RICE UNIVERSITY

Directional, Shift-Insensitive, Complex Wavelet Transforms with Controllable Redundancy

by

Felix C. A. Fernandes

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

Approved, Thesis Committee:

C. Sidney Burrus, Chairman
Dean of Engineering
Maxfield and Oshman Professor of
Electrical and Computer Engineering

Richard G. Baraniuk
Professor of Electrical and Computer
Engineering

Steven J. Cox
Professor of Computational and Applied
Mathematics

Houston, Texas

January, 2002

Directional, Shift-Insensitive, Complex Wavelet Transforms with Controllable Redundancy

Felix C. A. Fernandes

Abstract

Although the Discrete Wavelet Transform (DWT) is a powerful tool for signal and image processing, it has three serious disadvantages. First, the DWT is shift sensitive because input-signal shifts generate unpredictable changes in DWT coefficients. Second, the DWT suffers from poor directionality because DWT coefficients reveal only three spatial orientations. Third, DWT analysis lacks the phase information that accurately describes non-stationary signal behavior. To overcome these disadvantages, we introduce the notion of projection-based complex wavelet transforms.

These two-stage, projection-based complex wavelet transforms consist of a projection onto a complex function space followed by a DWT of the complex projection. Unlike other popular transforms that also mitigate DWT shortcomings, the decoupled implementation of our transforms has two important advantages. First, the controllable redundancy of the projection stage offers a balance between degree of shift sensitivity and transform redundancy. This allows us to create a directional, non-redundant, complex wavelet transform with potential benefits for image coding systems. To the best of our knowledge, no other complex wavelet transform is simultaneously directional and non-redundant. The second advantage of our approach is the flexibility to use *any* DWT in the transform implementation. We exploit this flexibility to create the Complex Double-density DWT (CDDWT): a shift-insensitive, directional, complex wavelet transform with a low redundancy of $\frac{3^m-1}{2^m-1}$ in m dimensions. To the best of our knowledge, no other transform achieves all these properties at a lower redundancy. Besides the mitigation of DWT shortcomings, our transforms have unique properties that will potentially benefit a variety of signal processing applications. As an example, we demonstrate that our projection-based complex wavelet transforms achieve state-of-the-art results in a seismic signal-processing application.

To my parents, Silvano and Angela Fernandes,

and my fiancée Helen.

Acknowledgments

I am deeply indebted to my advisor, Professor Sidney Burrus, for sharing his deep insight into signal processing and for his invaluable advice on every aspect of doing research. I am also especially grateful to Professor Richard Baraniuk for his contagious enthusiasm, useful suggestions and generous feedback. I would like to thank Professor Steven Cox for making functional analysis accessible and for his enlightening discussions.

I am fortunate to have had a fruitful collaboration with Rutger van Spaendonck, visiting scholar from Delft University of Technology. It is great fun to work with him and I am happy that he didn't pack his bags when things seemed really stooopid. I am also grateful to Professor Ivan Selesnick for his advice on various issues. His interest in my research has been very encouraging and I have learnt much from him. I would like to thank Professor Manos Papadakis for lengthy discussions on my work. His advice on mathematical matters is much appreciated. Thanks are due to Professor John Hempel for his unique treatment of the First Isomorphism Theorem. It inspired the theory behind the non-redundant projection. I am also grateful to Professor Stephen Semmes for various discussions and for pointing me to useful sources of information.

The Rice DSP group has been a great resource while doing research. I have had a lot of fun doing research in this environment. I am also fortunate to have been a member of Dr. Janice Hewitt's thesis-writing group. Thanks to her, I actually enjoyed writing this dissertation.

I owe a great debt to my parents for all the sacrifices that they made so that I could pursue my dream of doing research. Finally, I would like to thank Helen (and also Mimi and Toto) for her encouragement and support througout this endeavour.

Contents

Illustrations

Chapter 1

Compelling Reasons For Complex Wavelet Transforms

Although the Discrete Wavelet Transform (DWT) is a powerful tool for signal and image processing, it has certain disadvantages that undermine its usage in many applications. In this chapter, we shall briefly describe these shortcomings and survey various techniques that other researchers have adopted to overcome simultaneously some or all of these drawbacks. We shall then explain how our novel complex wavelet transforms simultaneously overcome all DWT shortcomings. Moreover, the implementation of our complex wavelet transforms offers the additional benefits of flexibility and controllable redundancy. These benefits enable the creation of new transforms with unprecedented properties.

In 1982, Grossman and Morlet introduced the Continuous Wavelet Transform in their seminal paper on function decompositions with atoms that they called "wavelets" [1]. This transform provides significant advantages over traditional Fourier and Gabor analysis. Unfortunately, due to its continuous-time ancestry, digital implementations of the Continuous Wavelet Transform on real-world signals are impractical. Then in 1988, Mallat and Meyer [2] combined ideas from the disciplines of harmonic analysis, signal processing, physics and computer vision. This cross pollination of concepts led to the multiresolution analysis, a new paradigm for the analysis and processing of functions.

To implement the multiresolution analysis on real-world signals and images, Mallat and Meyer devised a fast algorithm called the Discrete Wavelet Transform (DWT). Since the DWT is a critical sampling of the Continuous Wavelet Transform, it provides a practical technique for the wavelet analysis of real-world signals and images. In 1988, Daubechies [3, 4] designed a new family of wavelets that enables the DWT to provide a fast, local, sparse, orthogonal, multiresolution analysis of real-world signals and images. These important properties are responsible for the popularity of the DWT. Unfortunately, the DWT also has a few disadvantages that seriously undermine its application to signal- and image-processing applications. We explain these disadvantages in the following section. We then survey various techniques that mitigate some or all of these disadvantages simultaneously. In the remainder of the dissertation, we shall demonstrate persuasively that there are compelling reasons to use our complex wavelet transforms for the simultaneous mitigation of all DWT shortcomings.

1.1 Disadvantages of the DWT

Although the DWT is a powerful tool, it does have three disadvantages that undermine its application to certain signal- and image-processing tasks. We briefly explain these disadvantages below. In subsequent chapters, we shall dwell on these shortcomings in greater detail.

1. Shift Sensitivity: A transform is shift sensitive if an input-signal shift causes an unpredictable change in transform coefficients. Strang observed that the DWT is seriously disadvantaged by the shift sensitivity that arises from downsamplers in the DWT implementation [5]. Shift sensitivity is an undesirable property because it implies that DWT coefficients fail to distinguish between input-signal shifts.

2. Poor Directionality: An m-Dimensional transform suffers poor directionality when the transform coefficients reveal only a few feature orientations in the spatial domain. Since the separable 2D DWT partitions the frequency domain into three directional subbands, it enables distinction between only three

spatial-domain feature orientations: horizontal, vertical and diagonal. Natural images are comprised of smooth regions that are punctuated with edges at several different orientations; consequently, poor directionality compromises the optimality of the DWT representation of natural images.

3. Absence of Phase Information: In their seminal papers [6, 1], Grossman et al. claimed that they were motivated to develop the continuous wavelet transform because traditional Gabor analysis is incapable of decomposing functions using Hardy-space atoms. Such a decomposition provides phase information that greatly benefits signal-processing applications. Other researchers have elaborated on the advantages of this phase information. Granlund and Knutsson [7, Pg. 169] demonstrate that the phase information provides a description of the amplitude and local behavior of a function. Beylkin and Torresani [8] show that the amplitude-phase representation of a function is less oscillatory than the function itself. This representation imparts greater robustness to signal-processing algorithms. Unfortunately, most DWT implementations use filters associated with real-valued wavelets. Consequently, signal-processing algorithms are deprived of the significant benefits of phase information.

In the next section, we shall explore the approaches that various researchers have proposed to overcome simultaneously some or all of these DWT deficiencies. We shall then outline additional compelling reasons to use our complex wavelet transforms for the simultaneous mitigation of all DWT shortcomings.

1.2 Mitigating DWT Disadvantages

In the previous section, we discussed some of the disadvantages of the DWT. Several researchers have proposed different techniques to overcome these drawbacks and create a robust, informative transform for signal-processing applications. Since most methods tackle a single disadvantage, we shall group the methods according to their targeted disadvantages.

1.2.1 Reducing Shift Sensitivity

Since downsamplers in the DWT implementation create shift-sensitivity, Mallat [9], Beylkin [10] and Guo [11] devised the undecimated DWT, a wavelet transform without downsamplers. Although the undecimated DWT is shift insensitive, it has high transform redundancy due to the absence of downsamplers. Unfortunately, the high transform redundancy incurs a massive storage requirement that makes the undecimated DWT inappropriate for most signal processing applications. As an alternative to the undecimated DWT, Liang and Parks [12], Benno and Moura [13], and Bao et al. [14] used best-basis methods or optimal wavelet-designs to reduce DWT shift-sensitivity without any increase in transform redundancy. More recently, Selesnick and Sendur [15] formulated the Double-Density Wavelet Transform (DDWT), a low-redundancy DWT extension with reduced shift sensitivity. However the DDWT has poor directionality and offers no phase information. In Chapter 6 we shall demonstrate that the DDWT may be enhanced with improved directionality and phase information, while maintaining its reduced shift-sensitivity and low redundancy.

In 1992, Simoncelli et al. [16] confronted DWT shift sensitivity with a radically different approach. Instead of minimizing the unpredictable responses to input-signal shifts, they defined "shiftability", a new measure of reduced shift sensitivity. According to their definition, a transform is shiftable if and only if transform-subband energy is invariant under input-signal shifts. Shiftability is important because it is equivalent to interpolability, a property ensuring the preservation of transform-subband information under input-signal shifts. Simoncelli et al. and Wang [17] argued that since the Nyquist criterion is necessary and sufficient for shiftability, the only shiftable, non-redundant DWT is the DWT based on the Shannon wavelet. Since this particular DWT is practically unrealizable, Simoncelli et al. designed the steerable pyramid, a highly redundant, non-separable, directional, multiscale transform that attains approximate shiftability.

Abry [18] first demonstrated that approximate shiftability is possible in a DWT with a small, fixed amount of transform redundancy. He designed a pair of real wavelets such that one is approximately the Hilbert transform of the other. This wavelet pair defines a complex wavelet transform in the following sense. Consider the pair of DWT trees that implement the DWTs associated with the wavelet pair. A complex coefficient is obtained by interpreting the wavelet coefficient from one DWT tree as the real part of the complex coefficient, while the corresponding wavelet coefficient from the other tree is interpreted as the imaginary part. Magarey and Kingsbury [19, 20] showed that this complex wavelet coefficient can also be obtained using a single DWT tree based on a quasi-analytic wavelet. The complex-coefficient filters used to implement this DWT have approximately half the bandwidth of a DWT tree associated with a real-valued wavelet; consequently, the Nyquist criterion is approximately satisfied in each subband and so the DWT has approximate shiftability. However, since the quasi-analytic DWT is not an invertible transform, it can be used for analysis, but not for processing.

To overcome this disadvantage, Kingsbury [21] subsequently developed the Dual-Tree Wavelet Transform (DTWT), a quadrature pair of DWT trees similar to Abry's. Both the DTWT and Abry's transform are invertible and achieve approximate shiftability; however, the design of these quadrature wavelet-pairs is quite complicated. As an alternative, Selesnick [22, 23] recently formulated a general spectral-factorization method to design a quadrature wavelet-pair with specified length and vanishing-moment multiplicity. In Chapter 3, we prove that one of our complex wavelet transforms achieves approximate shiftability at the same redundancy as the DTWT. Then in Chapter 6, we demonstrate that another complex wavelet transform enjoys reduced shift sensitivity at lower redundancy than the DTWT.

1.2.2 Improving Directionality

It is widely known that a spatial feature with a well-defined orientation manifests itself as a wedge-shaped concentration of energy in the Fourier domain [7, Chapter 6]. Consequently, the intuitive approaches to improved directionality use non-separable filters with wedge-shaped passbands. We shall first examine these methods and then discuss separable transforms with improved directionality.

Non-separable Transforms with Improved Directionality

In 1992, Bamberger and Smith [24] introduced the Directional Filter Bank (DFB). This invertible, non-redundant transform uses an ingenious, non-separable filtering scheme to obtain wedge-shaped passbands that provide excellent directionality. Later, Park et al. [25] used linear phase filters to eliminate phase distortion in the DFB subbands. The DFB has two main disadvantages. First, the DFB basis functions are not smooth. Second, the DFB does not provide a multiscale decomposition: features with the same orientation but different scale are indistinguishable. Do and Vetterli [26] recently addressed both drawbacks. They incorporated regularity constraints into the DFB design to create smooth basis functions. By combining this version of the DFB with the Laplacian pyramid, they obtained the pyramidal directional filter bank, a multiscale implementation of the DFB.

The approximately shiftable, steerable pyramid described in Section 1.2.1 is another non-separable, multiscale transform that uses wedge-shaped passbands to improve directionality. To provide approximate shiftability, the steerable pyramid must have higher transform redundancy than the non-redundant DFB. Unlike the DFB, the spatial-domain implementation of the steerable pyramid does not offer perfect reconstruction: the inverse steerable-pyramid transform reconstructs only an approximation to the steerable-pyramid input.

Both the DFB and the steerable pyramid improve directionality by using non-separable filters with wedge-shaped passbands. A different approach to improving directionality is to design transforms that provide optimal representations for edges. Donoho [27] and Claypoole [28] defined wedgelets, an overcomplete collection of indicator functions at various locations, scales and orientations. These basis functions provide optimal representations for certain classes of edges. To obtain an optimal representation for lines in images, Candès [29] and Donoho [30]used a novel Radon-domain wavelet decomposition to define a set of analysis functions called ridgelets. Since ridgelets do not provide a multiscale decomposition, they then defined a new set of atoms that they call curvelets [31, 32]. These atoms enable a multiscale ridgelet-based analysis that yields optimal representations for objects that are smooth away from discontinuities across curves. Note that the pyramidal directional filter bank that we described earlier in this section uses curvelet-like atoms, yet enjoys a practical digital implementation.

Separable Transforms with Improved Directionality

The non-separable transforms described in the preceding subsection have excellent directionality. Unfortunately, such transforms are more difficult to design and implement than are separable transforms. Consequently, to improve directionality using separable transforms, Watson [33] and Burns [34] demonstrated that separable analytic[1] filters provide a Fourier-plane tiling that is denser than the tiling obtained with real-coefficient filters. This denser tiling yields a large number of directional subbands that can be used to discriminate between several different feature-orientations. Hence separable transforms that use analytic filters can provide improved directionality. In the quasi-analytic DWT that we discussed in Section 1.2.1, Magarey and Kingsbury exploited this technique to provide six directional subbands. In contrast, the DWT has only three directional subbands. Kingsbury and Selesnick also used this technique to improve the directionality of the Dual Tree Wavelet Transform (DTWT) described in Section 1.2.1. In Chapters 5 and 6 we shall demonstrate that separable analytic

[1] An analytic filter is a complex-coefficient filter that has a stopband in the interval $[-\pi, 0)$.

filters may be used to improve the directionality of our complex wavelet transforms as well.

1.2.3 Providing Phase Information

As explained in Section 1.1, most DWT implementations use real-coefficient filters associated with real-valued wavelets. When processing real-valued signals, these DWTs produce real-valued transform coefficients and hence phase information is absent. This is a significant disadvantage because the phase of complex transform-coefficients offers a wealth of information that may be exploited by signal processing applications. Consequently, many researchers have explored the use of multiscale transforms with complex-valued coefficients. We shall now outline some of their approaches.

The initial DWTs with complex-coefficient filters were obtained as new complex solutions to existing wavelet systems. In 1993, Lawton [35] argued that orthogonal Daubechies wavelets cannot use symmetric filters unless those filters have complex-valued coefficients. He obtained a complex symmetric solution for Daubechies scaling and wavelet filters and showed that these filters enable a DWT implementation with reduced boundary effects. Subsequently, Lina [36] provided a complete parameterization for all complex Daubechies wavelets and then used the associated complex filters to implement a DWT. He demonstrated that the imaginary part of the complex scaling coefficients from this DWT is a smoothed version of the Laplacian of the input. He exploited this fact in various image-processing applications. In 1999, Zhang et al. [37] provided a complete parameterization for all complex-valued, orthogonal wavelets. They also showed that symmetric, complex-coefficient filters do not have linear phase.

Other researchers recognized that analytic[2] wavelets are extremely useful in applications. Cohen [38, 39] designed orthogonal, complex, quasi-analytic scaling functions and wavelets. Unlike the quasi-analytic scaling filters and wavelets used by

[2]Analytic functions have frequency responses that vanish over negative frequencies.

Magarey and Kingsbury in Section 1.2.1, Cohen's scaling and wavelet filters do not have reduced bandwidths; therefore, Cohen's quasi-analytic DWT is not shiftable. Nevertheless, he proposed that the quasi-analytic DWT is useful for the analysis of the analytic part of an acoustic signal. Later, Bernard [40] proved that the smooth envelope of an analytic wavelet enables the stable estimation of optical flow. He demonstrated that a wavelet decomposition with a quasi-analytic wavelet may be obtained by filtering the real-valued, DWT bandpass subbands with a quasi-analytic filter.

The approximately shiftable, directional Dual Tree Wavelet Transform (DTWT) introduced in Section 1.2.1 provides phase information in its complex transform coefficients. This phase information is associated with a quasi-analytic wavelet decomposition and has been exploited for various image-processing applications. Romberg et al. [41] created a Hidden Markov Tree model with the complex DTWT coefficients and performed image denoising and texture classification with the model. De Rivaz and Kingsbury [42] exploited the similarity between the analytic complex wavelets and Gabor wavelets for image retrieval. Reeves and Kingsbury [43] devised a coarse-to-fine scale-prediction algorithm for complex wavelet coefficients; this algorithm has potential benefits for resolution enhancement systems.

The steerable pyramid introduced in Section 1.2.1 uses nonseparable filters with real coefficients to attain approximate shiftability and improved directionality. To obtain phase information, Simoncelli et al. [16] and van Spaendonck and Baraniuk [44] created complex steerable pyramids that use analytic non-separable filters. Portilla and Simoncelli [45] developed a model that incorporated the cross-scale phase statistics of complex steerable-pyramid coefficients. This model proved useful in a texture synthesis application. Van Spaendonck and Baraniuk [44] demonstrated that the smooth envelope of the complex steerable-pyramid basis functions is useful for seismic signal processing. In Chapter 2, we shall prove that our complex wavelet transforms provide useful phase information that may be easily exploited for signal-processing applications.

1.3 Compelling Reasons for Complex Wavelet Transforms

In the previous section, we examined several different methods that overcome some of the DWT disadvantages. Two of these methods, the complex steerable pyramid and the dual tree wavelet transform, simultaneously overcome all three disadvantages that we enumerated earlier. We shall review these methods now and then briefly explain the complex wavelet transforms that are the subject of this dissertation. These complex wavelet transforms also simultaneously overcome all the disadvantages of the DWT. Moreover, they also enjoy the additional benefits of flexibility and controllable redundancy.

The complex steerable pyramid is approximately shiftable, directional and provides useful phase information. Nevertheless, it is disadvantaged by high transform redundancy and the lack of perfect reconstruction in its non-separable, spatial-domain implementation. Like the complex steerable pyramid, The Dual Tree Wavelet Transform (DTWT) also simultaneously overcomes all three DWT disadvantages: shift sensitivity, poor directionality and lack of phase information. In contrast to the complex steerable pyramid, it is perfectly reconstructing and has a small fixed amount of redundancy (a factor of four in two dimensions). Nevertheless, even this small amount of redundancy makes the DTWT unsuitable for data compression, an application in which a non-redundant transform is essential. Furthermore, the DTWT can only provide an analysis using specific quadrature wavelet-pairs: it cannot be used to provide the wavelet decomposition associated with an arbitrary wavelet.

In Section 1.2.1, we explained that the DTWT coefficients may be interpreted as arising from the DWT associated with an analytic wavelet. In fact, this is the rationale behind the quasi-analytic wavelet transform of Magarey and Kingsbury in Section 1.2.1; however, the DTWT is preferable to the quasi-analytic wavelet transform because the latter transform is not invertible. We now point out that there is yet another interpretation for the DTWT coefficients: they may be interpreted as the DWT of the analytic signal associated with the input signal. For an $L^2(\mathbb{R} \rightarrow \mathbb{R})$ input signal, the analytic signal is the Hardy-space image of the input signal. Therefore, the DTWT coefficients may also be interpreted as arising from the DWT of the Hardy-space projection of the input signal. In Chapter 2, we shall explain that the Hardy-space projection of an $L^2(\mathbb{R} \rightarrow \mathbb{R})$ signal is impossible to compute. To circumvent this obstacle, we define a new function space called the Softy space. The Softy space is an approximation to Hardy space and, moreover, the Softy-space projection of an $L^2(\mathbb{R} \rightarrow \mathbb{R})$ signal may be computed using digital filters. In this dissertation, we shall refer to the DWT of the Softy-space projection of the input signal as the Complex Wavelet Transform (CWT) of the input signal. In subsequent chapters we shall explain that the CWT of $L^2(\mathbb{R} \rightarrow \mathbb{R})$ signals overcomes all the enumerated DWT disadvantages: it is shiftable, directional and offers useful phase information. However, unlike the complex steerable pyramid and the DTWT, the CWT has other significant benefits that we outline below.

1. *Controllable Redundancy:* The CWT consists of the DWT of the Softy-space projection of an $L^2(\mathbb{R} \rightarrow \mathbb{R})$ signal. In Chapter 5, we shall show that the 2D CWT has a transform-redundancy factor of four. Note that the DTWT also has the same amount of redundancy. It is this redundancy that enables the DTWT and the CWT to have approximate shiftability. However, in Chapter 4 we define the non-redundant projection, an alternative to the Softy-space projection. Unlike the Softy-space projection, the non-redundant projection introduces no transform redundancy although it is otherwise similar to the Softy-space projection. We refer to the DWT of the non-redundant projection as the non-redundant CWT. Although more shift sensitive than the CWT, the non-redundant CWT is, to the best of our knowledge, the only non-redundant, directional, complex wavelet transform. This unique combination of properties implies that the non-redundant CWT has significant potential benefits for im-

age coding systems. Hence the ability to control the redundancy of our complex wavelet transforms allows for exciting new possibilities.

2. *Flexibility:* Unlike the complex steerable pyramid and the DTWT, our complex wavelet transforms do not require the design of new wavelets. The complex wavelet coefficients are easily generated by projecting real-valued input signals onto a complex subspace and then computing the DWT of the projection using *any* desired DWT. In Chapter 6, we shall invoke this flexibility to create the Complex Double-Density Wavelet Transform (CDDWT), a new transform based on Selesnick's Double-Density DWT [15]. This particular complex wavelet transform enjoys reduced shift sensitivity at a low redundancy factor of 2.67 in 2D. We obtained excellent results when we used the CDDWT in a seismic signal processing application, as discussed in Chapter 7.

In the remainder of the thesis, we shall develop the theory and application of our complex wavelet transforms in greater detail. In Chapter 2, we rigorously define Softy space and the Complex Wavelet Transform (CWT). In Chapter 3, we prove that the CWT is approximately shiftable under certain conditions. Then, to demonstrate the notion of controllable redundancy, we define the non-redundant CWT in Chapter 4. Next we explain the improved directionality of the complex wavelet transforms in Chapter 5. We exploit the flexibility of the complex wavelet transform implementation to create the complex double-density wavelet transform, in Chapter 6. Subsequently, we showcase a seismic signal-processing application of our complex wavelet transforms, in Chapter 7. Finally in Chapter 8, we make some concluding remarks and sketch directions for future work on complex wavelet transforms. Before proceeding with these chapters, we present the reader with the notation and acronyms that will be employed.

1.4 Notation and Acronyms

- An infinite-length sequence of numbers will be represented by a lowercase letter. For example, c will represent an entire sequence of numbers. For some integer n, the notation $c(n)$ will denote the number indexed by n in the sequence c.

- The Z-transform of the sequence c will be denoted by $C(z)$. Recall that the evaluation of $C(z)$ on the Z-plane unit circle yields the Fourier transform denoted in standard notation as $C(e^{j\omega})$. We shall however adopt the reduced notation for the Fourier transform of a sequence and represent it by $C(\omega)$.

- The above convention for infinite-length sequences of numbers will also apply to finite-length vectors. For example, the filter-coefficient vector of an FIR filter may be represented by the lower-case letter h. The notation $h(n)$ will represent the nth element of the vector h. The Z-transform and reduced-notation Fourier transform of the vector h will be denoted by $H(z)$ and $H(\omega)$ respectively.

- A function whose name begins with an upper-case letter represents the Fourier transform of the function whose name begins with the corresponding lower-case letter. For example, $F(\omega)$ denotes the Fourier transform of the function f.

- The sequence of functions $f_1, f_2, \ldots$ will be denoted by $\{f_l\}_{l \in \mathbb{N}}$.

- The set-difference operator will be represented by "$-$". Therefore if A and B are arbitrary sets,

$$A - B \triangleq \{x : x \in A \text{ and } x \notin B\}.$$

- The uppercase letter I represents the identity matrix.

- The lowercase letter j represents the imaginary unit $\sqrt{-1}$.

- The "*" superscript represents the complex conjugate operator.

- The "$_*$" subscript represents the complex conjugation of the coefficients of a Z-transform.
- The "$*$" symbol represents the convolution operator.
- The "^{T}" superscript represents the matrix transpose operator.
- The "^{H}" superscript indicates the matrix conjugate-transpose operator "^{T*}".
- The "$\downarrow_2$" symbol represents the operation of downsampling by a factor of two. When applied to the discrete signal x, the output of the downsampler will be represented by $\downarrow_2 x$. This signal will consist of all the even-indexed samples of x.
- If the downsampling factor is not mentioned, it will be assumed to be equal to two.
- The Dirac delta function will be represented by $\delta(\cdot)$.
- We shall denote the indicator function of the interval $[a, b)$ by $\chi_{[a,b)}$ defined as

$$\chi_{[a,b)}(x) = \begin{cases} 1, & \forall x \in [a, b), \\ 0, & \text{otherwise.} \end{cases}$$

- Define

$$\chi^P_{[a,b)}(\omega) \triangleq \sum_{k \in \mathbb{Z}} \chi_{[a+Pk,b+Pk)}(\omega), \forall \omega,$$

to be the P-periodic indicator function of the set $\cup_{k \in \mathbb{Z}} [a + Pk, b + Pk)$.

- Usually, the notation $L^2(\mathbb{R})$ represents the space of square-integrable, complex-valued functions on the real line. To differentiate between complex-valued and real-valued functions, we shall denote this functional space by $L^2(\mathbb{R} \to \mathbb{C})$ and we shall use the notation $L^2(\mathbb{R} \to \mathbb{R})$ to represent the subspace of $L^2(\mathbb{R} \to \mathbb{C})$ that is comprised of real-valued functions. Similarly, $l^2(\mathbb{R} \to \mathbb{C})$ and $l^2(\mathbb{R} \to \mathbb{R})$ represent complex-valued and real-valued sequences that are square-summable.

- The overline denotes the closure of a set with respect to the $L^2(\mathbb{R} \to \mathbb{C})$ metric space. For example, $\overline{X}$ is the closure of the set X in $L^2(\mathbb{R} \to \mathbb{C})$.

- If f and g are functions in $L^2(\mathbb{R} \to \mathbb{C})$, then the notation $f \approx g$ indicates that the $L^2(\mathbb{R} \to \mathbb{C})$ distance between f and g is approximately zero.

- Consider the filter h and the input vector x. Let H denote the Toeplitz matrix defined by $H_{m,n} = h(m-n)$. Then Hx represents the output vector obtained by filtering x with h [46, Chapter 2].

- In this thesis, the acronym DWT (Discrete Wavelet Transform) will refer to Mallat's algorithm operating on the scaling coefficients of an $L^2(\mathbb{R} \to \mathbb{C})$ function. Throughout, we shall employ only real filters in Mallat's algorithm; therefore, the associated DWT wavelets and scaling functions will always be real. However since the scaling coefficients of a complex-valued $L^2(\mathbb{R} \to \mathbb{C})$ function may be complex, DWT coefficients may be complex in general.

- We shall use the acronym RWT (Real Wavelet Transform) to refer specifically to the DWT of a real-valued function in $L^2(\mathbb{R} \to \mathbb{R})$. Since we consider only DWTs implemented with real filters, RWT coefficients are always real.

- The acronyms IDWT, IRWT refer to Inverse Discrete Wavelet Transform and Inverse Real Wavelet Transform, respectively.

Chapter 2

The CWT: A Wavelet Transform in Softy Space

In Chapter 1, we explained that our approach to mitigating DWT shortcomings enables the creation of various complex wavelet transforms with differing properties. The purpose of the present chapter is to define our first complex wavelet transform comprehensively. We shall refer to this particular transform as "the complex wavelet transform" or "CWT".

As mentioned briefly in the preceding chapter, the CWT of an $L^2(\mathbb{R} \to \mathbb{R})$ function is defined as the DWT of the Softy-space projection of the function. To elucidate this definition, we shall start with the necessary mathematical prerequisites. Then we shall define the Softy space and examine its properties. Our objective is twofold. First, we shall demonstrate that the Softy-space projection and its inverse are both realizable with digital filters. This makes the CWT practically feasible. Second, we shall prove that the Softy space approximates the Hardy space. This makes the CWT useful because it enables the mitigation of DWT shortcomings, as we will show in subsequent chapters.

