reconstruction than using

$$f(t) = \sum_{j,k} \langle f, \psi_{j,k} \rangle \widetilde{\psi}_{j,k}. \tag{2.5}$$

  This also points out that dual wavelets rather than their counterpart wavelets should always be used in the analysis for yielding better physics and for possible computational efficiency. We note that a check of the shapes of all the listed bi-orthogonal wavelets clearly shows that these wavelets with small x and y are quite impractical (such as those shown in figure 3.1). Furthermore, for these bi-orthogonal wavelets, we doubt that there is going to be any significant improvement by further extending the support width related to y without extending support width related to x. Since, according to Daubechies [12], increasing the width (y) from some point on, gives no effect on the shapes of the dual wavelets (such as $y = 7$ or 9 for $x = 3$), and since it is the dual wavelet not the counterpart wavelet that provides better approximation.

- Entropy values of all groups of orthonormal wavelets do not fall to those of other

wavelets. Besides, difference in entropy values of long and short supports can barely be seen with only very slight indication that entropy related to longer support perhaps being smaller.

- For all orthonormal wavelets, the Meyer wavelet is infinitely smooth, the B&L is second order differentiable, and the others have various degrees of differentiability or regularity [12]; but, there is not much difference in entropy among all these orthonormal wavelets. It is therefore concluded that regularity or differentiability may be only of analytical interests for currently available orthonormal wavelets; since non of the orthonormal wavelets distinguish itself from various categories, and since no tendency within any subgroup is clearly observable.

- The most striking feature is that the dual Cubic B-spline wavelet yields obviously a far smaller entropy, even slightly better than that from spectral coefficients. A comparison of the cumulative probability distribution curves for several wavelets, including that of spectral entropy, is given in figure 2.4. The striking feature is reflected by the extreme flatness of the SO3D curve, nearly horizontal up until 90 percent of energy ratio. At about 96 percent energy ratio there is a crossing between spectral curve and the SO3D curve. Figure 2.5 shows the reconstructions of a section of a signal from its spectral and SO3D wavelet coefficients by keeping 35 percent of the coefficients. It is seen that the wavelet provides a better description of the details. The reasons for SO3D's strong performance might be associated with the following characters: total positivity of the scaling function, i.e., no oscillation or zero crossing of scaling function; and, complete oscillation of the wavelet, i.e., no unnecessary oscillation – no oscillation without zero crossing. And this, in other words, practically implies that our laboratory water waves are far less transient when compared with orthonormal or bi-orthogonal wavelets, and also indicates that the life span description of the waves is more likely to be up to the physics.

- We mentioned that we also test wavelet packet category for both best basis and best level. We probably have got an idea that chance is small for getting a better basis. The reason should be due to the inherent limitation of wavelet packet transform since they must use orthonormal wavelets as their primitive analyzing functions. And since orthonormal wavelets do not have good performances as stated above, it is therefore hard to anticipate the same strong performance as that from semi-orthogonal wavelets; nevertheless, wavelet packets do show improvements. Figure 2.4(b) gives the wavelet packet best bases using B&L and Meyer's wavelets, they show improvements when compared with figure 2.4(a), but certainly not to the degree of semi-

orthogonal wavelet or that of the spectral curve.

- Figure 2.6 shows cumulative distribution curves of wavelet packet coefficients at various levels as well as best level, best basis, and also the corresponding orthonormal wavelet basis (all using ON77S). Again wavelet packet best basis and best level yield smaller entropy than that of counterpart wavelet basis, but still their curves are far away from that of SO3D. And the curve for the best level comes quite close to that for the best basis.

- Up until now we haven't seen clear differences arising from different degrees of symmetry (least asymmetric, most asymmetric, and symmetric) among orthonormal wavelets. However, semi-orthogonal and bi-orthogonal are symmetric or antisymmetric and they have lower entropy levels. Since symmetry or antisymmetry implies linear phase of the two-scale sequence. This probably indicates linear phase is desired and without linear phase filtering visual impairment may occur, such as the non-symmetric distribution of time-frequency windows shown in figures 1.1 and 1.2. But it's degree of influence is still not clear at present stage and should be case dependent. (e.g., Meyer and B&L wavelets are also symmetric but their entropy values are not comparable to that of the ideal one).

2.5 Summary

Using entropy statistics of transform coefficients, it is found that, except the semi-orthogonal wavelets, there is not a wavelet basis tested here which can reach the level of approximation given by spectral analysis for our laboratory water waves. Still, many of the properties of these wavelets are more of analytical interests and hard to be physical significant. The strong performance of the semi-orthogonal wavelets indicates the usefulness of a modulated Gaussian wavelet or the Morlet wavelet, which are to be used in an adapted continuous wavelet transform for our signals. However, without a few features specific to CWT to be studied in next chapter the closeness between the best Riesz wavelet and our signals will not fulfill the final maximum usefulness.

Table 2.1: Entropy of orthonormal and semi-orthogonal wavelet coefficients as well as spectral coefficients under various statistic criterions.

Wavelet	L**2 coefficient entropy (0 division)	L**2 probability entropy (300 divisions)	L**1 probability entropy (300 divisions)	L**2 probability entropy (200 divisions)	Theotetical dimension (L**2 300 divisions)
B&L	4.691	1.330	3.417	1.179	3.782
Meyer	4.647	1.294	3.365	1.132	3.646
SO3O	4.833	1.669	3.756	1.488	5.307
SO3D	1.823	0.219	1.306	0.172	1.245
Spectrum	2.809	0.270	3.044	0.244	1.310
ON22A	4.993	1.761	3.891	1.516	5.815
ON33A	4.773	1.384	3.499	1.225	3.975
ON44A	4.790	1.517	3.596	1.363	4.559
ON55A	4.819	1.553	3.631	1.367	4.727
ON66A	4.790	1.373	3.456	1.203	3.946
ON77A	4.675	1.355	3.461	1.203	3.877
ON88A	4.645	1.229	3.283	1.082	3.418
ON99A	4.719	1.412	3.501	1.252	4.106
ON00A	4.787	1.423	3.511	1.244	4.149
ON44S	4.835	1.461	3.557	1.281	4.311
ON55S	4.758	1.492	3.576	1.298	4.426
ON66S	4.754	1.402	3.501	1.225	4.065
ON77S	4.751	1.336	3.331	1.188	3.804
ON88S	4.714	1.366	3.481	1.224	3.918
ON99S	4.755	1.469	3.570	1.288	4.345
ON00S	4.635	1.278	3.378	1.134	3.591
ON11C	4.938	1.696	3.832	1.457	5.452
ON22C	4.827	1.468	3.520	1.284	4.342
ON33C	4.756	1.488	3.573	1.333	4.427
ON44C	4.690	1.297	3.337	1.157	3.658
ON55C	4.644	1.309	3.405	1.154	3.703

Table 2.2: Entropy of bi-orthogonal wavelet coefficients under various statistic criterions.

Wavelet	L**2 coefficient entropy (0 division)	L**2 probability entropy (300 divisions)	L**1 probability entropy (300 divisions)	L**2 probability entropy (200 divisions)	Theoretical dimension (L**2 300 divisions)
BO11O	5.395	2.623	4.502	2.299	13.777
BO11D	5.395	2.623	4.502	2.299	13.777
BO13O	4.943	1.806	3.883	1.627	6.084
BO13D	5.266	2.371	4.373	2.053	10.708
BO15O	4.866	1.678	3.755	1.495	5.357
BO15D	5.227	2.291	4.327	1.987	9.882
BO22O	5.282	2.362	4.363	2.083	10.609
BO22D	4.434	1.181	3.284	1.034	3.257
BO24O	4.963	1.862	3.985	1.634	6.438
BO24D	4.359	1.090	3.220	0.962	2.975
BO26O	4.881	1.703	3.835	1.492	5.490
BO26D	4.332	1.064	3.174	0.940	2.899
BO28O	4.857	1.624	3.782	1.452	5.073
BO28D	4.318	1.069	3.157	0.941	2.914
BO31O	5.824	3.174	4.741	2.835	23.894
BO31D	4.377	1.058	2.655	0.936	2.880
BO33O	5.084	2.001	4.062	1.756	7.393
BO33D	4.205	1.102	2.827	0.965	3.011
BO35O	4.850	1.697	3.847	1.506	5.457
BO35D	4.125	1.026	2.776	0.908	2.789
BO37O	4.790	1.658	3.821	1.442	5.247
BO37D	4.106	0.986	2.737	0.873	2.679
BO39O	4.776	1.660	3.835	1.432	5.258
BO39D	4.098	0.967	2.713	0.866	2.629

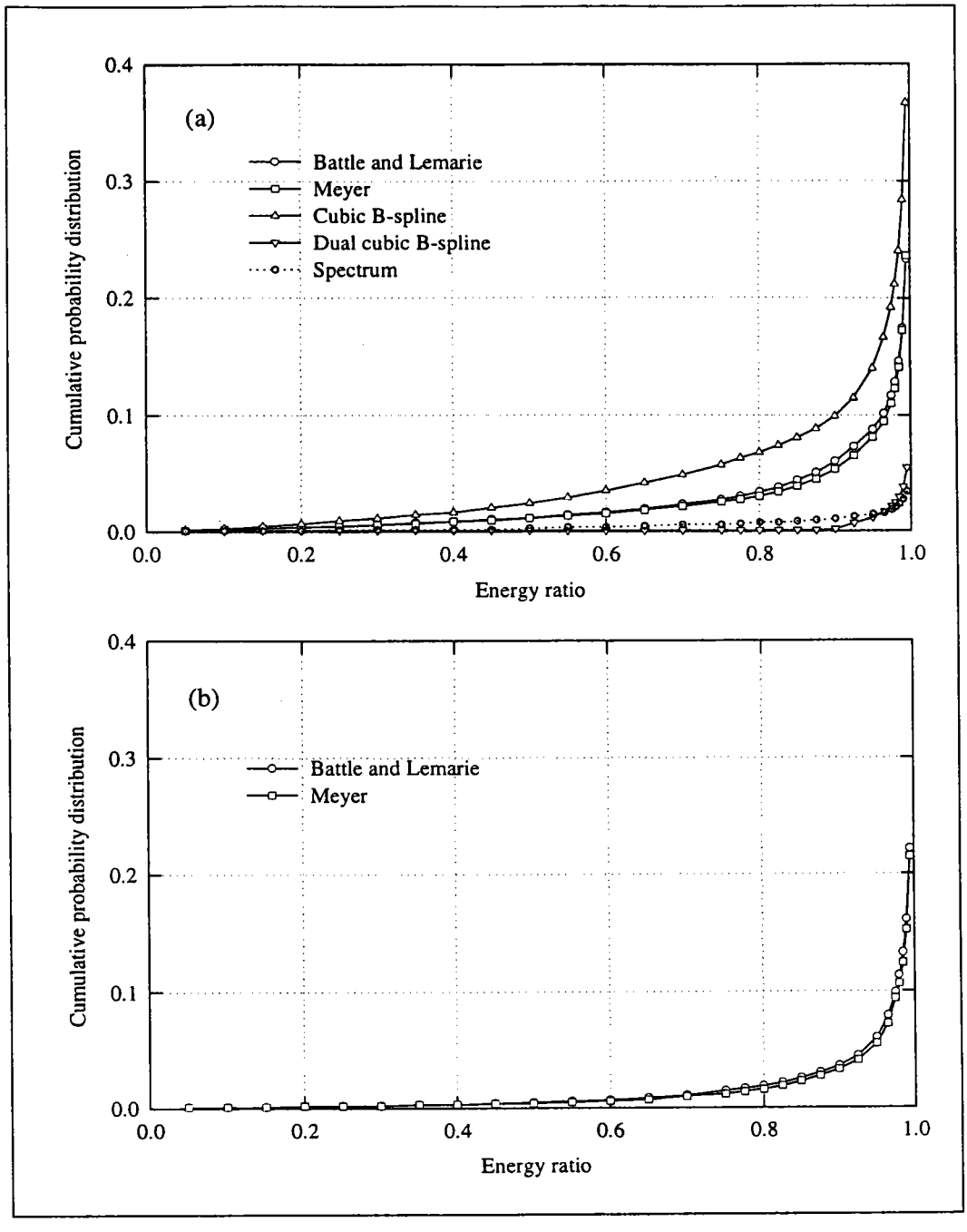

Figure 2.4: The cumulative probability distribution curves for various sets of transform coefficients derived from different function bases as indicated in the figures. The top figure shows those of the wavelet group as well as a curve for spectral coefficients; the bottom figure shows those of wavelet packets.

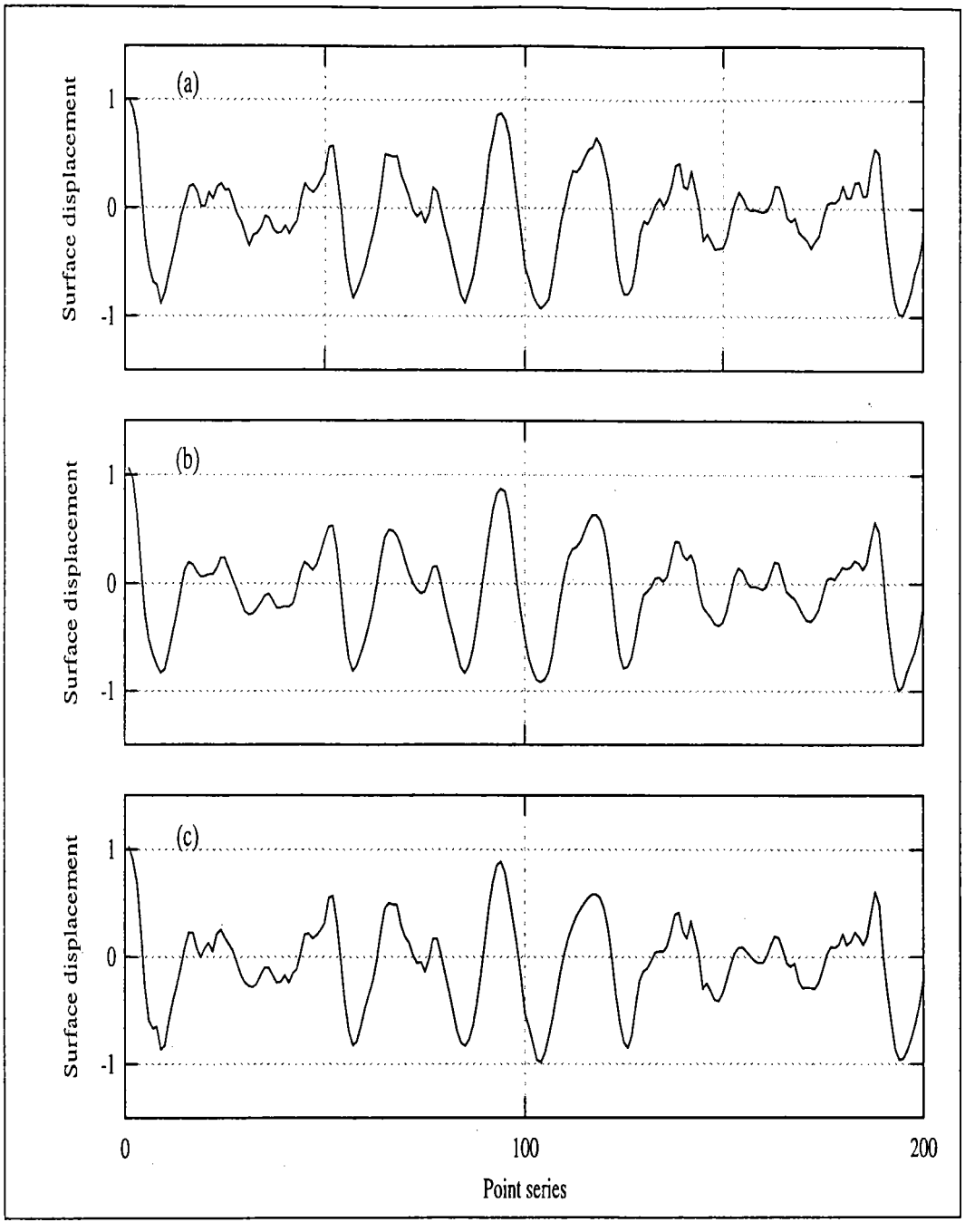

Figure 2.5: Reconstructed signals using truncated transform coefficients. Where 35% of the coefficients are kept. The original signal is shown in (a). Spectral coefficients in (b) and SO3D wavelet coefficients in (c). The semi-orthogonal wavelet is seen to better portrait the original signal, especially the small scale transient features.

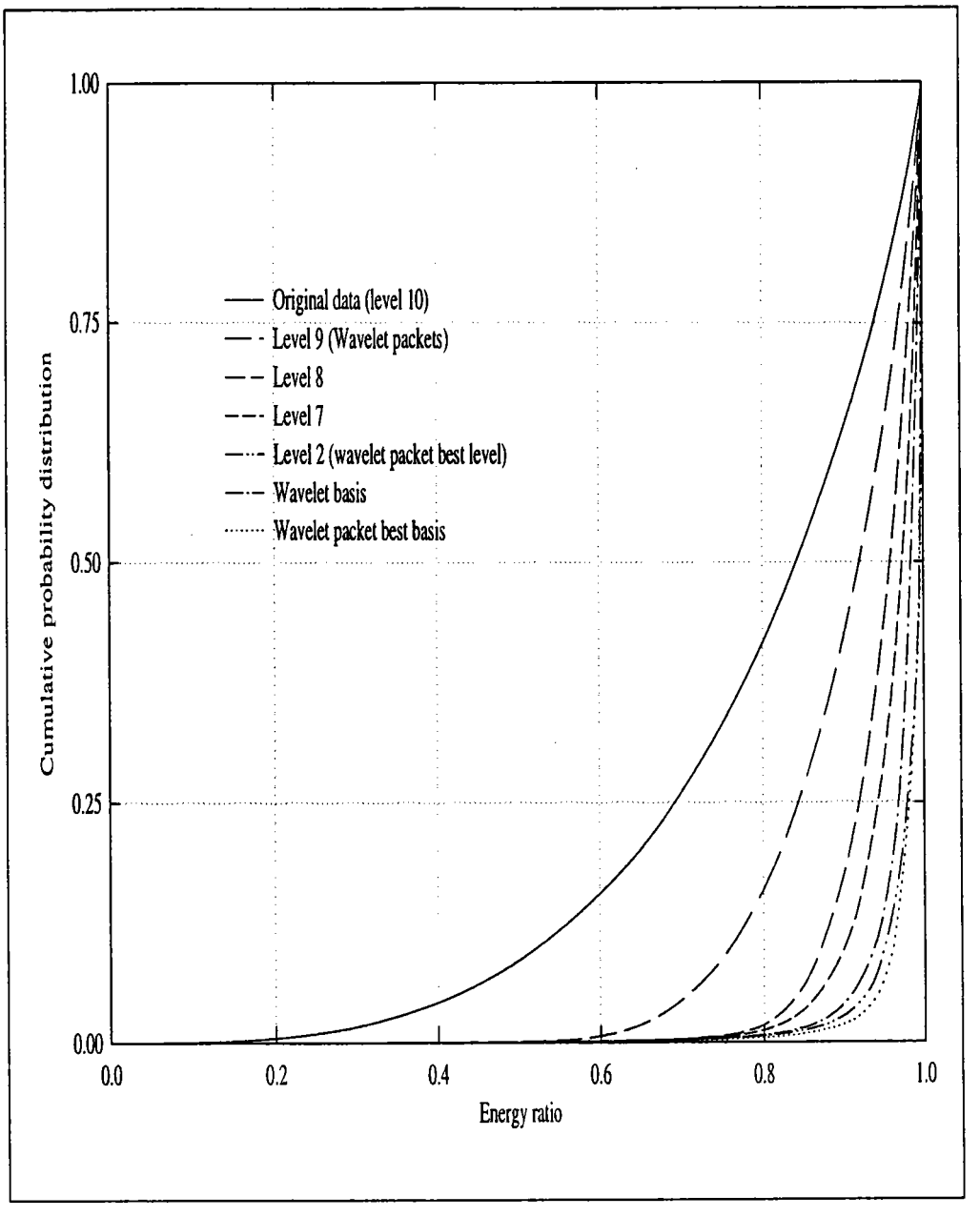

Figure 2.6: The cumulative probability distribution curves of the sorted wavelet and wavelet packet coefficients (L^2-norm squared, i.e., energy content) for various bases which all originate from a single mother wavelet. These bases include various wavelet packet levels, wavelet packet best basis, as well as the originating wavelet basis ON77S; as indicated in the legend.

Chapter 3

The Continuous Wavelet Transform Using Adapted Time-Frequency Windows

The fundamental thought for this chapter is rather simple: if time-frequency windows of fixed shape and size (STFT) are less suitable for characterizing transient signals or multi-scale phenomena than time-frequency windows of fixed size but with flexible shape (WT), then time-frequency windows with flexibility both in shape and in size should provide even better adaptiveness. In fact, if we are willing to disregard a lot of seemly complicated and impractical analytical points of wavelet analysis, particularly those of DWT, we are just as free as an artist to portrait our desired wave forms; and might not ever need to link this chapter to the previous one, but the two chapters do have links regarding the best analyzing method. The links enrich our fundamental thought and lead to usefulness.

3.1 Introduction

In the previous chapter we identified the most suitable Riesz wavelet basis for use with our data. However, it is not the wavelet that we directly use in our data analysis. The main differences are in two aspects. First, the previous chapter concerns the discrete wavelet transform which has a translation step of an integer multiple of the dilation scale and the dilations are in logarithmic measurement; while the actual scheme we use concerns the continuous

wavelet transform which might have a translation step as small as the sampling interval for all scales and the scales can be specified almost arbitrarily. Second, the wavelets in the previous chapter handle bases with frame bounds that are either tight or relatively tight, while the wavelet used here does not involve frame bounds (and might not have frame bounds at all when applied in the sense of discrete wavelet transform, i.e., not even qualified to be a Riesz wavelet). Nevertheless, we shall illustrate that the wavelet we use is a close proximity of the best Riesz basis identified. In this chapter we put efforts on clarifying these two points and try to illustrate the related advantages and disadvantages. Furthermore, we do this through the devising of a refined process of continuous wavelet transform using adapted time-frequency windows, which put emphases on its possible enhancements to physical implications.

3.2 Adaptability in association with the redundancy

If a function $\psi(t)$ is to be qualified as a wavelet of CWT, then the only requirement is that $\psi(t)$ meets the "admissibility condition,"

$$2\pi \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(\omega)|^2}{|\omega|} d\omega = C_\psi, \tag{3.1}$$

where C_ψ is a constant dependent only on ψ, and $\widehat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$. Here, among the several definitions of the Fourier transform pairs, the following system: $\widehat{\psi}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(t) e^{-i\omega t} dt$ and $\psi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{\psi}(\omega) e^{i\omega t} d\omega$ is chosen. The admissibility condition is the integration of power spectrum weighted by the inverse of the absolute value of frequency; therefore, it indicates that the wavelet should have little power at low frequency and is total nil at zero frequency, i.e., the area between the curve and the abscissa integrates to zero. This feature of reasonable decay or finite support is the reason why wavelet is called wavelet. The dilated and translated versions of this wavelet are $\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi(\frac{t-b}{a})$, where $a > 0, a, b \in \mathbf{R}$ are the dilation and translation parameter, respectively, and $\frac{1}{\sqrt{a}}$ is normalization factor for L^2-norm. The $\psi_{a,b}$ satisfies admissibility condition too. The admissibility condition is a very loose constrain; however, it does not provide a clear concept of the redundancy associated with applying CWT to discretely sampled signals. To illustrate this redundancy let us use the discrete wavelet frame (since frame wavelet certainly qualifies as a wavelet for CWT): $\psi_{a_0,b_0;j,k}(t) = {a_0}^{-j/2} \psi({a_0}^{-j} t - k b_0)$, where we have restricted a to set of discrete dilations a_0^j and b to set of discrete translations $a_0^j k b_0$, and $j, k \in \mathbf{Z}$, and $a_0 \neq 1$ and $b_0 > 0$ are fixed positive constants. For such a discrete wavelet frame we need to impose a more restrictive condition on $\psi(t)$ for it being admitted,

i.e., the stability condition,

$$b_0 A \leq 2\pi \sum_{j \in \mathbf{Z}} |\widehat{\psi}(a_0{}^j \omega)|^2 \leq b_0 B, \tag{3.2}$$

where A and B are positive constants and $0 < A \leq B < \infty$. The fixed constants b_0 and 2π are kept since we are dealing with normalized wavelet basis and since the magnitudes of A and B are related to the redundancy of the basis. The stability condition may look quite abstract, but We give its physical implication as: in order for a function to be able to be reconstructed from its wavelet coefficients, i.e., the operation is reversible, we need a process which is convergent when summing all its scales or frequency components. It is therefore necessary that the sum of the power of all the constituent elements can neither be nil or infinity. If it sums to zero, then the elements are all of zero measure — nothing exists. If it sums to infinity, then the elements are significantly overlapping in time and in frequency — there is either too much dependence or too much ambiguity and tangling (just as two vectors which are too parallel to each other are not belonging to a good choice of basis vectors). Speaking of reconstruction of a function from its wavelet coefficients one always involves a dual wavelet except for orthonormal basis where the wavelet is its own dual (self-dual). But since the roles of a wavelet and its dual can always be interchanged in both decomposition and reconstruction, the above statements apply equally well for dual wavelet (but their bounds will generally be different since the sets of coefficients are different as hinted by the different entropy values studied in the previous chapter).

If the basis functions are normalized and the inequality of the stability condition are optimized for both the greatest lower bound and the lowest upper bound, i.e., when A, B are defined as

$$A = \inf \left[\frac{2\pi}{b_0} \sum_{j \in \mathbf{Z}} |\widehat{\psi}(a_0{}^j \omega)|^2 \right], \tag{3.3}$$

$$B = \sup \left[\frac{2\pi}{b_0} \sum_{j \in \mathbf{Z}} |\widehat{\psi}(a_0{}^j \omega)|^2 \right], \tag{3.4}$$

then an indication of the redundancy is the average value or A and B, $\frac{A+B}{2}$, supposed that A and B are close to each other (almost tight). We elucidate the possible extreme redundancy of CWT as follows. If the dilated and translated versions a function originating from a certain set of discrete steps (a_0, b_0) constitute a frame with frame bounds A and B, then the frame bounds of the basis using the same function but with finer discrete steps, say

$a_0/2$ and $b_0/2$, are going to contain the bounds of coarser discrete steps; therefore, the new lower and upper bounds both grow together. This nested relation can be extended infinitely and in the limit sense it is included in the algorithm of CWT. This is the reason why there is no practical value of numerically reconstruction in CWT, although CWT is reversible analytically. Another intuitive explanation is even easier to comprehend: when apply CWT to discretely sampled signal, since for each scale there is the same number of wavelet coefficients as the number of data points and since we can specify scales in whatever resolution we like, we virtually have an unlimited number of wavelet coefficients. The sum of the powers of these coefficients can be unimaginatively huge, or even unbounded; On the other hand, the sum of signal energy is fixed. The ratio between the two sums indicates the degree of redundancy. But we guess that the information content or usefulness associated with the redundancy may behave like a cumulative pdf curve of a Gauss function which may saturate at a later stage.

Redundancy may be a nuisance in certain applications, such as those for the purpose of perfect reconstruction of signal or for the efficiency of coding and decoding; however redundancy has also shown the promising characters in several applications. First, redundancy does not mean that a whole bunch of coefficients are needed to represent the original signal, that is to say, that significant signal contents can be retrieved from only a comparatively amount of coefficients to that of tight or almost tight wavelet frame. Second, redundancy means that effects of noise either due to those embedded in the sampled signal or those arising from the nature of numerical processes (such as frequency leakage in discrete transform) might be reduced by taking advantage of the vast sample space of transform coefficients. One example related to the first character is Mallat and Zhong's [13, 25] signal reconstruction from local maxima using a quadratic spline wavelet. In fact, the wavelet they used is basically a loose wavelet (i.e., not really qualified as a wavelet), but they were able to recover images quite well from local peak values of wavelet coefficients and the transform was only performed at dyadic scales. The results of the present study on the coherence features in the wind, wave, and rain coupling system can serve as an example related to the second character.

To be more specific, a few questions which are natural extensions of the two previous characters and related to the objectives of the present work are stated as: (1) Can we utilize this redundancy to improve the relationship between the analytical form and its physical interpretability? (2) How well does the adaptation preserve the information content of the signal studied? (3) Is the adaptation efficient and easy to implement? Question one will be dealt in the next section. A short answer to question two in practical terms (proof for a special case using Morlet wavelet will be given in a later section) is that, if one just apply the

adaptation to finite range(s) of scale, then what is lost or unaccounted for in the adaptation process can still be recovered from some dilated and translated versions of some finer scale wavelets originating from the same $\psi(t)$ in the CWT. Another explanation is the implication of Mallat and Zhong's case just mentioned. Question three is dependent on the adaptation scheme. For now, based on the somewhat intuitive adaptation as related to question one, it is stated that nothing complicate is introduced. Furthermore, one practical point is: when analyzing signal we are almost always interested in only finite scale range(s), so what is really needed is to implement the adaptation locally. Hence the point is: use a scheme that is numerically with the same easiness and physically more sound, and don't be stuck with stubborn time-frequency windows.

Comparing the admissibility condition of CWT with the stability condition of DWT one is likely to acknowledge the great difference in flexibility between the two. In addition, the stability condition is a necessary condition, and not all choices for ψ, a_0, b_0 lead to wavelet frames; furthermore, stability may not guarantee good numerical behaviors. Figure 3.1 shows the results of numerical experiments on the blow-ups of wavelet curves and it indicates such a phenomenon. In the figure blow-ups of two bi-orthogonal wavelets around respective points using refinement cascade show the possible intrinsic absurdity for wavelets with peculiar analytical properties. The top sub-figure indicates a case where the DWT fails numerically to characterize itself even though its associated wavelet frame qualifies theoretically as a Riesz basis. The bottom sub-figure shows strange inclinations of wavelet curves. In fact, this figure also illustrates the point that most of the fancy wavelets with bizarre wave forms are not of our choice for studying water wave related physics, as is also indicated by their entropy values. Here two bi-orthogonal wavelets with four and twenty filter weights, respectively and constructed from quadratic spline scaling function [12] are used.

3.3 Degrees of freedom and the uncertainty relation

The flexibility of constructing wavelet function basis, i.e., the possibility of the adaptation, is associated with the number of degrees of freedom of the time-frequency window. The degrees of freedom of the time-frequency window is generally defined as the total area of the phase plane divided by the area of the time-frequency window corresponding to that determined by the Heisenberg uncertainty relation (or Heisenberg's inequality) [3]. Here we shall see that, even though it is impossible to increase the limiting degrees of freedom, there is no further limitation imposed upon present adaptation. In fact the present section also serves two purposes: (1) provide the description that illustrates the basic functionality of the modulation mechanism for a STFT, which in turn is conceptually the same as the

dilation mechanism in WT. (2) outline the relation between redundancy and the Heisenberg uncertainty using distribution of time-frequency windows within a phase plane.

The uncertainty relation states that the product of bandwidth Δ_ω and duration Δ_t of a signal cannot be less a minimum value of $\frac{1}{2}$. The Δ_t and $\Delta\omega$ are the standard deviations, with reference to their centroids, of packet energy $|f(t)|^2$ and power spectrum $|\widehat{f}(\omega)|^2$, respectively,

$$\Delta_t{}^2 = \frac{\int_{-\infty}^{\infty}(t-\bar{t})^2|f(t)|^2dt}{\|f(t)\|^2}, \tag{3.5}$$

$$\Delta_\omega{}^2 = \frac{\int_{-\infty}^{\infty}(\omega-\overline{\omega})^2|\widehat{f}(\omega)|^2d\omega}{\|\widehat{f}(\omega)\|^2}, \tag{3.6}$$

where $\bar{t} = \int_{-\infty}^{\infty} t|f(t)|^2dt/\|f(t)\|$ and $\overline{\omega} = \int_{-\infty}^{\infty}\omega|\widehat{f}(\omega)|^2d\omega/\|\widehat{f}(\omega)\|$. The way to get round of the uncertainty relation is to go through a modulation process (i.e., multiplying a basis function with a complex exponential). Since a modulation in one domain corresponds to a shift in the other domain in Fourier analysis, the new variance Δ_ω increases dramatically. Figure 3.2 shows the mechanism for this operation, where it is seen that $\Delta_t\Delta_\omega \gg \Delta_t D_\omega$. Now that the new value of $\Delta_t\Delta_\omega$ is significantly large than the limiting value of Heisenberg uncertainty relation for the un-modulated envelope, we therefore have quite a lot of flexibility to devise the time-frequency windows. Furthermore, as pointed out by Bracewell, there exists no theorem depicting the lower limit of $\Delta_t D_\omega$, i.e., no new restriction is imposed on D_ω. Therefore the uncertainty relation will not induce further restriction on the number of the degrees of freedom. That is to say, we have the same freedom to draw time-frequency windows which generally do not violates the uncertainty relation when we express a signal in its two dimensional phase plane, especially when considering the similarity between the Morlet wavelet (which is used for the present adaptation) and the modulated Gaussian, as will also be stated later.

3.4 Time-frequency windows of flexible size and the physics

An easier way to describe the algorithm of adaptation of time-frequency windows is going through examples. Here we choose the Morlet wavelet as our analyzing wavelet. The Morlet wavelet is almost identical to a modulated Gaussian. And a modulated Gaussian matches almost exactly with cardinal B-spline wavelet of order greater than or equal to 3 (quadratic spline-wavelet). As is given in Chui [6], the cardinal spline (scaling function) and its associated wavelets possess the nice properties of "total positivity" and "complete oscillation", respectively. We note that these two properties physically imply that its wave form

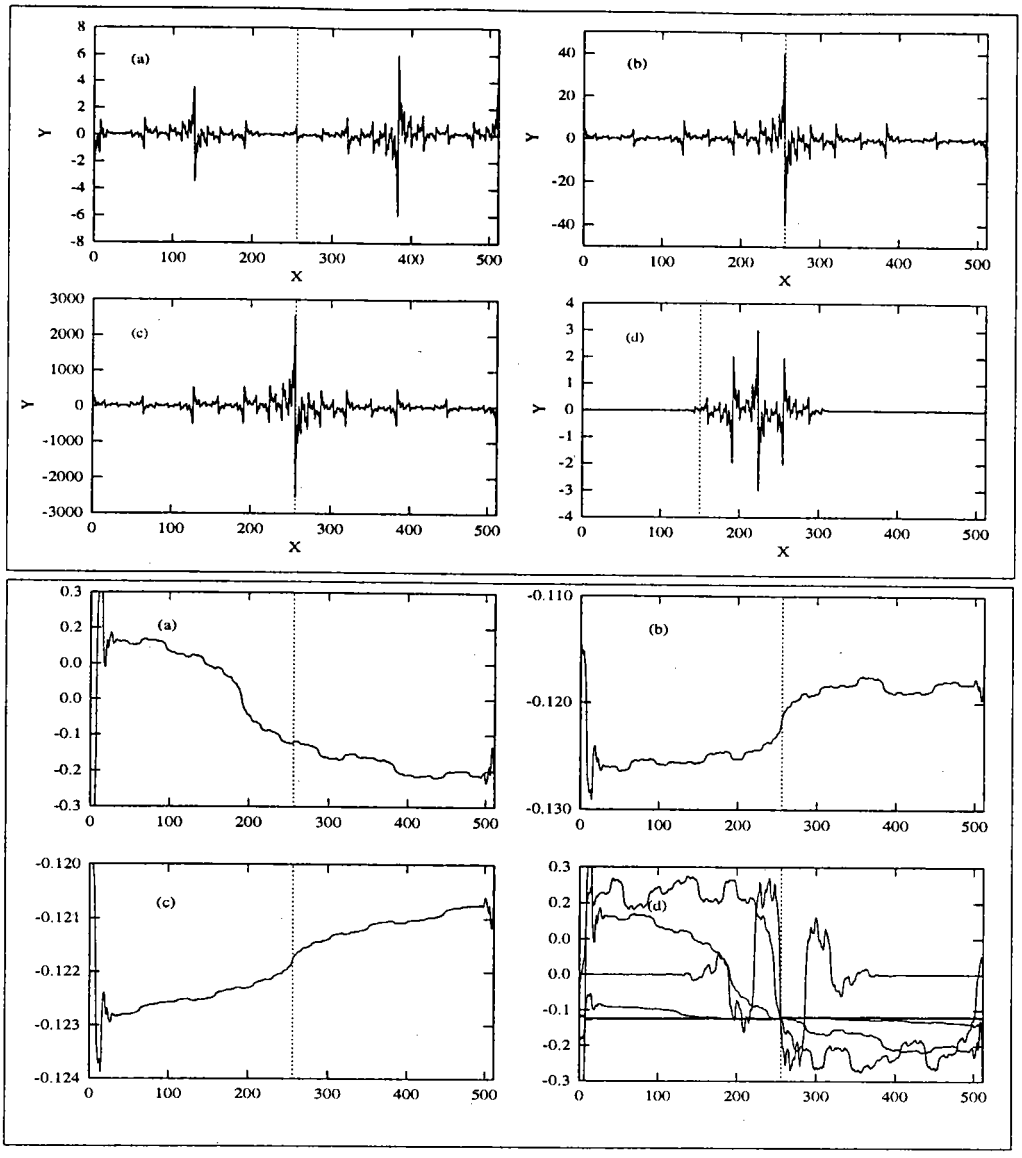

Figure 3.1: Wavelets with fancy analytical properties are often of bizarre wave forms and not of our choice for studying water wave related physics, either judging from they entropy values or form their stability condition — the blow-ups of bi-orthogonal wavelet BO31O (Top) and BO35O (Bottom). Related data for BO31O is: {Blow-up point: 150 (located at the dotted line in figure (d)); Origin: level 2, position 12 (i.e., U_2^{12} in figure 2.3); Length: 512 (the curve in figure (d)). Figures (a), (b), and (c) show successive blow-up scale of 2^6. The blow-ups diverge rapidly, i.e., the wavelet fails to identify itself numerically in refinement cascade.} Related data for BO35O is: {Blow-up point: 256 (located at the dotted line in figure (d)); Origin: level 2, position 12 (i.e., U_2^{12}); Length: 512 (the curve in figure (d) that occupies part of the abscissa). Figures (a), (b), and (c) show successive blow-up scale of 2^6. The blow-ups converge but with peculiar inclinations.}