2.1 Mathematical Prerequisites

2.1.1 The Multiresolution Analysis and Wavelet Transforms

Let h_s be the real-coefficient scaling filter that engenders the real scaling function $\phi(x)$ in the dilation equation

$$\phi(x) = \sum_n h_s(n)\phi(2x - n).$$

Define V_{j+1} as the real scaling space spanned by integer translates of $\phi(2^j x)$. Then $\{V_j\}_{j\in\mathbb{Z}}$ represents a nested sequence of scaling spaces satisfying

$$L^2(\mathbb{R} \to \mathbb{R}) \supset \ldots \supset V_1 \supset V_0 \supset V_{-1} \supset \ldots$$

in Mallat's multiresolution framework. Now suppose f is a real function in $L^2(\mathbb{R} \to \mathbb{R})$ and let f_j denote the projection of f onto V_j. We may write $f_1(x) = \sum_n c_1(n)\phi(x-n)$, or equivalently

$$F_1(\omega) = C_1(\omega)\Phi(\omega),$$

where c_1 is the scaling coefficient sequence representing f_1. Recall that the RWT of f_1 is the projection of f_1 onto the wavelet spaces $V_j - V_{j-1}$ for $j \leq 1$. Since these projections are performed by real-coefficient filters acting on real scaling coefficients to obtain real wavelet coefficients, the RWT provides no phase information. In subsequent chapters we shall show that the RWT is also shift sensitive and has poor directionality.

2.1.2 Hardy-Space Images: Theoretically Desirable, Practically Unrealizable

The classical Hardy space $H^2(\mathbb{R} \to \mathbb{C})$ is defined by

$$H^2(\mathbb{R} \to \mathbb{C}) \triangleq \{f \in L^2(\mathbb{R} \to \mathbb{C}) : F(\omega) = 0 \text{ for a.e. } \omega < 0\}.$$

Usually, the Hardy space is interpreted as a vector space over the complex field $\mathbb{C}$. However, in this thesis, we shall consider the Hardy space to be a vector space over the real field $\mathbb{R}$, for the following reason. Most signal-processing applications involve functions in the $L^2(\mathbb{R} \to \mathbb{R})$ function space, a vector space over the real field $\mathbb{R}$. If we also consider $H^2(\mathbb{R} \to \mathbb{C})$ to be a vector space over the real field $\mathbb{R}$, then it is possible to relate $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$. Later, we will demonstrate that this relationship is advantageous.

To determine the nature of the relationship between $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$, refer to the appendix for a proof of the isometry between these two function spaces.

We establish this isometry with respect to the norm induced by the $L^2(\mathbb{R} \to \mathbb{C})$ inner product. The isometry implies that a distance-preserving map exists between the function spaces. Our objective in relating $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$ is to enable us to work with the complex-valued $H^2(\mathbb{R} \to \mathbb{C})$ functions instead of the corresponding real-valued $L^2(\mathbb{R} \to \mathbb{R})$ functions. The isometry ensures that the distance between a pair of $L^2(\mathbb{R} \to \mathbb{R})$ functions is identical to the distance between the Hardy-space images of these functions. However many signal-processing applications require the computation of inner products between pairs of $L^2(\mathbb{R} \to \mathbb{R})$ functions. For such applications, mere distance preservation is insufficient; inner-product preservation is desirable. Inner-product preservation entails that the inner product between a pair of $L^2(\mathbb{R} \to \mathbb{R})$ functions is identical to the inner product between their Hardy-space images. Unfortunately, inner-product preservation fails for the $L^2(\mathbb{R} \to \mathbb{C})$ inner product. However, if we define a new inner product $\langle \cdot, \cdot \rangle_H$ as the real part of the $L^2(\mathbb{R} \to \mathbb{C})$ inner product $\langle \cdot, \cdot \rangle$, then a unitary (linear, bijective, inner-product preserving) map exists between $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$. We prove this result in the appendix. Function spaces that are related through a unitary map are said to be isomorphic to each other. We shall now explain why the isomorphism between $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$ is useful.

Let the superscript "^{H}" denote the isomorphic Hardy-space image of an $L^2(\mathbb{R} \to \mathbb{R})$ function map. Then, as shown in the appendix, the Hardy-space image of the $L^2(\mathbb{R} \to \mathbb{R})$ function f is given by

$$F^H(\omega) \triangleq F(\omega)\chi_{[0,\infty)}(\omega), \tag{2.1}$$

and the inverse map is specified by

$$F(\omega) \triangleq F^H(\omega) + F^H(-\omega)^* - \chi_{[0,0]}(\omega) \left[\frac{F^H(\omega) + F^H(-\omega)^*}{2} \right]$$

The isomorphism between $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$ suggests that we may examine the Hardy-space image f^H rather than the function f. There is value in such a scheme since the Hardy-space image f^H is a complex-valued function that makes useful phase information explicit; moreover, the multiscale decomposition of f^H provides valuable information on signal behavior [8, 7]. In subsequent chapters we shall demonstrate that the DWT of f^h is also less shift sensitive than the DWT of f. In addition, the DWT of f^h has higher directionality than the DWT of f.

With reference to Fig. 2.1, we suggest the following program to avail of the benefits of a multiresolution analysis in the Hardy space. Given an $L^2(\mathbb{R} \to \mathbb{R})$ function f, use the unitary map from $L^2(\mathbb{R} \to \mathbb{R})$ to Hardy space to compute f^H. Then compute the DWT of f^H to obtain a wavelet transform with explicit phase information, reduced shift sensitivity and improved directionality. Process the wavelet coefficients as desired and then perform the inverse DWT to get the processed Hardy-space function $\hat{f}^H$. Finally, exploit the inverse map from Hardy space to $L^2(\mathbb{R} \to \mathbb{R})$ to obtain the processed $L^2(\mathbb{R} \to \mathbb{R})$ function $\hat{f}$.

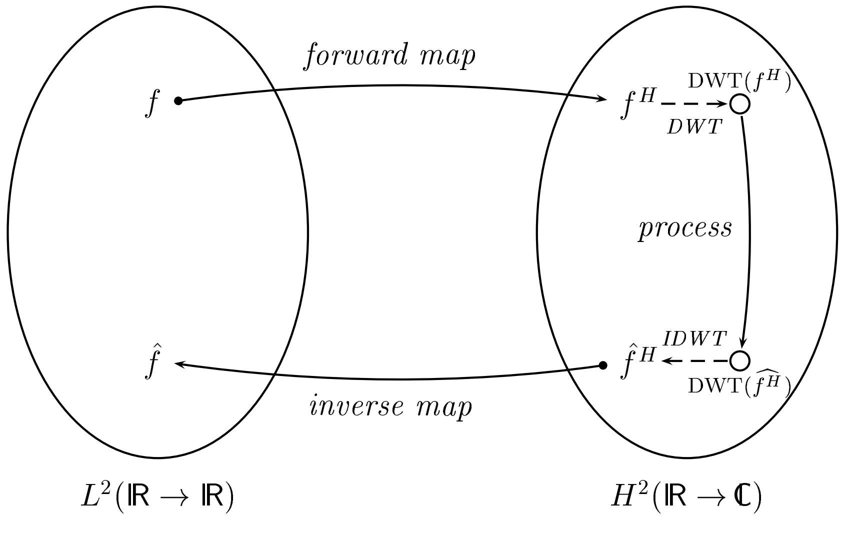


Figure 2.1 : Program for DWT in Hardy space.

Since we intend to compute the DWT of f^H, let us examine how f^H may be obtained from f prior to a multiresolution analysis. Let f be the projection of an

$L^2(\mathbb{R} \to \mathbb{R})$ function onto the scaling space V_0 so that

$$f(x) = \sum_n c(n)\phi(x-n),$$

or equivalently

$$F(\omega) = C(\omega)\Phi(\omega).$$

As dictated by Equation 2.1, the Hardy-space image f^H is given by

$$F^H(\omega) = \chi_{[0,\infty)}(\omega)C(\omega)\Phi(\omega). \tag{2.2}$$

We propose two methods to implement Equation 2.2. The first method applies the indicator function to ϕ to obtain its Hardy-space image ϕ^H,

$$F^H(\omega) = C(\omega)\Phi^H(\omega). \tag{2.3}$$

Observe that f^H is still represented by the scaling-coefficient sequence c associated with f. Walter [47] shows that since ϕ^H satisfies the same dilation equation as ϕ, the filter bank implementation of the DWT of f^H in Equation 2.3 is the same as that used to compute the DWT of f. As explained in Chapter 1, the DWT of f will still suffer from shift sensitivity and poor directionality. Consequently, the DWT of f^H obtained using Equation 2.3 has limited practical importance. Hence we discard the first method and proceed to the second method.

The second method to obtain f^H from Equation 2.2 involves the application of the indicator function to the scaling-coefficient sequence c,

$$F^H(\omega) = C^H(\omega)\Phi(\omega). \tag{2.4}$$

In later chapters we shall demonstrate that the suppressed negative frequencies of $C^H(\omega)$ in Equation 2.4 reduce shift sensitivity and improve directionality in the DWT of f^H. Hence Equation 2.4 would be preferable to Equation 2.3 for the computation of f^H from f. Unfortunately, since $\chi_{[0,\infty)}(\omega)$ is not 2π-periodic, it cannot be applied to the scaling-coefficient sequence c using a digital filter. Therefore neither method may be used to project an $L^2(\mathbb{R} \to \mathbb{R})$ function onto Hardy space prior to a multiresolution analysis with reduced shift sensitivity and improved directionality.

Since we cannot effectively compute Hardy-space images of $L^2(\mathbb{R} \to \mathbb{R})$ functions, we are motivated to define the Softy space, a new function space with properties similar to those of the Hardy space. In particular, the DWT of the Softy-space image of an $L^2(\mathbb{R} \to \mathbb{R})$ function will offer explicit phase information, reduced shift sensitivity and improved directionality. The important distinction between Softy space and Hardy space is that the unitary map between Softy space and $L^2(\mathbb{R} \to \mathbb{R})$ will be implementable with digital filters. We depict the relationship between these function spaces in Fig. 2.2.

The layout of this chapter relates directly to the three edges of the triangle in Fig. 2.2. In this section we described the $L^2(\mathbb{R} \to \mathbb{R})$—$H^2(\mathbb{R} \to \mathbb{C})$ edge. This first edge motivates the program for the DWT in Hardy space; unfortunately the benefits of this scheme are unattainable since this first edge is unrealizable. Next we shall characterize the $L^2(\mathbb{R} \to \mathbb{R})$—Softy-space edge. This second edge guarantees that we can perform the DWT in a function space that is isomorphic to $L^2(\mathbb{R} \to \mathbb{R})$. This has limited practical value prior to the introduction of the $H^2(\mathbb{R} \to \mathbb{C})$—Softy-space edge at the end of this chapter. This third edge describes the approximation of Softy space to Hardy space; this property implies that the DWT in the Softy space will inherit all the formerly unattainable benefits of the DWT in the Hardy space. We now proceed with the description of the $L^2(\mathbb{R} \to \mathbb{R})$—Softy-space edge. In the next section, we introduce the digital filters used to implement the realizable isomorphism between $L^2(\mathbb{R} \to \mathbb{R})$ and the Softy space. We shall formally define the Softy space after elaborating the properties of these filters.

2.2 Projection Filters

In the previous section, we explained that computing f^H via the method of Equation 2.4 offers significant benefits. However this method cannot be implemented prac-

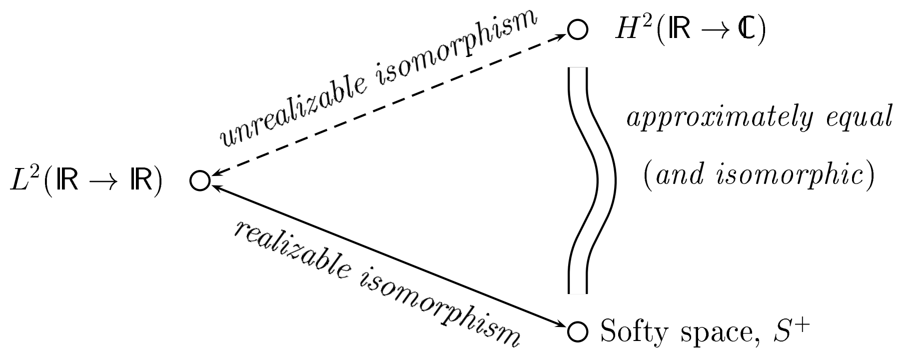


Figure 2.2 : Relationship between $L^2(\mathbb{R} \to \mathbb{R})$, Hardy space and Softy space.

tically because the frequency-domain operator $\chi_{[0,\infty)}(\omega)$ is not 2π-periodic. We can circumvent this problem and obtain an approximation to f^H by using the 2π-periodic operator $\chi^{2\pi}_{[0,\pi)}(\omega)$ instead of $\chi_{[0,\infty)}(\omega)$. We shall implement $\chi^{2\pi}_{[0,\pi)}(\omega)$ approximately using a digital filter h^+ with a passband over $[0, \pi)$ and a stopband over $[-\pi, 0)$. The "+" superscript indicates that the filter h^+ retains positive frequencies while attenuating negative frequencies. After filtering the real scaling-coefficient sequence $c \in V_0$ with h^+, we obtain the complex scaling-coefficient sequence c^+, as shown in Fig. 2.3.

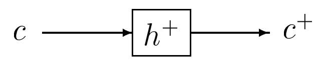


Figure 2.3 : The projection filter h^+.

We now create a new function space V_0^+ based on the scaling-coefficient sequences c^+:

$$
\begin{aligned}
V_0^+ \;\triangleq\; & \{f^+ : \forall f \in V_0, \text{where } F(\omega) = C(\omega)\Phi(\omega), \\
& \text{define } C^+(\omega) = H^+(\omega)C(\omega), \\
& F^+(\omega) = C^+(\omega)\Phi(\omega)\}. \qquad (2.5)
\end{aligned}
$$

Evidently the filter h^+ defines a linear, bijective map between the function spaces V_0 and V_0^+. Given $f \in V_0$, let $f^+ \in V_0^+$ denote the image of f under this map. If c and c^+ are the scaling-coefficient sequences associated with f and f^+ respectively, then the forward map from V_0 to V_0^+ is characterized by

$$C^+(\omega) = H^+(\omega)C(\omega),$$

and the inverse map is given by

$$C(\omega) = \begin{cases} G(\omega)C^+(\omega), & \omega \in [0, \pi), \\ G(-\omega)^* C^+(-\omega)^*, & \omega \in [-\pi, 0), \end{cases} \tag{2.6}$$

with

$$G(\omega) = \begin{cases} \frac{1}{H^+(\omega)}, & \omega \in [0, \pi), \\ 0, & \omega \in [-\pi, 0). \end{cases} \tag{2.7}$$

Observe that $\frac{1}{H^+(\omega)}$ is well-defined in the interval $[0, \pi)$ because the filter $H^+(\omega)$ has a passband in that interval. The existence of these maps allows us to propose the program in Figure 2.4 for the DWT in the function space V_0^+.

Unfortunately, although the forward map is implementable with the digital filter h^+, this program is still unimplementable since the implementation of the inverse map would involve a digital filter $G(\omega)$ with frequency response specified by Equation 2.7; unfortunately, only the zero filter can have zero frequency-response over an interval of non-zero measure. This observation shows that the inverse map corresponding to an arbitrary filter h^+ is unimplementable in general. To circumvent this obstacle, we shall introduce a class of filters h^+ for which the inverse map is also implementable. In the sequel, we shall label a bijective map as "realizable" iff the forward and inverse maps are both implementable with digital filters.

Proposition 1 *Let h_0 and g_0 be the lowpass analysis and synthesis filters of a two-band, real-coefficient, perfect-reconstruction filter bank. Create the projection and inverse-projection filters h^+ and g^+ by frequency shifting $H_0(\omega)$ and $G_0(\omega)$ by $\pi/2$,*

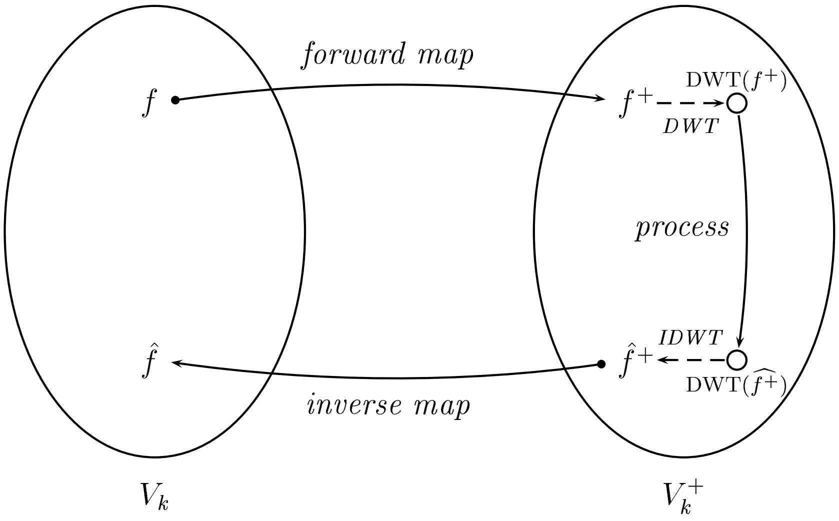


Figure 2.4 : Program for DWT in V_k^+, $\forall k \in \mathbb{Z}$.

so that

$$H^+(z) = H_0(-jz), \tag{2.8}$$

$$G^+(z) = G_0(-jz). \tag{2.9}$$

Let the scaling-coefficient sequence c *be associated with a function* f *in* V_0. *The "forward map" block in Figure 2.5 illustrates the forward map to the scaling-coefficient sequence* c^+ *associated with* $f^+ \in V_0^+$, *the image of* f. *The "inverse map" block shows the inverse map from* c^+ *to* $\hat{c}$. *Then we have* $\hat{c} = c$.

Proof: The subscript "$_*$" denotes complex conjugation of the coefficients of the Z-

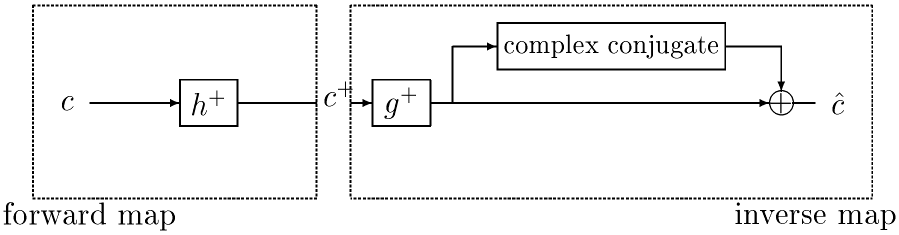


Figure 2.5 : A realizable bijective map.

transform.

$$
\begin{aligned}
\hat{C}(z) &= G^+(z)H^+(z)C(z) + G_*^+(z)H_*^+(z)C_*(z)], \\
&= (G^+(z)H^+(z) + G_*^+(z)H_*^+(z))C(z), \text{ since } c(n) \text{ is real}, \\
&= (G_0(-jz)H_0(-jz) + G_0(jz)H_0(jz))C(z), \\
&= (G_0(\nu)H_0(\nu) + G_0(-\nu)H_0(-\nu))C(j\nu), \quad \text{where } \nu = -jz, \\
&= C(z), \quad \text{since } G_0(\nu)H_0(\nu) \text{ is a halfband filter.} \quad \square
\end{aligned}
$$

Proposition 1 demonstrates that the linear, bijective map between V_0 and V_0^+ is implementable with the digital filters h^+, g^+. We now show that this concept extends easily from V_0 to $V_k,\ k \in \mathbb{Z}$. Given $f \in V_k$, the filter $H^+(2^{k-1}\omega)$ defines a new function space V_k^+:

$$
\begin{aligned}
V_k^+ \triangleq\ & \{f^+ : \forall f \in V_k, \text{where } F(\omega) = C(2^{k-1}\omega)\Phi(2^{k-1}\omega), \\
& \text{define } C^+(\omega) = H^+(2^{k-1}\omega)C(2^{k-1}\omega), \\
& F^+(\omega) = C^+(\omega)\Phi(2^{k-1}\omega)\}, \forall k \in \mathbb{Z}.
\end{aligned} \tag{2.10}
$$

Using the filters $H^+(2^{k-1}\omega), G^+(2^{k-1}\omega)$ instead of $H(\omega), G(\omega)$ respectively in Proposition 1 provides an implementation of the linear, bijective map between V_k and V_k^+.

The filter h^+ introduced in Proposition 1 is called a "projection" filter since it enables the projection of functions from V_k onto V_k^+. Since g^+ accomplishes the inverse projection from V_k^+ to V_k, we call it the "inverse-projection" filter. We can now perform the DWT in V_k^+, as proposed earlier in Fig. 2.4.

For many transform-domain signal-processing applications, it is crucial that the transform should preserve inner products. We shall now examine whether the transform-domain signal-processing scheme in Fig. 2.4 meets this criterion. Let f_1, f_2 be arbitrary functions in V_k. The scheme proposes to process the wavelet coefficients of f_1^+, f_2^+ as determined by a signal-processing application. Inner-product preservation requires that the inner product between f_1 and f_2 should be identical to that between the wavelet transforms of f_1^+ and f_2^+. Recall that orthogonal DWTs preserve inner products. Hence from Fig. 2.4, it is sufficient to ensure that the linear, bijective map from V_k to V_k^+ should also preserve inner products. Note that a linear, bijective, inner-product preserving map is called a unitary map. In the next proposition, we prove that the only digital filter that preserves inner products is the allpass filter.

Proposition 2 *The digital filter $H(z)$ preserves inner products iff it is an allpass filter.*

Proof: Suppose x_1, x_2 are arbitrary, real input sequences that create the output sequences y_1, y_2 when fed to $H(z)$. Let H denote the Toeplitz matrix that represents filtering with $H(z)$. Then $y_i = Hx_i,\ i = 1, 2$. Now $\langle y_1, y_2 \rangle = y_1^H y_2 = x_1^T H^H H x_2 = x_1^T x_2 = \langle x_1, x_2 \rangle$, iff $H^H H = I$. Careful examination of the matrix H^H reveals that it is the Toeplitz matrix that implements the digital filter $H_*(z^{-1})$. This implies that

$$\begin{aligned} H^H H = I \quad &\Leftrightarrow \quad H_*(z^{-1})H(z) = 1, \\ &\Leftrightarrow \quad H(\omega)^* H(\omega) = 1, \forall \omega \\ &\Leftrightarrow \quad |H(\omega)|^2 = 1, \forall \omega \\ &\Leftrightarrow \quad H \text{ is an allpass filter.} \quad \square \end{aligned}$$

Recall that the filter $H^+(z)$ cannot be an allpass filter as it has a stopband over the interval $[-\pi, 0)$. Therefore Proposition 2 implies that the forward map of Fig. 2.4

cannot be a unitary map if implemented with a digital filter. While proving this result, we assumed that all inner products are computed with the usual $l^2(\mathbb{R} \to \mathbb{C})$ inner product. However, we shall now prove that a simple redefinition of the inner product enables the digital-filter implementation of a unitary map. Recall that we were also forced to redefine the Hardy-space inner product to establish the unitary map between $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$

Let d_1, d_2 be scaling-coefficient sequences associated with arbitrary functions in V_k^+. Define the inner product between d_1 and d_2 as

$$\langle d_1, d_2 \rangle_+ \triangleq \mathcal{R}e\left\{d_1^H d_2\right\}. \tag{2.11}$$

We shall use the notation $||\cdot||_+$ to denote the norm induced by this redefined inner product. In the next proposition we establish a realizable unitary map between the scaling-coefficient sequences of functions in V_k and V_k^+. Inner products between scaling-coefficient sequences associated with V_k^+ functions will be computed with respect to the redefined inner product of Equation 2.11.

Proposition 3 *Let h_0 be an orthogonal lowpass filter that satisfies all the conditions in Proposition 1. Create the projection filter h^+ from h_0 by the operation $H^+(z) = H_0(-jz)$. Now suppose that c_1, c_2 are scaling-coefficient sequences associated with arbitrary real functions in V_k. Filter c_1 (c_2) with h^+ to obtain the scaling-coefficient sequence d_1 (d_2) associated with a function in V_k^+. Then*

$$\langle c_1, c_2 \rangle \;=\; \langle d_1, d_2 \rangle_+ .$$

Proof: We have

$$\begin{aligned}
\langle d_1, d_2 \rangle_+ &= \mathcal{R}e\left\{d_1^H d_2\right\}, \\
&= \mathcal{R}e\left\{c_1^T (H^+)^H H^+ c_2\right\}, \\
&= c_1^T \mathcal{R}e\left\{(H^+)^H H^+\right\} c_2, \text{ since } c_1, c_2 \text{ are real sequences,} \\
&= c_1^T c_2, \text{ since } \mathcal{R}e\left\{(H^+)^H H^+\right\} = I, \text{ as shown in the following lemma,} \\
&= \langle c_1, c_2 \rangle . \quad \square
\end{aligned}$$

Lemma 1 *Let H^+ be the Toeplitz matrix representing the filter h^+ in Proposition 3. Then*

$$\mathcal{R}e\left\{(H^+)^H H^+\right\} = I.$$

Proof: We have

$$H^+(z) = \sum_m h_0(m) e^{j\frac{\pi}{2}m} z^{-m},$$

and

$$H_*^+(z^{-1}) = \sum_m h_0(m) e^{-j\frac{\pi}{2}m} z^m.$$

Recall that the Toeplitz matrix $(H^+)^H H^+$ is represented by the autocorrelation filter $H_*^+(z^{-1})H^+(z)$ given by

$$\begin{aligned} H_*^+(z^{-1})H^+(z) &= \sum_m h_0(m) e^{-j\frac{\pi}{2}m} z^m \sum_n h_0(n) e^{j\frac{\pi}{2}n} z^{-n}, \\ &= \sum_m h_0(m) \sum_n h_0(n) e^{j\frac{\pi}{2}(n-m)} z^{-(n-m)}, \\ &= \sum_k z^{-k} e^{j\frac{\pi}{2}k} \sum_n h_0(n-k) h_0(n), \text{ with } k = n-m. \end{aligned} \tag{2.12}$$

Remember that $h_0(n)$ is the lowpass analysis filter from an orthogonal, two-band filter bank. Therefore the autocorrelation filter

$$H_0(z^{-1})H_0(z) = \sum_k z^{-k} \sum_n h_0(n-k) h_0(n),$$

is a halfband filter. This implies that

$$\sum_n h_0(n-k) h_0(n) = \begin{cases} 1, & \text{for } k = 0, \\ 0, & \text{for } k \text{ even}, \\ r_k \in \mathbb{R}, & \text{for } k \text{ odd}. \end{cases}$$

Using this in Equation 2.12, we obtain

$$H_*^+(z^{-1})H^+(z) = 1 + \sum_{k \text{ odd}} z^{-k} j^k r_k. \tag{2.13}$$

Equation 2.13 indicates that all the off-diagonal entries in the Toeplitz matrix $(H^+)^H H^+$ are either zero entries or purely imaginary entries. The diagonal entries are all equal to one. We conclude that

$$\mathcal{R}e\left\{(H^+)^H H^+\right\} = I. \quad \square$$

Proposition 3 uses the redefined inner product in Equation 2.11 to establish a realizable unitary map between the scaling-coefficient sequences associated with functions in V_k and V_k^+. As explained earlier, this property is beneficial to certain applications.

In some applications, it may be necessary to use the standard $l^2(\mathbb{R} \rightarrow \mathbb{C})$ inner product instead of the redefined inner product. In such cases, Proposition 2 ensures that a realizable unitary map between V_k and V_k^+ does not exist. In the following proposition, we demonstrate that although the realizable map between scaling coefficient sequences associated with functions in V_k and V_k^+ is not unitary with respect to the $l^2(\mathbb{R} \rightarrow \mathbb{C})$ inner product, it is however isometric (linear, bijective and norm-preserving).

Proposition 4 *Let h_+ be a projection filter that satisfies all the conditions in Proposition 3. Suppose that c is a scaling-coefficient sequence associated with an arbitrary real function in V_k. Filter c with h^+ to obtain the scaling-coefficient sequence d associated with a function in V_k^+. Then*

$$||d||_2^2 = \frac{1}{2}||c||_2^2$$

where $||\cdot||_2$ is the norm induced by the usual $l^2(\mathbb{R} \rightarrow \mathbb{C})$ inner product.

Proof: We have

$$\begin{aligned}
||d||_2^2 &= \int_{-\pi}^{\pi} |D(\omega)|^2 d\omega, \\
&= \int_{-\pi}^{\pi} |H^+(\omega)C(\omega)|^2 d\omega, \\
&= \int_{-\pi}^{\pi} |H_0(\omega - \frac{\pi}{2})|^2 |C(\omega)|^2 d\omega, \\
&= \int_{-\pi}^{0} |H_0(\omega - \frac{\pi}{2})|^2 |C(\omega)|^2 d\omega + \int_{0}^{\pi} |H_0(\omega - \frac{\pi}{2})|^2 |C(\omega)|^2 d\omega, \\
&= \int_{-\pi}^{0} |H_0(\omega - \frac{\pi}{2})|^2 |C(-\omega)|^2 d\omega + \int_{0}^{\pi} |H_0(\omega - \frac{\pi}{2})|^2 |C(\omega)|^2 d\omega, \\
&\quad \text{since } C(\omega) \text{ has conjugate symmetry,} \\
&= \int_{0}^{\pi} |H_0(-\xi - \frac{\pi}{2})|^2 |C(\xi)|^2 d\xi + \int_{0}^{\pi} |H_0(\omega - \frac{\pi}{2})|^2 |C(\omega)|^2 d\omega, \\
&\quad \text{with } \xi = -\omega \text{ in the first integral,} \\
&= \int_{0}^{\pi} |H_0(\xi + \frac{\pi}{2})|^2 |C(\xi)|^2 d\xi + \int_{0}^{\pi} |H_0(\omega - \frac{\pi}{2})|^2 |C(\omega)|^2 d\omega, \\
&\quad \text{since } H_0(\omega) \text{ has conjugate symmetry,} \\
&= \int_{0}^{\pi} |C(\omega)|^2 \left(|H_0(\omega + \frac{\pi}{2})|^2 + |H_0(\omega - \frac{\pi}{2})|^2| \right) d\omega, \\
&= \int_{\frac{\pi}{2}}^{3\frac{\pi}{2}} |C(\nu - \frac{\pi}{2})|^2 \left(|H_0(\nu)|^2 + |H_0(\nu - \pi)|^2 \right) d\nu, \text{ with } \nu = \omega + \frac{\pi}{2}, \\
&= \int_{0}^{\pi} |C(\omega)|^2 d\omega, \text{ since } H_0(z)H_0(z^{-1}) \text{ is a halfband filter,} \\
&= \frac{1}{2}||c||_2^2, \text{ since } C(\omega) \text{ has conjugate symmetry.} \quad \square
\end{aligned}$$

We now emphasize an important consequence of the above result. Let c^+ be the scaling-coefficient sequence associated with a function in V_k^+. Then

$$||c^+||_2^2 = ||c^+||_+^2.$$

This means that the norms induced by the $l^2(\mathbb{R} \to \mathbb{C})$ inner product and the redefined inner product are equal. This is true because both norms are equal to $||c||_2^2$ where c is the inverse projection of c^+ onto V_k. Note that for a pair of vector spaces over the complex field $\mathbb{C}$, the polarization identity ensures that a norm-preserving map between the vector spaces must also preserve the corresponding inner-product [48, Chapter 1]. However, since we define $H^2(\mathbb{R} \to \mathbb{C})$ and $L^2(\mathbb{R} \to \mathbb{R})$ as vectors over the real field $\mathbb{R}$, the polarization identity doesn't hold; therefore a norm preserving map between these vector spaces does not necessarily preserve the $L^2(\mathbb{R} \to \mathbb{C})$ inner-product.