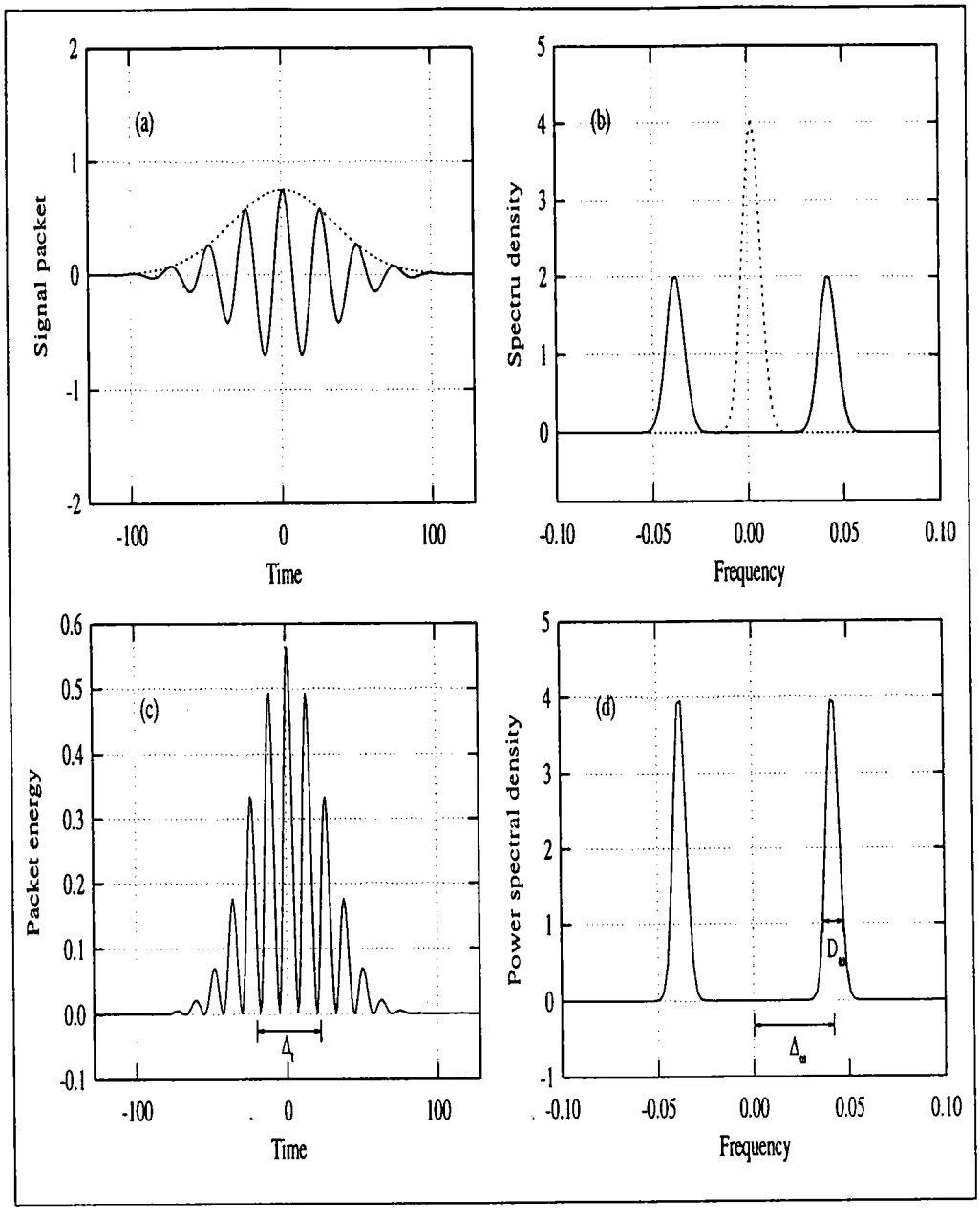

Figure 3.2: The uncertainty relation and the modulation versus shift property (adapted from Bracewell 1986). It is seen that a modulation process renders $\Delta_t \Delta_\omega \gg \Delta_t D_\omega$. The new value of $\Delta_t \Delta_\omega$ is significantly larger than the limiting value of Heisenberg uncertainty relation for the un-modulated envelope; we therefore have quite a lot of flexibility to devise the time-frequency windows.

are relatively smooth and without extremely rapid transient variations when compared with some fancy wavelets such as that shown in figure 3.1. Moreover, the cardinal B-wavelet is either symmetric or anti-symmetric; therefore, it benefits from linear-phase filtering (this physically implies that a slight difference in wavelet coefficients will not cause significant difference in their reconstructed wave forms, i.e., the variations of the wave forms are intuitively more natural or its impact to human visualization is smaller. Figure 1.1 shows the phase plane associated with wavelet without linear phase filtering.) Due to these reasons, when applied to discretely sampled data using CWT, the Morlet wavelet is likely to have better computational efficiency (in the sense of wavelet coefficients with smaller entropy) and lesser distortion (when some coefficients are dropped) in comparison with other wavelets. The Morlet wavelet is given by

$$\psi(t) = \pi^{-1/4}(e^{-i\omega_0 t} - e^{-\omega_0^2/2})e^{-t^2/2}, \tag{3.7}$$

where the term $e^{-\omega_0^2/2}$ is for justifying the condition of admissibility. Its Fourier transform is almost a shifted Gaussian and is given by

$$\widehat{\psi}(\omega) = \pi^{-1/4}[e^{-(\omega-\omega_0)^2/2} - e^{-\omega^2/2}e^{-\omega_0^2/2}], \tag{3.8}$$

in which ω_0 is a constant. The ω_0 is the modulation frequency. And it also has the physical implication of amplitude ratio r that is the relative magnitude of the second highest peak to highest peak of $\psi(t)$, i.e., $r = \psi(t_2)/\psi(0)$, in which t_2 is the abscissa of the second highest peak. A fairly good estimation of t_2 can be obtained by dropping the second term in the above equation (the exact value involves solving transcendental equation numerically) since the second term is generally of five order of magnitude less than the maximum value of the first term. Therefore,

$$\omega_0 \approx \frac{2\pi}{t_2} \approx \pi\left(-\frac{2}{\ln r}\right)^{1/2}. \tag{3.9}$$

The higher the ω_0 the smaller the ratio r. If ω_0 is constant, then the ratio r for different wavelet dilations or scales keeps constant too. Here one naturally asks: is it true that constituent wave components of different scales or time spans all possess this fixed decay feature?

To show that this is not true, let examine water wave decay due to viscous damping. For deep water waves with a clean surface, the energy losses due to viscous dissipation arise almost entirely from the straining of the irrotational motion in the water column, and the part of contribution from viscous stresses in the surface layer is negligible. It was shown

[28] that the time rate of change of the energy density is

$$\dot{E} = -2\mu\sigma^2 {a_w}^2 k, \tag{3.10}$$

where μ, σ, a_w and k are the dynamic viscosity of the water, the wave frequency, wave amplitude, and wave number, respectively. Since in deep water $E = (2k)^{-1}\rho\sigma^2 {a_w}^2$, where ρ is the water density, the attenuation coefficient

$$\gamma_\nu = -\frac{\dot{E}}{2E} = 2\nu k^2, \tag{3.11}$$

where ν is the kinematic viscosity of the water. Therefore the energy density of the wave evolves as

$$E = C_1 e^{-2\gamma_\nu t}, \tag{3.12}$$

where C_1 is a constant, and the amplitude decreases with time in accordance with

$$a_w = \sqrt{\frac{C_1 2k}{\rho\sigma^2}} e^{-\gamma_\nu t} = C_2 e^{-2\nu k^2 t}, \tag{3.13}$$

where C_2 is a constant if σ does not vary. Comparing the decay of wave amplitude of Morlet wavelet with the decay of the physical model, one see both similarity and dissimilarity. The similarity is that the attenuation coefficients in both models have inverse square dependence on scales (the former in $(1/a)^2$, the latter in k^2). The dissimilarity is the time dependence of the exponent in the exponential: in Morlet wavelet it is in t^2 dependence, while the physical model has linear dependence. It is therefore anticipated that a fixed shape Morlet wavelet is not going to have good universal similarity to water waves of different scales and it either overestimates the viscous decay of water waves at lower frequency end or under-estimates at high frequency end. This effect is probable more significant for waves with longer life spans. The simple reasoning provides the basis for the present adaptation: by adjusting the amplitude ratio r in the Morlet wavelet we are trying to adapt the evolutions of waves of different scales to flexible time-frequency windows. This in turn uses a variable ω_0. The general guideline is to use a larger ω_0 with narrower frequency band for waves of longer time support; and vice versa, a smaller ω_0 with wider frequency band for shorter life span. Here ω_0 is assumed to be a function of scale, i.e., $\omega_0 = \omega_0(a)$. The varying shapes and sizes of the time-frequency windows can be seen from

$$\psi_a(\frac{t-b}{a}) = \pi^{-1/4}\left[e^{-i\frac{\omega_0(a)}{a}(t-b)} - e^{-\omega_0(a)^2/2}\right] e^{-\frac{(t-b)^2}{2a^2}}. \tag{3.14}$$

Another advantage of using Morlet wavelet is that one is easy to have the physical perception of the size or property of a "scale". We have seen a lot of ambiguity or abstraction in a few studies for not involving the physical perception of scales, even when they were dealing with the simple case of Morlet wavelet. Earlier we mentioned that the wavelet coefficient generally refers to "scale" not "frequency". Scale does not have dimension, but frequency has precise physical unit. Furthermore, scale might be corresponding to complicate combination of several frequency bands, such as the compactly supported orthogonal wavelets used in figure 1.1. In order to have a clear picture of a "scale" one needs to know: what the basic wavelet looks like; what the actual support is; and, what the sampling interval is. All These tangle one's thoughts a lot, and we get lost easily. Take as an example: the numerical processes for both discrete Fourier transform and DWT care nothing about the units, only the index is important; however, there is an easy conversion from index to frequency for discrete Fourier transform, whereas there is no easy way to visualize the corresponding object from the index of a wavelet coefficient. This difficulty is avoided for Morlet wavelet since it is almost a modulated Gaussian and a modulated Gaussian is associated with another precise and physical unit named "carrier frequency", which for our adaptation is easily seen to be $\omega = \omega_0(a)/a$, supposing that $\omega_0(a)$ is large enough, say above 5. This further points out that using scale parameter a can be confusing and misleading, since the same a may correspond to different scales or frequencies when different adaptations or different wavelets are used.

The remaining problem is how to define a suitable decay parameter ω_0. As stated in the previous section that the present adaptation can always be applied to finite scale range(s) and the transform only needs to be implemented for scale range(s) that we are interested in. Still, since one can always regards that the set of sampling points of a signal is associated with, or derived from, a certain function, but there are basically infinitely many functions which can pass all these sampling points and since these functions may be either band-limited or band-unlimited but sampled signal are always band-limited (i.e., it is always true that numerical analysis is always associated with finite scale range); therefore, this indicates that we are quite free to make adaptation for ω_0. This in fact also hints the origin of the possible redundancy when CWT is applied to the sampled signals.

Both numerical simulation and experimental data are used to illustrate the adaptation. For the simulated data we use a parabolic chirp where the frequency range of interest covers the whole range of simulation, i.e., from almost zero to that corresponding to Nyquist sampling rate. For the experimental data the water wave measured in the wind blowing oval tank is used, in which the reasonable frequency range should be about 1.5 to 10 Hz

For the chirp signal we assume a linear variation of $\omega_0(a)$ from 10 to 7 in contrast to

the commonly adopted value of 5.3 (corresponding to $r \approx 0.5$). Figure 3.3 shows such a comparison.

It is seen that the adapted one gives better frequency localization for almost all frequencies except the lowest one (in fact it can be further adjusted). A phase plane of the complex wavelet coefficients is also shown.

Earlier we mentioned that the Morlet wavelet is likely to overestimate, for the wave tank signals, the decay of longer waves in the long run; therefore, relative to higher frequency waves, we should reduce the decay parameter ω_0 for low frequency ones . We heuristically assume

$$\mathrm{Erfc}\left(\frac{4}{10}\left(\frac{\omega_0}{a}-2\right)3+5\right)=a\omega \tag{3.15}$$

where Erfc is the complimentary error function and ω is the carrier frequency. This equation may be modified according to the type of signal studied or according to the frequency range of one's interest. Figure 3.4 shows the curve of the function. The logic for the choice of its constants is self explained in the attached program piece. Figure 3.5 shows the time-frequency maps of the signal. Again there is less smearing effects at the lower portion of the phase plane, since we mainly adjust the decay parameter for low frequency end. A few results and associated physics can be explained from the figure: (1) The dominant (carrier) frequency is about 2.5 Hz; (2) Waves of all frequencies keep constantly evolving, since there are always light and dark interlaces; (3) There are significant grouping effects, and the life spans of waves of significant energy are seen more enduring (the duration of the darker bands is longer); (4) There is obvious bifurcation among scales, especially for the intermediate frequency range of about 3 to 5 Hz. This indicates the phenomenon of energy cascade from where the energy concentrates to higher frequencies. Judging from these characters it seems that the energy phenomenon in a wave field is quite similar to that in a turbulent flow field [31].

3.5 Existence of admissibility condition

Earlier we gave physical description on how the present adaptation manages to provide an almost "loseless" operation. Loseless means that the full information of a function is preserved during the transform and we can recover the function from its wavelet coefficients, i.e., there exists a reverse operation. In the following we give a formalism through checking its admissibility condition, which is equivalent to verifying the existence of the resolution of the identity. For earlier adaptation we have a family of wavelets which originates from the $\psi(t)$ and has the parameter ω_0 in it. Where ω_0 is basically adjusted according to dilation

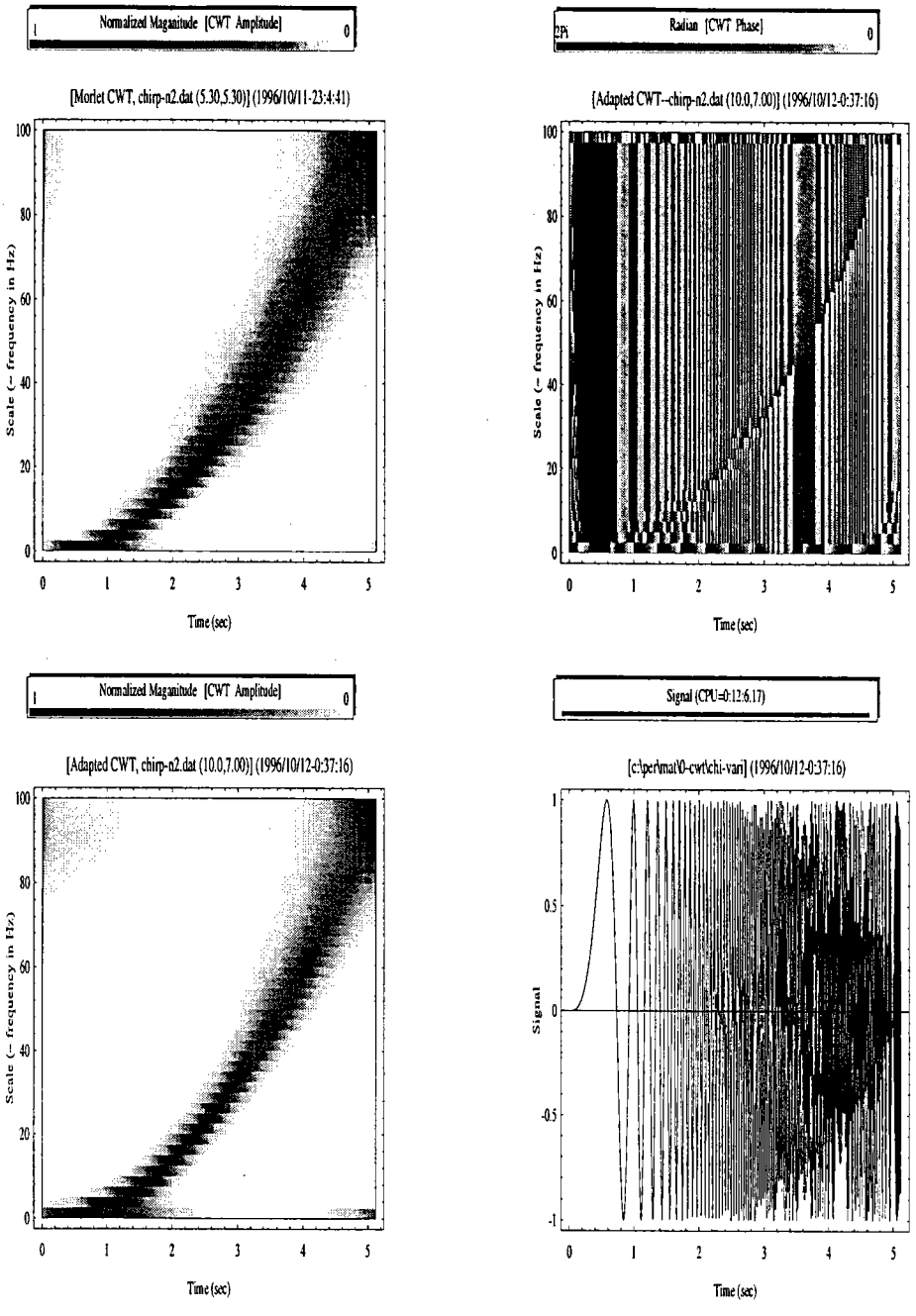

Figure 3.3: Phase plane characters for a parabolic chirp (bottom right) with (top left) and without (bottom left) adaptation of time-frequency windows. Top right shows the map of phase using a newly devised wavelet variant, which has properties quite in contrast to those of Morlet wavelet and has refined ridge extraction capability (under preparation).

```
obeg=11.;  oend=5.;
fcenter=2.5;  fdilation=10/4;   fshift=2. ;

perfc=Plot[ Erfc[(1/fdilation)* (freq +fcenter)-fshift]*
                          (obeg-oend)/2+ oend, {freq, -2.5, 8.5} ]
```

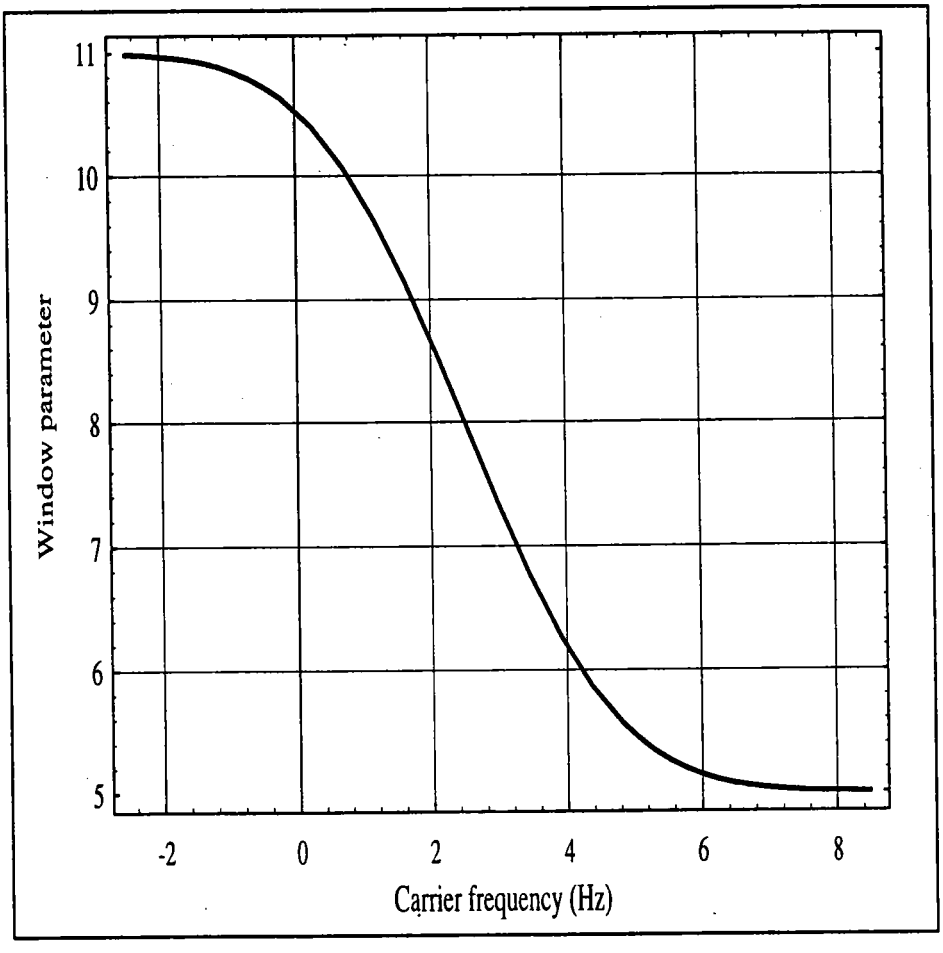

Figure 3.4: The assumed wave decay parameter ω_0 as a function of carrier frequency. The curve can be adjusted according to several parameters: approximate peak frequency, significant range of frequency, range of decay parameter, as well as a shift adjustment parameter; as are indicated in the attached program piece.

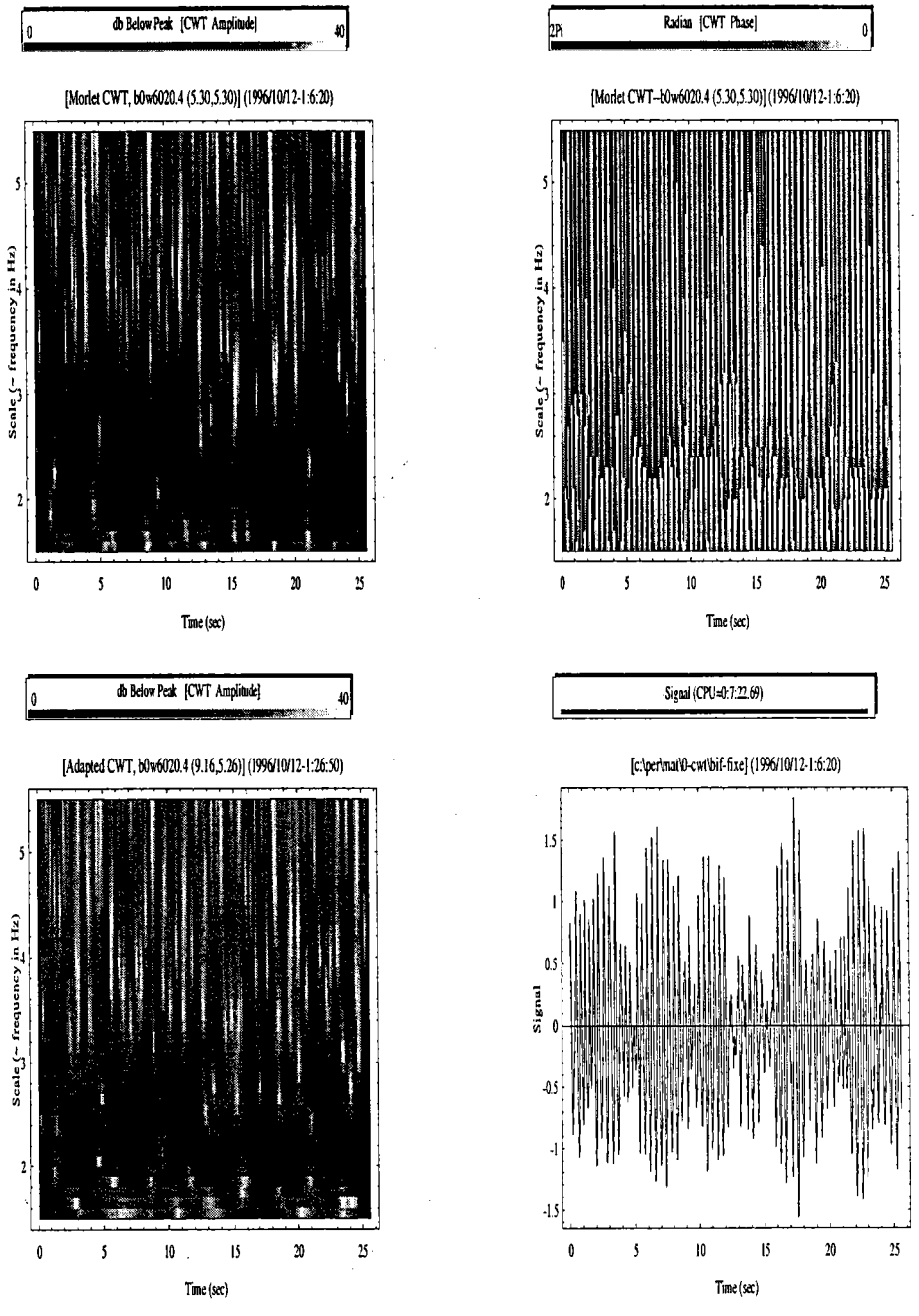

Figure 3.5: Phase plane characters of a water wave signal measured in a wind blowing oval tank (Top left: without adaptation; Bottom left: with adaptation; Top right: phase map; Bottom right: wind wave signal.) Since the assumed adaptation mainly adjusts the decay coefficients for low frequency part, one sees there is less smearing there. The phase map using the wavelet variant (cf. figure 3.3) provides clearer identification of ridges of main power; as is almost impossible using phase map of Morlet wavelet.

parameter a (depending on one's preference, one may specify a first or ω first); however, judging from the fact, as explained in the previous section, that the carrier frequency $\omega = \omega_0/a$, i.e., $\omega_0 = a\omega$, we therefore assume $\omega_0 = \omega_0(a\omega)$ for reasons as needed in the following verification of formulas. Moreover, we regard using a single variable $(a\omega)$ which combines $a\omega$ and $\omega_0(a\omega)$ should have more freedom in choosing the adaptation than using two variables in $\omega_0(a, \omega)$. So,

$$\psi_{\omega_0}(t) = \psi(t\,;\, \omega_0(a\omega)). \tag{3.16}$$

Its dilated and translated versions are given by

$$\psi_{a,b;\omega_0}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a};\, \omega_0(a\omega)\right). \tag{3.17}$$

And the wavelet coefficients of a function $f(t)$ are given by

$$\begin{aligned} \mathrm{W}f_{\omega_0}(a, b) &= \langle f, \psi_{a,b;\omega_0}\rangle \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{|a|}} f(t)\overline{\psi_{\omega_0}\left(\frac{t-b}{a}\right)}dt \\ &= \int_{-\infty}^{\infty} \sqrt{|a|}\,\widehat{f}(\omega)\overline{\widehat{\psi}_{\omega_0}(a\omega)}e^{-ib\omega}d\omega, \end{aligned} \tag{3.18}$$

in which $\widehat{\psi_{\omega_0}}(\omega) = \widehat{\psi}(\omega;\, \omega_0(a\omega))$ since $\omega_0(a\omega)$ is independent of t. We follow the formalism to check that the inner product two functions f and g, $\langle f, g\rangle$, can be recovered from the integration of the projection of $\mathrm{W}f(a, b;\, \omega_0)$ into $\mathrm{W}g(a, b;\, \omega_0)$ along both real lines of dilation and translation variables. That is, if

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{a^2}\mathrm{W}f(a, b;\, \omega_0(a\omega)) \\ &\qquad \overline{\mathrm{W}g(a, b;\, \omega_0(a\omega))}dadb = C_{\psi_{\omega_0}}\langle f, g\rangle, \end{aligned} \tag{3.19}$$

where $C_{\psi_{\omega_0}}$ is a fixed constant, then, when g is taken as a Gaussian function with variance approaching the limit zero and being almost the delta distribution $\delta(t)$, the inner product $\langle f(t'), g(t'-t)\rangle = \langle f(t'), \delta(t'-t)\rangle$ recovers $f(t)$. The right hand side of the above equation equals

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{a^2}\left[\int_{-\infty}^{\infty} \sqrt{|a|}\,\widehat{f}(\omega)e^{-ib\omega}\overline{\widehat{\psi}(a\omega;\, \omega_0(a\omega))}d\omega\right] \times \\ &\qquad \left[\int_{-\infty}^{\infty} \sqrt{|a|}\,\overline{\widehat{g}(\omega')}e^{ib\omega'}\widehat{\psi}(a\omega';\, \omega_0(a\omega'))d\omega'\right]dadb. \end{aligned} \tag{3.20}$$

Let define

$$\begin{aligned}
&\widehat{F_a}(t, \omega_0(a\omega)) \\
&\quad = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-it\omega} \sqrt{|a|}\, \widehat{f}(\omega) \overline{\widehat{\psi}(a\omega; \omega_0(a\omega))} d\omega \\
&\quad = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-it\omega} F_a(\omega; \omega_0(a\omega)) d\omega, \qquad (3.21)
\end{aligned}$$

$$\begin{aligned}
&\overline{\widehat{G_a}(t, \omega_0(a\omega))} \\
&\quad = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{it\omega} \sqrt{|a|}\, \overline{\widehat{g}(\omega)} \widehat{\psi}(a\omega; \omega_0(a\omega)) d\omega \\
&\quad = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{it\omega} \overline{G_a(\omega; \omega_0(a\omega))} d\omega. \qquad (3.22)
\end{aligned}$$

One has

$$\begin{aligned}
&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{2\pi}{a^2} \widehat{F_a}(t; \omega_0(a\omega)) \overline{\widehat{G_a}(t; \omega_0(a\omega))} da dt \\
&\quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{2\pi}{a^2} F_a(\omega; \omega_0(a\omega)) \\
&\qquad\qquad \overline{G_a(\omega; \omega_0(a\omega))} da d\omega \\
&\quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{2\pi}{|a|} \widehat{f}(\omega) \overline{\widehat{g}(\omega)} \\
&\qquad\qquad |\widehat{\psi}(a\omega; \omega_0(a\omega))|^2 da d\omega \\
&\quad = 2\pi \int_{-\infty}^{\infty} f(t) \overline{g(t)} dt \int_{\infty}^{\infty} \frac{|\widehat{\psi}(a\omega; \omega_0(a\omega))|^2}{|a|} da \\
&\quad = 2\pi \langle f, g \rangle C_{\psi_{\omega_0}}, \qquad (3.23)
\end{aligned}$$

where the resolution of identity is given by the condition

$$\int_{-\infty}^{\infty} \frac{|\widehat{\psi}(a\omega; \omega_0(a\omega))|^2}{|a|} da = C_{\psi_{\omega_0}}. \qquad (3.24)$$

This condition is more restrictive than equation (1) in that $\widehat{\psi}(0, \omega_0(u)) = 0$ for all $u \in \mathbf{R}$; Otherwise there is no other restriction since what is changed in the integration is limited to finite range and is anticipated to be finite. Fortunately, the above example using Morlet wavelet satisfies this condition. In fact, we even believe other wavelets might be adapted either in the present sense or in a sense similar to that using multi-voice or multi-wavelet in

DWT.

3.6 Summary

Based on the minimum entropy Riesz wavelet we look for its CWT counterpart for better descriptions of our water wave signals. With the intent to enhance its significance in physical implications we exploit the redundant (or non-orthogonal) feature of CWT coefficients and propose flexible constructions of the time-frequency windows with better adaptation to physics. An example of wave decay was use to justify the concept of this adaptation. Numerical results using chirp signal and experimental data were used to show the overall improvement of time-frequency resolutions in their phase planes. With slight modification to the original prove we verify that the resolution of the identity is generally obeyed under the adaptation, at least for the Gaussian-modulated wavelet used here; that is to say, there is no loss of information using this adaptation. And using Morlet wavelet as the analyzing wavelet we illustrate that the decay parameter ω_0 or carrier frequency can be modified for different scales in virtually unlimited ways. Even though at the current stage the forms of adaptation are mostly intuitive, they are conceptually reasonable and fit into the nature of wave modulations in multi-scale systems. In short this chapter illustrate the additional flexibility of wavelet analysis as well as the possible usefulness of the redundant nature of the CWT coefficients, and these lead to the success of its application to the investigation of signal coherences in the wave and current fields.

Chapter 4

Wavelet Coherences and Spectral Coherences in the Wave and Current Fields

When we embarked on the wavelet studies for the interests of physics regarding water wave related signals we did not foresee most of the results of the current three main subject topics; not to mention that we had any idea about the close links among them. In fact, for the first subject topic on wavelet bases, we initially focused almost entirely on seeking meaningful, and somewhat "fancy", wavelet packet phase plane representations of our signals, which turned out to be an effort in vain from physics point of view up until now. And, for a period of time, the semi-orthogonal wavelet with minimum entropy did not catch much of our attention since only orthonormal wavelets have two dimensional phase plane representations and the cardinal B-spline wavelets simply don't have. For the second subject topic, we were only wondering that the following sequence should be a very natural evolution process regarding adapting time-frequency windows — an adaptation form fixed shape and size to varying shape with fixed size, and further to varying shape with changing size. And this process should be working for any wavelet not just for the Morlet wavelet. The reason why it was chosen was because it is simple and the scale information it conveys has the most easily understandable physical unit as the carrier frequency. For the third subject topic, we were just curious to clarify our doubts on the presentations of a paper [21] that we read. In that paper the definition of coherence and the explanation of the results look awkward to us and we somewhat got lost.

To us, significant information comes most often in surprise. It seems that the intricate links of the three subject topics come to our minds all of a sudden. Of course, this should be attributed to the accumulation of further understandings when more or repeated readings have been gone through.

4.1 Wavelet coherence and spectral coherence

The coherences among different target quantities manifest the degree of mutual interactions in a multi-scale coupling system. By studying their variations for different experimental setups or parameters it is possible to identify the evolutions of different scales and to isolate key influencing factors and their main effects.

In this chapter both spectral and wavelet coherences will be used to study the scale features in the wind and current fields in a wind, wave, and rain coupling system. A few points on the energy phenomena in the system will also be explained. Let first have a few background information of the two methods and highlight their advantages and disadvantages.

Earlier we mentioned the effects of non-stationary or local transient variations. And in a related study focusing on Hilbert transform's viewpoint [19] the repeatability problems of Fourier spectra were also addressed . Further evidences of the drawback caused by this phenomenon will also be seen here by comparing the performances of spectral coherence with those of wavelet coherence. It is a bit to our surprise that wavelet coherences are able to provide much better information that is unambiguous in outcomes and economical in amount of data needed. Numerical results to be shown later shall state these all. But let first cite the basic differences in two aspects. First, from the viewpoint of their origins, wavelet coherence has a closer similarity to its analytical counterpart than spectral coherence does. Or stated more specifically, the wavelet coherence is a direct and natural extension of the wavelet resolution of identity and, therefore, involves less artificial intervention. Second, the wavelet coherence is derived from an extremely redundant set of coefficients; while spectral coherence is associated with a set of coefficients derived form orthonormal basis functions. The redundancy provides a fine resolution of scale as well as a huge population space. These factors minimize noise effects and yield clear coherent tendencies. In the following we illustrate these points by going through a few related formulations.