2.2.1 Data Redundancy Introduced by Projection Filters

Refer to Fig. 2.6. Consider the compactly supported function f in V_1. We represent

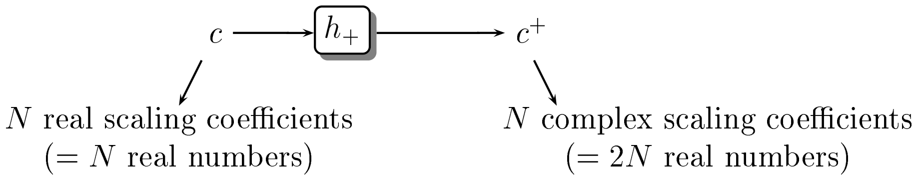


Figure 2.6 : Data redundancy introduced by h^+.

f digitally by its scaling-coefficient sequence c consisting of N real coefficients that will be represented by N real numbers in a digital computer. After filtering with the projection filter h^+, we obtain the scaling-coefficient sequence c^+ that also has N coefficients. However, since the frequency response of h^+ is not conjugate symmetric, the projection-filter coefficients are complex-valued. This implies that c^+ consists of N complex coefficients that will be represented by $2N$ real numbers in a digital computer. Therefore the amount of storage space for the scaling-coefficient sequence c^+ is twice that required for the scaling-coefficient sequence c. We refer to this phenomenon as the data redundancy introduced by the projection filter h^+.

In Chapter 3, we shall demonstrate that this data redundancy enables the reduction in shift sensitivity, a significant advantage of the CWT. However, in certain applications the benefit of a non-redundant transform may be more important than that of reduced shift sensitivity. For such applications, we shall introduce the non-redundant CWT in Chapter 4. Having described the relevant properties of the projection filter h^+, we are now ready to proceed to the next section in which we shall define Softy spaces.

2.3 Softy Space

As explained in Section 2.1.2, the DWT of the Hardy-space image of an $L^2(\mathbb{R} \to \mathbb{R})$ function promises explicit phase information, reduced shift sensitivity and improved directionality. Unfortunately, the unitary map between $L^2(\mathbb{R} \to \mathbb{R})$ and Hardy space is unrealizable with digital filters. This motivates the definition of the Softy space, a new function space with properties similar to those of the Hardy space. The important distinction between Softy space and Hardy space is that the unitary map between Softy space and $L^2(\mathbb{R} \to \mathbb{R})$ is implementable with the projection filters described in the previous section.

Definition 1 *Let $V_k,\ k \in \mathbb{Z}$ be the sequence of real nested scaling spaces associated with a real scaling filter h_s of length M. Suppose that h^+ is the length-N projection filter described in Proposition 3. As explained in Section 2.2, the projection filter h^+ unitarily maps real scaling-coefficient sequences associated with functions in V_k onto complex scaling-coefficient sequences associated with functions in the function space $V_k^+,\ \forall k \in \mathbb{Z}$. We define the* ***Softy space*** S^+ *associated with h_s and h^+ as*

$$S_{M,N}^+ \triangleq \overline{\cup_{k\in\mathbb{Z}} V_k^+}. \tag{2.14}$$

We wish to stress that the preceding definition creates a *family* of function spaces parameterized by scaling and projection filters of lengths M and N respectively. On selecting specific scaling and projection filters of lengths M_0 and N_0, we define the Softy space S_{M_0,N_0}^+ associated with these particular filters. For brevity, we will sometimes omit the subscripts in this notation and write S^+; this notation still refers to the Softy space associated with specific scaling and projection filters that are determined from the context.

We devote the remainder of this chapter to proving that S^+ and $L^2(\mathbb{R} \to \mathbb{R})$ are isomorphic Hilbert spaces, and that Softy space approximates Hardy space. First, however, we must define the inner product for functions in S^+. We remind the reader that a similar redefinition of the inner product was also required to show that Hardy space and $L^2(\mathbb{R} \to \mathbb{R})$ are isomorphic.

Suppose that f^+, g^+ are arbitrary functions in S^+. Then the Softy-space inner product is defined as

$$\left\langle f^+, g^+ \right\rangle_+ \;\stackrel{\triangle}{=}\; \mathcal{R}e\left\{ \int_{-\infty}^{\infty} f^+(x) g^+(x)^* dx \right\}, \;\forall f^+, g^+ \in S^+. \tag{2.15}$$

We choose this definition for the Softy-space inner product since it is compatible with our earlier redefinition of the inner product between scaling-coefficient sequences in V_k^+ for some integer k. This is evident from the following analysis. Let f^+, g^+ be arbitrary functions in V_k^+ for some integer k. We can express f^+ and g^+ as

$$f^+(x) = \sum_{n \in \mathbb{Z}} c_f^+(n) \phi(2^k x - n),$$

$$g^+(x) = \sum_{n \in \mathbb{Z}} c_g^+(n) \phi(2^k x - n),$$

where c_f^+, c_g^+ are scaling-coefficient sequences. Now

$$
\begin{aligned}
\left\langle f^+, g^+ \right\rangle_+ &= \mathcal{R}e\left\{\int_{-\infty}^{\infty} f^+(x) g^+(x)^* dx\right\}, \\
&= \mathcal{R}e\left\{\int_{-\infty}^{\infty} \sum_{n\in\mathbb{Z}} c_f^+(n)\phi(2^k x - n) \sum_{m\in\mathbb{Z}} c_g^+(m)^*\phi(2^k x - m) dx\right\}, \\
&= \mathcal{R}e\left\{\int_{-\infty}^{\infty} \sum_{n\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} c_f^+(n)\phi(2^k x - n) c_g^+(m)^*\phi(2^k x - m) dx\right\}, \\
&= \mathcal{R}e\left\{\sum_{n\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} c_f^+(n) c_g^+(m)^* \int_{-\infty}^{\infty} \phi(2^k x - n)\phi(2^k x - m) dx\right\}, \\
&= \mathcal{R}e\left\{\sum_{n\in\mathbb{Z}} c_f^+(n) c_g^+(n)^*\right\}, \\
&\quad \text{due to the orthogonality of the translates of } \phi, \\
&= \mathcal{R}e\left\{(c_g^+)^H c_f^+\right\}, \\
&= \left\langle c_f^+, c_g^+ \right\rangle_+ . \qquad (2.16)
\end{aligned}
$$

Therefore the Softy-space inner product between functions in V_k^+ is equal to the redefined inner product between their scaling-coefficient sequences. This fact will be crucial while establishing the isomorphism between $L^2(\mathbb{R} \to \mathbb{R})$ and S^+.

In the previous section we showed that the scaling-coefficient sequence norms induced by the $l^2(\mathbb{R} \to \mathbb{C})$ inner product and the redefined inner product are equal for the scaling-coefficient sequence c^+ associated with a function f^+ in V_k^+. Hence

$$||c^+||_2^2 = ||c^+||_+^2.$$

A similar result holds for the associated function f^+. Since f^+ is in the Hilbert space $L^2(\mathbb{R} \to \mathbb{C})$, we have

$$||f^+||_2^2 = ||c^+||_2^2.$$

From Equation 2.16, we get

$$||f^+||_+^2 = ||c^+||_+^2.$$

We therefore deduce that

$$||f^+||_2^2 = ||f^+||_+^2.$$

The continuity of the inner product allows us to extend this result from functions in V_k^+ to all functions in S^+. Therefore the norm induced by the $L^2(\mathbb{R} \to \mathbb{C})$ inner product is equal to that induced by the S^+ inner product for functions in S^+.

Earlier we motivated the program in Fig. 2.1 for obtaining a DWT with explicit phase information, reduced shift sensitivity and improved directionality. Unfortunately, this program is unrealizable since a projection onto Hardy space is impossible with digital filters. Our objective is to prove that Softy space may be used instead of Hardy space to implement the program in Fig. 2.1. To do this we will first show that Softy space is isomorphic to $L^2(\mathbb{R} \to \mathbb{R})$. This ensures that the Softy-space image of an $L^2(\mathbb{R} \to \mathbb{R})$ function is well-defined. Next we will show that the properties of Softy space are similar to those of Hardy space. This guarantees that we will obtain a DWT with explicit phase information, reduced shift sensitivity and improved directionality when we project an $L^2(\mathbb{R} \to \mathbb{R})$ function onto Softy space and compute its DWT.

2.3.1 The Isomorphism Between S^+ and $L^2(\mathbb{R} \to \mathbb{R})$

To exhibit an isomorphism between two Hilbert spaces, we must verify that they are indeed Hilbert spaces and that a unitary map exists between them.

It is well-known that $L^2(\mathbb{R} \to \mathbb{R})$ is a Hilbert space. So we must prove that S^+ is also a Hilbert space. We shall first prove that S^+ is a subspace of $L^2(\mathbb{R} \to \mathbb{C})$.

Both results rely on the map $\mathcal{H}^+ : L^2(\mathbb{R} \to \mathbb{R}) \to S^+$ defined as follows.

$$\forall f \in \cup_{k\in\mathbb{Z}} V_k,\ f^+ = \mathcal{H}^+ f \stackrel{\triangle}{=} C^+(\omega)\Phi(\omega), \tag{2.17}$$
$$\text{where } C^+(\omega) = H^+(\omega)C(\omega),\ F(\omega) = C(\omega)\Phi(\omega),\ f \in V_k,\ \text{some } k \in \mathbb{Z}.$$

$$\forall f \in \overline{\cup_{k\in\mathbb{Z}} V_k} - \cup_{k\in\mathbb{Z}} V_k,\ \exists \{f_l\}_{l\in\mathbb{Z}} \subset \cup_{k\in\mathbb{Z}} V_k \text{ such that } \lim_{l\to\infty} f_l = f.$$

Using Equation 2.17, define $f_l^+ = \mathcal{H}^+ f_l,\ \forall l \in \mathbb{Z}$. Since S^+ is closed and $\mathcal{H}^+$ is norm-preserving in Equation 2.17, $\lim_{l\to\infty} f_l^+$ exists in S^+. Denoting this limit by f^+, we have

$$\mathcal{H}^+ f \stackrel{\triangle}{=} f^+, \ \forall f \in \overline{\cup_{k\in\mathbb{Z}} V_k} - \cup_{k\in\mathbb{Z}} V_k. \tag{2.18}$$

Equations 2.17, 2.18 define $\mathcal{H}^+$ for all $f \in L^2(\mathbb{R} \to \mathbb{R}) = \overline{\cup_{k\in\mathbb{Z}} V_k}$.

Proposition 5 *S^+ is a subspace of $L^2(\mathbb{R} \to \mathbb{C})$.*

Proof: First we show that S^+ is a subset of $L^2(\mathbb{R} \to \mathbb{C})$. Recall that

$$S^+ = \overline{\cup_{k\in\mathbb{Z}} V_k^+}.$$

Now, for any $f^+ \in S^+$, define

$$f \stackrel{\triangle}{=} (\mathcal{H}^+)^{-1} f^+.$$

Two cases are possible. Either

1. $f \in V_k$, some $k \in \mathbb{Z}$, or

2. $\exists \{k_l\}_{l\in\mathbb{Z}}$, $f = \lim_{l\to\infty} f_{k_l}$, where $f_{k_l} \in V_{k_l}$.

We consider these cases separately.

Case 1: $f \in V_k \subset L^2(\mathbb{R} \to \mathbb{C})$. Clearly $h^+ f = f^+ \in L^2(\mathbb{R} \to \mathbb{C})$ since filtering by h^+ involves only scalar multiplication and addition, and $L^2(\mathbb{R} \to \mathbb{C})$ is a vector space.

Case 2: By the argument in Case 1, $f^+_{k_l} \in L^2(\mathbb{R} \to \mathbb{C}), \forall l \in \mathbb{Z}$, where $f^+_{k_l} \stackrel{\triangle}{=} \mathcal{H}^+ f_{k_l}$. Since $L^2(\mathbb{R} \to \mathbb{C})$ is closed, $f^+ \in L^2(\mathbb{R} \to \mathbb{C})$.

In both cases, $f^+ \in L^2(\mathbb{R} \to \mathbb{C})$. This proves that

$$S^+ \subset L^2(\mathbb{R} \to \mathbb{C}). \tag{2.19}$$

Now $L^2(\mathbb{R} \to \mathbb{R})$ is a vector space over $\mathbb{R}$, not $\mathbb{C}$, since it is not closed under scalar multiplication with complex numbers. Since we intend to exhibit an isomorphism between $L^2(\mathbb{R} \to \mathbb{R})$ and S^+, we shall show now that S^+ is a vector space over $\mathbb{R}$.

$\forall f_1^+, f_2^+,\ \forall r_1, r_2 \in \mathbb{R}$,

$$\begin{aligned} r_1 f_1^+ + r_2 f_2^+ &= r_1 \mathcal{H}^+ f_1 + r_2 \mathcal{H}^+ f_2, \text{ where } f_1, f_2 \in L^2(\mathbb{R} \to \mathbb{R}), \\ &= \mathcal{H}^+[r_1 f_1 + r_2 f_2], \text{ by the linearity of } \mathcal{H}^+, \\ &= \mathcal{H}^+ f_3, \text{ where } f_3 \in L^2(\mathbb{R} \to \mathbb{R}), \\ &\quad \text{since } L^2(\mathbb{R} \to \mathbb{R}) \text{ is a vector space over } \mathbb{R}, \\ &= f_3^+ \in S^+. \end{aligned}$$

Therefore S^+ is a real vector space. With Equation 2.19, this proves that S^+ is a subspace of $L^2(\mathbb{R} \to \mathbb{C})$.

Proposition 6 *S^+ is a Hilbert space.*

Proof:

$$\mathcal{H}^+ : L^2(\mathbb{R} \to \mathbb{R}) \to S^+ \text{ is a norm-preserving, bijective map.} \tag{2.20}$$

With respect to the metric induced by its inner product,

$$L^2(\mathbb{R} \to \mathbb{R}) \text{ is a complete metric space.} \tag{2.21}$$

Statements 2.20 and 2.21 imply that S^+ is a complete metric space with respect to the norm induced by the Softy-space inner product. Hence S^+ is a Hilbert space. □

Proposition 7 *$\mathcal{H}^+ : L^2(\mathbb{R} \to \mathbb{R}) \to S^+$ is a unitary map.*

Proof: $\mathcal{H}^+$ is linear and bijective. Due to the continuity of the inner product, $\mathcal{H}^+$ preserves the Softy space inner product. Since $\mathcal{H}^+$ is linear, bijective and preserves the Softy space inner product, we conclude that it is a unitary map. □

Proposition 7 shows that the Hilbert spaces $L^2(\mathbb{R} \to \mathbb{R})$ and S^+ are isomorphic. It is easy to show that this isomorphism is not the trivial isomorphism that exists between a Hilbert space and itself. From the definition of S^+ in Equation 2.14, S^+ contains complex-valued functions. On the other hand, by definition $L^2(\mathbb{R} \to \mathbb{R})$

contains no complex-valued functions. Hence $S^+ \neq L^2(\mathbb{R} \to \mathbb{R})$. Therefore, the isomorphism established in Proposition 7 is non-trivial. This completes the description of the $L^2(\mathbb{R} \to \mathbb{R})$—Softy-space edge in Fig. 2.2; this edge characterizes a realizable isomorphism between these two function spaces. We conclude this chapter by describing the $H^2(\mathbb{R} \to \mathbb{C})$—Softy-space triangle-edge. We shall prove that Softy space is approximately equal to Hardy space. This gives us a realizable isomorphism from $L^2(\mathbb{R} \to \mathbb{R})$ to Softy space, a space that is approximately equal to Hardy space. Due to this approximation of Softy space to Hardy space, the DWT of Softy-space images of $L^2(\mathbb{R} \to \mathbb{R})$ functions will exhibit explicit phase information, reduced shift sensitivity and improved directionality.

2.3.2 Softy Space: An Approximation to Hardy Space

In the previous section, we established a realizable isomorphism between $L^2(\mathbb{R} \to \mathbb{R})$ and the Softy space S^+. We shall now show that the Softy-space image of an $L^2(\mathbb{R} \to \mathbb{R})$ function is approximately equal to its Hardy-space image. It is precisely this property that imparts value to the realizable S^+ image of an $L^2(\mathbb{R} \to \mathbb{R})$ function. Since the DWT of the Hardy-space image of an $L^2(\mathbb{R} \to \mathbb{R})$ function exhibits explicit phase information, reduced shift sensitivity and improved directionality, we can expect the same advantages from the realizable S^+ image of the $L^2(\mathbb{R} \to \mathbb{R})$ function provided that Softy space approximates Hardy space. We remind the reader that the Softy space S^+ is associated with specific scaling and projection filters, h_s and h^+ respectively. Later we shall examine the approximation of S^+ to Hardy space when the scaling and projection filters are derived from length-M and length-N Daubechies scaling filters respectively. However, to simplify the proof of the following proposition, we shall consider idealized versions of h_s and h^+. Since all scaling filters and projection filters approximate the idealized versions described below, the result applies for the Softy space associated with any scaling and projection filter.

Proposition 8 *Let S^+ be the Softy space associated with the following filters:*

1. *The scaling filter h_s engenders a scaling function ϕ whose Fourier transform decays rapidly so that $\Phi(\omega)$ is small for $|\omega| > \pi$,*

2. *The projection filter h^+ has the ideal frequency response $\chi^{2\pi}_{[0,\pi]}(\omega)$.*

Suppose that $f \in L^2(\mathbb{R} \to \mathbb{R})$. Then $f^+ \approx f^H$.

Proof: Either $f \in \cup_{k\in\mathbb{Z}} V_k$ or $f \in \overline{\cup_{k\in\mathbb{Z}} V_k} - \cup_{k\in\mathbb{Z}} V_k$.
Case 1: We have $f \in \cup_{k\in\mathbb{Z}} V_k$. Therefore $f \in V_k$ for some integer k. To incur minimal notational overhead, we first prove that $f^+ \approx f^H$ for $k = 1$. We shall then explain how this proof can be generalized to obtain the proof of $f^+ \approx f^H$ for $k \neq 1$.

We have $f \in V_1$. Therefore

$$F(\omega) = C(\omega)\Phi(\omega), \tag{2.22}$$

where c is the scaling-coefficient sequence associated with f. To create f^+, we input c to the projection filter h^+ that outputs the scaling-coefficient sequence c^+ associated with f^+. We have

$$\begin{aligned} C^+(\omega) &= H^+(\omega)C(\omega), \\ &= C(\omega)\chi^{2\pi}_{[0,\pi)}(\omega). \end{aligned}$$

Therefore

$$F^+(\omega) = C^+(\omega)\Phi(\omega), \tag{2.23}$$

$$ = C(\omega)\Phi(\omega)\chi^{2\pi}_{[0,\pi)}(\omega). \tag{2.24}$$

Since $\Phi(\omega)$ is small for $|\omega| > \pi$,

$$\begin{aligned} \Phi(\omega)\chi^{2\pi}_{[0,\pi)}(\omega) &\approx \Phi(\omega)\chi_{[0,\infty)}(\omega) \\ &= \phi^H(\omega), \text{ the Hardy-space image of } \phi. \end{aligned}$$

Therefore we may rewrite Equation 2.24 as

$$\begin{aligned} F^+(\omega) &\approx C(\omega)\phi^H(\omega), \\ &= F^H(\omega), \text{ for } f \in V_1. \end{aligned}$$

Now when $f \in V_k - V_1$, observe that $F(\omega)$ may now be expressed as

$$F(\omega) = C(2^{k-1}\omega)\Phi(2^{k-1}\omega),$$

which is simply Equation 2.22 with $C(\omega), \Phi(\omega)$ replaced by their dilations $C(2^{k-1}\omega), \Phi(2^{k-1}\omega)$ respectively. Hence we can adopt the same line of reasoning as in the preceding proof for $k = 1$. The important difference is that all instances of $C(\omega), \Phi(\omega), H^+(\omega)$ will be replaced by their dilations $C(2^{k-1}\omega), \Phi(2^{k-1}\omega), H^+(2^{k-1}\omega)$ respectively. This will prove that

$$f^+ \approx f^H, \forall f \in V_k, \forall k \in \mathbb{Z}. \tag{2.25}$$

Case 2: We have $f \in \overline{\cup_{k\in\mathbb{Z}} V_k} - \cup_{k\in\mathbb{Z}} V_k$. There must exist a sequence of functions $\{f_{k_l}\}_{l\in\mathbb{N}}$ where $f_{k_l} \in V_{k_l}, l \in \mathbb{N}$ and

$$\lim_{l\to\infty} f_{k_l} = f. \tag{2.26}$$

Since Case 1 applies to $f_{k_l}, \forall l \in \mathbb{N}$,

$$f_{k_l}^+ \approx f_{k_l}^H, \forall l \in \mathbb{N}. \tag{2.27}$$

Recall that $L^2(\mathbb{R} \to \mathbb{R}), H^2(\mathbb{R} \to \mathbb{C})$ and S^+ are isomorphic Hilbert spaces. Therefore f^H and f^+, the isomorphic $H^2(\mathbb{R} \to \mathbb{C})$– and S^+–images of f, satisfy

$$\lim_{l\to\infty} f_{k_l}^H = f^H, \tag{2.28}$$

and

$$\lim_{l\to\infty} f_{k_l}^+ = f^+. \tag{2.29}$$

Equations 2.27, 2.28, 2.29 and the isomorphism between $L^2(\mathbb{R} \to \mathbb{R}), H^2(\mathbb{R} \to \mathbb{C})$ and S^+ imply that

$$f^+ \approx f^H, \ \forall f \in \overline{\cup_{k\in\mathbb{Z}} V_k} - \cup_{k\in\mathbb{Z}} V_k.$$

From Cases 1 and 2, we conclude that

$$f^+ \approx f^H, \ \forall f \in L^2(\mathbb{R} \to \mathbb{R}). \quad \square$$

The above proposition describes the $H^2(\mathbb{R} \to \mathbb{C})$—$S^+$ edge of the triangle in Fig. 2.2. We now know that the realizable S^+ image of an $L^2(\mathbb{R} \to \mathbb{R})$ function is approximately equal to the unrealizable $H^2(\mathbb{R} \to \mathbb{C})$ image of the function. Hence we can replace the Hardy space in Fig. 2.1 with the Softy space, to obtain the program in Fig. 2.7; the preceding result will ensure that the DWT of f^+ will exhibit the advantages associated with the DWT of f^H in the unrealizable program of Fig. 2.1: explicit phase information, reduced shift sensitivity and improved directionality.

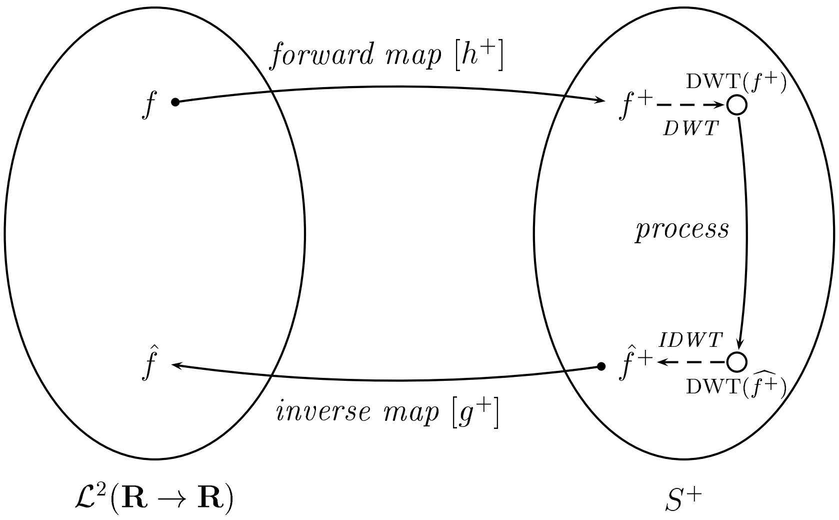


Figure 2.7 : Program for DWT in Softy space.

Let us now consider the computation of the DWT in Fig. 2.7. We argue that the Mallat algorithm may be used to calculate the DWT of functions in Softy space. Consider the function f in V_k. In the preceding proof, we demonstrated that f^+, the S^+ image of f, is expressible as

$$f^+(x) = \sum_n c^+(n)\phi(2^{-k}x - n).$$

Remember that the Mallat algorithm employs filters associated with ϕ to evaluate the DWT of an arbitrary $L^2(\mathbb{R} \to \mathbb{C})$ function $g(x) = \sum_n c(n)\phi(2^{-k}x - n)$, given c.

In Proposition 5 we demonstrated that S^+ is a subspace of $L^2(\mathbb{R} \to \mathbb{C})$. Therefore the Mallat algorithm using scaling and wavelet filters associated with ϕ will indeed provide the DWT of f^+. This DWT will exhibit explicit phase information, reduced shift sensitivity and improved directionality, because f^+ approximates f^H, the Hardy-space image of f.

In Proposition 8 we showed that the Softy space associated with idealized scaling and projection filters approximates the Hardy space. This ensured that the DWT of the S^+ image of an $L^2(\mathbb{R} \to \mathbb{R})$ function would inherit the advantages enjoyed by the DWT of the Hardy-space image of the function. In the final section of this chapter we will characterize this approximation more explicitly. We shall achieve this by examining the Softy space associated with the length-M Daubechies scaling filter, and the projection filter derived from the length-N Daubechies scaling filter. Apart from providing a better understanding of the Softy-space approximation to Hardy space, this analysis will also validate the applicablity of the CWT to natural images as we shall explain in the following section.

2.4 Asymptotic Approximation Performance for Polynomials

In 1988 Daubechies [3] designed new wavelets and scaling functions with vanishing moment properties. The corresponding scaling spaces contain all polynomials with degrees equal to the vanishing-moment multiplicity of the scaling filter. Since natural images are well approximated by piecewise polynomial functions, Daubechies' wavelet basis provides a sparse representation for natural images. Daubechies thus showed that the performance of a transform on natural images may be inferred from its performance on polynomials.

In an earlier section, we introduced the CWT consisting of a Softy-space projection followed by the DWT of the projection. Ideally, we would like to obtain phase information from the phase of the Hardy-space projection [8, 7]; in practice, the

Softy-space projection is an approximation to the Hardy-space projection. Therefore to gauge the accuracy of phase information in the CWT of a natural image, we shall provide upper bounds for the approximation error between Softy-space and Hardy-space projections of polynomials. As explained above, the performance of a transform on polynomials reveals its performance on natural images.

Let h represent the Daubechies scaling filter with N vanishing moments. h generates the scaling function ϕ via the ubiquitous scaling equation

$$\phi(x) = \sum_n h(n)\phi(2x - n).$$

Suppose f is a polynomial of degree N. Then there exists $C(\omega)$ such that

$$F(\omega) = C(\omega)\Phi(\omega).$$

We may write

$$F^H(\omega) = C(\omega)\Phi(\omega)\chi_{[0,\infty)}(\omega) \tag{2.30}$$

where f^H is the analytic image of the polynomial f. The next two subsections provide upper bounds for the approximation error between f^h and f^+. Section 2.4.1 assumes an ideal projection filter while Section 2.4.2 does not.

2.4.1 Approximation Error for Ideal Softy-Space Projection of Polynomials

Suppose f_∞^+ represents the function associated with the ideal (infinite-length) projection filter. Then we have

$$F_\infty^+(\omega) = C(\omega)\chi_{[0,\pi)}^{2\pi}\Phi(\omega). \tag{2.31}$$

From Equations 2.30, 2.31

$$
\begin{aligned}
||f^H - f_\infty^+||_2^2 &= \int_{-\infty}^{\infty} |C(\omega)\Phi(\omega)\chi_{[0,\infty)}(\omega) - C(\omega)\Phi(\omega)\chi_{[0,\pi)}^{2\pi}(\omega)|^2 d\omega \\
&\leq \sup_{\omega\in[0,2\pi)}|C(\omega)|^2 \int_{-\infty}^{\infty} |\Phi(\omega)|^2 |\chi_{[0,\infty)}(\omega) - \chi_{[0,\pi)}^{2\pi}(\omega)|^2 d\omega, \qquad (2.32) \\
&\quad \text{since } C(\omega) \text{ is } 2\pi\text{-periodic}, \\
&\leq C\int_{-\infty}^{\infty} \frac{|\chi_{[0,\infty)}(\omega) - \chi_{[0,\pi)}^{2\pi}(\omega)|}{(1+|\omega|)^{0.415N}} d\omega, \text{ as explained in [49, Equation 7.1.23]}, \\
&\leq C\int_{0}^{\infty} \frac{|\chi_{[0,\infty)}(\omega) - \chi_{[0,\pi)}^{2\pi}(\omega)|}{(1+\omega)^{0.415N}} d\omega, \text{ since the integrand is an even function}, \\
&\leq C\int_{0}^{\infty} \frac{\chi_{[\pi,\infty)}(\omega)}{(1+\omega)^{0.415N}} d\omega, \\
&\leq C\int_{\pi}^{\infty} \frac{d\omega}{(1+\omega)^{0.415N}}, \\
&\leq \frac{C}{(\pi+1)^{0.415N}}. \qquad (2.33)
\end{aligned}
$$

Equation 2.33 shows that the L^2-error between the analytic image f^H and the ideal Hardy-space projection f_∞^+ decays exponentially in the number of vanishing moments N that describe the Daubechies scaling space. This equation indicates that rapid frequency-domain decay of $\Phi(\omega)$ reduces the L^2-error.