The cross correlation function of two functions $g(t)$ and $h(t)$ is the following inner product $c(t)$

$$c(t) = \langle g(t+\tau), h(\tau) \rangle, \tag{4.1}$$

where τ is a dummy variable. The correlation coefficient function $r_s(t)$ is

$$r_s(t) = \frac{c(t)}{\|g(t)\| \|h(t)\|}. \tag{4.2}$$

For real $g(t)$ and $h(t)$, its Fourier transform is

$$\frac{\mathcal{F}[\langle g(t+\tau), h(t) \rangle]}{\|g(t)\| \|h(t)\|} = \frac{G(\omega)\overline{H(\omega)}}{\|G(\omega)\| \|\overline{H(\omega)}\|}. \tag{4.3}$$

The artifacts to be introduced in spectral coherence are associated with the form of expected values as well as the introduction of a normalization as given by

$$R_s^2(\omega) = \frac{|\mathbf{E}[G(\omega)\overline{H(\omega)}]|^2}{\left(\mathbf{E}[|G(\omega)|^2]\mathbf{E}[|H(\omega)|^2]\right)^{1/2}}, \tag{4.4}$$

where the symbol $\mathbf{E}$ stands for taking expected value. This equation is identically unity for all frequencies if data sequence is not segmented and arranged in an array with one additional dimension, since expected values take no action without introducing one more dimension. The process of this segmentation is just like that commonly implemented in calculating the power spectrum. The purpose is to reduce the uncertainty or standard deviation of the spectrum. There is no doubt that the inherent properties of discrete Fourier analysis are imposing similar limitations to the conclusiveness of spectral coherence.

As to wavelet coherence the derivation is even simpler along with fewer artifacts. The wavelet resolution of identity of two functions is

$$\langle g, h\rangle = \frac{1}{c_\psi}\int_0^\infty \frac{1}{a^2}\int_{-\infty}^\infty \langle g, \psi_{a,b}\rangle\overline{\langle h, \psi_{a,b}\rangle} db da. \tag{4.5}$$

For a fixed scale a

$$\langle g_a, h_a\rangle = \frac{1}{c_\psi}\frac{1}{a^2}\int_{-\infty}^\infty \langle g, \psi_{a,b}\rangle\overline{\langle h, \psi_{a,b}\rangle} db. \tag{4.6}$$

Here the integration with respect to the translation parameter b is physically as well as intuitively similar to the operation of taking an expected value by summing up the elements in the population space. It is therefore quite straightforward to define the wavelet coherence as the the natural extension of the normalized equation of resolution of identity:

$$R_w^2(a) = \frac{|\mathbf{E}_b[\langle g, \psi_{a,b}\rangle\overline{\langle h, \psi_{a,b}\rangle}]|^2}{\left(\mathbf{E}_b[|\langle g, \psi_{a,b}\rangle|^2]\mathbf{E}_b[|\langle h, \psi_{a,b}\rangle|^2]\right)^{1/2}}, \tag{4.7}$$

where the subscript b in $\mathbf{E}_b$ stands for taking average with respect to the translation parameter. It is clear that wavelet coherence has a more direct linkage to its analytical relevance than does spectral coherence. Unlike spectral coherence, there is no need to segment the data. The expected values can be obtained in a sense of summing up the results of simple convolution through proper imaging and sign change of the functions and variables. Therefore the population size of the sample space of wavelet coefficients is generally two or three order of magnitude larger than that for spectral coherence. That is to say, for almost any practical data acquisition scheme and any specific scale the amount of summable coefficients is much bigger.

For wavelet coherence we can focus only on the portion of scale range that is significant or meaningful to us. But for spectral coherence we have no control at all over the frequency range of interest, so a great portion of the spectral results might be entirely irrelevant to our interests. Judging from the fact that for all practical cases we generally only want to, and are just able to, focus a finite range of frequency, we know that spectral approach wastes resources in the unwanted while wavelet coherence does just the most right, even though non-orthonormal cases take much longer computation time. Finally, we like to mention again that the wavelet used here is the adapted modulated Gaussian wavelet discussed in the previous chapter.

4.2 Coherence in the wind, wave and rain coupling system

4.2.1 Experiments

The experiments were performed in an oval tank equipped with wind-blowing facilities and mechanic wave generator. The tank is 35cm wide and 45cm height with a straight observation segment of 5m. The water depth is 24 cm and a two-meter rain module was mounted atop the tank (figure 4.1). Major experimental setups involve changing the combination of wind speeds, rain intensities, and mechanic wave parameters. Three kinds of data were measured: average wind speeds, wave heights, and LDA aqueous flows. Further details of measurement will be mentioned when case requires.

4.2.2 Wind wave cases

The wavelet and the spectral coherences between waves and aqueous flows at different depths with and without rain (68 mm hr^{-1} rain with 6.0 m sec^{-1} wind) for a data length of 1024-point are shown in figure 4.2. Figures 4.3 and 4.4 show results using 2048- and 4096-point data length, respectively. It is quite amazing that, even for such non-stationary wind wave conditions the curves of wavelet coherence using 1024-data data length have yielded quite consistent shapes as those using longer data length. It is also interesting to note that, when extremely long data lengths are used the spectral coherence curves are seen to approximate the wavelet coherence curves, as can be seen from figure 4.5 when compared to the previous figure. Besides, these curves further evidences that the improvements with the lengthening of data length for spectral curves are very slow, especially for scale ranges with smaller energy contents. Moreover, judging from the fact that the two function bases are different, but they somewhat close to each other, we regard that our water waves are

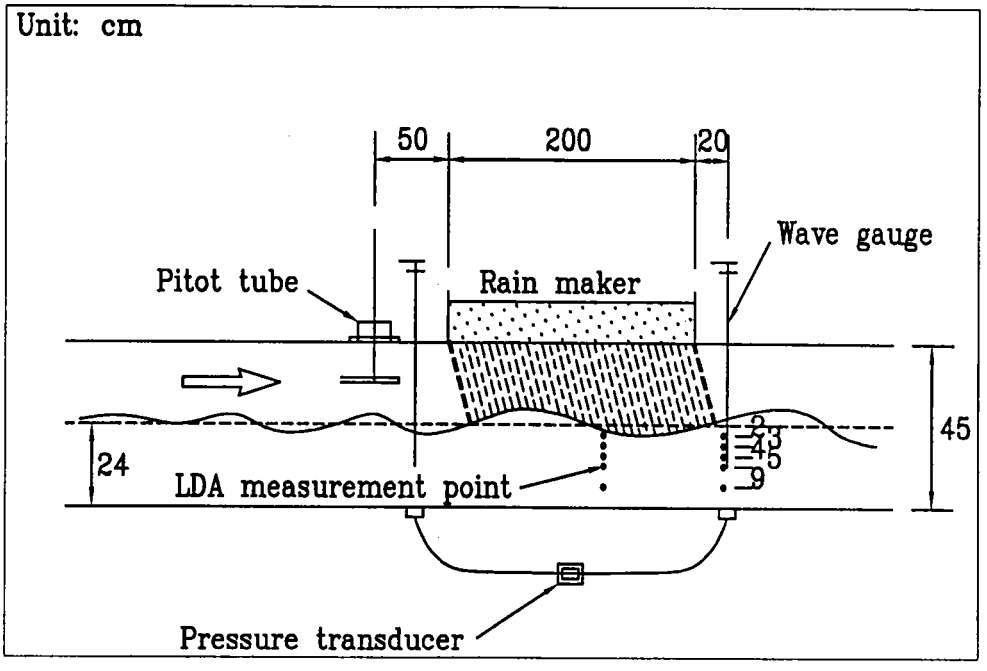

Figure 4.1: The experimental setup

somewhat relatively "regular" when compared with other wavelet basis functions, such as what was said in figure 3.1. It has been clear that these figures demonstrate the usefulness of the redundancy of wavelet coefficients associated with a continuous transform when applied to discretely sampled signals. In an alternative perspective, they point out the large variance values and slow convergence when orthonormal function basis is used. In fact, keeping in mind that the actual frequencies (or scales) of main power may not locate exactly at discrete resolution points in the frequency axis, these disadvantages can be expected form the serious shift-non-invariant property shown in 2.1, where the wavelet transforms are associated with either orthonormal (tight) function basis or relatively tight frame.

The above statements do not focus on the factor of rain. To see more clearly the effects of rain on the wave-current coherences for aqueous flow measured at different depths, let compare the curves for each individual depth as shown in figure 4.6. It is seen that rain has greater influence at high frequency region (about > 3.5 Hz) and significantly lower the coherences in this frequency range for aqueous flows measured in the near-surface region (2 and 3 cm below the still water surface); while for deeper depths (4 and 5 cm below the still water surface) the coherence curves widens and their values somewhat increase. However, at further depth (9cm) the overall coherence is significantly reduced. This strongly indicates that, under the action of rain, scales of main power contents do not as well penetrate as those without rain do, and also implies that the scales in deep region are de-tuned. Another set

of figures similar to figure 4.6 but under a different wind speed of 5.1 m sec^{-1} is shown in figure 4.7. It shows basically the same tendencies.

Figure 4.11 shows wave-current coherences under different wind speeds for depths 2 and 5 cm with and without rain. It is seen that for higher wind speed the peak coherence increases and also shifts toward lower frequency. The reasons can be explained from the point of view of wave development. Since these are measured at the same cross section in the oval tank higher wind induces a more mature stage of development.

4.2.3 Stokes wave cases

The coherent features for Stokes wave cases also provide interesting and informative accounts for the evolutions of different scales within the coupling system. Figures 4.12 shows the wave-current coherences under different rain intensities for Stokes wave with fundamental harmonic at 1.4 Hz and wave slope of 0.06, as are judged from Fourier spectrum shown in 4.9 and envelope curve for the band of fundamental harmonic shown in figure 4.10. It is seen that rain enhances and broadens wave-current coherences near the fundamental harmonic for aqueous flows measured at various depths. For the second harmonic, the coherence curves generally somewhat broaden too and also shift to the right. These tendencies indicate the diversification of scales due to rain within a relatively simple wave field. And the phenomena agree with the argument that the impact of rain provides the impetus of tuning and de-tuning among waves of different scales within the interaction system.

It should also be noted that, even for this nearly stationary wave field, spectral approach yields extremely poor behaviors of coherences as shown in figure 4.14, where the data and its length are the same as the previous ones. Examples here further strengthen our understanding about the usefulness of the redundancy related to the non-orthonormal function basis. Most importantly, they show that wavelet viewpoint provides better descriptions of the intrinsic physics even for stationary signals. And this should formerly be attributed to the appropriateness of depicting waves as modulated wave forms.

The wave-current coherences for aqueous flows measured at different depths (2, 3, 4, and 5 cm below still water surface) for Stokes wave with fundamental harmonic of 2.7 Hz and a high wave slope value of 3.0 is shown in the right hand part of figure 4.14 (the left hand part shows those of the previous Stokes waves). And the top and bottom sub-figures are without and with rain, respectively. For the right hand part sub-figures, due to the high wave slope the coherences for aqueous flows measured in the near-surface region are significant lower than those measured in deeper region. This is probably related to the high non-linear effects due to the Benjamen and Feir side-band instability [19]. Another significant feature due to rain for the high wave slope cases is the obvious increase of wave-current coherences

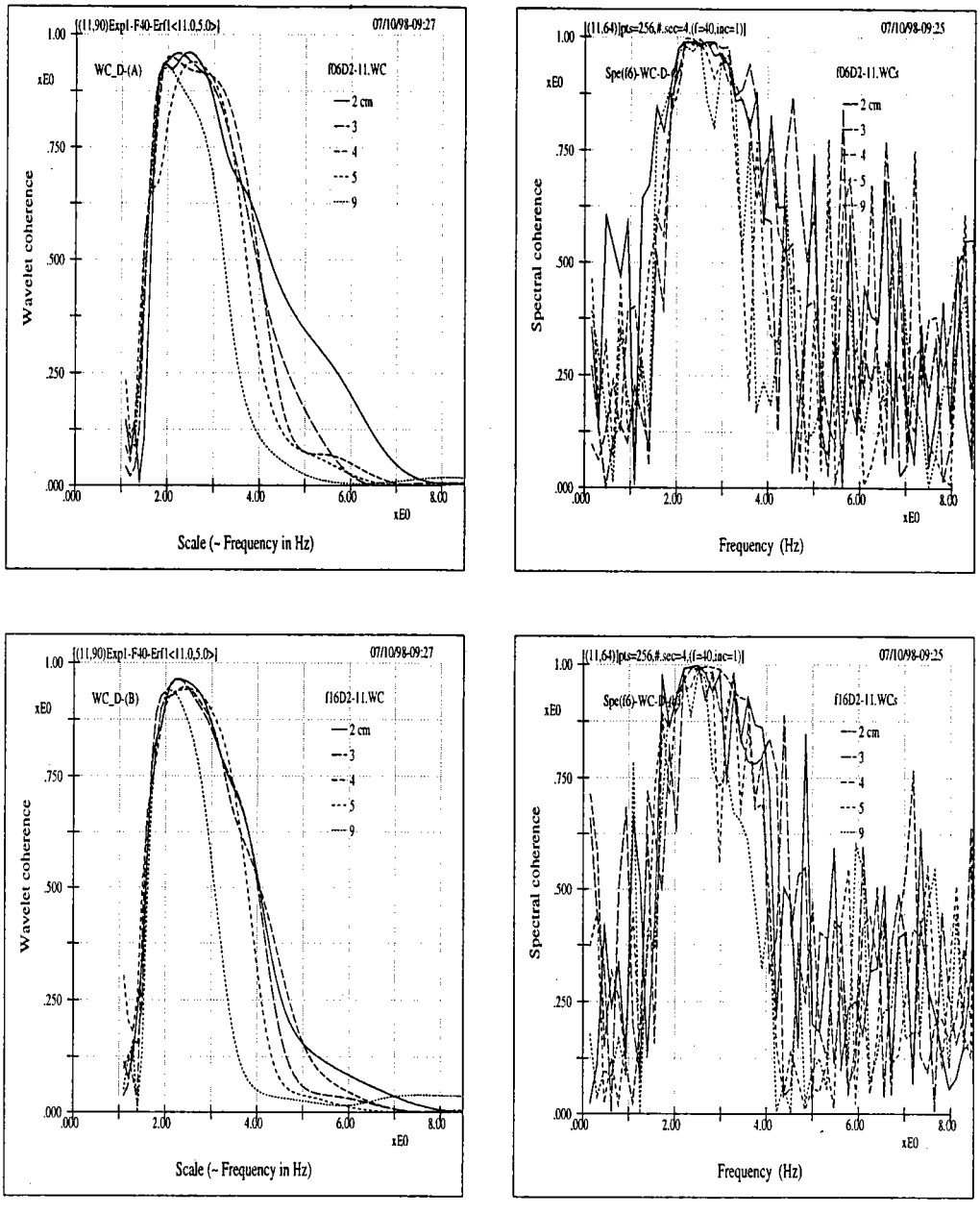

Figure 4.2: The wavelet (Left) and spectral (Right) coherences between wave and aqueous flow at each individual depth as indicated in the figures without (Top) and with (Bottom) rain (68 mm hr^{-1}) under 6.0 m sec^{-1} wind. Here the length of the data segment is 1024-point. In reference to the two subsequent figures using longer data lengths the usefulness of the redundancy of CWT is most obviously seen since the wavelet coherence curves have shown extremely good proximity to those using longer data lengths.

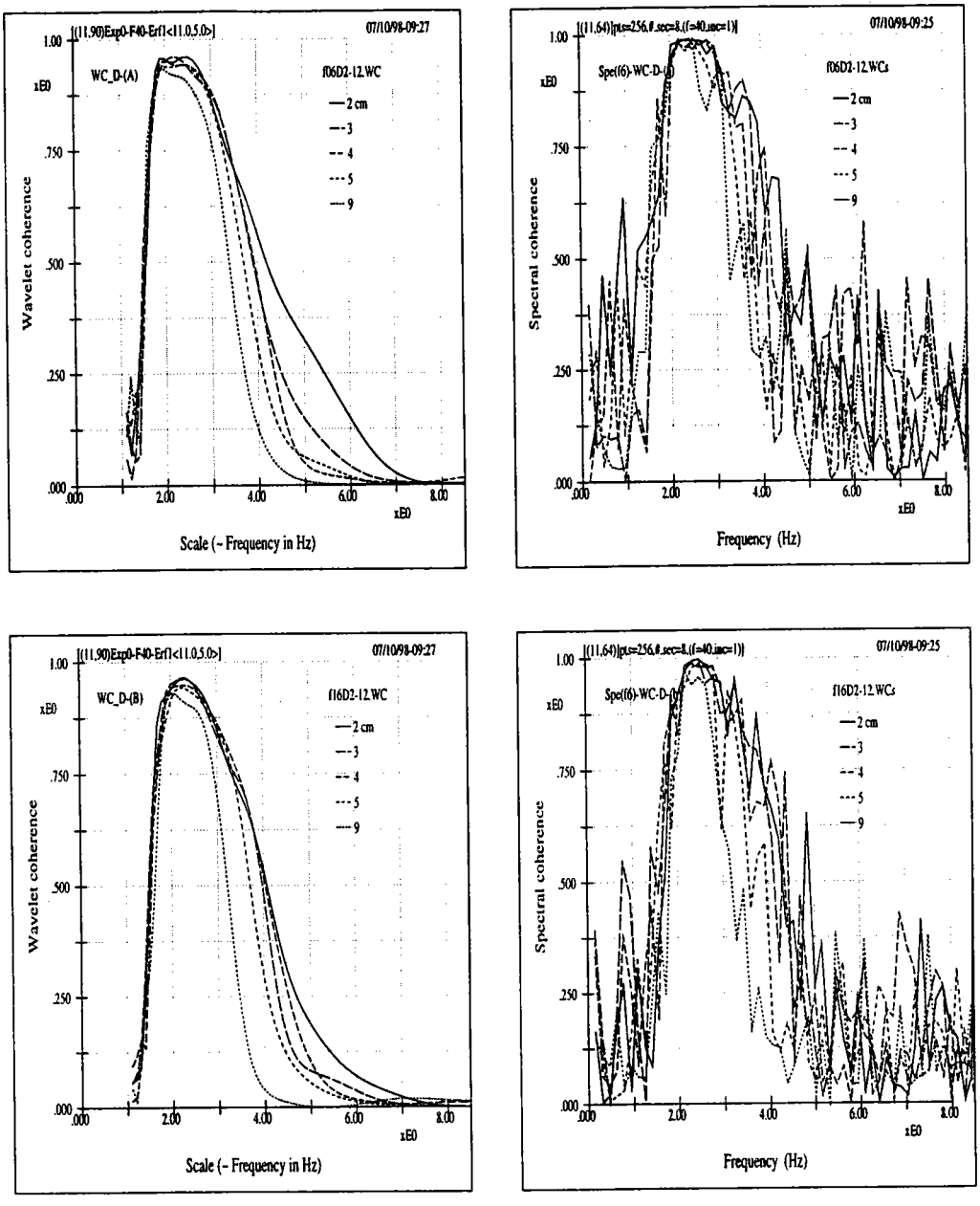

Figure 4.3: The wavelet (Left) and spectral (Right) coherences between waves and aqueous flows at different depths without (Top) and with (Bottom) rain (68 mm hr^{-1}) under 6.0 m sec^{-1} wind. Here the length of the data segment is 2048-point.

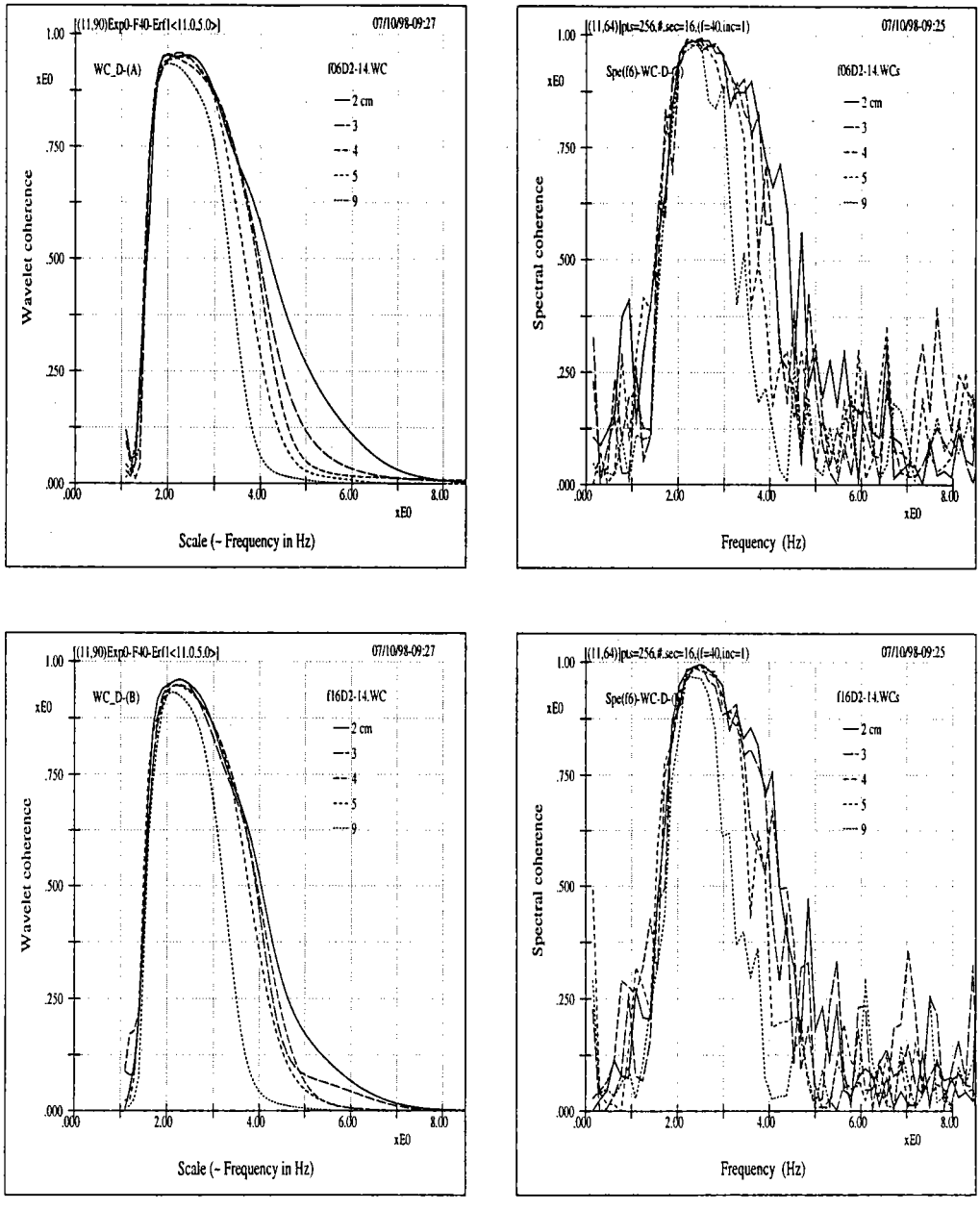

Figure 4.4: The wavelet (Left) and spectral (Right) coherences between waves and aqueous flows at different depths without (Top) and with (Bottom) rain (68 mm hr^{-1}) under 6.0 m sec^{-1} wind. Here the length of the data is 4096-point.

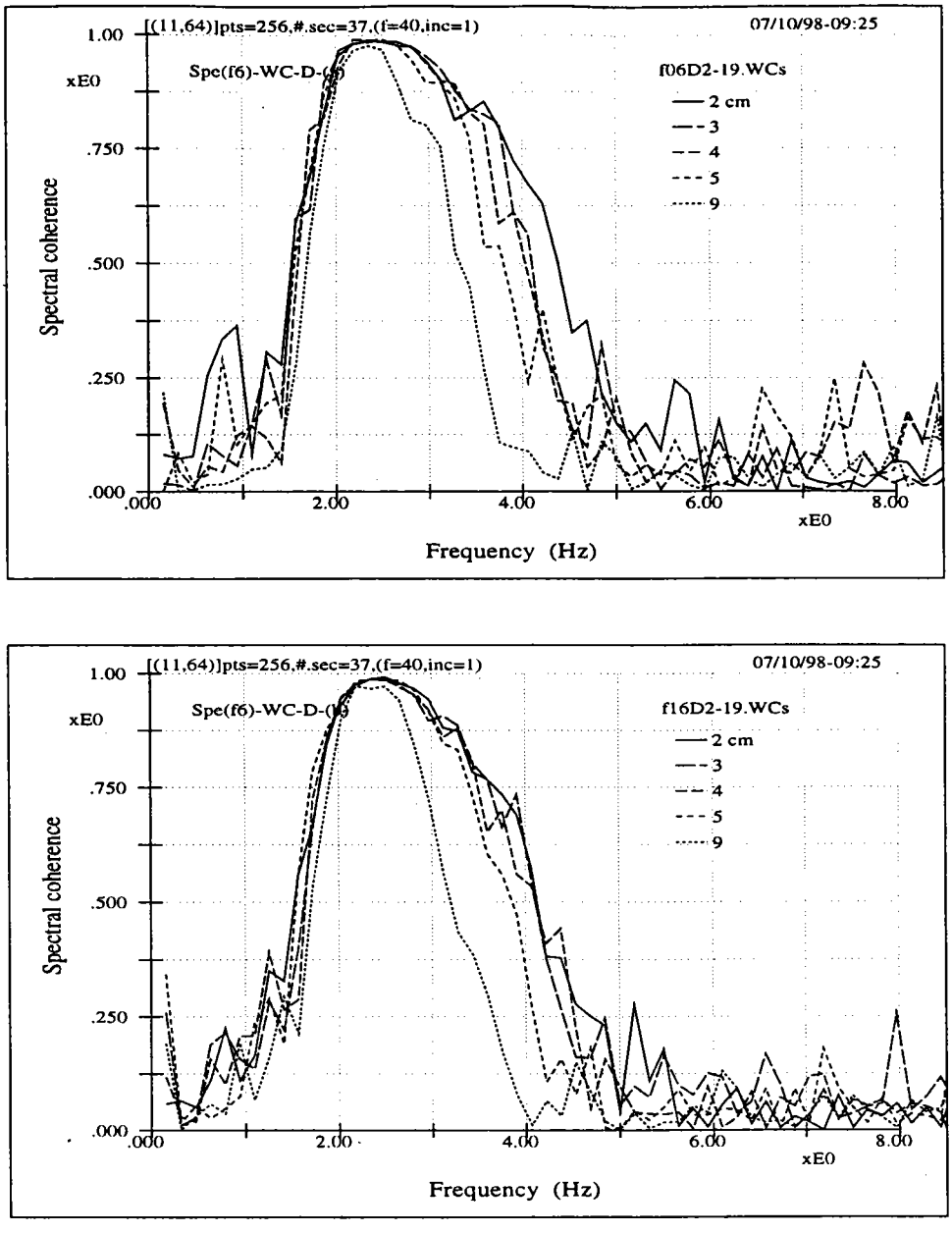

Figure 4.5: The spectral coherences between wave and aqueous flow at different depths without (Top) and with (Bottom) rain (68 mm hr^{-1}) under 6.0 m sec^{-1} wind. Here the length of the data is 9472-point. Even though the function bases of spectral and wavelet coherences are different the spectral coherence curves are seen to approximate those of wavelet coherences; while, there still exists wriggling except in the peak coherence regions.

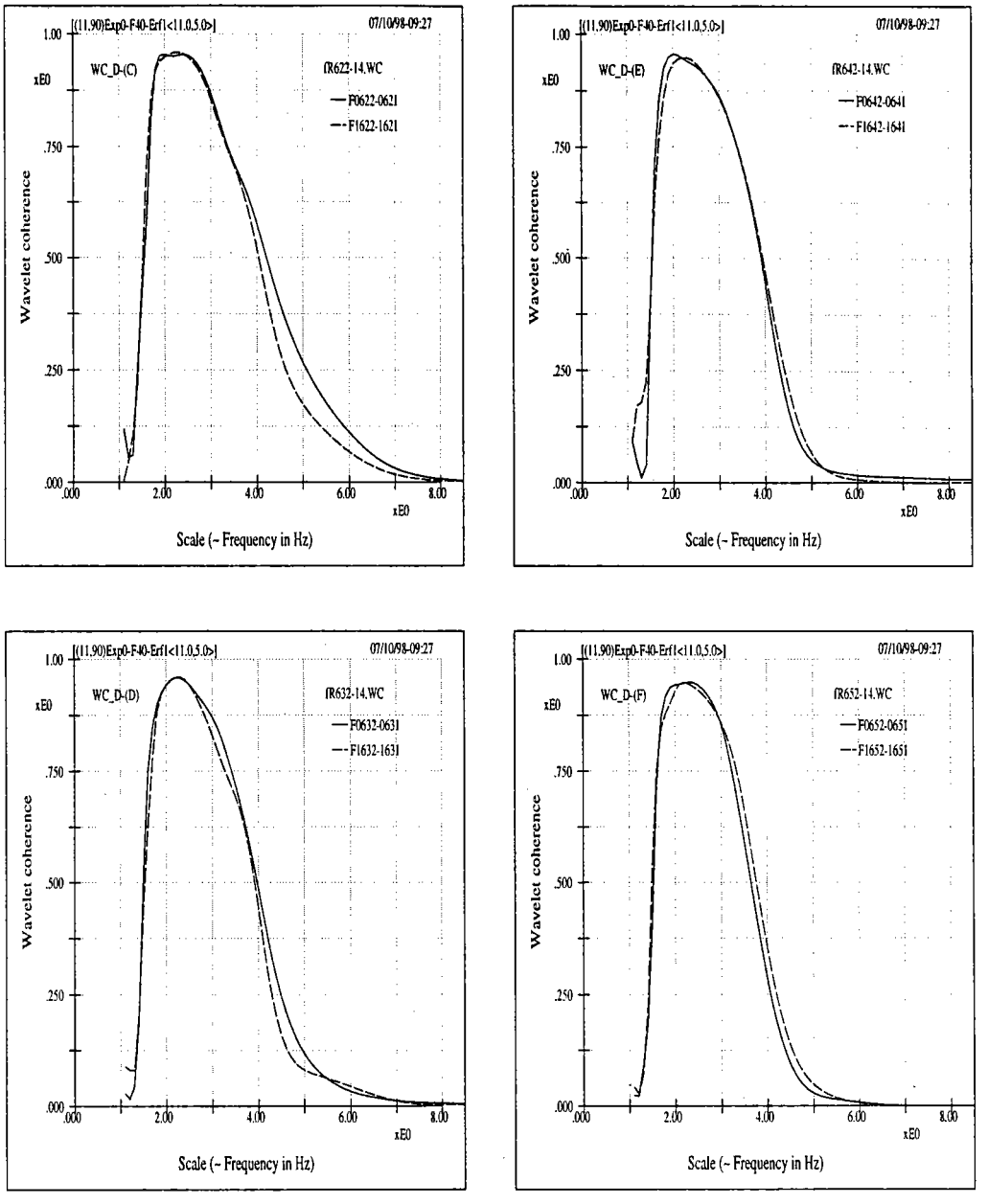

Figure 4.6: Wavelet coherences between wave and aqueous flow with (Dashed) and without (Solid) rain. The rain intensity is 68 mm hr^{-1}, wind speed 6.0 m sec^{-1}, and the aqueous flows are measured at 2 (Top Left), 3 (Bottom Left), 4 (Top Right) , 5 (Bottom Right) cm below the still water surface.

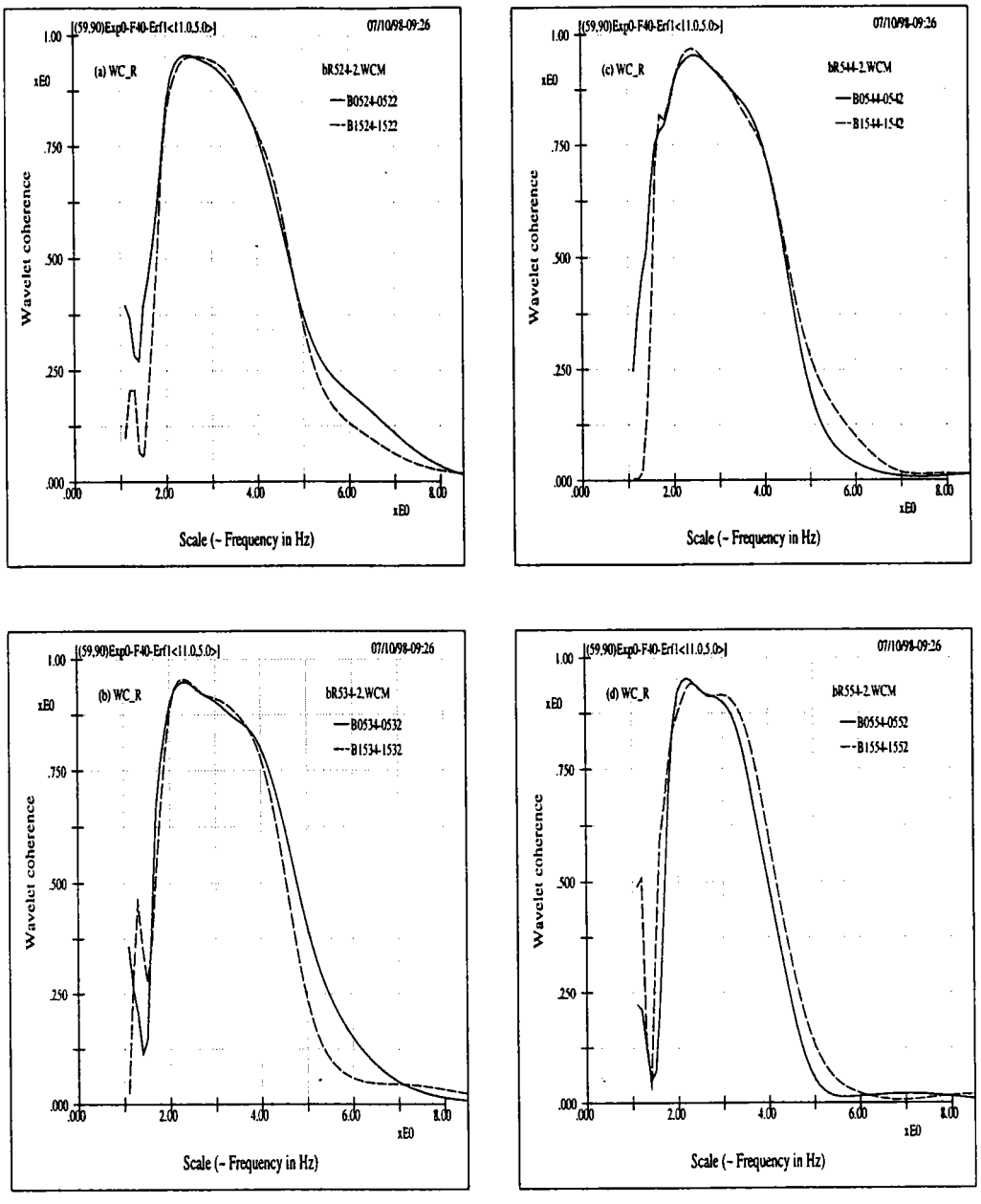

Figure 4.7: Wavelet coherences between wave and aqueous flow with (Dashed) and without (Solid) rain. The rain intensity is 68 mm hr^{-1}, wind speed 5.1 m sec^{-1}, and the aqueous flows are measured at 2 (Top Left), 3 (Bottom Left), 4 (Top Right) , 5 (Bottom Right) cm below the still water surface.

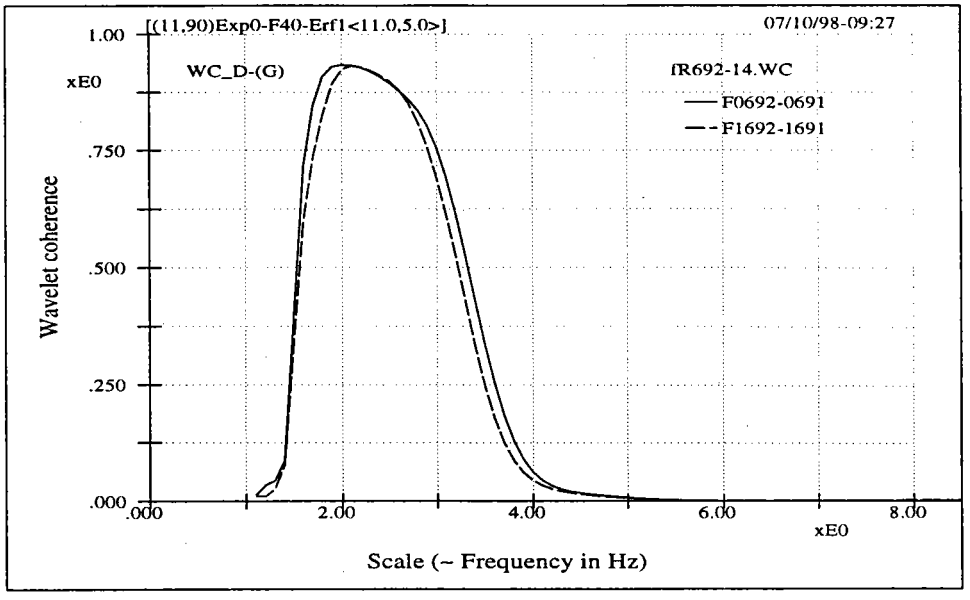

Figure 4.8: Rain's effects on wave-current coherence for aqueous flow measured at 9cm below still water surface. It is seen that rain reduces the overall coherence in deeper region.

at low frequency side for aqueous flow measured at depth 5cm. This probably indicates that rain causes energy to propagate into deep region and since only internal waves can transfer energy in the vertical direction we argue that there is a chance for the growth of internal waves.