2.4.2 Approximation Error for Non-Ideal Softy-Space Projection of Polynomials

The ideal Softy-space projection analyzed in the previous subsection is impossible to implement because it would require an infinite-length, non-causal projection filter. In this section, we analyze the approximation error when the projection filter h_M^+ is created by frequency-shifting the Daubechies lowpass filter of length $M+1$. Denote this lowpass filter by h_M. Given function f with Hardy-space projection f^H, let f_M^+ represent the Softy-space projection provided by the projection filter h_M^+. We have

$$F_M^+(\omega) = C(\omega)\Phi(\omega)H_M^+(\omega) \qquad (2.34)$$

From Equations 2.30, 2.34, we may express the approximation error between the Hardy- and Softy-space projections by

$$\begin{aligned}
||f^H - f_M^+||_2^2 &= \int_{-\infty}^{\infty} |C(\omega)\Phi(\omega)\chi_{[0,\infty)}(\omega) - C(\omega)\Phi(\omega)H_M^+(\omega)|^2 d\omega, \\
&\leq \sup_{\omega\in[0,2\pi)}|C(\omega)|^2 \int_{-\infty}^{\infty} |\Phi(\omega)|^2|\chi_{[0,\infty)}(\omega) - H_M^+(\omega)|^2 d\omega, \\
&\quad \text{since } C(\omega) \text{ is } 2\pi\text{-periodic}, \\
&\leq C\int_{-\infty}^{\infty} \frac{|\chi_{[0,\infty)}(\omega) - H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega, \\
&\quad \text{as explained in [49, Equation 7.1.23]}. \\
&= C\int_{-\infty}^{0} \frac{|H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + C\int_{0}^{\infty} \frac{|1-H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega, \\
&= C\int_{-\infty}^{-\pi} \frac{|H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + C\int_{-\pi}^{0} \frac{|H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + \\
&\quad C\int_{0}^{\pi} \frac{|1-H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + C\int_{\pi}^{\infty} \frac{|1-H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega, \\
&\leq C\int_{-\infty}^{-\pi} \frac{1}{(1+|\omega|)^{0.415N}} d\omega + C\int_{-\pi}^{0} \frac{|H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + \\
&\quad C\int_{0}^{\pi} \frac{|1-H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + C\int_{\pi}^{\infty} \frac{1}{(1+|\omega|)^{0.415N}} d\omega, \\
&= C\int_{-\infty}^{-\pi} \frac{1}{(1+|\omega|)^{0.415N}} d\omega + C\int_{-\pi}^{0} \frac{|H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + \\
&\quad C\int_{0}^{\pi} \frac{|1-H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega,
\end{aligned}$$

since the first and third integrals in the previous expression are identical. Therefore

$$||f^H - f_M^+||_2^2 \leq \frac{C}{(\pi+1)^{0.415N}} + C\int_{-\pi}^{0} \frac{|H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega + C\int_{0}^{\pi} \frac{|1-H_M^+(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega. \tag{2.35}$$

To evaluate the above integral, we express H_M^+ in terms of Strang and Nguyen's linear model for the frequency-domain characteristics of H_M [46, Chapter 5]. As shown in Figure 2.8, this model assumes that the transition-band width of h_M is proportional

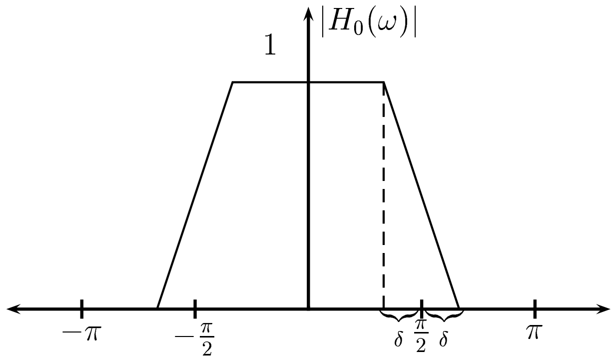


Figure 2.8 : The Strang-Nguyen model. $\delta = \frac{2}{\sqrt{M}}$.

to $\sqrt{M}$. The model allows us to approximate $H_M^+(\omega)$ as

$$H_M^+(\omega) = \begin{cases} -\sqrt{M}/4\omega + (0.5 - \pi\sqrt{M}/4), & \omega \in [-\pi, -\pi + 2/\sqrt{M}], \\ 0, & \omega \in [-\pi + 2/\sqrt{M}, -2/\sqrt{M}], \\ \sqrt{M}/4\omega + 0.5, & \omega \in [-2/\sqrt{M}, 2/\sqrt{M}], \\ 1, & \omega \in [2/\sqrt{M}, \pi - 2/\sqrt{M}], \\ -\sqrt{M}/4\omega + (0.5 + \pi\sqrt{M}/4), & \omega \in [\pi - 2/\sqrt{M}, \pi]. \end{cases}$$

Now define

$$\Delta(\omega) = \begin{cases} H_M^+(\omega), & \omega \in [-\pi, 0], \\ 1 - H_M^+(\omega), & \omega \in [0, \pi], \end{cases}$$

so that we may rewrite 2.35 as

$$\begin{aligned} ||f^H - f_M^+||_2^2 &\leq \frac{C}{(\pi+1)^{0.415N}} + C\int_{-\pi}^{\pi} \frac{|\Delta(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega, \\ &\leq \frac{C}{(\pi+1)^{0.415N}} + C\int_{-\pi}^{0} \frac{|\Delta(\omega)|^2}{(1+|\omega|)^{0.415N}} d\omega, \end{aligned}$$

since $\Delta(\omega)$ is an even function. Thus

$$\begin{aligned}
||f^H - f_M^+||_2^2 \;\leq\; & \frac{C}{(\pi+1)^{0.415N}} + C\overbrace{\int_{-\pi}^{-\pi+2/\sqrt{M}} \frac{|-\sqrt{M}/4\omega + 0.5 - \pi\sqrt{M}/4|^2}{(1-\omega)^{0.415N}} d\omega}^{A_1} \\
& + \underbrace{C\int_{-2/\sqrt{M}}^{0} \frac{|\sqrt{M}/4\omega + 0.5|^2}{(1-\omega)^{0.415N}} d\omega}_{A_2}. \qquad (2.36)
\end{aligned}$$

Denote the first and second integrals in the above expression by A_1 and A_2 respectively. We have

$$A_2 = C\int_{-2/\sqrt{M}}^{0} \frac{|\sqrt{M}/4\omega + 0.5|^2}{(1-\omega)^{0.415N}} d\omega,$$

Using Matlab, we evaluate A_2 as

$$\begin{aligned}
A_2 \;=\; & C(-M^{-\frac{1}{2}} - \frac{83}{160}(1+2\,M^{-\frac{1}{2}})^{\frac{83}{200}N}N + \\
& 3/4\,(1+2\,M^{-\frac{1}{2}})^{\frac{83}{200}N}\sqrt{M} + \frac{6889}{160000}(1+2\,M^{-\frac{1}{2}})^{\frac{83}{200}N}N^2 - 3/4\,\sqrt{M} + \\
& 1/8\,(1+2\,M^{-\frac{1}{2}})^{\frac{83}{200}N}M - 1/8\,M - 3/2 + \\
& 3/2\,(1+2\,M^{-\frac{1}{2}})^{\frac{83}{200}N} - \frac{83}{800}(1+2\,M^{-\frac{1}{2}})^{\frac{83}{200}N}N\sqrt{M})(\frac{913}{200}N - \frac{20667}{20000}N^2 - \\
& 6 + \frac{571787}{8000000}N^3)^{-1}((1+2\,M^{-\frac{1}{2}})^{\frac{83}{200}N})^{-1} \qquad (2.37)
\end{aligned}$$

We wish to study the separate effects of M and N on the asymptotic behavior of A. Hence we first fix N and analyze each term in Equation 2.37. On dividing out the denominator in this equation and representing $(1+2\,M^{-\frac{1}{2}})^{-\frac{83}{200}N}$ by the first three terms of its series expansion,the equation may be rewritten as the sum of various powers of $M^{-\frac{1}{2}}$. Hence the term with the slowest decay in A_2 is $M^{-\frac{1}{2}}$. Similar analysis reveals that for fixed N, the term with the slowest decay in A_2 is N^{-1}. These observations may be summarized by

$$A_2 \sim \begin{cases} O(M^{-\frac{1}{2}}), & \text{for fixed } N, \\ O(N^{-1}), & \text{for fixed } M. \end{cases} \qquad (2.38)$$

Next, Matlab evaluation of A_1 in Equation 2.36 gives

$$\begin{aligned} A_1 &= [-\frac{6889}{80000}(1+\pi)^{\frac{83}{200}N}(-\pi+2M^{-\frac{1}{2}})N^2\sqrt{M}(1/2-1/4\,\pi\sqrt{M}) \\ &\quad +39 \text{ other terms}]((1+\pi-2M^{-\frac{1}{2}})^{\frac{83}{200}N})^{-1} \\ &\quad (\frac{913}{200}N-\frac{20667}{20000}N^2+\frac{571787}{8000000}N^3-6)^{-1}((1+\pi)^{\frac{83}{200}N})^{-1} \end{aligned} \tag{2.39}$$

For fixed N, each of the 40 terms in Equation 2.39 is a polynomial in $\sqrt{M}$ with degree less than 3. As $M \to \infty, A_1$ is an integral on a set of measure zero; therefore $\lim_{M\to\infty} A_1 = 0$. Hence for fixed N, A_1 decays to zero no slower than $M^{-\frac{1}{2}}$. Termwise examination of Equation 2.39 also reveals that for fixed M, A_1 decays no slower than C^{-N}. These observations may be summarized by

$$A_1 \sim \begin{cases} O(M^{-\frac{1}{2}}), & \text{for fixed } N, \\ O(c^{-N}), & \text{for fixed } M. \end{cases}$$

Finally we combine Equations 2.36, 2.38 and 2.40 to obtain the following description for the approximation of f_M^+ to f^H,

$$||f^H - f_M^+||_2^2 \sim \begin{cases} O(M^{-\frac{1}{2}}), & \text{for fixed } N, \\ O(N^{-1}), & \text{for fixed } M. \end{cases}$$

Observe that, for fixed N, the asymptotic decay is dictated by the transition-band width of the projection filter h_M^+. In the above analysis, we generated h_M^+ from a Daubechies lowpass filter with a wide transition-band width proportional to $M^{-\frac{1}{2}}$. However, if we generate h_M^+ from an equiripple lowpass filter with transition-band width that is proportional to M^{-1}, we would derive the approximation of f_M^+ to f^H as

$$||f^H - f_M^+||_2^2 \sim \begin{cases} O(M^{-1}), & \text{for fixed } N, \\ O(N^{-1}), & \text{for fixed } M. \end{cases}$$

These results demonstrate that the approximation of Softy-space to Hardy space improves asymptotically in the lengths of the scaling and projection filters. In fact, this property of Softy space inspired its name: as the projection and scaling filters increase in length, the associated Softy space gets "harder" as it approaches the Hardy space. (We offer our apologies to the late Prof. G. H. Hardy for this silly pun.)

Chapter 3

Reducing Shift Sensitivity With The CWT

Shift sensitivity is a major disadvantage of the RWT, as explained in Chapter 1. Our objective in this chapter is to prove that the CWT is less shift sensitive than the RWT. We shall proceed as follows. Initially, we examine shift sensitivity in detail. Next, we shall delve into the proof for the reduced shift sensitivity of the CWT. We then conclude the chapter with experimental evidence for this important property of the CWT.

3.1 Shift Sensitivity in the RWT

In systems theory, a transform is shift invariant iff the transformation operator commutes with the shift operator. The magnitude-Fourier transform is an example of a shift-invariant transform. On the other hand, the RWT is not shift-invariant. To illustrate this fact, first consider Fig. 3.1(b) that shows the lowpass-RWT subband created from the input signal in Fig. 3.1(a). On shifting this signal by one sample we obtain Fig. 3.1(c). The lowpass-RWT subband generated from the shifted input signal is shown in Fig. 3.1(d). Observe the disparity between the lowpass-RWT subbands before shifting and after shifting. The RWT is clearly shift sensitive because RWT coefficients behave unpredictably under signal shifts.

It is widely appreciated that shift-invariance is impossible for practical implementations of the critically sampled RWT. Indeed, this is a primary RWT disadvantage [5, 16] that seriously undermines the application of the transform to several signal-processing problems. To circumvent this disadvantage, Mallat [9], Beylkin [10] and Guo [11] proposed the undecimated RWT that achieves shift-invariance at the expense

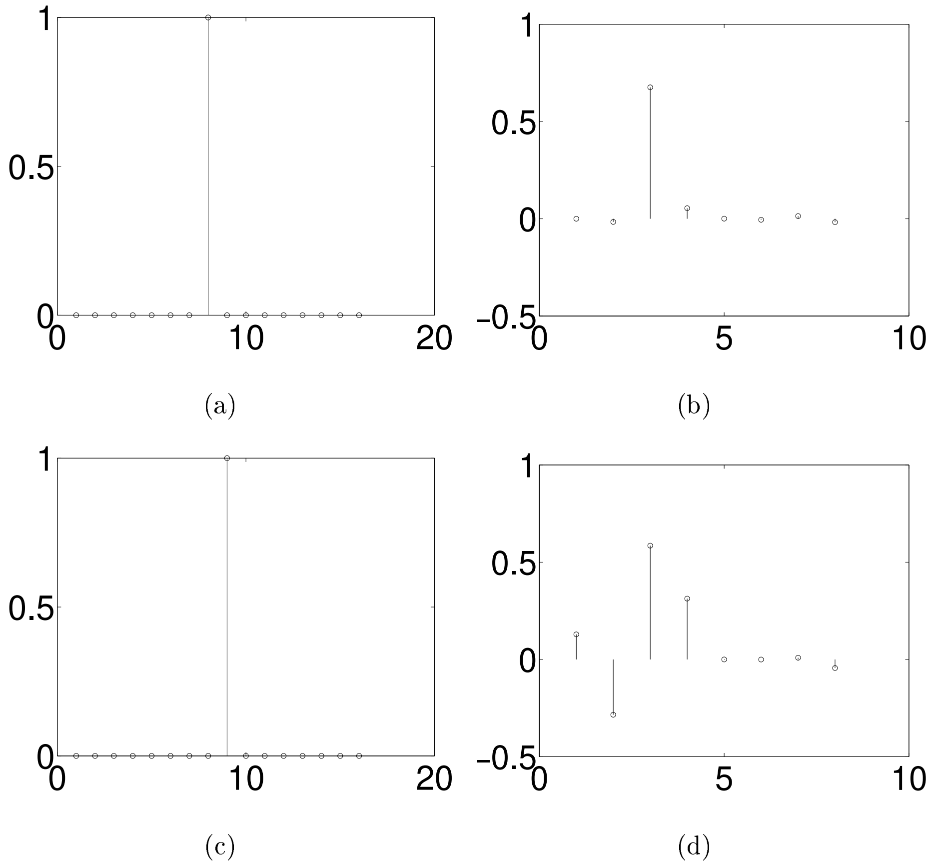


Figure 3.1 : (a) RWT input signal before shifting, (b) Output of RWT level-1 lowpass subband given unshifted input signal on the left, (c) RWT input signal generated by shifting the above input signal, (d) Output of RWT level-1 lowpass subband given shifted input signal on the left.

of an explosion in the storage-space required for the transformed data.

To reduce shift sensitivity and storage-space requirements simultaneously, Simoncelli et al. [16] defined "shiftability", a new benchmark for shift sensitivity. According to their definition, a transform is shiftable iff the coefficient energy in each transform subband is conserved under input-signal shifts. Although a weaker property than shift-invariance, shiftability is important for applications because it guarantees that the information in a transform subband is confined within that subband even when the input signal is shifted in position. In particular, shiftability ensures intra-subband interpolability, a transform property that allows us to determine the subband coefficients for an arbitrary shift of the input-signal given the subband coefficients for any other shift of the input signal.

The necessary and sufficient condition for shiftability is that each transform subband must satisfy the Nyquist criterion. Since this criterion is violated by the critically sampled RWT, this transform is not shiftable. This is evident from Fig. 3.2 that shows significant variations in the energy of the lowpass RWT subband under shifts of the input signal. By eliminating the downsamplers in highpass subbands and as-

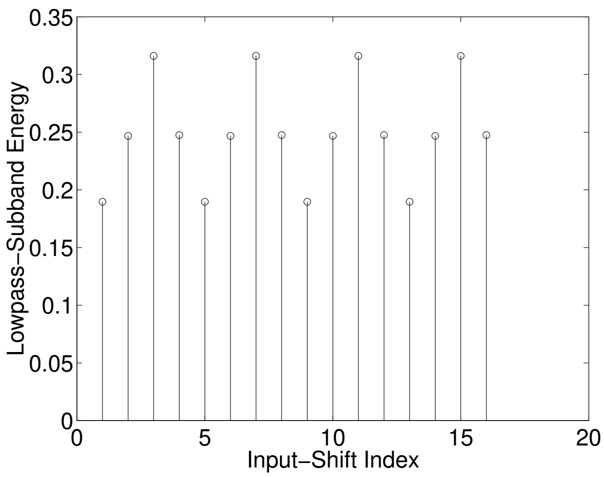


Figure 3.2 : Absence of shiftability in RWT lowpass subband.

suming negligible aliasing in the lowpass subband, Simoncelli et al. [16] designed the steerable pyramid, a transform that approximately preserves subband energy under input-signal shifts. Hence the steerable pyramid achieves approximate weak shift invariance with slightly less redundancy than the redundant RWT. In the next section we show that the CWT has approximate shiftability even though it has much less redundancy than the steerable pyramid.

3.2 The CWT: A Low-Redundancy Transform with Approximate Shiftability

In the preceding section we explained that satisfying the Nyquist criterion in each transform subband is a necessary and sufficient condition for shiftability. Since RWT subband filters are non-ideal, the signal bandwidth in each RWT subband of an $L^2(\mathbb{R} \to \mathbb{R})$ function f is larger than that required by the Nyquist criterion. Hence the RWT is not shiftable. We shall prove this rigorously in Proposition 9.

On the other hand, if f^H is the Hardy-space image of f, then f^h has half the bandwidth of f; hence the bandwidth in each DWT subband of f^H is half the bandwidth in the corresponding RWT subband of f. Consequently the Nyquist criterion is satisfied in each DWT subband of f^H; therefore the DWT of f^H enjoys shiftability. Observe that we have exploited the low bandwidth of the Hardy-space image to mitigate a problem caused by the large bandwidth of a real-valued function. As an aside, we point out that this concept also motivates single-sideband modulation in telecommunications.

In Chapter 2, we argued that it is impossible to obtain the Hardy-space image of an $L^2(\mathbb{R} \to \mathbb{R})$ function. We showed that the Softy-space image is a practical approximation to the Hardy-space image. Since the CWT is an approximation to the DWT of the Hardy-space image, the CWT will have approximate shiftability: the energy in CWT subbands will remain approximately constant under input-signal shifts. Remember that the CWT transform-redundancy is the small amount of redundancy introduced by the projection filter. Consequently, the CWT achieves approximate shiftability at lower redundancy than the steerable pyramid. We shall prove this result in the next section.

In this section, we argued intuitively that reduced bandwidth in transform subbands creates approximate shiftability. In the next section we prove that the CWT has this important property.

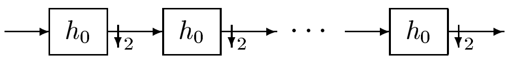


Figure 3.3 : M iterations of the scaling filter and downsampler.

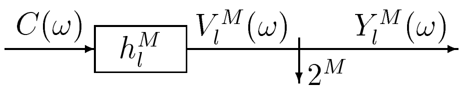


Figure 3.4 : Level-M lowpass subband.

3.3 Shiftability of the CWT

To prove that the CWT has shiftability, we must show that CWT subband energy is approximately constant under input-signal shifts. Recall that the CWT filter-bank implementation consists of a DWT filter bank preceded by the projection filter h^+. For an M-level DWT, the lowpass subband is as shown in Fig. 3.3. It consists of M iterations of the scaling filter $H_0(\omega)$ and downsampler. Define the filter $H_l^M(\omega)$ by

$$H_l^M(\omega) \triangleq H_0(\omega)H_0(2\omega)\ldots H_0(2^{M-1}\omega), \;\; M \geq 1, \tag{3.1}$$

so that we may exploit the Noble identity to represent the lowpass subband by the filter $H_l^M(\omega)$ followed by a downsampler with decimation factor 2^M as shown in Fig. 3.4. This representation corresponds to the parallel-structured filter bank [50] for Mallat's algorithm and will simplify subband-energy computations. Similarly, the level-M bandpass subband of the DWT has the parallel-structured representation shown in Fig. 3.5 where

$$H_h^M(\omega) \triangleq \begin{cases} H_1(\omega), & M = 1 \\ H_0(\omega)H_0(2\omega)\ldots H_0(2^{M-2}\omega)H_1(2^{M-1}\omega), & M \geq 2. \end{cases} \tag{3.2}$$

and $H_1(\omega)$ is a wavelet filter.

In the next proposition, we prove that the CWT will have approximate shiftability provided that sufficiently long filters are used in the implementation. To prove this

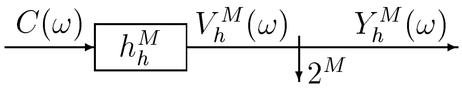


Figure 3.5 : Level-M bandpass subband.

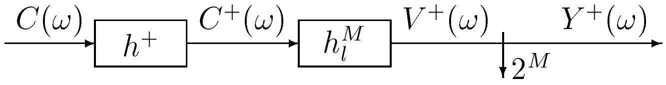


Figure 3.6 : Level-M CWT lowpass subband.

lemma we need explicit characterizations for the stopbands of the scaling and wavelet filters. Hence we shall assume that Daubechies scaling and wavelet filters are used to compute the DWT. Note that the proposition will hold in general for any choice of sufficiently long scaling and wavelet filters. The critical scaling-filter length for shiftability will be determined by the transition bandwidth: shorter bandwidths imply that shiftability is possible with shorter filters.

Proposition 1 *Let $H^+(\omega)$ be a projection filter with a transition bandwidth of δ_+. Consider the CWT implemented by placing $H^+(\omega)$ before an M-level DWT that uses Daubechies scaling and wavelet filters of length $N+1$. This CWT has approximate shiftability if $N \geq 15$ and*

$$\frac{8}{\sqrt{N}} + 2^M \delta_+ \leq 2\pi. \tag{3.3}$$

Proof:

Let x^+ be the Softy-space image of x, an $L^2(\mathbb{R} \to \mathbb{R})$ function. If c^+ and c are the scaling-coefficient sequences associated with x^+ and x, then $C^+(\omega) = H^+(\omega)C(\omega)$. The level-$M$ CWT lowpass subband is shown in Fig. 3.6. The filtered output will be labeled v^+ and the subband output after downsampling will be denoted by y^+. We have

$$Y^+(\omega) = \sum_{m=0}^{2^M-1} V^+\left(\frac{\omega - 2\pi m}{2^M}\right),$$

where

$$V^{+}(\omega) = H_l^M(\omega)C^{+}(\omega). \tag{3.4}$$

We now wish to study the effect of an input-signal shift on subband energy. We shall consider only integer-shifts of the input signal x. Shifting x by an integer k causes the scaling-coefficient sequence $c(n),\ \forall n \in \mathbb{Z}$ to shift to $c(n-k), \forall n \in \mathbb{Z}$. In response to this shift, the scaling-coefficient spectrum $C(\omega)$ will become $C(\omega)e^{jk\omega}$, denoted by $C_k(\omega)$. The scaling-coefficient spectrum $C^{+}(\omega)$ changes to $C_k^{+}(\omega) = C^{+}(\omega)e^{jk\omega}$, the filtered output $V^{+}(\omega)$ changes to $V_k^{+}(\omega) = V^{+}(\omega)e^{jk\omega}$ and $Y^{+}(\omega)$ changes to

$$Y_k^{+}(\omega) = \sum_{m=0}^{2^M-1} V^{+}\left(\frac{\omega - 2\pi m}{2^M}\right) e^{jk\frac{\omega-2\pi m}{2^M}}.$$

The L^2-energy of $Y_k^{+}(\omega)$ is

$$||y_k^{+}||_2^2 \ = \ \int_{-2^M\pi}^{2^M\pi} |Y_k^{+}(\omega)|^2 d\omega.$$

This implies that

$$2^M||y_k^{+}||_2^2 \ = \ \int_{-2^M\pi}^{2^M\pi} \left(\sum_{m=0}^{2^M-1} \left| V^{+}\left(\omega - \frac{2\pi m}{2^M}\right)\right|^2 + 2 \sum_{m,n=0;m\neq n}^{2^M-1} \Delta_{m,n}^{+}(k) \right) d\omega, \tag{3.5}$$

where

$$\Delta_{m,n}^{+}(k) \triangleq \mathcal{R}\mathrm{e}\left\{ V^{+}\left(\omega - \frac{2\pi m}{2^M}\right) V^{+}\left(\omega - \frac{2\pi n}{2^M}\right)^{*} e^{jk\frac{2\pi m}{2^M}(n-m)} \right\}. \tag{3.6}$$

Finally, in Lemma 4 we prove that $\Delta_{m,n}^{+}(k) \approx 0, \forall m,n \in \{0,\ldots,2^M-1\}$ if $N \geq 15$ and

$$\frac{8}{\sqrt{N}} + 2^M\delta_{+} \leq 2\pi.$$

Hence Equation 3.5 implies that

$$2^M||y_k^{+}||_2^2 \approx \int_{-2^M\pi}^{2^M\pi} \sum_{m=0}^{2^M-1} \left| V^{+}\left(\omega - \frac{2\pi m}{2^M}\right)\right|^2 d\omega.$$

Since the subband-energy $||y_k^{+}||_2^2$ is almost constant with respect to an input shift k, the level-M CWT lowpass subband has approximate shiftability. Similarly we can use the expression for $H_h^M(\omega)$ in Equation 3.2 to show that the level-M CWT bandpass subband-energy is approximately preserved under input-signal shifts. Hence the CWT has approximate shiftability if $N \geq 15$ and Condition 3.3 holds. $\square$

The above proposition establishes the conditions for approximate shiftability in the CWT. We defer the proof for $\Delta_{m,n}(k) \approx 0,\ \forall m, n \in \{0, \ldots, 2^M - 1\}$ to Lemma 4. To prove that result we shall first prove two lemmas that will simplify Expression 3.5. First, in Lemma 2 we show that the double summation in the second term of Equation 3.5 collapses into a single summation. Then, in Lemma 3 we provide approximations for the subband transfer functions $H_l^M(\omega), H_h^M(\omega)$ that constitute $\Delta_{m,n}(k)$. These approximations will enable us to prove Lemma 4.

Lemma 2 *Equation 3.5 may be rewritten as*

$$2^M ||y_k^+||_2^2 = \int_{-2^M\pi}^{2^M\pi} \left(\sum_{m=0}^{2^M-1} \left| V^+ \left(\omega - \frac{2\pi m}{2^M} \right) \right|^2 + 2 \sum_{p=0}^{2^M-1} \alpha_p \Delta_p^+(k) \right) d\omega, \quad (3.7)$$

where α_p is an integer indicating the multiplicity of terms $\Delta_{m,n}^+$ satisfying $n - m = p$ and $\Delta_p^+(k) \triangleq \Delta_{0,p}^+(k)$.

<u>Proof:</u> From Equation 3.6 we have

$$\int_{-\pi}^{\pi} \Delta_{m,n}^+(k) d\omega = \int_{-\pi}^{\pi} \mathcal{R}\mathrm{e}\{V^+ \left(\omega - \frac{2\pi m}{2^M} \right) V^+ \left(\omega - \frac{2\pi n}{2^M} \right)^* e^{jk\frac{2\pi m}{2^M}(n-m)}\} d\omega.$$

Let us substitute ξ for $\left(\omega - \frac{2\pi m}{2^M}\right)$ to get

$$\int_{-\pi}^{\pi} \Delta_{m,n}^+(k) d\omega = \int_{-\pi-\frac{2\pi m}{2^M}}^{\pi-\frac{2\pi m}{2^M}} \mathcal{R}\mathrm{e}\{V^+(\xi) V^+ \left(\xi - \frac{2\pi(n-m)}{2^M} \right)^* e^{jk\frac{2\pi}{2^M}(n-m)}\} d\xi.$$

Since the integrand is 2π-periodic in ξ and the integral is over a single period, we may replace the interval of integration with the period $[-\pi, \pi)$. On doing this and writing ω for ξ, we obtain

$$\int_{-\pi}^{\pi} \Delta_{m,n}^+(k) d\omega = \int_{-\pi}^{\pi} \mathcal{R}\mathrm{e}\{V^+(\omega) V^+ \left(\omega - \frac{2\pi(n-m)}{2^M} \right)^* e^{jk\frac{2\pi}{2^M}(n-m)}\} d\omega.$$

Clearly, $\Delta^{+}_{m,n}(k)$ is a function of $n-m$. This allows us to express Equation 3.5 as

$$2^M||y_k^+||_2^2 \;=\; \int_{-2^M\pi}^{2^M\pi}\left(\sum_{m=0}^{2^M-1}\left|V^+\left(\omega-\frac{2\pi m}{2^M}\right)\right|^2+2\sum_{p=0}^{2^M-1}\alpha_p\Delta_p^+(k)\right)d\omega, \quad (3.8)$$

where α_p is an integer indicating the multiplicity of terms $\Delta^{+}_{m,n}$ satisfying $n-m=p$ and $\Delta^{+}_{p}(k) \triangleq \Delta^{+}_{0,p}(k)$. $\square$

Lemma 2 enables us to reduce the number of terms to be analyzed later in the proof of Lemma 4. We shall show that it is sufficient to prove that $\Delta^{+}_{1}(k) \approx 0$. From Equations 3.6 and 3.4, we see that $\Delta^{+}_{1}(k)$ is dependent on $H_l^M(\omega)$. In Lemma 3 we shall provide simple expressions for $H_l^M(\omega)$ and $H_l^M(\omega)$ that will facilitate the analysis of $\Delta^{+}_{1}(k)$.

To obtain explicit expressions for $H_l^M(\omega)$ and $H_l^M(\omega)$ we assume that Daubechies scaling and wavelet filters of length $N+1$ are used to implement the DWT. We can simplify the expressions for $H_l^M(\omega)$ and $H_h^M(\omega)$ by using Strang and Nguyen's model for the scaling filter $H_0(\omega)$ [46, Theorem 5.6]. According to this model,in the length $N+1$ Daubechies scaling filter $|H_0(\omega)|$, the transition from $|H_0(\omega)| = 0.98$ to $|H_0(\omega)| = 0.02$ occurs over a bandwidth of length $\frac{4}{\sqrt{N}}$. Approximating 0.98 and 0.02 by 1 and 0 respectively, we can use Fig. 3.7 to depict the model for the Daubechies scaling filter.

Our subband-energy analysis will require a clear distinction between the frequencies where $|H_0(\omega)|$ is zero and those frequencies where it is non-zero. To make this distinction explicit, we shall use the unit circle in the complex plane to distinguish between the stopband and non-stopband frequencies. Fig. 3.8 shows the unit circle representation of $H_0(\omega)$ under Strang and Nguyen's model for the stopband. Similarly, Fig. 3.9 is the unit circle representation of $H_1(\omega)$.

In the following lemma, we demonstrate how the Strang and Nguyen model may be used to simplify the expressions for $H_l^M(\omega)$ and $H_h^M(\omega)$. These simplified expressions will allow us to prove Lemma 4 that eliminates shift-dependent terms in the expression for CWT lowpass subband energy.

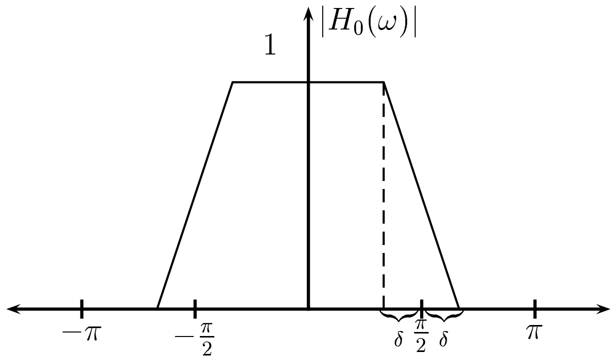


Figure 3.7 : Strang-Nguyen model for frequency-magnitude response of length-N Daubechies scaling filter. $\delta = \frac{2}{\sqrt{N}}$.

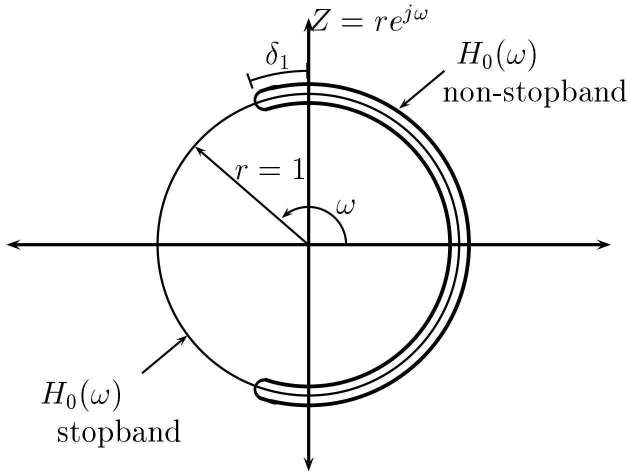


Figure 3.8 : Unit-circle representation for $H_0(\omega)$. $\delta_1 = \frac{2}{\sqrt{N}}$.

Lemma 3 *Suppose that $H_0(\omega)$ and $H_1(\omega)$ are length $N+1$ Daubechies scaling and wavelet filters respectively. Assume that $H_l^M(\omega)$ and $H_h^M(\omega)$ are given by Equa-*

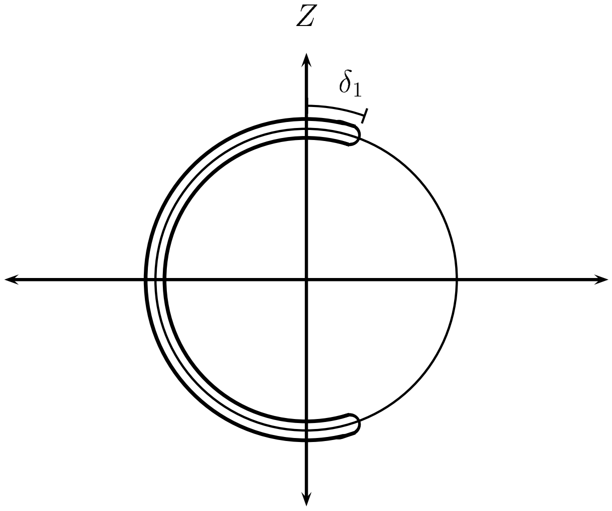


Figure 3.9 : Unit-circle representation for $H_1(\omega)$. $\delta_1 = \frac{2}{\sqrt{N}}$.

tions 3.1 and 3.2 respectively. Define

$$\begin{aligned}
\delta_M &\triangleq \frac{2}{2^{M-1}\sqrt{N}}, \\
\Omega_l &\triangleq \left[-\frac{\pi}{2^M} - \delta_M,\ \frac{\pi}{2^M} + \delta_M\right), \\
\Omega_h &\triangleq \left[\frac{\pi}{2^M} - \delta_M,\ \frac{\pi}{2^{M-1}} + \delta_M\right) \cup \left[-\frac{\pi}{2^{M-1}} - \delta_M,\ -\frac{\pi}{2^M} + \delta_M\right).
\end{aligned}$$

If $N \geq 15$ then $\forall M \geq 1$,

$$H_l^M(\omega) \approx \begin{cases} H_0(2^{M-1}\omega), & \omega \in \Omega_l, \\ 0, & \omega \in [-\pi, \pi) - \Omega_l, \end{cases} \tag{3.9}$$

$$H_h^M(\omega) \approx \begin{cases} H_1(2^{M-1}\omega), & \omega \in \Omega_h, \\ 0, & \omega \in [-\pi, \pi) - \Omega_h. \end{cases} \tag{3.10}$$

These expressions allow us to use the unit-circle representations in Figs. 3.10 and 3.11 for $H_l^M(\omega)$ and $H_h^M(\omega)$ respectively.

<u>Proof:</u> For $M = 1$, the result follows immediately from the definitions. We proceed

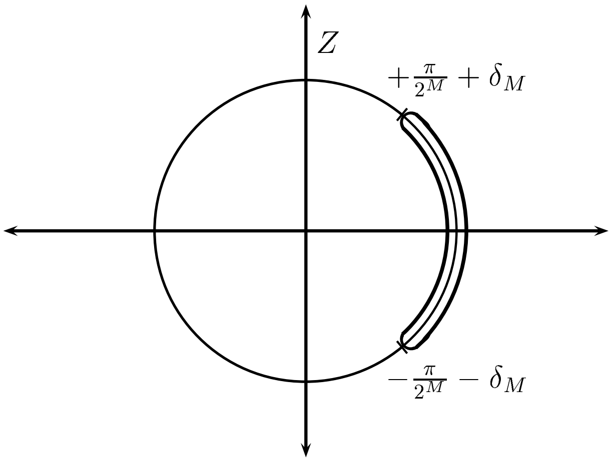


Figure 3.10 : Unit-circle representation for $H_l^M(\omega)$. $\delta_M = \frac{2}{2^{M-1}\sqrt{N}}$.

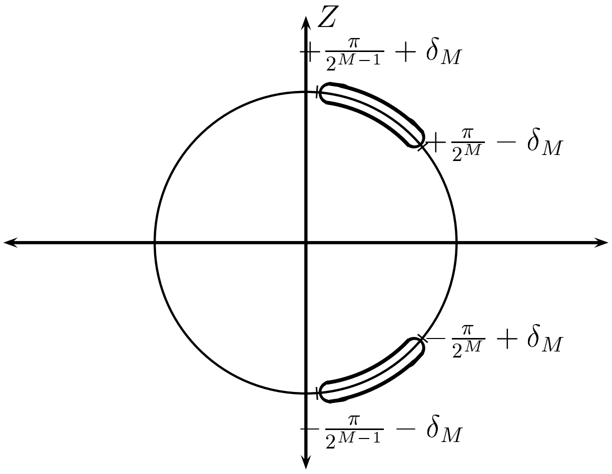


Figure 3.11 : Unit-circle representation for $H_h^M(\omega)$. $\delta_M = \frac{2}{2^{M-1}\sqrt{N}}$.

by induction on M for $M \geq 2$. When $M = 2$, we have

$$H_l^M(\omega) = H_0(\omega)H_0(2\omega).$$

The unit-circle representations of the filters $H_0(\omega)$ and $H_0(2\omega)$ are shown in Figs. 3.8 and 3.12 respectively. Observe that Equation 3.9 holds provided that the second

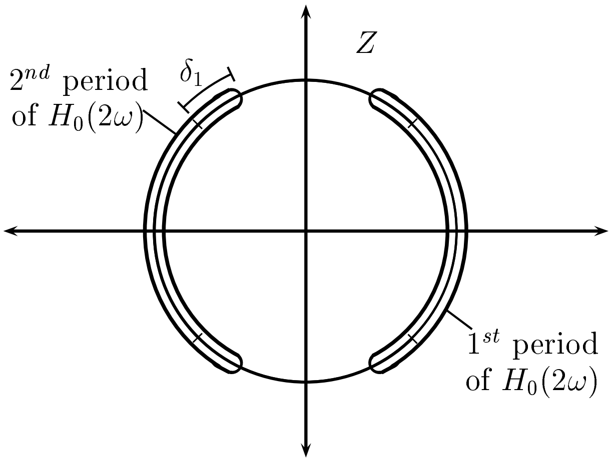


Figure 3.12 : Unit-circle representation for $H_0(2\omega)$. $\delta_1 = \frac{2}{\sqrt{N}}$.

period of $H_0(2\omega)$ doesn't overlap with $H_0(\omega)$. This condition holds if

$$
\begin{aligned}
\delta_1 + \delta_2 &\leq \frac{\pi}{4}, \\
\frac{2}{\sqrt{N}} + \frac{2}{2\sqrt{N}} &\leq \frac{\pi}{4}, \\
N &\geq 15.
\end{aligned}
$$

Now assume that Equation 3.9 is true when $M = K - 1$ for some $K \geq 3$. In the following inductive step, we shall show that Equation 3.9 holds for $M = K$. We have

$$
\begin{aligned}
H_l^K(\omega) &= H_0(\omega)H_0(2\omega)\ldots H_0(2^{K-2}\omega)H_0(2^{K-1}\omega), \text{ by definition,} \\
&= H_l^{K-1}(\omega)H_0(2^{K-1}\omega).
\end{aligned}
$$

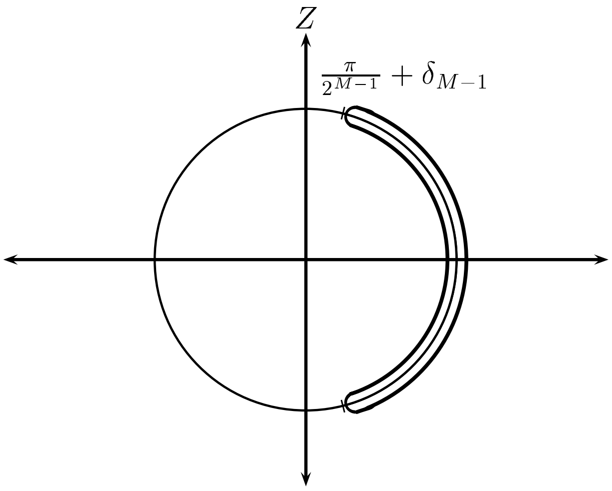


Figure 3.13 : Unit-circle representation for $H_l^{M-1}(\omega)$. $\delta_M = \frac{2}{2^{M-1}\sqrt{N}}$.

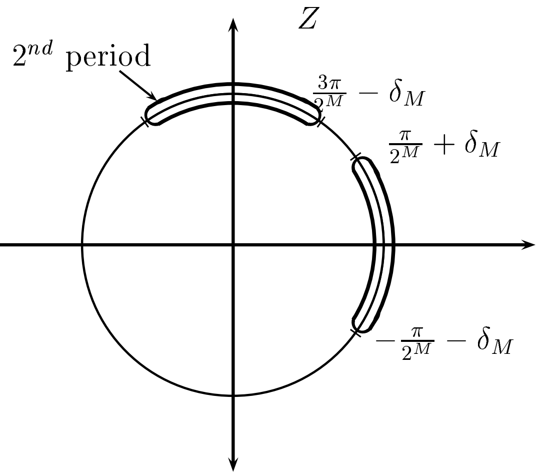


Figure 3.14 : Unit-circle representation for $H_0(2^{M-1}\omega)$. $\delta_M = \frac{2}{2^{M-1}\sqrt{N}}$.

Figs. 3.13 and 3.14 show unit-circle representations for the filters $H_l^{K-1}(\omega)$ and $H_0(2^{K-1}\omega)$ respectively. Equation 3.9 holds for $M = K$ provided that the second period of $H_0(2^{K-1}\omega)$ does not overlap with $H_l^{K-1}(\omega)$. Evidently, this is true if

$$\begin{aligned} \frac{3\pi}{2^M} - \delta_M &\geq \frac{\pi}{2^{M-1}} + \delta_{M-1}, \\ \frac{3\pi}{2^M} - \frac{2}{2^{M-1}\sqrt{N}} &\geq \frac{\pi}{2^{M-1}} + \frac{2}{2^{M-2}\sqrt{N}}, \\ N &\geq 15. \end{aligned}$$

Hence Equation 3.9 holds for all $M \geq 1$, by induction.

We shall now prove Equation 3.10. This equation is easily verified when $M = 1$. For $M \geq 2$,

$$\begin{aligned} H_h^M(\omega) &= H_0(\omega)H_0(2\omega)\ldots H_0(2^{M-2}\omega)H_1(2^{M-1}\omega), \text{ by definition,} \\ &= H_l^{M-1}(\omega)H_1(2^{M-1}\omega). \end{aligned}$$

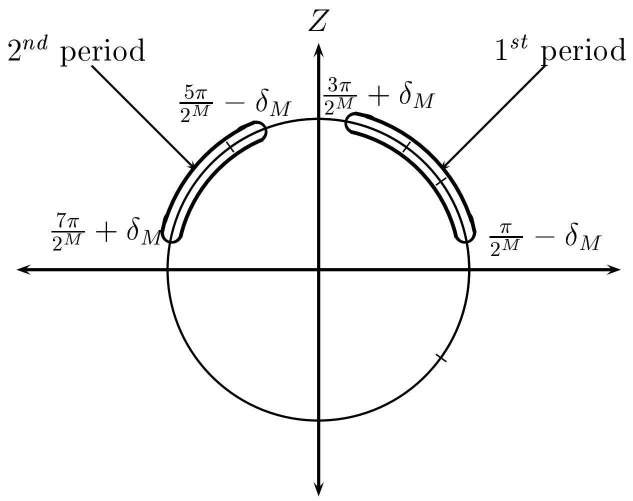


Figure 3.15 : Unit-circle representation for $H_1(2^{M-1}\omega)$. $\delta_M = \frac{2}{2^{M-1}\sqrt{N}}$.

The unit-circle representations for the filters $H_l^{M-1}(\omega)$ and $H_1(2^{M-1}\omega)$ are shown in Figs. 3.13 and 3.15. From these figures, it is evident that Equation 3.10 holds if the second period of $H_1(2^{M-1}\omega$ does not overlap with $H_l^{M-1}(\omega)$. This happens when

$$\begin{aligned}
\frac{\pi}{2^{M-1}} + \delta_{M-1} &\leq \frac{5\pi}{2^M} - \delta_M, \\
\frac{\pi}{2^{M-1}} + \frac{2}{2^{M-2}\sqrt{N}} &\leq \frac{5\pi}{2^M} - \frac{2}{2^{M-1}\sqrt{N}}, \\
N &\geq 15. \quad \square
\end{aligned}$$

We are now ready to eliminate the shift-dependent terms in the subband-energy expression of Equation 3.5.

Lemma 4 *Using the same definitions as in Proposition 1 we have* $\Delta^+_{m,n}(k) \approx 0$, *for* $M > 1, \forall m, n \in \{0, \ldots, 2^M - 1\}, m \neq n$. *if* $N \geq 15$ *and*

$$\frac{8}{\sqrt{N}} + 2^M \delta_+ \leq 2\pi. \tag{3.11}$$

Proof: Lemma 2 allows us to write

$$\Delta^+_p(k) = \Delta^+_{m,n}(k), \text{ where } p = n - m.$$

We have

$$\begin{aligned}
\Delta^+_p(k) &= \mathcal{R}\mathrm{e}\{H_l^M(\omega)X^+(\omega)H_l^M(\omega-\theta_p)^* X^+(\omega-\theta_p)^* e^{jk\theta_p}\}, \text{ where } \theta_p = \frac{2\pi p}{2^M}, \\
&= \mathcal{R}\mathrm{e}\{H_l^M(\omega)H_l^M(\omega-\theta_p)^* X(\omega)H^+(\omega)X(\omega-\theta_p)^* H^+(\omega-\theta_p)^* e^{jk\theta_p}\}, \\
&= \mathcal{R}\mathrm{e}\{H_l^M(\omega)H_l^M(\omega-\theta_p)^* H^+(\omega)H^+(\omega-\theta_p)^* X(\omega)X(\omega-\theta_p)^* e^{jk\theta_p}\},
\end{aligned}$$

To show that $\Delta^+_p(k) \approx 0$, we will demonstrate that

$$H_l^M(\omega)H_l^M(\omega-\theta_p)^* H^+(\omega)H^+(\omega-\theta_p)^* \approx 0, \ \forall \omega, \ p \in \{0, \ldots, 2^M - 1\}. \tag{3.12}$$

This is easily done by examining the unit-circle filter representations described in Lemma 3. Recall that these representations are valid for $N \geq 15$. Since the non-stopband regions of these filters will have maximal overlap for $p = 1$, it is sufficient to prove Equation 3.12 for $p = 1$.

Figures 3.10, 3.16, 3.17 show $H_l^M(\omega), H_l^M\left(\omega - \frac{2\pi}{2^M}\right), H_l^M(\omega)H_l^M\left(\omega - \frac{2\pi}{2^M}\right)$ respectively and Figures 3.18, 3.19, 3.20 show $H^+(\omega), H^+\left(\omega - \frac{2\pi}{2^M}\right), H^+(\omega)H^+\left(\omega - \frac{2\pi}{2^M}\right)$ respectively. Evidently, from Figures 3.17 and 3.20 the non-stopband regions of $H_l^M(\omega)H_l^M\left(\omega - \frac{2\pi}{2^M}\right)$ and $H^+(\omega)H^+\left(\omega - \frac{2\pi}{2^M}\right)$ have no overlap if

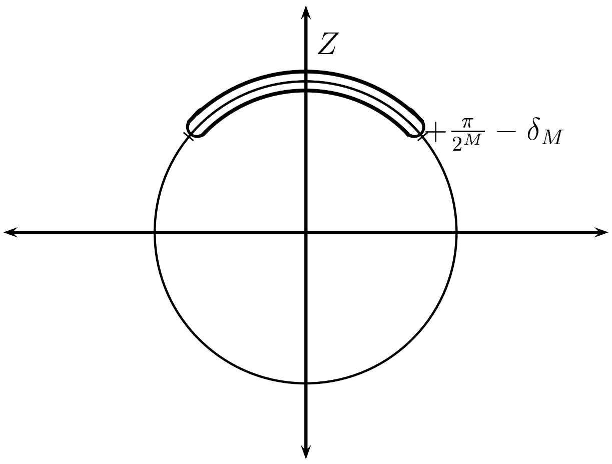


Figure 3.16 : Unit-circle representation for $H_l^M(\omega - \frac{2\pi}{2^M})$. $\delta_M = \frac{2}{2^{M-1}\sqrt{N}}$.

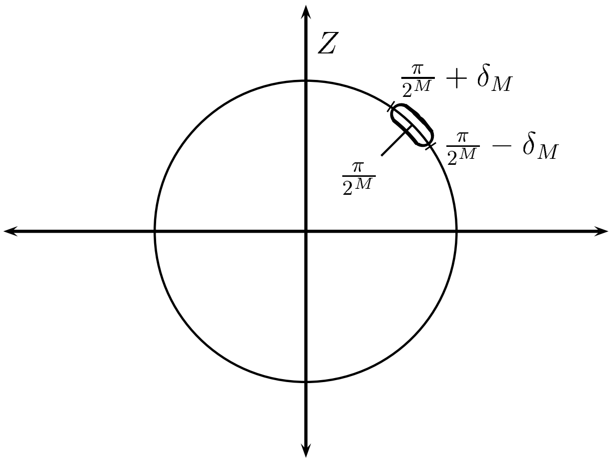


Figure 3.17 : Unit-circle representation for $H_l^M(\omega)H_l^M(\omega - \frac{2\pi}{2^M})$. $\delta_M = \frac{2}{2^{M-1}\sqrt{N}}$.

$$\frac{\pi}{2^M} + \delta_M \leq \frac{2\pi}{2^M} - \frac{\delta_+}{2},$$

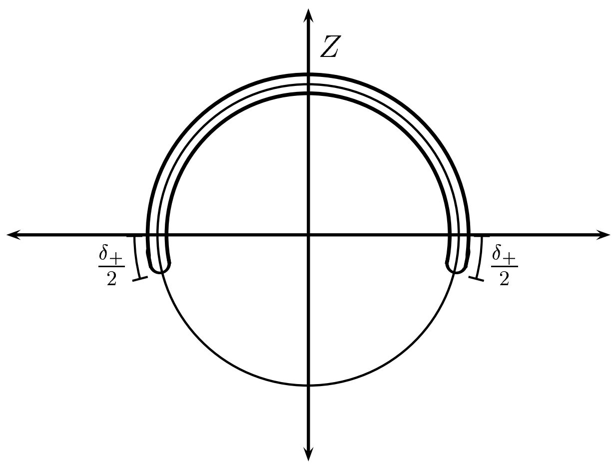


Figure 3.18 : Unit-circle representation for $H^+(\omega)$.

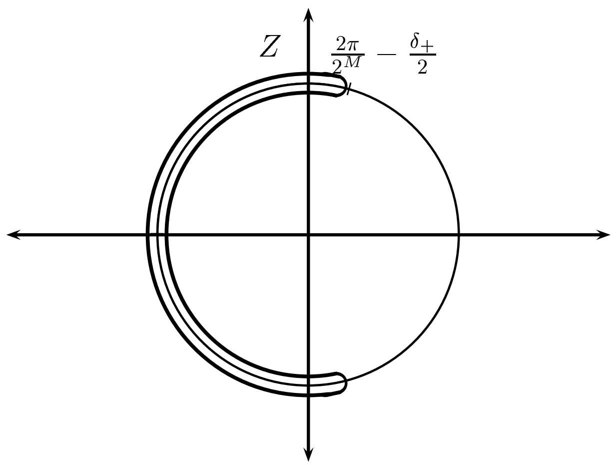


Figure 3.19 : Unit-circle representation for $H^+(\omega - \frac{2\pi}{2^M})$.

$$\frac{2}{2^{M-1}\sqrt{N}} + \frac{\delta_+}{2} \leq \frac{\pi}{2^M},$$

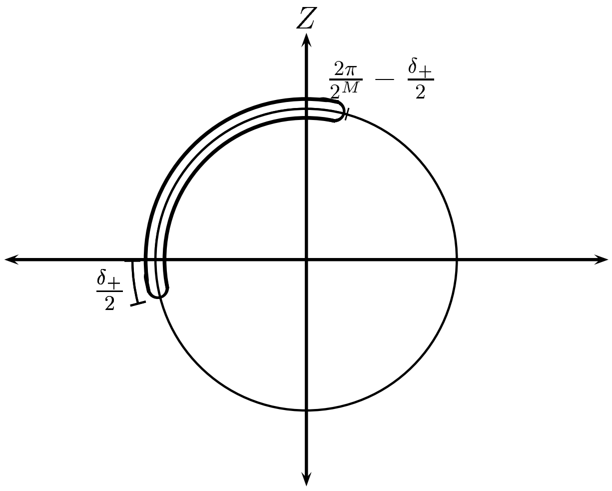


Figure 3.20 : Unit-circle representation for $H^+(\omega)H^+(\omega - \frac{2\pi}{2^M})$.

$$\frac{8}{\sqrt{N}} + 2^M\delta_+ \leq 2\pi.$$

Hence the result. $\Box$

Now that we have demonstrated that the CWT has approximate shiftability under certain conditions, it will be instructive to use our definitions to show that the RWT is not approximately shiftable. We perform this analysis in the proof of the following proposition.

Proposition 9 *The RWT is not approximately shiftable.*

Proof: Let x be an $L^2(\mathbb{R} \to \mathbb{R})$ function with scaling-coefficient sequence c. Fig. 3.3 shows the level-M RWT lowpass subband with input $C(\omega)$. Proceeding as in Proposition 1, we obtain the following expression for the level-M subband-energy under an input-signal shift by an integer k

$$2^M||y_k^M||_2^2 = \int_{-2^M\pi}^{2^M\pi} \left(\sum_{m=0}^{2^M-1} \left| V_l^M \left(\omega - \frac{2\pi m}{2^M} \right) \right|^2 + 2 \sum_{p=0}^{2^M-1} \alpha_p \Delta_p^M(k) \right) d\omega, \quad (3.13)$$

where

$$\Delta_p^M(k) = \mathcal{R}\mathrm{e}\{H_l^M(\omega)H_l^M(\omega-\theta_p)^* C(\omega)C(\omega-\theta_p)^* e^{jk\theta_p}\},$$

and α_p is an integer indicating the multiplicity of terms $\Delta_p^M(k)$ for each p. Now $\Delta_p^M(k) \approx 0$ iff $H_l^M(\omega)H_l^M(\omega-\theta_p)^* \approx 0,\ \forall\omega$. The frequency responses $H_l^M(\omega)$ and $H_l^M(\omega-\theta_p)^*$ will have maximal overlap for $p=1$. Hence it is sufficient to show that

$$H_l^M(\omega)H_l^M\left(\omega - \frac{2\pi}{2^M}\right)^* \approx 0,\ \forall\omega.$$

Figures 3.10 and 3.16 show the unit-circle filter representations for $H_l^M(\omega)$ and $H_l^M(\omega - \frac{2\pi}{2^M})^*$ respectively. Clearly, these frequency responses have no overlap iff

$$\begin{aligned} \frac{\pi}{2^M} + \delta_M &\leq \frac{\pi}{2^M} - \delta_M, \\ \frac{1}{2^M\sqrt{N}} &\leq 0. \end{aligned}$$

Since this condition is never satisfied for finite M, N, we find that $\Delta_p^M(k) \neq 0$. Hence the subband energy $||y_k^M||_2^2$ is dependent on the input-signal shift k. Consequently the RWT is not approximately shiftable. $\square$

We conclude this chapter with experimental evidence for the approximate shiftability of the CWT. Consider the input signal in Fig. 3.1(a). We fed this signal to a two-level RWT and also to a two-level CWT. Then for both transforms we computed the subband energy over 16 circular shifts of the input signal. Fig. 3.21 shows the transform-subband energies as a function of input-signal shifts. Observe the large oscillations in the RWT subband energies. In contrast, the CWT subband energy remains approximately constant over input-signal shifts. This corroborates our proof of the approximate shiftability in the CWT.