4.2.4 Non-concurrent or different localities cases

All the above figures focus on wave-current coherences using data acquired at a cross section right behind the raining segment and the wave and aqueous flow are measured simultaneously. Let examine a few different kind of coherences where data are either not measured simultaneously or not at the same cross section. Such as the cases of current-current coherences between different depths and wave-wave coherences between different wave gauge data (the results of the later are not shown here, since they mainly reveal the trivial information of negligible coherences). Figure 4.15 shows the wavelet coherence between the aqueous flows measured at different depths with and without rain for two different wind speeds. Here the section of measurement is right under the raining segment (70 cm ahead of the end of rain module). It is interesting to note that without rain the coherences basically are unimportant and there exists no significant feature; however, under the action of

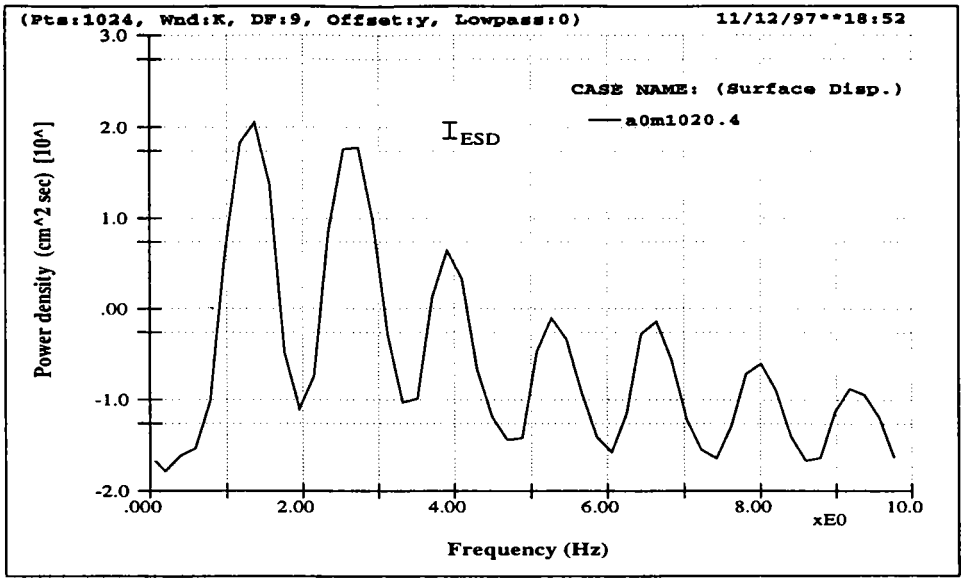

Figure 4.9: Power spectrum of Stokes wave with fundamental harmonic at about 1.4 Hz and corresponding wave slope of 0.06. Its related wave-current coherences are shown in figures 4.12 and 4.13.

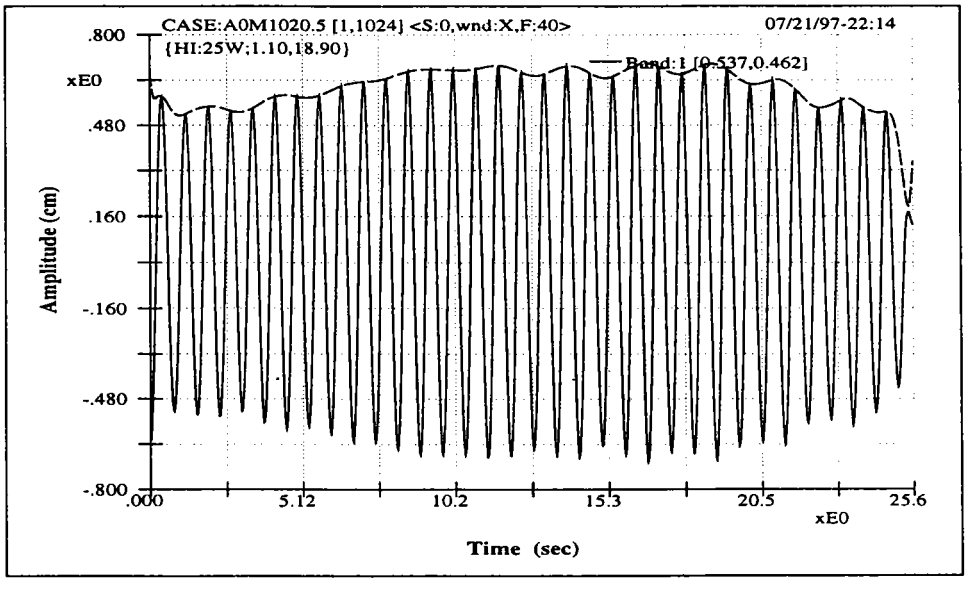

Figure 4.10: Amplitude modulations for the frequency band corresponding to fundamental harmonics of the stokes wave shown in figure 4.9. Representative value of wave slope is estimated from the envelope curve.

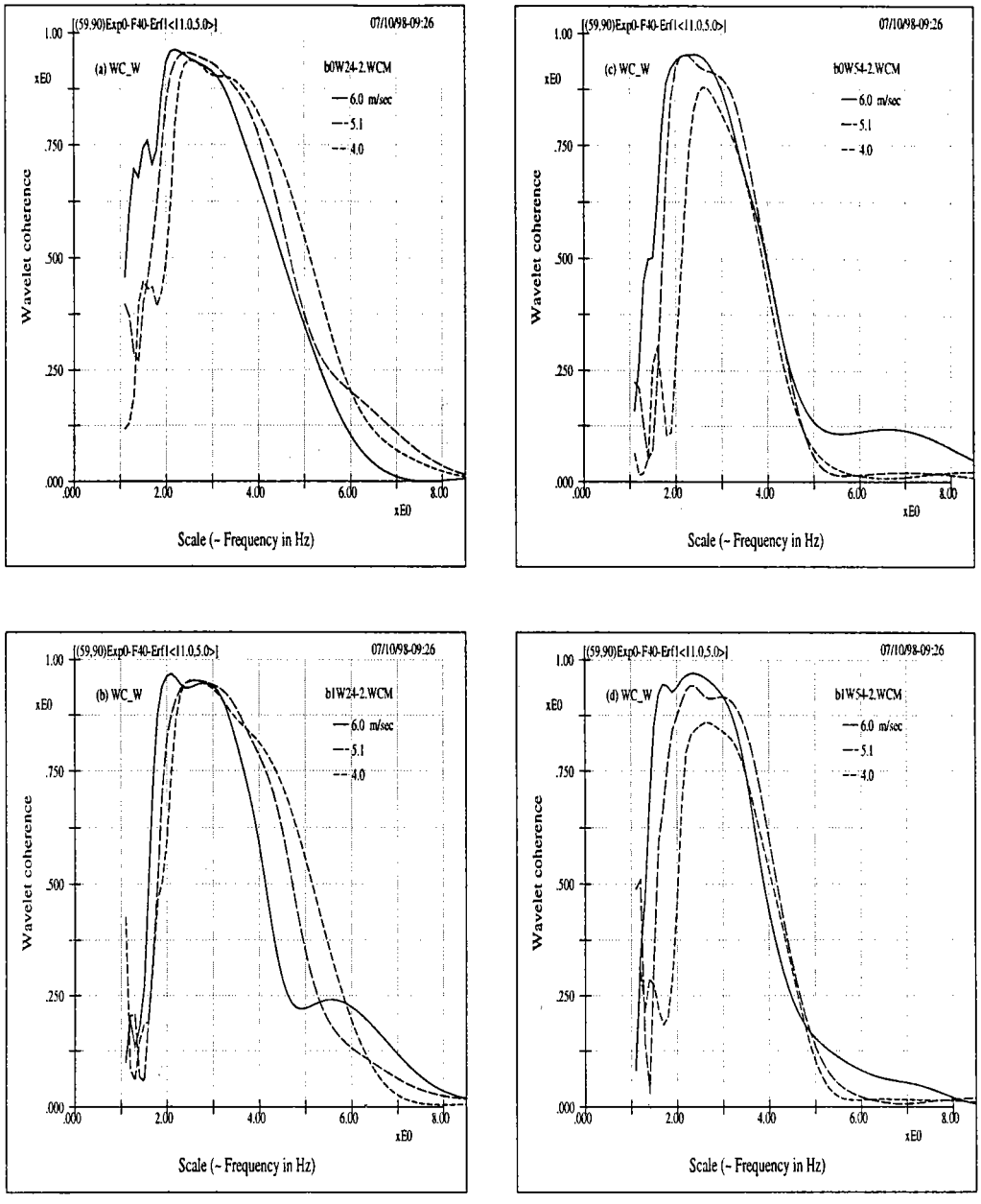

Figure 4.11: Wavelet coherences between wave and aqueous flow under different wind speeds without (Top) and with (Bottom) rain (68 mm hr^{-1}) for aqueous flows measured at depths of 2 (Left) and 5 (Right) cm. Since under higher wind speed the wave and current fields are more developed or in a more mature stage, the peak coherence increases and shifts to lower frequency.

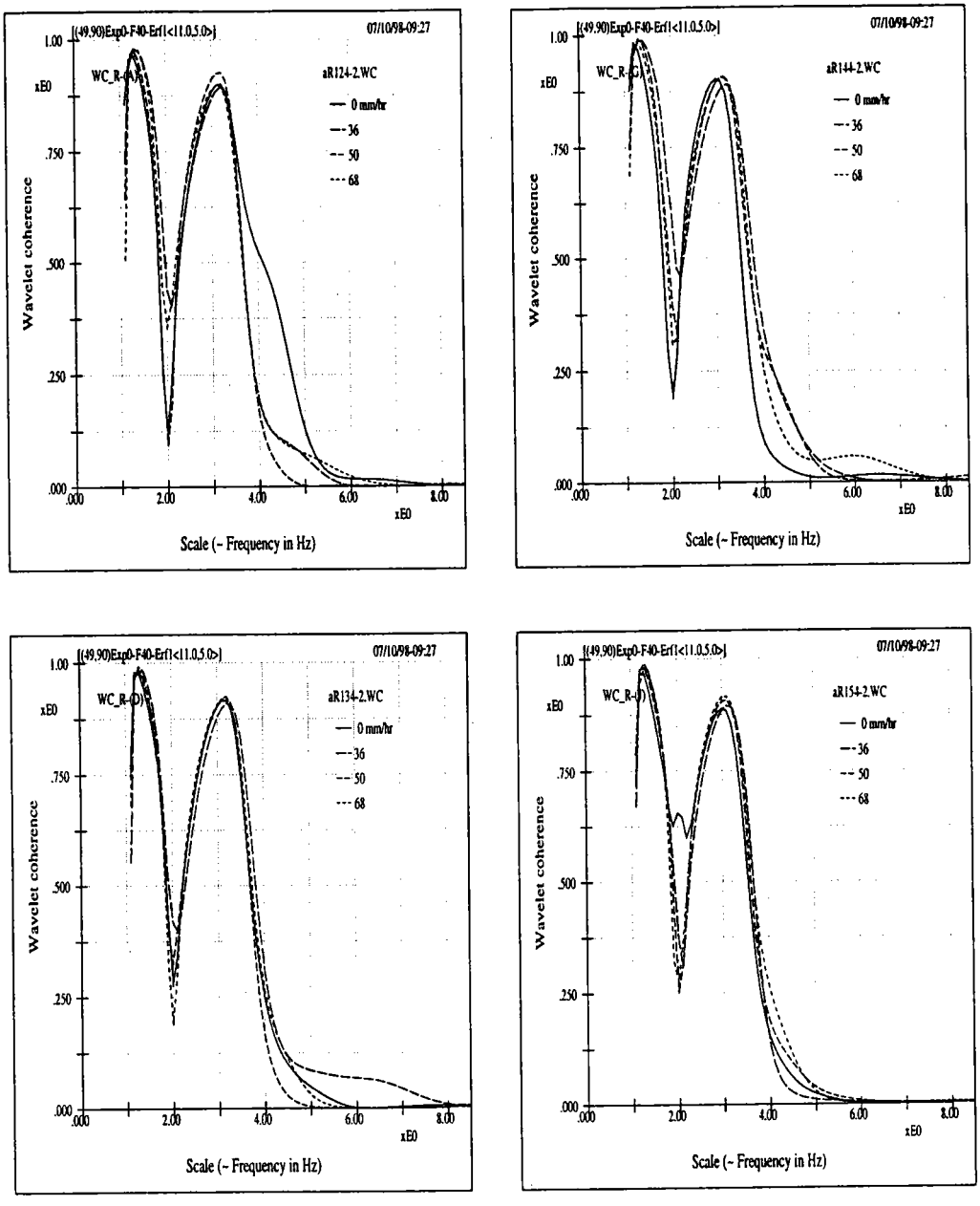

Figure 4.12: The wavelet coherences between wave and aqueous flow under different rain intensities for Stokes wave with fundamental harmonic at 1.4 Hz and wave slope of 0.06. The aqueous flows were measured at 2 (Top Left), 3 (Bottom Left), 4 (Top Right), and 5 (Bottom Right) cm below the still water surface.

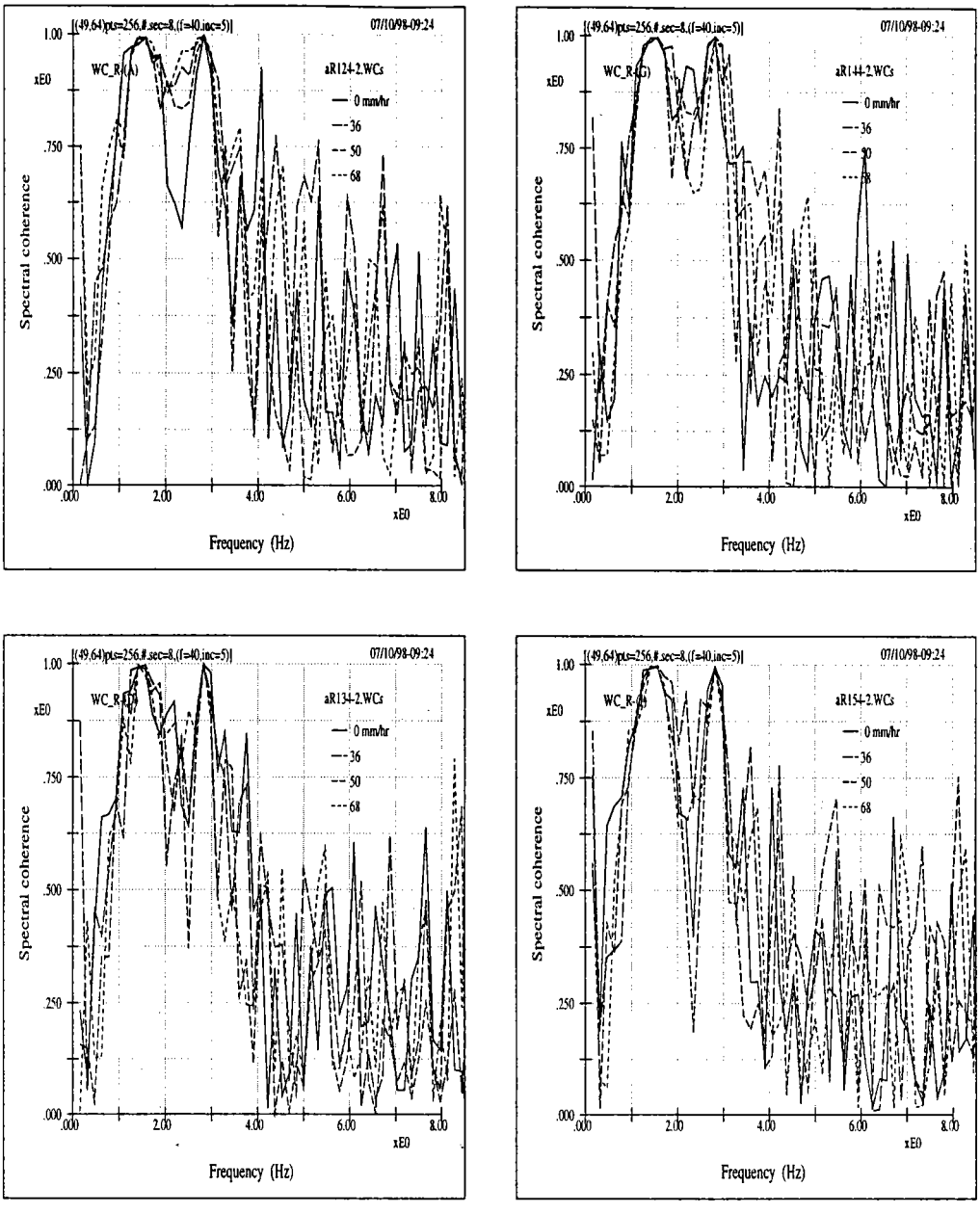

Figure 4.13: Curves of spectral coherence using the same data set as last figure. It is interesting to note that even for such an extremely stationary condition the performance of spectral coherence is relatively poor. Together with the previous figure it can be concluded that a redundant system has obvious advantages over an orthonormal system when studying the coherence statics. Most importantly, the two figures together point out that wavelet viewpoint can still provides better descriptions of the intrinsic physics even for stationary signals, and this must be related to the appropriateness of the modulation nature of waves.