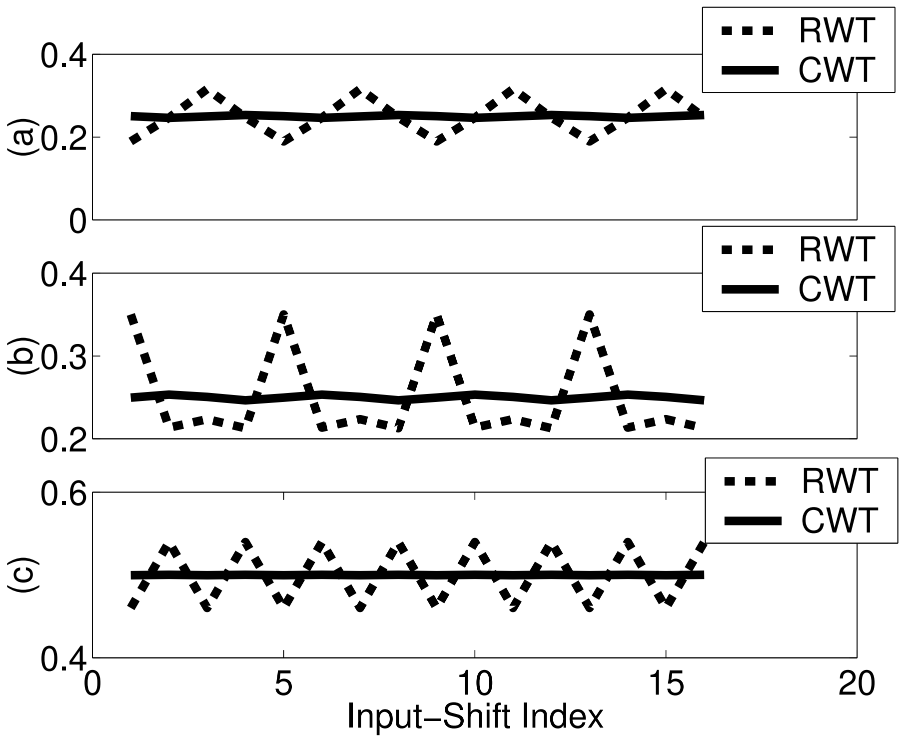


Figure 3.21 : Subband-energy\input-shift for RWT and CWT subbands. (a) Level-2 lowpass subband-energy, (b) Level-2 bandpass subband-energy, (c) Level-1 bandpass subband-energy. Both the RWT and CWT used length-16 Daubechies scaling and wavelet filters. The CWT projection filter was generated from a length-16 Daubechies scaling filter.

Chapter 4

The Non-Redundant CWT

In Chapter 2 we introduced the CWT as the DWT of the Softy-space projection of an $L^2(\mathbb{R} \to \mathbb{R})$ function. Since the Softy-space projection of an $L^2(\mathbb{R} \to \mathbb{R})$ function approximates the Hardy-space projection of that function, the CWT exhibits explicit phase information, reduced shift sensitivity and improved directionality. These advantages are obtained at the expense of a small amount of data redundancy introduced by the projection filter. In this chapter, we shall first develop the non-redundant projection: a projection that incurs no data redundancy while projecting an $L^2(\mathbb{R} \to \mathbb{R})$ function onto a function space that approximates Softy space. We shall refer to the DWT of this projected function as "the non-redundant CWT". We shall then demonstrate that the non-redundant CWT offers useful phase information and improved directionality; however, it does not enjoy reduced shift sensitivity. This transform will benefit applications such as data compression in which a non-redundant transform is more important than reduced shift sensitivity.

Apart from creating the non-redundant CWT, the non-redundant projection has another significant application. In Chapter 6, we shall demonstrate that the non-redundant projection may be used to enhance directionality and provide useful phase information without increasing the redundancy and while preserving the shift-invariance properties of transforms such as the undecimated DWT and the Double-Density DWT [15].

4.1 Defining the Non-Redundant Projection

In Chapter 2 we defined the projection of an $L^2(\mathbb{R} \to \mathbb{R})$ function onto Softy space using the projection filter h^+. As explained in Fig. 2.6 the complex coefficients of h^+ introduce data redundancy into the CWT. To eliminate this data redundancy, we now propose the non-redundant projection scheme shown in Fig. 4.1. As shown in

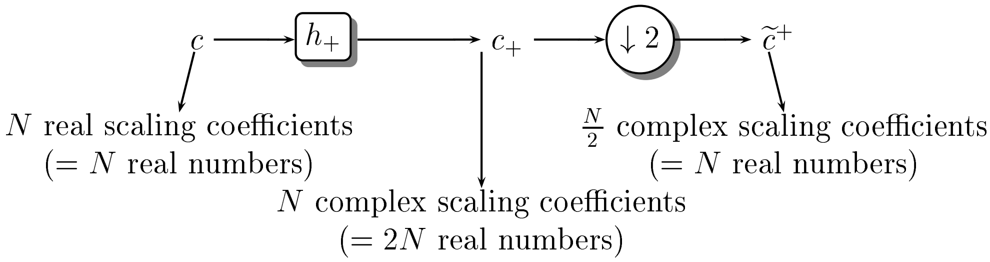


Figure 4.1 : The non-redundant projection.

the figure, we define the non-redundant projection as the concatenation of a projection filter and a downsampler (elimination of odd-indexed scaling coefficients). The downsampler eliminates the redundancy created by the projection filter that generates complex scaling coefficients from real scaling coefficients. Observe that the scaling-coefficient sequences c and $\widetilde{c}^+$ can both be represented by N real numbers within a digital computer; therefore there is no data redundancy in the scaling-coefficient sequence $\widetilde{c}^+$.

4.2 Characterizing the Non-Redundant Projection

Given the scaling-coefficient sequence c associated with a function f in the scaling space V_1, we know that the scaling-coefficient sequence c^+ is associated with f^+ in V_1^+, where f^+ is the S^+ image of f. Our immediate task is to characterize the function space containing the functions that are associated with the scaling-coefficient sequences $\widetilde{c}^{+}$. Let us denote that function space by $\widetilde{V}_0^{+}$ where

$$\begin{aligned} \widetilde{V}_0^{+} \triangleq \Big\{ \widetilde{f}^{+} : \widetilde{F}^{+}(\omega) &= \widetilde{C}^{+}(2\omega)\Phi(2\omega), \text{ where } \widetilde{c}^{+} = \downarrow_2 c^{+}, \\ \text{and } C^{+}(\omega) &= H^{+}(\omega)C(\omega), \text{ with } F(\omega) = C(\omega)\Phi(\omega), \forall f \in V_1 \Big\}. \end{aligned}$$

By definition, the non-redundant projection maps V_1 onto $\widetilde{V}_0^{+}$. Denote this map by $\widetilde{\mathcal{H}}^{+}$. Later in this section we shall prove that $\widetilde{\mathcal{H}}^{+}$ is a unitary map. First however, we shall use $\widetilde{\mathcal{H}}^{+}$ to demonstrate a link between $\widetilde{V}_0^{+}$ and the Softy space defined in Chapter 2. This link will enable us to exploit the non-redundant projection to obtain some of the advantages associated with the CWT.

We shall now introduce some notation for the function spaces used to demonstrate the link between the non-redundant projection and Softy space. Refer to Fig. 4.2 for the inter-relationships between these function spaces. Let V_1^L denote the subset of

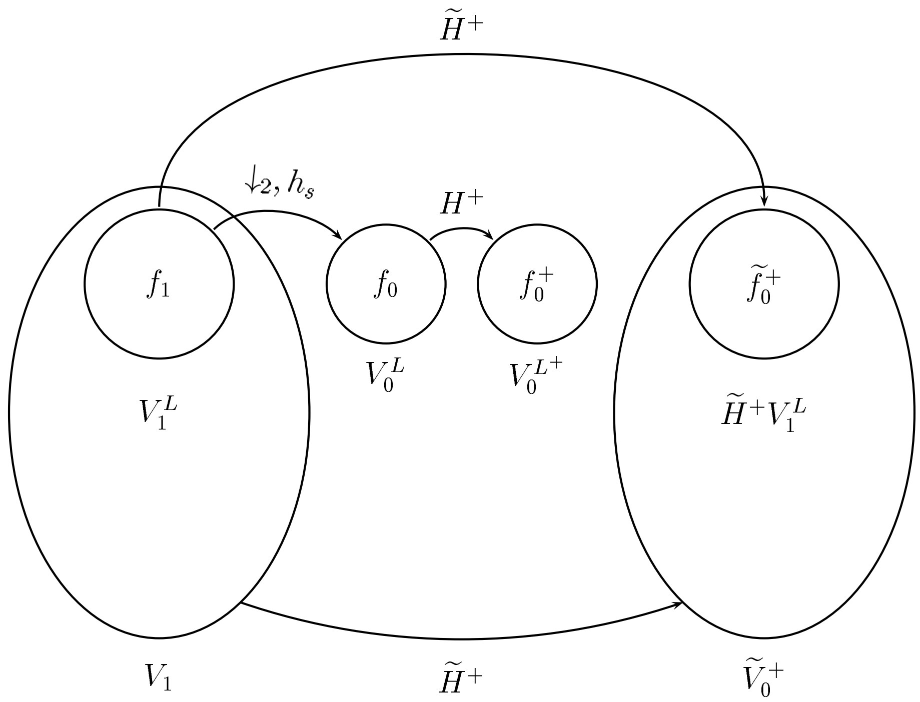


Figure 4.2 : The action of $\widetilde{H}^{+}$ and its significance.

lowpass functions in V_1 so that $\forall f \in V_1^L$ the scaling-coefficient sequences c associated with f satisfy the condition that $|C(\omega)|$ is small for $|\omega| \in [\pi/2, \pi)$. The function space V_1^L is mapped onto $\widetilde{\mathcal{H}}^+ V_1^L$ by the non-redundant projection. Let us now consider the action of the DWT on V_1^L. In the multiresolution analysis that underlies the DWT, we usually use V_0 to denote the lower resolution scaling space obtained by filtering and downsampling scaling-coefficient sequences associated with V_1 functions. We shall now use V_0^L to denote the subset of V_0 that is obtained from V_1^L functions by filtering and downsampling. Let V_0^{L+} represent the function space comprised of the Softy-space images of functions in V_0^L. We are now ready to demonstrate the link between functions in $\widetilde{\mathcal{H}}^+ V_1^L$ and V_0^{L+}, a subset of the Softy space.

Proposition 5 *Refer to Fig. 4.2. Consider an arbitrary function $f_1 \in V_1^L$, where f_1 is represented by the scaling-coefficient sequence c_1. Let f_0 denote the image of f_1 in the function space V_0^L. The function f_0 is represented by the scaling-coefficient sequence c_0 that is obtained after filtering c_1 with the scaling filter h_s and downsampling. Let f_0^+ be the Softy-space image of f_0. The function f_0^+ lies in the function space V_0^{L+}, and is represented by the scaling-coefficient sequence c_0^+. Now suppose that $\widetilde{f}_0^+$ is the image of f_1 under the non-redundant projection map $\widetilde{\mathcal{H}}^+$. The function $\widetilde{f}_0^+$ belongs to the function space $\widetilde{\mathcal{H}}^+ V_1^L$ and is represented by the scaling-coefficient sequence $\widetilde{c}_0^+$. Then the non-redundant projection $\widetilde{f}_0^+$ is approximately equal to the Softy-space function f_0^+.*

<u>Proof:</u>

Before rigorously proving the result, we would like to explain the reasoning behind the proof. Since f_1 has a lowpass characteristic, $|C_1(\omega)|$ may be represented as in Fig. 4.3. To obtain the non-redundant projection of f_1 we first filter $C_1(\omega)$ with $H^+(\omega)$ to get $C_1^+(\omega)$ sketched in Fig. 4.4. Next we downsample $C_1^+(\omega)$ to obtain $\widetilde{C}_0^+(\omega)$, the scaling-coefficient spectrum associated with the the non-redundant projection $\widetilde{f}_0^+$. We sketch $|\widetilde{C}_0^+(\omega)|$ in Fig. 4.5. Observe that the lowpass frequencies of $C_1(\omega)$ in

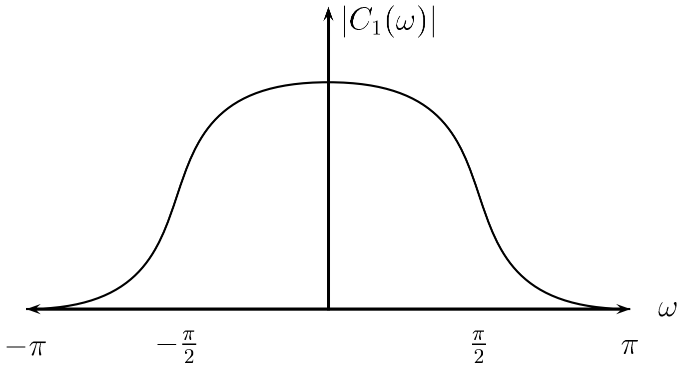


Figure 4.3 : $|C_1(\omega)|$

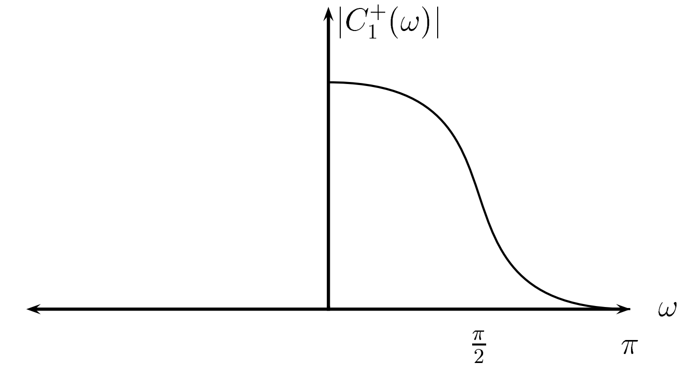


Figure 4.4 : $|C_1^+(\omega)|$

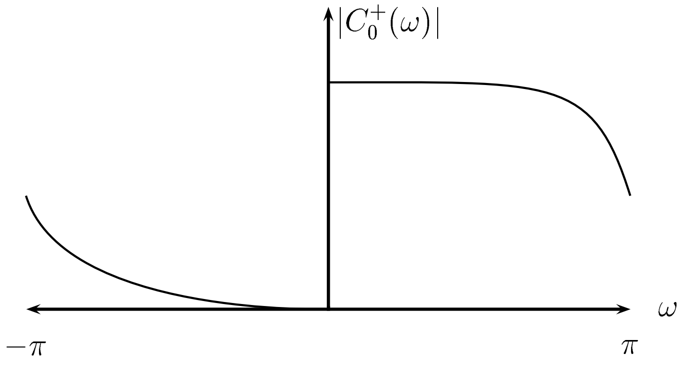


Figure 4.5 : $|C_0^+(\omega)|$

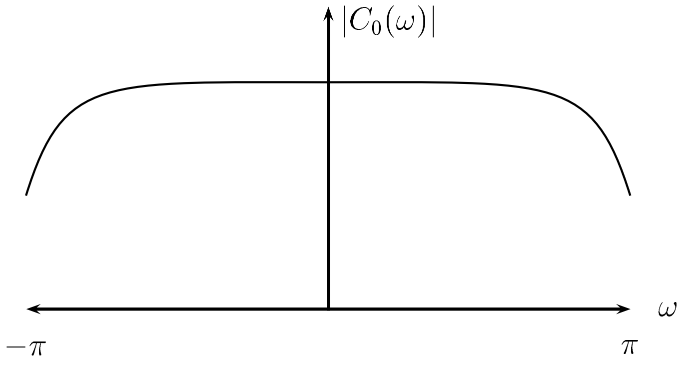


Figure 4.6 : $|C_0(\omega)|$

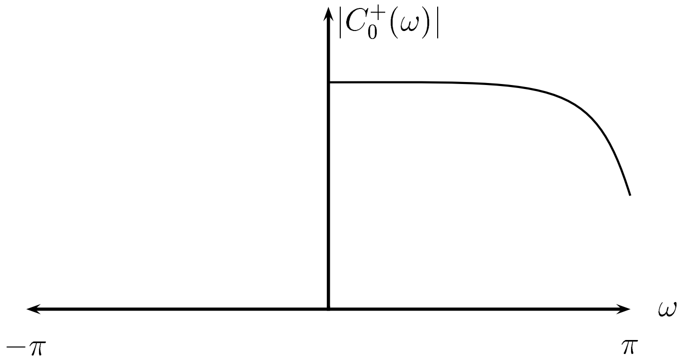


Figure 4.7 : $|C_0^+(\omega)|$

the interval $[0, \pi/2)$ occupy the interval $[0, \pi)$ in $\widetilde{C}_0^+(\omega)$. The highpass frequencies of $C_1(\omega)$ in the interval $[\pi/2, \pi)$ occupy the interval $[-\pi, 0)$ in $\widetilde{C}_0^+(\omega)$. Remember that $\widetilde{C}_0^+(\omega)$ is associated with the non-redundant projection $\widetilde{f}_0^+$.

Next we sketch scaling-coefficient spectra associated with f_0^+. After passing $C_1(\omega)$ through the scaling filter $H_s(\omega)$ and downsampling, we obtain $C_0(\omega)$ sketched in Fig. 4.6. Filtering $C_0(\omega)$ with the projection filter $H^+(\omega)$ gives $C_0^+(\omega)$ depicted in Fig. 4.7. Note that the lowpass frequencies of $C_1(\omega)$ now occupy the interval $[0, \pi)$ in $C_0^+(\omega)$ and that $|C_0^+(\omega)|$ is small in the negative-frequency interval $[-\pi, 0)$. Comparing $\widetilde{C}_0^+(\omega)$ and $C_0^+(\omega)$ in Figures 4.5, 4.7 we see that $\widetilde{C}_0^+(\omega) \approx C_0^+(\omega)$ provided that the negative frequencies in $|\widetilde{C}_0^+(\omega)|$ are small. As pointed out earlier, this condition is satisfied if the highpass frequencies in $|C_1(\omega)|$ are small. Therefore we conclude that $\widetilde{C}_0^+(\omega) \approx C_0^+(\omega)$ and hence $\widetilde{f}_0^+(\omega) \approx f_0^+(\omega)$ provided $f_1 \in V_1^L$.

Let us now proceed to the rigorous proof based on the above argument. We have

$$F_1(\omega) = C_1(\omega)\Phi(\omega), \forall\omega,$$

and

$$F_0(\omega) = C_0(2\omega)\Phi(2\omega), \forall\omega, \tag{4.1}$$

where c_0 and c_1 are related by

$$C_0(\omega) = C_1(\omega/2)H_s(\omega/2) + C_1(\omega/2+\pi)H_s(\omega/2+\pi), \forall\omega.$$

Since $f_1 \in V_1^L$, $|C_1(\omega)|$ is small in the range $|\omega| \in [\pi/2, \pi)$, and so $\forall\omega \in [-\pi, \pi]$, the term $C_1(\omega/2+\pi)H_s(\omega/2+\pi)$ is also small in magnitude. Hence

$$C_0(\omega) \approx C_1(\omega/2)H_s(\omega/2), \quad \forall\omega \in [-\pi, \pi].$$

Note that since $C_0(\omega)$ is still 2π-periodic, we may describe it for all ω as

$$C_0(\omega) \approx C_1\left(\frac{[\omega+\pi]\text{mod}2\pi - \pi}{2}\right) H_s\left(\frac{[\omega+\pi]\text{mod}2\pi - \pi}{2}\right), \forall\omega. \tag{4.2}$$

Now let us write $\widetilde{c}_0^+$ as,

$$\begin{aligned}
\widetilde{C}_0^+(\omega) &= C_1(\omega/2)H_+(\omega/2) + C_1(\omega/2+\pi)H_+(\omega/2+\pi), \forall\omega \in [-\pi, \pi] \\
&\approx C_1(\omega/2)H_+(\omega/2), \forall\omega \in [-\pi, \pi], \text{ since } C_1(\omega) \text{ has a lowpass characteristic.} \\
&\approx C_1(\omega/2)H_s(\omega/2)H^+(\omega/2), \forall\omega \in [-\pi, \pi], \text{ since } H_s(\omega/2) \approx 1, \ \forall\omega \in [-\pi, \pi].
\end{aligned}$$

Since $\widetilde{C}_0^+(\omega)$ is also 2π-periodic, we may describe it for all ω by

$$\begin{aligned}
\widetilde{C}_0^+(\omega) &\approx C_1\left(\frac{[\omega+\pi]\text{mod}2\pi - \pi}{2}\right) H_s\left(\frac{[\omega+\pi]\text{mod}2\pi - \pi}{2}\right) H^+(\omega), \forall\omega, \\
&\approx C_0(\omega)H^+(\omega), \forall\omega, \text{ from Equation 4.2,} \\
&= C_0^+(\omega), \ \forall\omega.
\end{aligned} \tag{4.3}$$

We conclude that $\widetilde{f}_0^+ \approx f_0^+$. $\square$

We have just proved that the non-redundant projection of an $L^2(\mathbb{R} \to \mathbb{R})$ function with a lowpass characteristic is approximately equal to the Softy-space projection of a low resolution version of the function. This result is significant because it empowers the non-redundant projection with some of the properties associated with projection filters that map $L^2(\mathbb{R} \to \mathbb{R})$ functions onto the Softy space.

Since $\widetilde{C}_0^+(\omega) \approx C_0^+(\omega)$ as proved above and sketched in Figures 4.5, 4.7 we can expect the DWT of $\widetilde{C}_0^+(\omega)$ to also provide useful phase information and improved directionality. Unfortunately, since the highpass frequencies of $C_1(\omega)$ fold into the negative frequencies, Strang and Nguyen's model is not applicable to $\widetilde{C}_0^+(\omega)$, unless these frequencies have extremely small magnitudes; hence we cannot expect the DWT of the non-redundant projection to enjoy reduced shift sensitivity.

The preceding proposition showed that the non-redundant projection of an $L^2(\mathbb{R} \to \mathbb{R})$ function approximates the Softy-space image of the low resolution version of the function provided that the function has a lowpass characteristic. Does this assumption really hold for real-world signals and images? When dealing with bandlimited digital signals, we can assume that they have a lowpass characteristic if they are sampled at twice the Nyquist rate. For natural images we are on firmer ground since it is widely accepted that although natural images are usually not bandlimited, they have a lowpass characteristic with low energy in the high-frequency interval [7, Pg. 176]. Having established the importance and applicability of the non-redundant projection, we shall now describe its properties.

4.3 The Non-Redundant Projection: A Realizable Unitary Map

In this section, we shall show that the non-redundant projection defines a realizable unitary map between V_1 and $\widetilde{V}_0^+$. To do this we must show that the map $\widetilde{\mathcal{H}}^+$ is linear, realizably bijective and inner-product preserving. The bijectiveness of the map will ensure that the non-redundant projection is invertible. We shall also demonstrate that the inverse of the non-redundant projection is realizable with digital filters. These properties ensure that the inverse of the non-redundant CWT is practically implementable. It is important to show that the non-redundant projection preserves inner-products because this allows the non-redundant CWT to be an orthogonal transform, an important requirement in many transform-based signal-processing applications.

Let us now begin the analysis of $\widetilde{\mathcal{H}}^+$. To demonstrate linearity, observe that this map is applied to functions in V_1 by filtering and downsampling the associated scaling-coefficient sequences. Both of these operations are linear. It follows that the non-redundant projection is also linear.

4.3.1 The Non-Redundant Projection: A Realizable Bijection

In the following proposition, we prove that the non-redundant projection is bijective.

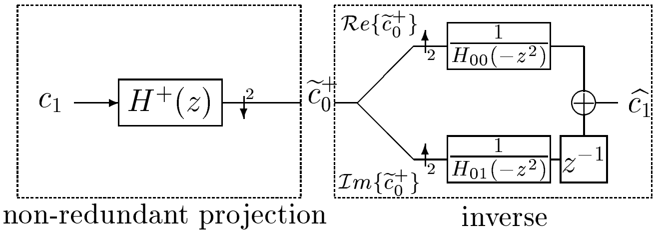


Figure 4.8 : The non-redundant projection and its inverse.

Proposition 6 *Let h_{00} and h_{01} be real-coefficient, allpass filters selected to be the even- and odd-polyphase components of a lowpass filter h_0. Create a projection filter h^+ by frequency shifting the lowpass frequency response $H_0(\omega)$ by $\pi/2$, so that $H^+(z) = H_0(-jz)$. Suppose that the scaling-coefficient sequence c_1 is associated with an arbitrary function f_1 in V_1. The "non-redundant projection" block in Fig. 4.8 illustrates the forward map $\widetilde{\mathcal{H}}^+$ to the scaling-coefficient sequence $\widetilde{c}_0^+$ associated with the non-redundant projection $\widetilde{f}_0^+$. The "inverse" block shows the inverse map $(\widetilde{\mathcal{H}}^+)^{-1}$ from $\widetilde{c}_0^+$ to $\hat{c}_1$. We have $\hat{c}_1 = c$.*

Proof: Selesnick [51] describes a design method for the allpass filters h_{00} and h_{01} so that

$$H_0(z) = H_{00}(z^2) + z^{-1}H_{01}(z^2)$$

is indeed a lowpass filter. He also argues that $H_0(z)$ is a root halfband filter that satisfies

$$H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) = 1.$$

Now we have $H_{+0}(z) = H_{00}(-z)$ and $H_{+1}(z) = jH_{01}(-z)$, where $H_{+0}(z)$ and $H_{+1}(z)$ are the even- and odd-polyphase components of the projection filter h_+. If the even- and odd-polyphase components of the scaling-coefficient sequence c_1 are c_{10} and c_{11}, then we can express $\widetilde{C}_0^+(z)$ as

$$\begin{aligned} \widetilde{C}_0^+(z) &= H_{+0}(z)C_{10}(z) + z^{-1}H_{+1}(z)C_{11}(z), \\ &= H_{00}(-z)C_{10}(z) + jz^{-1}H_{01}(-z)C_{11}(z). \end{aligned} \tag{4.4}$$

Now the kernel of the map $\widetilde{\mathcal{H}}^+$ consists of those scaling-coefficient sequences c_1 for which $\widetilde{c}_0^+ = 0$. Therefore, from Equation 4.4, we deduce that

$$\begin{aligned} \ker(\widetilde{\mathcal{H}}^+) &= \{C_1(z) : H_{00}(-z)C_{10}(z) = -jz^{-1}H_{01}(-z)C_{11}(z)\}, \\ &= \{0\}, \end{aligned}$$

since c_1 is a real-valued sequence and $H_{00}(-z), H_{01}(-z)$ are real-coefficient allpass filters that are non-zero on the unit circle. We now know that $\widetilde{\mathcal{H}}^+$ is invertible because $\ker(\widetilde{\mathcal{H}}^+) = \{0\}$. The inverse mapping will allow us to reconstruct c_1 from $\widetilde{c}_0^+$. From Equation 4.4, observe that

$$\begin{aligned} \mathcal{R}e\left\{\widetilde{C}_0^+(z)\right\} &= H_{00}(-z)C_{10}(z), \\ \mathcal{I}m\left\{\widetilde{C}_0^+(z)\right\} &= H_{01}(-z)C_{11}(z), \end{aligned}$$

since c_1 is a real-valued sequence and $H_{00}(-z), H_{01}(-z)$ are real-coefficient filters. Now from Fig. 4.8, we have

$$\begin{aligned}\hat{C}_1(z) &= \frac{1}{H_{00}(-z)}\mathcal{R}e\left\{\widetilde{C}_0^+(z^2)\right\} + z^{-1}\frac{1}{H_{01}(-z)}\mathcal{I}m\left\{\widetilde{C}_0^+(z^2)\right\}, \qquad (4.5)\\ &= \frac{1}{H_{00}(-z^2)}H_{00}(-z^2)C_{10}(z^2) + z^{-1}\frac{1}{H_{01}(-z^2)}H_{01}(-z^2)C_{11}(z^2),\\ &= C_{10}(z^2) + z^{-1}C_{11}(z^2),\\ &= C_1(z).\end{aligned}$$

Hence $\hat{c}_1 = c_1$. We conclude by emphasizing that the inverse map $(\widetilde{\mathcal{H}}^+)^{-1}$ is indeed realizable since

$$\begin{aligned}\frac{1}{H_{00}(-z)} &= H_{00}(-z^{-1}),\\ \frac{1}{H_{01}(-z)} &= H_{01}(-z^{-1}),\end{aligned}$$

because $H_{00}(-z), H_{01}(-z)$ are allpass filters. $\square$

The above proposition proves that the realizable forward map and inverse map of Fig. 4.8 implement the non-redundant projection $\widetilde{\mathcal{H}}^+$ and its inverse $(\widetilde{\mathcal{H}}^+)^{-1}$ on the scaling-coefficient sequences associated with V_1 functions. Hence $\widetilde{\mathcal{H}}^+$ is a realizable bijective map.

Allpass polyphase filters of the type described in the above proposition have also been applied to subband image coding by Smith and Eddins [52]. They stress the implementation efficiency and excellent frequency-response characteristics of these filters. In particular, an L order filter can be implemented with only L multiplies and $2L$ adds. This implementation provides a filter requiring only about one multiply per output sample to provide a magnitude response comparable to that of a 32-tap power complementary FIR filter. However, for some applications it is desirable to use FIR filters exclusively. We examine this issue in a later section. We shall now consider the preservation of inner-products by the realizable, linear, bijective map $\widetilde{\mathcal{H}}^+$.

4.3.2 Preservation of Inner Products

In Chapter 1, we showed that although Hardy-space projections and Softy-space projections preserve the norm induced by the $L^2(\mathbb{R} \to \mathbb{C})$ inner product, they do not, however, preserve the $L^2(\mathbb{R} \to \mathbb{C})$ inner product itself. Consequently, we were forced to redefine the inner product to obtain a projection that preserves an inner-product. In this section, we shall demonstrate that the identical situation arises for the non-redundant projection. In Proposition 7, we prove that the non-redundant projection does not preserve the $L^2(\mathbb{R} \to \mathbb{C})$ inner product. Then, in Proposition 8, we show that the non-redundant projection does preserve the norm induced by the $L^2(\mathbb{R} \to \mathbb{C})$ inner product. Finally, in Proposition 9, we prove that the non-redundant projection preserves the Softy-space inner product, a re-defined inner product.

Proposition 7 *Suppose that h_+ is a projection filter created by shifting the frequency response of a lowpass filter $H_0(\omega)$ by $\pi/2$ so that $H_+(z) = H_0(-jz)$. Implement the non-redundant projection of Fig. 4.8 with the projection filter h_+. Then the non-redundant projection does not preserve the $L^2(\mathbb{R} \to \mathbb{C})$ inner product.*

Proof: Let $H_{+0}(z), H_{+1}(z)$ be the even- and odd-polyphase components of the projection filter $H_+(z)$ and let $H_{00}(z), H_{01}(z)$ be the even- and odd-polyphase components of the lowpass filter $H_0(z)$. We have

$$\begin{aligned} H_{+0}(z) &= H_{00}(-z), \quad (4.6) \\ H_{+1}(z) &= jH_{01}(-z). \quad (4.7) \end{aligned}$$

Suppose that c_0 and c_1 are the scaling-coefficient sequences associated with arbitrary functions $f_0, f_1 \in V_1$. Perform the non-redundant projection of Fig. 4.1 on c_0, c_1 to obtain the scaling-coefficient sequences d_0, d_1 associated with the functions $g_0, g_1 \in \widetilde{V}_0^+$. Strang and Nguyen [46, Pg. 115] demonstrate that the decimated output of the filter h_+ may be represented in polyphase notation using

$$D_i(z) = H_{+0}(z)C_{i0}(z) + z^{-1}H_{+1}(z)C_{i1}(z), \;\; i = 0, 1,$$

where $C_{i0}(z), C_{i1}(z)$ are the even- and odd- polyphase components of $C_i(z),\ i = 0, 1$. Using Equations 4.6 and 4.7 we get

$$D_i(z) = H_{00}(-z)C_{i0}(z) - jz^{-1}H_{01}(-z)C_{i1}(z),\ i = 0, 1.$$

This equation may be expressed concisely using matrix notation to obtain

$$d_i = H_{00}c_{i0} - jH_{01}c_{i1},\ i = 0, 1,$$

where H_{00} and H_{01} are the Toeplitz matrices associated with the filters $H_{00}(-z)$ and $z^{-1}H_{01}(-z)$ respectively. Now the inner product between g_0 and g_1 is

$$\begin{aligned}
\langle g_0, g_1 \rangle &= \langle d_0, d_1 \rangle, \\
&= d_1^H d_0, \\
&= \left[H_{00}c_{00} - jH_{01}c_{01}\right]^H \left[H_{00}c_{10} - jH_{01}c_{11}\right], \\
&= \left[H_{00}c_{00} + jH_{01}c_{01}\right]^T \left[H_{00}c_{10} - jH_{01}c_{11}\right], \\
&= c_{00}^T H_{00}^T H_{00} c_{10} + c_{01}^T H_{01}^T H_{01} c_{11} - j\left[c_{00}^T H_{00}^T H_{01} c_{11} - c_{01}^T H_{01}^T H_{00} c_{10}\right]. \qquad (4.8)
\end{aligned}$$

We can express the inner product between f_0 and f_1 as

$$\begin{aligned}
\langle f_0, f_1 \rangle &= \langle c_0, c_1 \rangle, \\
&= c_1^T c_0, \\
&= c_{00}^T c_{10} + c_{01}^T c_{11}. \qquad (4.9)
\end{aligned}$$

Comparing Equations 4.8 and 4.9, we find that $\langle g_0, g_1 \rangle = \langle f_0, f_1 \rangle$ iff

$$\begin{aligned}
H_{00}^T H_{00} &= H_{01}^T H_{01} = I, \qquad (4.10) \\
\text{and } H_{00}^T H_{01} &= H_{01}^T H_{00} = 0. \qquad (4.11)
\end{aligned}$$

Realize that Equation 4.10 expresses the following condition on the auto-correlation filters $H_{00}(z^{-1})H_{00}(z)$ and $H_{01}(z^{-1})H_{01}(z)$:

$$H_{00}(z^{-1})H_{00}(z) = H_{01}(z^{-1})H_{01}(z) = 1.$$

This implies that

$$|H_{00}(\omega)|^2 = |H_{01}(\omega)|^2 = 1, \ \forall\omega.$$

Hence

$$H_{00}(\omega), H_{01}(\omega) \text{ are allpass filters.} \tag{4.12}$$

Similarly Equation 4.11 implies that

$$H_{00}(\omega)^* H_{01}(\omega) = 0, \ \forall\omega.$$

This condition cannot be satisfied when $H_{00}(\omega)$ and $H_{01}(\omega)$ are both allpass filters as required by Equation 4.12. Hence Equations 4.10 and 4.11 contradict each other. Therefore $\langle g_0, g_1\rangle \neq \langle f_0, f_1\rangle$. We conclude that the non-redundant projection fails to preserve the $L^2(\mathbb{R} \to \mathbb{C})$ inner product. $\square$

In the following proposition, we establish an isometry (distance-preserving map) between the function spaces at the input and output of the non-redundant projection.

Proposition 8 *Let h_+ be a projection filter that is created by frequency shifting a lowpass filter $H_0(\omega)$ so that $H_+(z) = H_0(-jz)$. The non-redundant projection using h_+ preserves the norm induced by the $L^2(\mathbb{R} \to \mathbb{C})$ inner product iff h_0 is the sum of allpass polyphase components as described in Proposition 6.*

Proof: Let $H_{+0}(z), H_{+1}(z)$ be the even- and odd-polyphase components of the projection filter $H_+(z)$ and let $H_{00}(z), H_{01}(z)$ be the even- and odd-polyphase components of the lowpass filter $H_0(z)$. We have

$$H_{+0}(z) = H_{00}(-z), \tag{4.13}$$

$$H_{+1}(z) = jH_{01}(-z). \tag{4.14}$$

Suppose that c_1 is the scaling-coefficient sequence associated with an arbitrary function $f_1 \in V_1$. Perform the non-redundant projection of Fig. 4.1 on c_1 to obtain the scaling-coefficient sequence d_0 associated with the function $g_0 \in \widetilde{V}_0^+$. Strang and

Nguyen [46, Pg. 115] demonstrate that the decimated output of the filter h_+ may be represented in polyphase notation using

$$D_0(z) = H_{+0}(z)C_{10}(z) + z^{-1}H_{+1}(z)C_{11}(z),$$

where $C_{10}(z), C_{11}(z)$ are the even- and odd- polyphase components of $C_1(z)$. Using Equations 4.13 and 4.14 we get

$$D_0(z) = H_{00}(-z)C_{10}(z) - jz^{-1}H_{01}(-z)C_{11}(z).$$

This equation may be expressed concisely using matrix notation to obtain

$$d_0 = H_{00}c_{10} - jH_{01}c_{11},$$

where H_{00} and H_{01} are the Toeplitz matrices associated with the filters $H_{00}(-z)$ and $z^{-1}H_{01}(-z)$ respectively. Now

$$\begin{aligned}
\|g_0\|^2 = \|d_0\|^2 & \\
&= d_0^H d_0, \\
&= \left[H_{00}c_{00} - jH_{01}c_{01}\right]^H \left[H_{00}c_{00} - jH_{01}c_{01}\right], \\
&= \left[H_{00}c_{00} + jH_{01}c_{01}\right]^T \left[H_{00}c_{00} - jH_{01}c_{01}\right], \\
&= c_{00}^T H_{00}^T H_{00}c_{00} + c_{01}^T H_{01}^T H_{01}c_{01} - j\left[(H_{00}c_{00})^T H_{01}c_{01} - (H_{01}c_{01})^T H_{00}c_{00}\right], \\
&= c_{00}^T H_{00}^T H_{00}c_{00} + c_{01}^T H_{01}^T H_{01}c_{01}, \\
&= c_{00}^T c_{00} + c_{01}^T c_{01}, \text{ iff } H_{00}^T H_{00} = H_{01}^T H_{01} = I, \\
&= c_0^T c_0, \\
&= \|c_0\|^2, \\
&= \|f_0\|^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
\|g_0\|^2 = \|f_0\|^2 &\Leftrightarrow H_{00}^T H_{00} = H_{01}^T H_{01} = I, \\
&\Leftrightarrow \quad H_{00}(z), H_{01}(z) \text{ are allpass filters, as proved in Proposition 7.}
\end{aligned}$$

We conclude that the non-redundant projection using h_+ preserves the norm induced by the $L^2(\mathbb{R} \to \mathbb{C})$ inner product iff h_0 is the sum of allpass polyphase components. $\square$

We now demonstrate that the non-redundant projection preserves the Softy-space inner product.

Proposition 9 *Let h_+ be a projection filter that is created by frequency shifting a lowpass filter $H_0(\omega)$ so that $H_+(z) = H_0(-jz)$. The non-redundant projection using h_+ preserves the Softy-space inner product iff h_0 is the sum of allpass polyphase components as described in Proposition 6.*

Proof:

Using the same definitions as in the proof of Proposition 7, the Softy-space inner product between g_0 and g_1 is

$$\begin{aligned}
\langle g_0, g_1 \rangle_+ &= \langle d_0, d_1 \rangle_+ \,, \\
&= \mathcal{R}e\left\{d_1^H d_0\right\}, \\
&= \mathcal{R}e\left\{\left[H_{00}c_{00} - jH_{01}c_{01}\right]^H \left[H_{00}c_{10} - jH_{01}c_{11}\right]\right\}, \\
&= \mathcal{R}e\left\{\left[H_{00}c_{00} + jH_{01}c_{01}\right]^T \left[H_{00}c_{10} - jH_{01}c_{11}\right]\right\}, \\
&= c_{00}^T H_{00}^T H_{00} c_{10} + c_{01}^T H_{01}^T H_{01} c_{11}.
\end{aligned} \tag{4.15}$$

Now we can express the inner product between f_0 and f_1 as

$$\begin{aligned}
\langle f_0, f_1 \rangle &= \langle c_0, c_1 \rangle \,, \\
&= c_1^T c_0, \\
&= c_{00}^T c_{10} + c_{01}^T c_{11}.
\end{aligned} \tag{4.16}$$

Comparing Equations 4.15 and 4.16, we find that

$$\begin{aligned}
\langle g_0, g_1 \rangle = \langle f_0, f_1 \rangle &\Leftrightarrow H_{00}^T H_{00} = H_{01}^T H_{01} = I, \\
&\Leftrightarrow \quad H_{00}(z), H_{01}(z) \text{ are allpass filters, as proved in Proposition 7.}
\end{aligned}$$

We conclude that the non-redundant projection using h_+ preserves the Softy-space inner product iff h_0 is the sum of allpass polyphase components. $\square$

Alternate Proof: Using the same definitions as in the preceding proof, the polarization identity for inner-product spaces [48, Chapter 1] states that

$$\langle f_0, f_1 \rangle = \frac{1}{4}\left[\|f_0 + f_1\|^2 - \|f_0 - f_1\|^2 + j\|f_0 + jf_1\|^2 - j\|f_0 - jf_1\|^2\right].$$

But since f_0 and f_1 are real-valued functions, their $L^2(\mathbb{R} \to \mathbb{C})$ inner product must also be real-valued. Therefore

$$\langle f_0, f_1 \rangle = \frac{1}{4}\left[\|f_0 + f_1\|^2 - \|f_0 - f_1\|^2\right]. \tag{4.17}$$

Now

$$\begin{aligned}
\langle g_0, g_1 \rangle_+ &= \mathcal{R}e\left\{\langle g_0, g_1 \rangle\right\}, \\
&= \mathcal{R}e\left\{\frac{1}{4}\left[\|g_0 + g_1\|^2 - \|g_0 - g_1\|^2 + j\|g_0 + jg_1\|^2 - j\|g_0 - jg_1\|^2\right]\right\}, \\
&\quad \text{by the polarization identity,} \\
&= \frac{1}{4}\left[\|g_0 + g_1\|^2 - \|g_0 - g_1\|^2\right], \\
&= \frac{1}{4}\left[\|f_0 + f_1\|^2 - \|f_0 - f_1\|^2\right], \text{ using Proposition 8,} \\
&= \langle f_0, f_1 \rangle, \text{ using Equation 4.17.}
\end{aligned}$$

Recall that Proposition 8 holds iff h_0 is the sum of allpass polyphase components. We conclude that the non-redundant projection using h_+ preserves the Softy-space inner product iff h_0 is the sum of allpass polyphase components. $\square$

4.3.3 The Non-Redundant CWT

Now that we have defined the non-redundant projection and characterized its properties, we are ready to introduce the non-redundant CWT. We refer the reader to Figure 4.9. The non-redundant projection $\widetilde{\mathcal{H}}^+$ maps $f \in V_1^L$ to $\widetilde{f}^+ \in \widetilde{\mathcal{H}}^+ V_1^L$. The non-redundant CWT consists of the DWT of $\widetilde{f}^+$ and is represented by DWT($\widetilde{f}^+$) in the figure. We may perform application-dependent processing of the non-redundant CWT

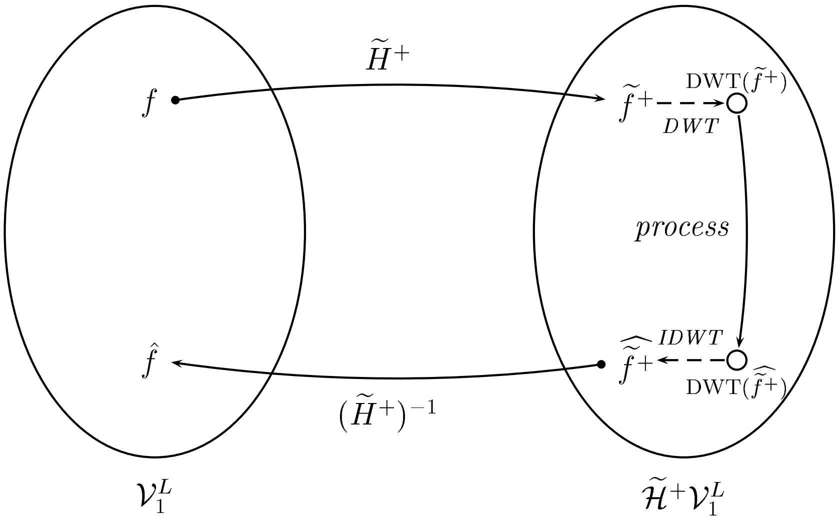


Figure 4.9 : The non-redundant CWT.

coefficients to obtain DWT($\widehat{\widetilde{f}^+}$). The inverse DWT yields $\widehat{\widetilde{f}^+}$, the non-redundant projection of the processed function $\widehat{f}$. To recover $\widehat{f}$, perform the inverse map $(\widetilde{\mathcal{H}}^+)^{-1}$ on $\widehat{\widetilde{f}^+}$.

4.4 Approximating Allpass Filters

In Propositions 8 and 9 we proved that the non-redundant projection preserves the norm and the Softy-space inner product iff the projection filter used to implement the projection is created from a lowpass filter that is the sum of allpass polyphase components. Since a non-trivial allpass filter has an infinite-length impulse response, the projection and inverse-projection filter will be IIR filters. In this section we explain the disadvantages of IIR filters in certain applications. We then examine various allpass-filter approximations that overcome these disadvantages.

4.4.1 Disadvantages of IIR Filters

We consider two disadvantages. The first disadvantage is associated with IIR filters in general, irrespective of the underlying application. To free the non-redundant CWT from this disadvantage, in Section 4.4.2 we avoid IIR filters altogether by using FIR approximations to the IIR filters in the implementation of the non-redundant projection and its inverse.

The second disadvantage affects applications in which it is essential to guarantee that processing artifacts are localized in the processed function $\widehat{f}$ of Fig. 4.9. However, the IIR filters in the implementation of $(\widetilde{\mathcal{H}}^{+})^{-1}$ tend to distribute processing artifacts globally in the processed function $\widehat{f}$. We stress that the severity of this disadvantage is reduced drastically by ensuring that the IIR filters have rapid spatial-domain decay. In Sections 4.4.4 and 4.4.5, we shall consider two different methods to avoid these artifacts by using FIR approximations to the IIR filters in the inverse projection.

4.4.2 Implementing the Non-Redundant Projection with FIR Filters

In this section we examine the first disadvantage of IIR filters: the presence of round-off error and the need to estimate initial conditions. These problems are associated with IIR filters irrespective of the underlying application. Several researchers have suggested techniques to mitigate round-off error. However, if round-off error and initial-condition estimation are unacceptable in a certain application, then FIR approximations to IIR filters must be used. We now explain how this may be done in the implementation of $\widetilde{\mathcal{H}}^{+}$ and $(\widetilde{\mathcal{H}}^{+})^{-1}$. Recall from Proposition 6 that the implementation of both maps require allpass IIR filters if $\widetilde{\mathcal{H}}^{+}$ preserves norms and the Softy-space inner product.

We shall first consider the consequences of FIR approximation of the IIR projection filter used in the non-redundant projection $\widetilde{\mathcal{H}}^{+}$. In Propositions 8 and 9 we proved that $\widetilde{\mathcal{H}}^{+}$ preserves norms and the Softy-space inner product iff the polyphase components of $H_{+}(z)$ are allpass IIR filters satisfying certain conditions. Hence if we use FIR-approximations to these IIR filters, the non-redundant projection will not preserve norms or Softy-space inner products. Nevertheless, it is possible to preserve norms in an approximate sense when the polyphase components of $H_+(z)$ are approximations to allpass filters. We explain this in detail in the following proposition.

Proposition 10 *Let h_0 be a lowpass filter created from the sum of polyphase components h_{00}, h_{01} that have frequency responses which are approximately allpass in nature. Create the projection filter h_+ from h_0 by frequency shifting $H_0(\omega)$ by $\frac{\pi}{2}$ so that $H_+(z) = H_0(-jz)$. Suppose that f_1 is an arbitrary function in V_1 and that $\widetilde{g}_0^+$ is the non-redundant projection of f_0 using the projection filter h_+. Then*

$$\beta\|f_0\|_2^2 \leq \|\widetilde{g}_0^+\|_2^2 \leq \alpha\|f_0\|_2^2,$$

where

$$\begin{aligned} \alpha &\triangleq \max\left(\sup_{\omega\in[-\pi,\pi)} |H_{00}(\omega)|^2, \sup_{\omega\in[-\pi,\pi)} |H_{01}(\omega)|^2\right), \\ \beta &\triangleq \min\left(\inf_{\omega\in[-\pi,\pi)} |H_{00}(\omega)|^2, \inf_{\omega\in[-\pi,\pi)} |H_{01}(\omega)|^2\right). \end{aligned}$$

Proof: Given an arbitrary scaling-coefficient sequence c, the norm of its non-redundant projection $\widetilde{d}^+$ is

$$\begin{aligned} \|\widetilde{d}^+\|_2^2 &= \sum_n \left|\widetilde{d}^+(n)\right|^2, \\ &= \sum_n \widetilde{d}_r^+(n)^2 + \widetilde{d}_i^+(n)^2, \text{ where } \widetilde{d}_r^+(n) = \mathcal{R}e\left\{\widetilde{d}^+(n)\right\}, \ \widetilde{d}_i^+(n) = \mathcal{I}m\left\{\widetilde{d}^+(n)\right\}, \\ &= \|\widetilde{d}_r^+\|_2^2 + \|\widetilde{d}_i^+\|_2^2, \\ &= \|h_{00} * c_0\|_2^2 + \|h_{01} * c_1\|_2^2, \\ &\quad \text{where } h_{00},\ h_{01} \text{ represent the even and odd polyphase components of } h_0, \\ &\quad \text{and } c_0,\ c_1 \text{ represent the even and odd polyphase components of } c, \\ &= \int_{-\pi}^{\pi} |H_{00}(\omega)C_0(\omega)|^2\, d\omega + \int_{-\pi}^{\pi} |H_{01}(\omega)C_1(\omega)|^2\, d\omega. \end{aligned}$$

This expression will help us to determine the unit-norm function f_{max} that maximizes the norm of its non-redundant projection g_{max}. Let f_{max} be associated with the scaling-coefficient sequence c_{max}. We may assume without loss of generality that $\alpha = \sup_{\omega\in[-\pi,\pi)} |H_{00}(\omega)|$. Defining

$$\omega_\alpha \stackrel{\triangle}{=} \arg \max_{\omega\in[-\pi,\pi)} |H_{00}(\omega)|^2,$$

it is clear from the previous expression that c_{max} may be described by its polyphase components $c_{max,0},\ c_{max,1}$ as follows, $\forall n \in \mathbb{Z}$,:

$$\begin{aligned} c_{max,0} &= \cos \omega_\alpha n, \\ c_{max,1} &= 0, \end{aligned}$$

so that

$$C_{max,0}(\omega) = \frac{1}{2}\delta(\omega_\alpha) + \frac{1}{2}\delta(-\omega_\alpha).$$

Suppose that $\widetilde{d}^+_{max}$ is the non-redundant projection of c_{max}. Then we have

$$\begin{aligned} \|\widetilde{d}^+_{max}\|_2^2 &= \int_{-\pi}^{\pi} |H_{00}(\omega)|^2 \left|\frac{1}{2}\delta(\omega_\alpha) + \frac{1}{2}\delta(-\omega_\alpha)\right|^2 d\omega, \\ &= \alpha. \end{aligned}$$

Hence for any unit-norm function f in V_1, the norm of $\widetilde{g}^+$, the non-redundant projection of f satisfies

$$\|\widetilde{g}^+\|_2^2 \leq \alpha.$$

This implies that the norm of $\widetilde{g}^+_0$, the non-redundant projection of an arbitrary function $f_0 \in V_1$ satisfies

$$\|\widetilde{g}^+_0\|_2^2 \leq \alpha\|f_0\|^2_.$$

A similar argument will show that

$$\beta\|f_0\|_2^2 \leq \|\widetilde{g}^+_0\|_2^2.$$

Hence the result. $\square$

Let $H^F_{00}(z)$ and $H^F_{01}(z)$ be FIR approximations to the allpass filters $H_{00}(z)$ and $H_{01}(z)$ that would be used to implement a norm-preserving non-redundant projection.

To appreciate the significance of the above result, observe that $\beta \to \alpha$ and $\alpha \to 1$ as the FIR approximations to the allpass filters $H_{00}(z)$ and $H_{01}(z)$ are improved. Hence Proposition 10 relates the allpass-filter approximation to the degree of norm preservation provided by the non-redundant projection.

We now consider the choice of inverse-projection filters that will invert the non-redundant projection implemented with a projection filter created from $H_{00}^F(z)$ and $H_{01}^F(z)$. From the proof of Proposition 6 it is clear that the allpole filters $1/H_{00}^F(-z), 1/H_{01}^F(-z)$ can be used as inverse-projection filters, provided that these filters have no poles on the unit circle. However, since we wish to avoid IIR filters, we shall approximate $1/H_{00}^F(-z)$ and $1/H_{01}^F(-z)$ with the FIR filters $G_{00}^F(z)$ and $G_{01}^F(z)$. Unfortunately, since $1/H_{00}^F(-z)$ and $1/H_{01}^F(-z)$ is the only filter pair that inverts the non-redundant projection, the FIR approximations $G_{00}^F(z)$ and $G_{01}^F(z)$ will not provide perfect reconstruction of the input sequence before non-redundant projection. The reconstruction error will be determined by the quality of the FIR approximation to the IIR inverse projection filters. For applications such as data compression, Near Perfect Reconstruction (NPR) is not a severe drawback.

In this subsection, we explained that roundoff error and initial-condition computation are inherent drawbacks of IIR filters. We then showed that the allpass IIR filters used to implement the non-redundant projection and its inverse may be replaced with FIR approximations, thereby eliminating the inherent drawbacks of IIR filters. The consequence of using FIR filters in the non-redundant projection is that norms are only preserved approximately. We proved Proposition 10 that made this notion more explicit. We also argued that replacing the inverse-projection IIR filters with FIR approximations would create an NPR system. In the next section we examine the second disadvantage of IIR filters: global distribution of processing artifacts. We shall see that it is possible to overcome this disadvantage without sacrificing the perfect reconstruction property.

4.4.3 Global Distribution of Processing Artifacts in the Non-Redundant CWT

In Propositions 8 and 9 , we proved that allpass IIR filters must be used to invert a non-redundant projection that preserves norms and the Softy-space inner product. It is well known that filter banks with IIR synthesis filters can create processing artifacts that are spread throughout a large region in the processed signal. We shall now use the notation in Fig. 4.9 to explain this phenomenon in the context of the non-redundant CWT.

Let $\widehat{c}$ and $\widehat{\widetilde{c}^{+}}$ be the scaling-coefficient sequences associated with the functions $\widehat{f}$ and $\widehat{\widetilde{f}^{+}}$ respectively. Now Equation 4.5 that relates these scaling-coefficient sequences states that $\widehat{c_0}$ and $\widehat{c_1}$, the even- and odd-polyphase components of $\widehat{c}$ are related to $\widehat{\widetilde{c}^{+}}$ by

$$\begin{aligned} \widehat{c_0} &= g_{00} * \mathcal{R}e\left\{\widehat{\widetilde{c}^{+}}\right\}, \\ \widehat{c_1} &= g_{01} * \mathcal{I}m\left\{\widehat{\widetilde{c}^{+}}\right\}, \end{aligned}$$

where g_{00} and g_{01} are the infinite-length impulse responses of the allpass filters $1/H_{00}(-z^2)$ and $1/H_{01}(-z^2)$. Clearly these infinite-length impulse responses make each coefficient $\widehat{c}$ dependent on infinitely many scaling-coefficients from the sequence $\widehat{\widetilde{c}^{+}}$. This implies that a processing artifact in any scaling-coefficient $\widehat{\widetilde{c}^{+}}$ will affect the entire sequence $\widehat{c}$ and will hence be visible throughout the processed function $\widehat{f}$.

Evidently the IIR inverse-projection filters tend to distribute processing artifacts throughout the processed function $\widehat{f}$. This phenomenon may be mitigated by ensuring that g_{00} and g_{01} decay rapidly. However in certain applications it will be necessary to ensure that processing artifacts are localized within the processed function $\widehat{f}$. We shall now outline two different methods explaining how this may done by using FIR inverse-projection filters.

4.4.4 Localizing Artifacts and Preserving Norms

We wish to use FIR inverse-projection filters to localize processing artifacts in the output of the non-redundant CWT. Suppose that h_{00} and h_{01} represent the allpass polyphase filters that are used to implement the non-redundant projection that preserves norms and the Softy-Space inner product, as described in Proposition 8. In Proposition 6, we proved that the IIR filters $1/H_{00}(-z)$ and $1/H_{01}(-z)$ will invert this non-redundant projection. Let $H^F_{00}(z)$ and $H^F_{01}(z)$ represent FIR approximations to the IIR filters $1/H_{00}(-z)$ and $1/H_{01}(-z)$. Using these FIR approximation filters as inverse-projection filters will result in a non-redundant CWT with NPR. However, since the non-redundant projection still uses a projection filter derived from allpass polyphase components, it will preserve norms and the Softy-space inner product. Any processing artifacts in the non-redundant CWT output will be localized since the inverse-projection filters now have finite-length impulse responses. Next we propose a different method to localize artifacts and achieve Perfect Reconstruction (PR).

4.4.5 Localizing Artifacts and Maintaining PR

In the previous subsection, we used FIR approximations to IIR inverse-projection filters to localize processing artifacts. We saw that PR could no longer be achieved with that method. We now suggest a different method that does maintain a PR non-redundant CWT. Let $H_{00}(z)$ and $H_{01}(z)$ be the allpass polyphase filters used to implement a norm-preserving non-redundant projection. Suppose that $H^P_{00}(z)$ and $H^P_{01}(z)$ are allpole approximations to $H_{00}(z)$ and $H_{01}(z)$ respectively. Then $1/H^P_{00}(-z)$ and $1/H^P_{01}(-z)$ are FIR filters that may be used to invert the non-redundant projection. Observe that this implementation results in a PR non-redundant CWT and that any processing artifacts will be localized since the inverse-projection filters have finite-length impulse responses. However, since the projection filter is derived from allpole approximations to allpass filters, the non-redundant projection will preserve norms in an approximate sense only. We made this notion more explicit in Proposition 10.

4.4.6 FIR Approximations to IIR Filters

The preceding subsections outlined various methods for the implementation of the non-redundant projection and its inverse with FIR approximations to the allpass IIR filters that are required for an inner-product preserving, PR, non-redundant projection. We shall now conclude this chapter by explaining briefly how these FIR approximations can be obtained.

The simplest method to obtain an FIR approximation to an IIR filter is to evaluate the infinite-length impulse response of the IIR filter and to use a finite-length truncation of this impulse response as the desired FIR approximation. The advantage of this method lies in its simplicity, while the main disadvantage is that long FIR filters are required to obtain good approximations.

To overcome this disadvantage, Beylkin [53] describes a different method that provides a factored FIR approximation to an IIR filter. The technique is based on replacing each pole of the IIR filter by a cascade of zeros. The factored FIR filter may be implemented with less computational complexity than a comparable FIR approximation obtained by impulse-response truncation.

Chapter 5

Improving Directionality

The separable two-dimensional (2D) RWT is a powerful image-processing tool, but in some applications its poor directionality is a serious disadvantage. The 2D RWT can distinguish between only three different orientations of spatial features: horizontal, vertical and diagonal. In the next section, we explain the poor directionality of the 2D RWT based on the Fourier-plane tiling associated with its subbands. We shall then extend our one-dimensional projections to two dimensions. Applying 2D DWTs to such projections creates 2D complex wavelet transforms. The Fourier-plane partitioning of these transforms is responsible for the improvement in directionality: it enables distinctions between spatial features oriented at $+15, +45, +75, -75, -45, -15$ degrees.