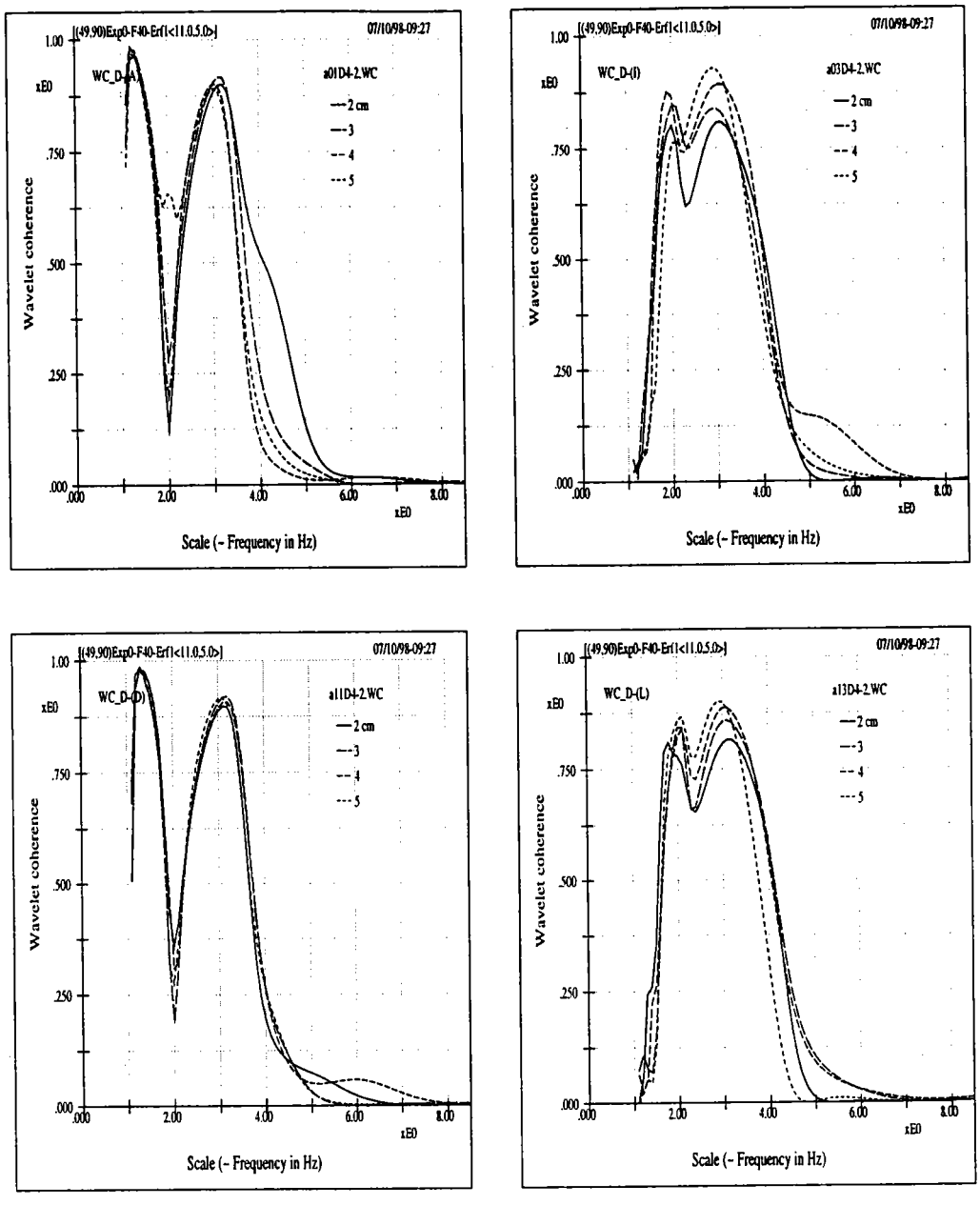

Figure 4.14: The wavelet coherences between wave and aqueous flow measured at the specified depth without (Top) and with (Bottom) 45 mm hr^{-1} rain for two Stokes waves, one with 1.4 Hz fundamental harmonic and 0.06 wave slope (Left), the other with 2.7 Hz fundamental harmonic and a 3.0 high wave slope (Right). The high wave slope significantly reduces the coherences near the surface regions.

rain there are very consistent and unambiguous features for both wind speeds, as seen from the two significant peaks corresponding to the main frequency of the wind waves. In the figure we see that the coherences of aqueous flows at 3 and 4 cm depths are quite small; while the coherences for 4 and 5 cm and for 3 and 5 cm both have similar sharp peaks for the two wind speeds. In contrast, for current-current coherences with measurement cross section right behind the raining segment these features disappear, as are shown in figure 4.16. Overall, these features provide further support of our argument that rain has the role of a beating process that induces interactions and causes the development of waves of certain scales. It is also noted that right now the level of coherence is very low when compared with that of simultaneous measurements. We also mention that, at first sight, it seems that the high coherence for aqueous flows at 3 and 5 may be intuitively dubious since the coherences between 3 and 4 cm are low. But we argue that the reasons should be attributed to the non-orthogonal wavelet constituents as opposed to the mutual exclusiveness of the orthogonal wave constituents, as related to an issue of the frequency leakage to be discussed in an further study on a topic regarding a wavelet variant for refined ridge extraction.

4.3 Summary

It is probably not necessary for us to give the many figures just for showing the serviceability of the chosen function basis and its associated method of analysis; but these figures reveal many reasonable aspects of physics and justify the reasoning for a few energy phenomena in the multi-scale coupling system concerning wind, wave, and rain interactions. Overall, they help materialize the links of the main three subject topics and validate the appropriateness of a best analyzing function basis. Most profoundly, we are further convinced that waves are intrinsically modulating and instability [4, 5, 19] should be a common nature of water waves; therefore, wavelet approach is a means with more natural perception even for stationary wave fields.

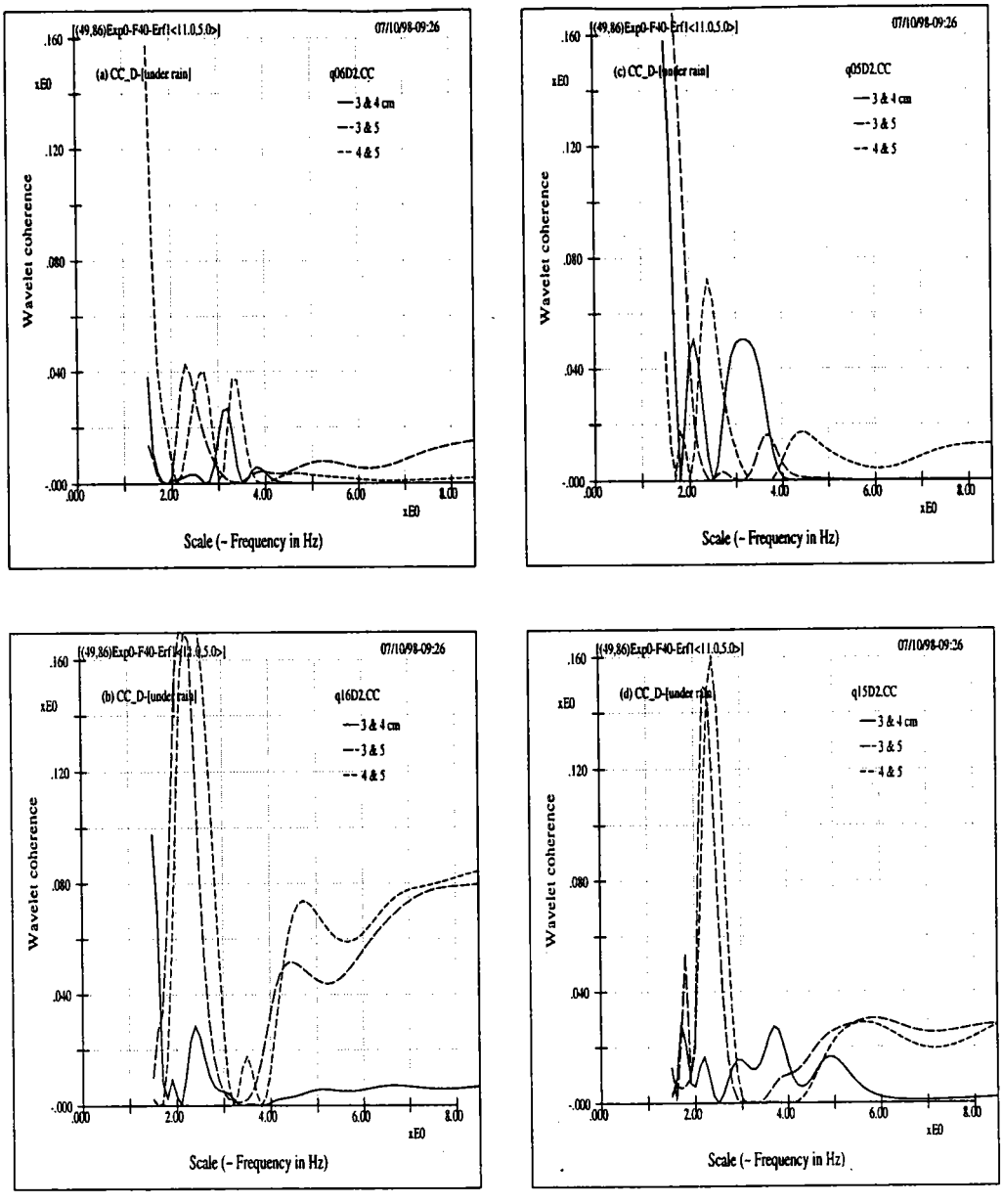

Figure 4.15: The wavelet coherences between aqueous flows measured at different depths without (Top) and with (Bottom) rain (68 mm hr^{-1}) under two different wind speeds of 5.8 (Left) and 5.0 (Right) m sec^{-1}. Here the measurement cross section is right under the raining segment. Since the aqueous flows were not measured simultaneously, the coherences are relatively low. However, the two significant peaks show unambiguous tendencies which indicate the tuning and de-tuning process related to rain drop's beating effects on the interaction process as explained in the text. Additional point regarding the effects of the non-orthogonal wavelet constituents as opposed to the mutual exclusiveness of the orthogonal wave constituents is stated in the text.

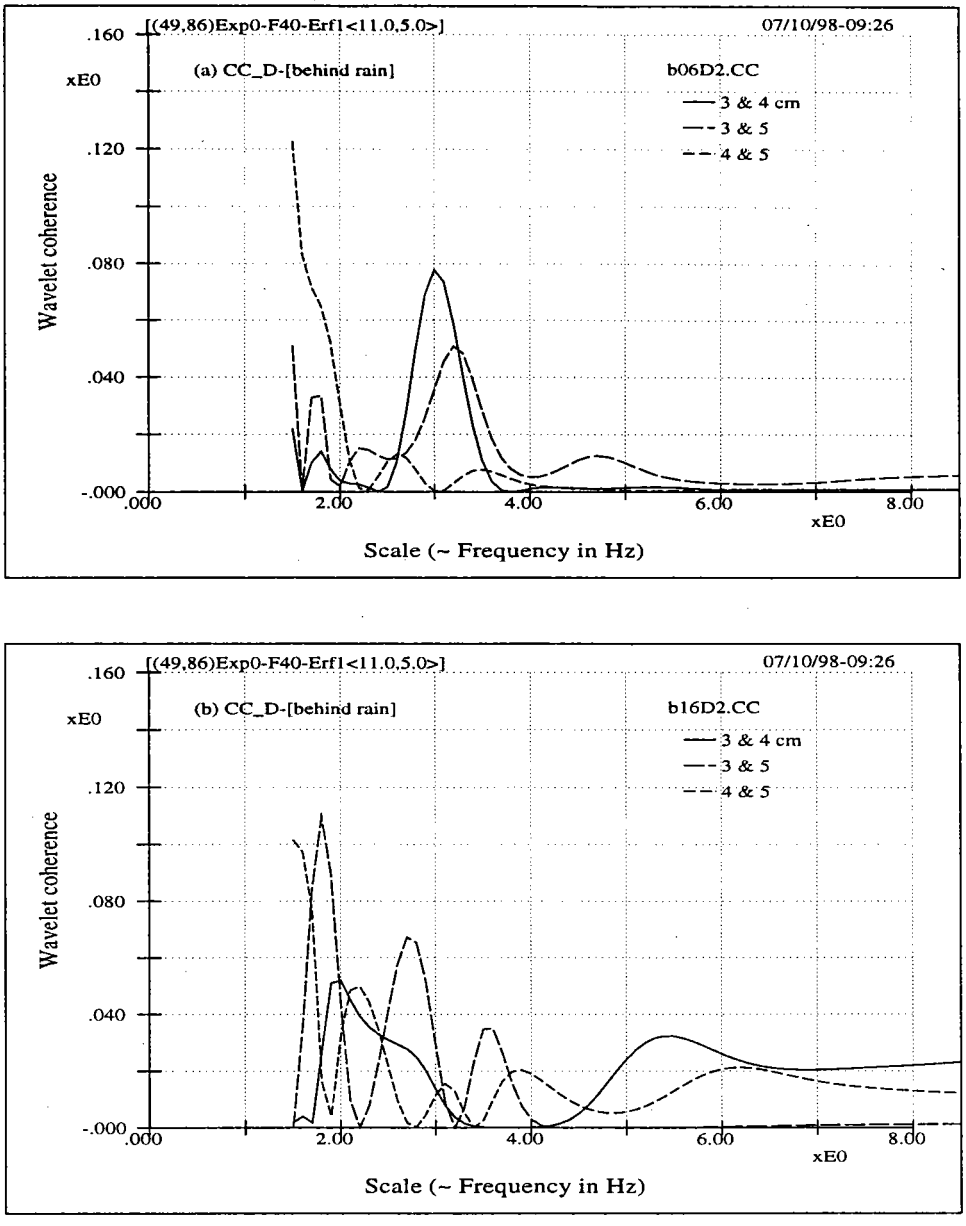

Figure 4.16: The wavelet coherences between aqueous flows measured at different depths without (Top) and with (Bottom) rain (68 mm hr^{-1}) under two different wind speeds of 5.8 (Left) and 5.0 (Right) m sec^{-1}. Here the measurement cross section is right behind the raining segment. In contrast to the cases where measurement cross section was right under the raining segment, the two peaks disappear and no trend is observed.

Chapter 5

Conclusions

Through the studies on the coherent features in the wave and current fields in a laboratory water tank we tie their significant outcomes to the results of two earlier papers, which are also mostly incorporated here, but with many refined statements. We explicate the close links among the three subject topics and disclose the advantages and disadvantage of various function bases, including Fourier basis and a vast array of wavelet bases. In all the context, we have tried to put the most practical and physical touches — that we are capable of — to the descriptions of the analytical counterparts and illustrate the reasons that facilitate the usefulness of the proposed analyzing function basis and method of analysis.

In short, we first identify among discrete Riesz wavelets the most intimate counterpart to the analyzing function we use, and then point out its affinity to the modulated Gaussian wavelet based on continuous wavelet transform using adapted time-frequency windows. The adaptation scheme serves two purposes: enhancing wavelet's affiliations with the physics of water waves and clarifying a few concepts related to intrinsic properties of wavelet analysis that lead to the success of the present application.

Besides, we are further convinced that waves are intrinsically modulating and instability of water waves is quite common to various physical processes — even for stationary conditions. Hence, descriptions based on localized wave constituents should reveal better physics.

In the interests of further studies we also like to point out that effects of phase noise for non-concurrent or displaced measurements are still serious and that the present adaptation of time-frequency windows is still rather heuristic.

Bibliography

[1] Auscher, P. Wavelet bases for l^2(R) with rational dilation factor. In M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, editor, *Wavelets and Their Applications*, pages 439–452. Jones and Bartlett Publishers, Boston, New York, USA, 1992.

[2] Battle, G. Cardinal spline interpolation and the block spin construction of wavelets. In C.K. Chui, editor, *Wavelets: A tutorial in Theory and Applications*, pages 73–90. Academic Press, Inc., San Diego, California, USA, 1992.

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

[4] Brooke Benjamin, T. Instability of periodic wave trains in nonlinear dispersive systems. *Proc. Roy. Soc. Lond.*, A299:59–75, 1967.

[5] Brooke Benjamin, T. and J. E. Feir. The disintegration of wave trains on deep water. *J. Fluid Mech.*, 27:417–430, 1967.

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

[7] Chui, C.K. On cardinal spline-wavelets. In M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, editor, *Wavelets and Their Applications*, pages 439–452. Jones and Bartlett Publishers, Boston, New York, USA, 1992.

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

[9] Coifman, R., Y.Meyer, and M.V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Tran. on Inform. Theory*, 38(2):713–718, 1992.

[10] Coifman, R., Y.Meyer, and M.V. Wickerhauser. Size properties of wavelet packets. In M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, editor, *Wavelets and Their Applications*, pages 453–470. Jones and Bartlett Publishers, Boston, New York, USA, 1992.

[11] Coifman, R., Y.Meyer, and M.V. Wickerhauser. Wavelet analysis and signal processing. In M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, editor, *Wavelets and Their Applications*, pages 153–178. Jones and Bartlett Publishers, Boston, New York, USA, 1992.

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

[13] Froment, J. and, S. Mallat. Second generation compact image coding with wavelets. In C.K. Chui, editor, *Wavelets: A tutorial in Theory and Applications*, pages 655–678. Academic Press, Inc., San Diego, California, USA, 1992.

[14] 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.

[15] 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.

[16] Lee, Y.R. Signal analysis from wave modulation perspective. Technical report, No.7, Institute of Harbor and Marine Technology, Taichung, Taiwan, 1997.

[17] 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.

[18] 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.

[19] Lee Y.R. and J. Wu. Time-frequency features and side band instability. *Proc. 19th Conf. on Coastal Engineering in Taiwan*, pages —, 1997.

[20] Lighthill, M.J. Some special cases treated by whitham theory. *Proc. Roy. Soc. Lond.*, A299:38–53, 1967.

[21] P. Liu. Wavelet spectrum analysis and ocean wind waves. In E.F. Georgiou, editor, *Wavelets in Geophysics*, pages 151–166. Academic Press, Inc., San Diego, California, USA, 1994.

[22] Longuet-Higgins, M.S. The instability of gravity waves of finite amplitude in deep water I, Superharmonics. *Proc. Roy. Soc. Lond.*, A360:471–488, 1978.

[23] Longuet-Higgins, M.S. The instability of gravity waves of finite amplitude in deep water II, Subharmonics. *Proc. Roy. Soc. Lond.*, A360:489–505, 1978.

[24] Mallat, S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Tran. on Pattern Anal. and Mach. Intel.*, 11(7):674–693, 1989.

[25] Mallat, S. and S. Zhong. Wavelet transform maxima and multiscale edges. In M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, editor, *Wavelets and Their Applications*, pages 67–104. Jones and Bartlett Publishers, Boston, New York, USA, 1992.

[26] Melville, W. K. The instability and breaking of deep water waves. *J. Fluid Mech.*, 115:165–185, 1982.

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

[28] Phillips, O.M. *The Dynamics of the Upper Ocean.* Cambridge University Press, 2 edition, 1977.

[29] 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.

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

[31] Tennekes, H., and J.L. Lumley. *A First Course in Turbulence.* The MIT Press, Cambridge, Massachusetts, USA, 1972.

[32] Whitham, G. B. Nonlinear dispersion of water waves. *J. Fluid Mech.*, 27:399–412, 1967.

[33] Wickerhauser, M.V. Acoustic signal compression with wavelet packets. In C.K. Chui, editor, *Wavelets: A tutorial in Theory and Applications*, pages 679–700. Academic Press, Inc., San Diego, California, USA, 1992.

[34] Wickerhauser, M.V. Comparison of picture compression methords: wavelet, wavelet packet, and local cosine. In C. K. Chui, editor, *Wavelets: Theory, Algorithms, and Applications*, pages 585–621. Academic Press, Inc., San Diego, California, USA, 1994.

[35] Yuen, H. C. and B. M. Lake. Nonlinear deep water waves: Theory and experiment. *Phys. Fluids*, 18:956–960, 1975.