5.1 Poor Directionality in the RWT

Figure 5.1 shows the Fourier-plane partitioning obtained from a separable, two-level 2D RWT. The HH_2 subband is associated with diagonally-oriented spatial features at scale 2 and concentrates its energy in the four blocks labeled HH_2. High energy at the output of the filter in the HH_2 subband indicates the presence of this class of features. The HH_2 blocks in the upper-left and lower-right corners indicate features whose gradients are oriented at -45 degrees, while those in the upper-right and lower-left corners indicate feature-gradient orientations of $+45$ degrees. Since all four blocks are associated with the output of a single filter in the RWT, we cannot differentiate between these two orientations. Similarly, feature-gradient orientations at $+15$ degrees and -15 degrees are indistinguishable from their wavelet coefficients

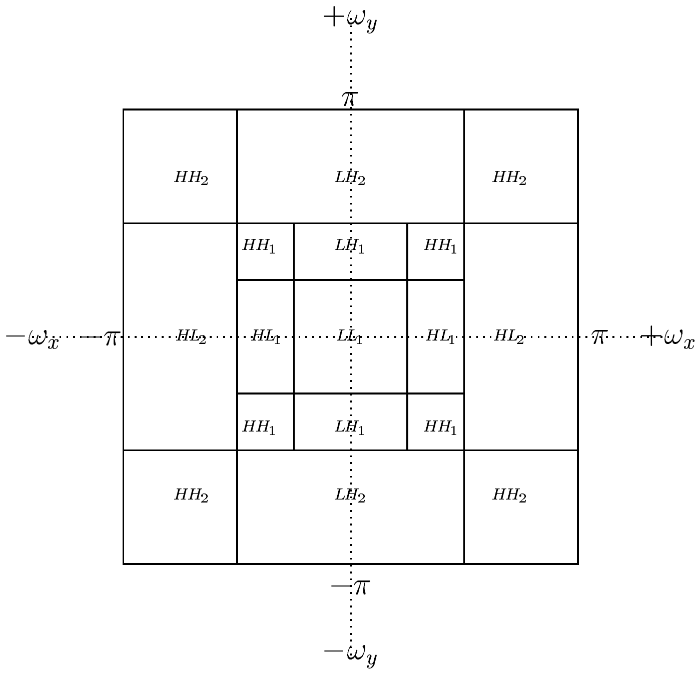


Figure 5.1 : Fourier-plane partitioning obtained from a separable two-level 2D RWT.

since they are both associated with energy in the HL_2 subband. Features oriented at +75 degrees and −75 degrees suffer a similar fate. It is evident that the poor directionality of the RWT is due to positively- and negatively-oriented blocks both being associated with the same subband. In the following section, we introduce the two-dimensional Softy space projection that will decouple these blocks, thereby enabling improved directionality.

5.2 The Two-Dimensional Softy-Space Projection

A one-dimensional $L^2(\mathbb{R} \to \mathbb{R})$ function is fully described by its positive frequencies in the Fourier domain. Since the negative frequencies carry redundant information they may be zeroed out thereby yielding the Hardy-space image of the function. We exploited this fact in Chapters 2 and 3 when we computed the Softy-space projection of an $L^2(\mathbb{R} \to \mathbb{R})$ function in order to obtain phase information and approximate shiftability in a wavelet transform. Similarly, for real-valued two-dimensional images,

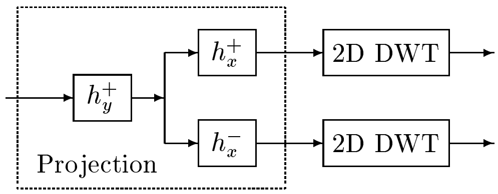


Figure 5.2 : The 2D CWT.

the lower half-plane of the Fourier transform carries redundant information. We may eliminate this redundant information to obtain a complex-valued spatial image with phase information by filtering the columns of the image with a projection filter h_y^+ as defined in Proposition 1 of Chapter 3. The subscript "y" indicates that filtering is performed along the image columns.

To eliminate the redundant information contained in the negative vertical frequencies, we use the projection filter h_y^+ to perform Softy-space projection of the image columns. In one dimension, elimination of redundant information along the negative half-line of the Fourier plane provides phase information and DWT shiftability. However, eliminating the negative vertical half-plane in the two-dimensional Fourier plane and then performing the DWT does not improve DWT directionality. To do so, we must decouple the positive, horizontal frequencies from the negative, horizontal frequencies. We show how this may be done in Figure 5.2. The h_y^+ projection filter retains the positive vertical frequencies. The h_x^+ filter is a projection filter that operates along the rows of the output from the h_y^+ filter, retaining the positive horizontal frequencies. Therefore the output from the h_x^+ filter is associated with the block labeled "h_x^+ output" in Fig. 5.3. This block contains the positive-vertical, positive-horizontal frequencies.

Recall that the h_x^+ projection filter is created by frequency-shifting the lowpass analysis filter from a two-band filter bank, so that h_x^+ retains positive frequencies and suppresses negative frequencies. The h_x^- projection filter is created similarly by fre-

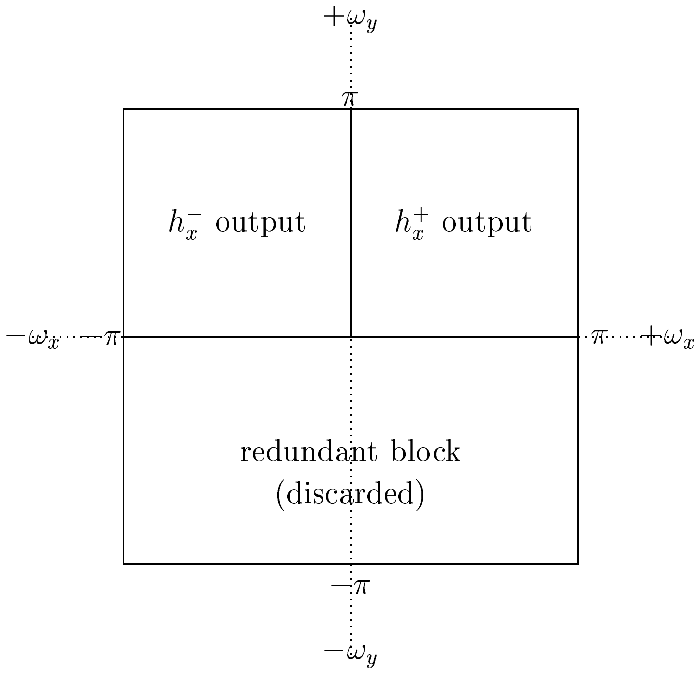


Figure 5.3 : Fourier-plane partitioning after 2D Softy-space projection.

quency shifting the highpass analysis filter corresponding to the above lowpass analysis filter. Therefore h_x^- retains negative horizontal frequencies and suppresses positive horizontal frequencies. The output from this filter is associated with the block labeled "h_x^- output" in Fig. 5.3. This block contains the positive-vertical, negative-horizontal frequencies. Therefore the "projection" module in Fig. 5.2 eliminates the redundant negative-vertical frequencies and decouples the positive horizontal frequencies from the negative horizontal frequencies, as is evident in Fig. 5.3.

Having defined the 2D extension of the Softy-space projection, we are ready to define the 2D CWT. This is easily done by performing 2D DWTs on the images output by the "Projection" module in Fig. 5.2. Applying a 2D two-level DWT to the output of the h_x^+ filter partitions the " h_x^+ output" block of Fig. 5.3 into the subbands with the "+" superscripts in Fig. 5.4. Similarly the 2D two-level DWT applied to the h_x^- filter output yields the subbands with the "-" superscripts in Fig. 5.4.

The 2D DWT of the 2D Softy-space projection of a real-valued image defines the 2D CWT. This transform has improved directionality because it provides six directional subbands at each scale. These subbands are the $HL^{+}, HH^{+}, LH^{+}, LH^{-}, HH^{-}, HL^{-}$ subbands that describe feature-gradients oriented at $+15, +45, +75, -75, -45, -15$ degrees. The directionality of the 2D CWT is significantly better than that of the 2D RWT with only three directional subbands: horizontal, vertical and diagonal.

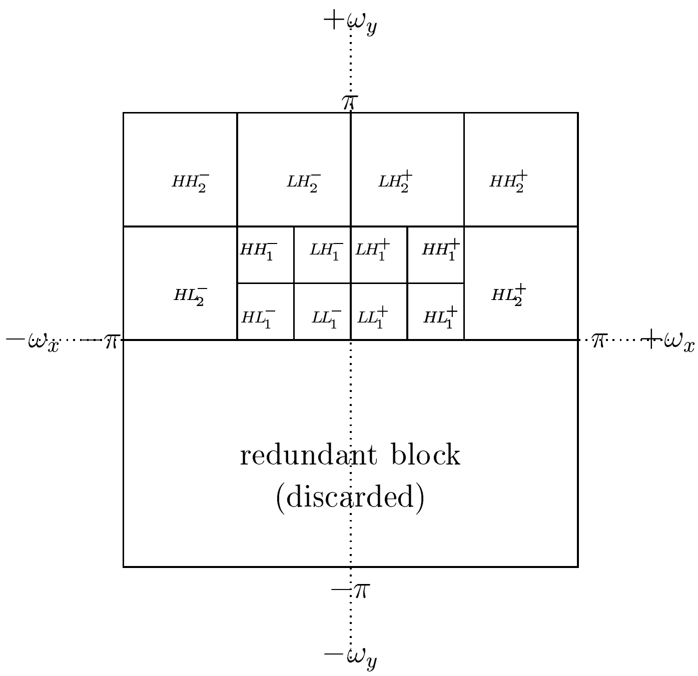
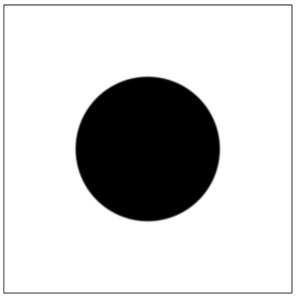


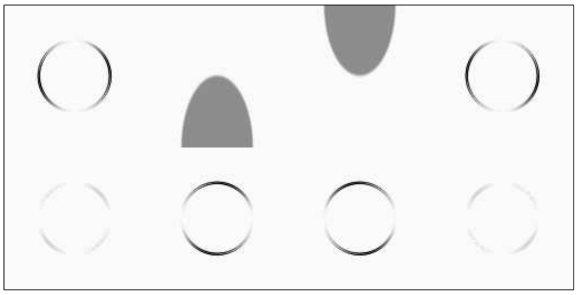
Figure 5.4 : Fourier-plane partitioning of the 2D two-level CWT.

Observe that the 2D DWTs of the decoupled positive-horizontal and negative-horizontal frequencies in Fig. 5.3 will provide two different lowpass subbands LL^{+} and LL^{-}. These subbands discriminate between low-scale features oriented at -45 and +45 degrees.

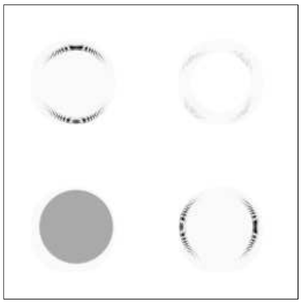
To compare the directional selectivity of various different transforms, we used the disc image displayed in Figure 5.5 (a). The separable, single-level 2D Daubechies DWT of the image is displayed in Figure 5.5 (c). The wavelet coefficients for both the $+45$ and -45 degree orientations appear in the diagonal-orientation subband and



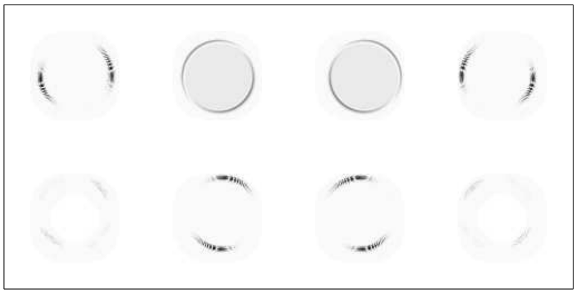
(a) Disc Image



(b) 2D DTWT



(c) 2D DWT



(d) 2D CWT

Figure 5.5 : **(a)** Disc image. **(b)** 2D DTWT using length-16 filters. Clockwise from upper-rightmost block: -75, +45, -15, +15, -45, +75 degree highpass subband, upper half of lowpass subband, lower half of lowpass subband. **(c)** 2D Daubechies DWT. Clockwise from upper-right block: vertical-, diagonal-, horizontal-orientation subbands and lowpass subband. **(d)** 2D CWT. Clockwise from upper-rightmost block: -75, +45, -15, +15, -45, +75 degree highpass subbands, -45, +45 degree lowpass subbands. Length-16 Daubechies filters are used for the wavelet decomposition and to generate the projection filter.

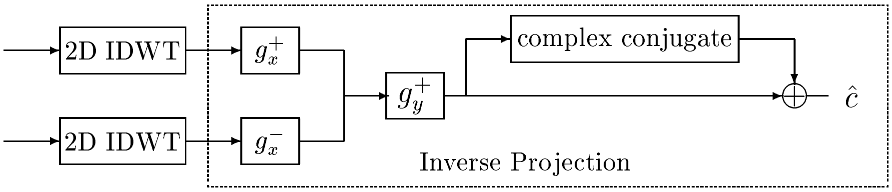


Figure 5.6 : The 2D ICWT.

cannot be distinguished. Figure 5.5 (b) shows the 2D, single-level Dual-Tree Wavelet Transform (DTWT) of the disc image. Figure 5.5 (d) presents the single-level 2D CWT obtained using the filter-bank structure of Figure 5.2. Both the 2D DTWT and the 2D CWT provide six highpass directional subbands and are four times redundant in two dimensions. We shall derive the redundancy for m-dimensional CWTs in Chapter 7. Observe that the 2D CWT also discriminates between features oriented at +45 and -45 degrees in the lowpass subband. The DTWT lowpass subband, however, is rather large and has been displayed in two halves.

5.3 The 2D Inverse CWT

In the previous section, we demonstrated that the 2D CWT has better directionality than the 2D RWT. In image-processing applications, we may process the 2D CWT coefficients based on the directional information contained in the subbands. To recover the directionally processed image, we must perform the 2D inverse CWT (2D ICWT) transform on the processed 2D CWT coefficients. In the absence of any processing, the 2D CWT followed by the 2D ICWT must perfectly reconstruct the original image.

Fig. 5.6 illustrates the 2D ICWT that inverts the 2D CWT thereby achieving perfect reconstruction. The blocks labeled "2D IDWT" invert the 2D DWTs in the 2D CWT of Fig. 5.2. The next step is to invert the h_x^+ and h_x^- filters that are applied to the image rows in the 2D DWT. Remember that h_x^+ and h_x^- are created by frequency shifting the analysis filters h_0 and h_1 by $\frac{\pi}{2}$. Now suppose that g_0 and g_1 are the synthesis filters associated with h_0 and h_1. Shift g_0 and g_1 by $\frac{\pi}{2}$ in the frequency domain to obtain g_x^+ and g_x^-. These filters constitute a 2-band synthesis filter bank that inverts the 2-band analysis filter bank consisting of h_x^+ and h_x^-. Therefore the sum of the outputs of g_x^+ and g_x^- in Fig. 5.6 recovers the output from h_y^+ in the 2D CWT of Fig. 5.2.

To invert the h_y^+ filter in the 2D CWT, recall that it is a projection filter operating on the input-image columns. In Proposition 1 in Chapter 2 we proved that the structure labeled "inverse map" in Fig. 2.5 would invert such a projection filter. Therefore, applying this structure to the image-columns as shown in Fig. 5.6 completes the perfect reconstruction of the input image by the 2D ICWT.

5.4 Approximate Shiftability in the 2D CWT

In Chapter 3 we proved that under certain conditions the CWT of $L^2(\mathbb{R} \to \mathbb{R})$ signals enjoys approximate shiftability, an important property that ensures that information is conserved within transform subbands under input-signal shifts. The proof is based on the reduced bandwidth of the Softy-space projection, due to the suppression of negative frequencies. The 2D CWT is a separable extension of the 1D CWT. Consequently, it is possible to generalize our earlier proof for approximate shiftability. We shall not provide the rigorous proof here. Instead we shall explain the reduction in bandwidth after the 2D Softy-space projection and indicate how this leads to approximate shiftability in the 2D CWT.

From Fig. 5.2 it is evident that the inputs to the 2D DWTs are images with frequency-domain supports labeled by the blocks "h_x^+ output" and "h_x^- output" in Fig. 5.3. Therefore each 2D DWT processes an image with bandwith π along the horizontal and vertical frequencies. In contrast, the 2D RWT operates on the real-valued image with bandwidth 2π along the vertical and horizontal frequencies. Due to this large bandwidth, the Nyquist criterion is violated in each 2D RWT subband and hence shiftability is not achieved. On the other hand, the reduced bandwidth of the inputs to the 2D DWTs in the 2D CWT enables the Nyquist criterion to be satisfied under certain conditions. This is an intuitive explanation for the approximate shiftability of the 2D CWT.

We used the following experiment to illustrate the approximate shiftability of the 2D CWT. Consider the disc image in Figure 5.5. We calculated the energy in the RWT level-2 HH subband and the CWT level-2 HH^+ subbands over 16 input-image shifts. Each shift was by one pixel rightward and one pixel downward. Figure 5.7 shows the subband-energy variations versus shift for the two transforms. The RWT subband energy is highly variable with respect to shift. In contrast, CWT subband energy remains approximately constant due to the approximate shiftability of the 2D CWT.

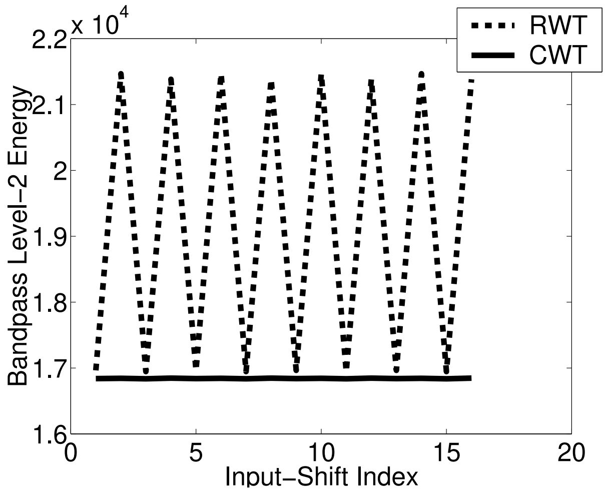


Figure 5.7 : Subband-energy\Input-shift for RWT level-2 HH and CWT level-2 HH^+ subbands.

5.5 The 2D Non-Redundant CWT

In Chapter 4 we defined the non-redundant CWT, a wavelet transform that offers phase information without incurring any data redundancy. We shall now extend this

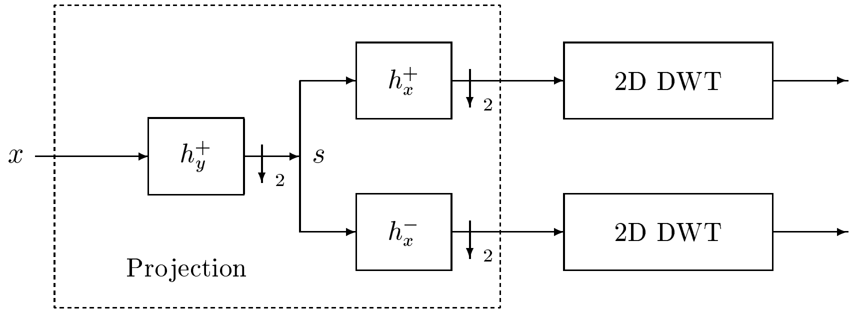


Figure 5.8 : The non-redundant 2D CWT

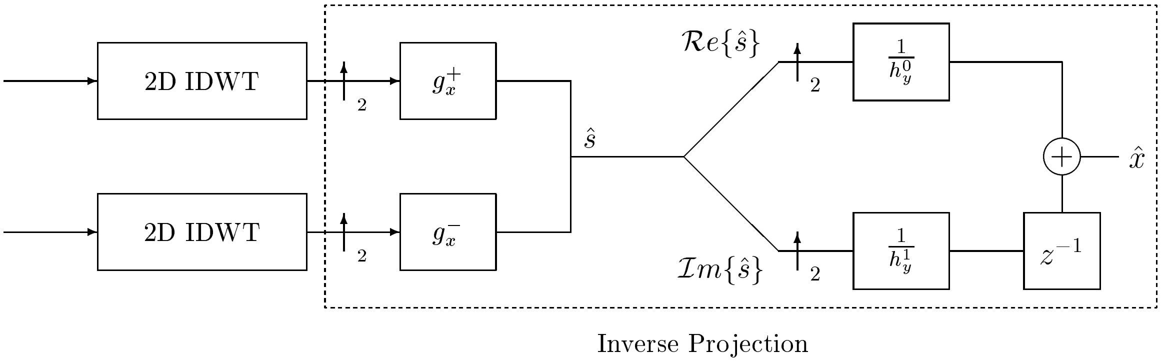


Figure 5.9 : The non-redundant 2D ICWT. The symbols $\frac{1}{h_y^0}, \frac{1}{h_y^0}$ represent the filters $\frac{1}{H_{00}(-z^2)}, \frac{1}{H_{00}(-z^2)}$ operating along the image columns.

transform to two dimensions so that it may be applied to images. This will provide improved directionality without any additional redundancy.

In the previous section, we created the directional 2D CWT by first applying projection filters separably to the rows and columns of a real-valued image and then computing 2D DWTs of the projected images. The projection filters are used to eliminate redundant negative vertical frequencies and to decouple positive horizontal frequencies from negative horizontal frequencies. The non-redundant 2D CWT is defined in a similar manner: non-redundant projections are applied separably to the rows and columns of a real-valued image to eliminate redundancy and decouple frequencies prior to 2D DWT computation. This is illustrated in Fig. 5.8. A non-redundant projection is first applied along the image columns. This is done by filtering the columns with h_y^+ and then downsampling them. The next step is to decouple positive horizontal frequencies from negative horizontal frequencies using non-redundant projections that maintain the non-redundant nature of the transform. Row filtering by h_x^+ and h_x^- followed by downsampling along rows decouples positive horizontal frequencies from negative horizontal frequencies in a non-redundant manner.

It is possible to generalize Proposition 5 of Chapter 4 to show that for images with a lowpass frequency-domain characteristic, the above 2D extension of non-redundant projection retains positive lowpass vertical frequencies and decouples positive lowpass horizontal frequencies from negative lowpass horizontal frequencies. Therefore the application of the 2D DWTs to the decoupled images will provide a non-redundant directional wavelet decomposition of the image provided that the above condition holds. To demonstrate the directionality of the non-redundant 2D CWT, we decomposed the disc image of Figure 5.5(a) using the single-level non-redundant 2D CWT. Figure 5.10 shows the directional subbands thus obtained.

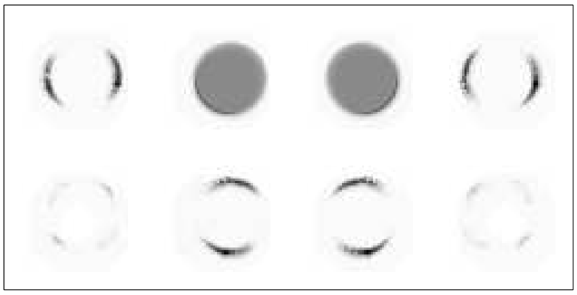


Figure 5.10 : The non-redundant 2D CWT.

To invert the non-redundant 2D CWT, we propose the structure shown in Fig. 5.9. The blocks labeled "2D IDWT" invert the 2D DWTs that were applied to the decoupled images. Next we must invert the non-redundant projections applied along the image rows. Recall that, the filters h_x^+ and h_x^- constitute the analysis bank of a 2-band decimated filter bank. Therefore the synthesis-bank filters of this filter bank will invert the analysis bank. Let g_x^+ and g_x^- denote these synthesis bank filters acting on the rows of the outputs from the 2D IDWTs after upsampling. Denote the sum of the outputs of these filters by $\hat{s}$. Note that $\hat{s} = s$ where s is defined in Fig. 5.8. To invert the non-redundant projection applied to the columns of the input image in Fig. 5.8, we separate the columns of $\hat{s}$ into real and imaginary parts and filter these images separately using the block labeled "synthesis" in Fig. 4.8. This block will invert the non-redundant projection applied to the columns of the input image, thereby recovering the original image.

5.6 The 2D Non-Redundant Post-Projection CWT

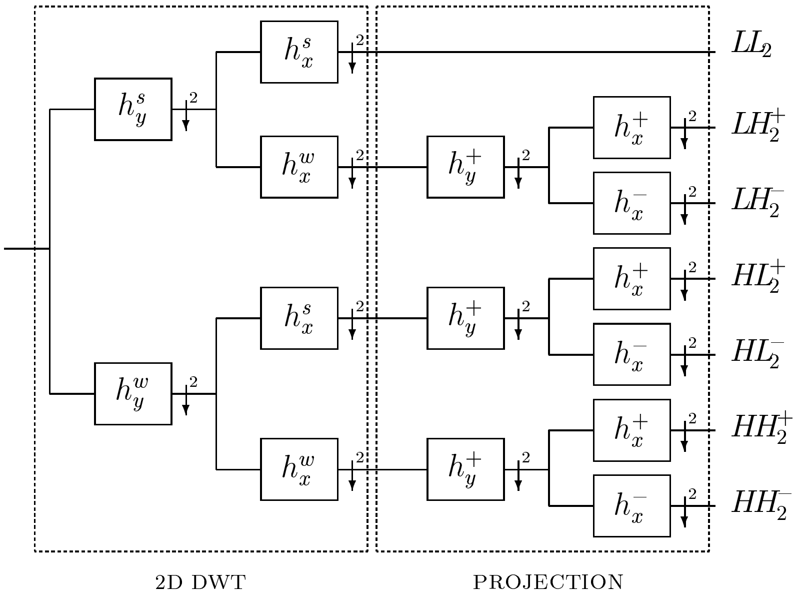


Figure 5.11 : Non-redundant post-projection analysis filter-bank structure for separable, directional, complex 2D DWT . The y (x) subscripts indicate filtering along the columns (rows) of an image. The $+$ ($-$) superscripts indicate projection filters that attenuate negative (positive) frequencies.

In the previous section, we showed that the separable extension of the non-redundant projection can be used to eliminate redundant negative vertical frequencies and to decouple positive horizontal frequencies from negative horizontal frequencies. This decoupling produces two images that are decomposed separately by 2D DWTs to obtain directional subbands that constitute the non-redundant 2D CWT. In this section, we show that it is possible to perform the non-redundant 2D projection after the 2D DWT to obtain directional subbands. We shall refer to this transform as the non-redundant 2D post-projection CWT since the projection is performed after the 2D DWT. Post-projection is important since it is the basis for the complex double density wavelet transform that will be introduced in the following chapter.

Consider Fig. 5.11 that shows the filter bank implementation of the non-redundant 2D post-projection CWT. The first two columns of filters contain scaling and wavelet filters denoted by "s" and "w" superscripts respectively. The "x" and "y" subscripts indicate whether a filter operates along image rows or image columns. These scaling and wavelet filters compute the first level of the DWT and output the subbands with the subscripts "2" in Figure 5.4. The filters in the third and fourth columns perform the 2D non-redundant projection on each subband of the 2D DWT. Since the 2D DWT subbands are real-valued, they contain redundant negative vertical frequencies. The 2D non-redundant projection eliminates the redundant vertical frequencies and decouples positive horizontal frequencies from negative horizontal frequencies, thus creating directional subbands. The following example will explain this concept in greater detail.

Consider the subband corresponding to the two blocks labeled LH_2 in Figure 5.1. If we filter this subband in the y direction (along columns) with a filter h_y^+ as shown in Figure 5.11, we obtain a new subband containing the two blocks labeled LH_2^- and LH_2^+ in Figure 5.4. We obtain a decoupled subband LH_2^+ by filtering this new subband in the x direction (along rows) with h_x^+. The other decoupled subband LH_2^- is similarly obtained by row filtering with h_x^-. The decoupled subband LH_2^- indicates feature gradients at -75 degrees, while the other decoupled subband LH_2^+ indicates feature gradients at +75 degrees.

By applying the projection filters to the other subbands in Figure 5.1, we can now distinguish between features oriented in 6 directions at any particular scale. At scale 2, these 6 directions correspond to the 6 directional subbands labeled HL_2^+, HH_2^+, LH_2^+, LH_2^-,HH_2^- and HL_2^- in Figure 5.4. These directional subbands correspond to feature gradients oriented at +15, +45, +75, -75, -45 and -15 degrees respectively. To get the directional subbands at scale 1 in Figure 5.4, we iterate the processing in Figure 5.11 on the LL_2 subband. We can then use the projection filters to decouple the lowpass subband LL_1 into LL_1^- and LL_1^+ blocks that discriminate between low-scale feature gradients oriented at -45 and +45 degrees.

Finally, let us consider the inverse non-redundant 2D post-projection CWT. Observe that the non-redundant 2D post-projection CWT applies the projection block of Fig. 5.8 to each subband of the 2D DWT. Recall that this projection block is inverted by the inverse-projection block of Fig. 5.9. Therefore to invert the non-redundant 2D post-projection CWT, we invert each projection block with an inverse-projection block. This step recovers the 2D DWT subbands. We then use the 2D IDWT to reconstruct the original image from the 2D DWT subbands.

Chapter 6

The Complex Double-Density Wavelet Transform

To overcome the shift sensitivity of the RWT, Selesnick et al. [15] recently proposed the Double-Density Wavelet Transform (DDWT), a multiscale, almost shift-invariant transform with a low transform-domain redundancy of 2.67 in 2D. The DDWT produces real-valued coefficients and hence does not provide any phase information. Moreover, the 2D DDWT has poor directionality: like the 2D RWT it can distinguish between only three different spatial feature orientations. In this chapter, we use the non-redundant post projection of Chapter 5 to create a Complex, directional Double-Density Wavelet Transform (CDDWT) from the DDWT coefficients. The CDDWT not only retains the low-redundancy (2.67) and shift-invariance properties of the DDWT, but also provides valuable phase information and superior directionality.

6.1 The Double-Density Wavelet Transform

Selesnick et al. [15] introduce the DDWT as the tight-frame equivalent of Daubechies' orthonormal wavelet transform; the wavelet filters are of minimal length and satisfy certain important polynomial properties in an oversampled framework. Because the DDWT, at each scale, has twice as many wavelets as the DWT, it achieves lower shift sensitivity than the DWT. The DDWT is implemented on discrete-time signals using the oversampled analysis and synthesis filter bank shown in Figure 6.1. The frequency responses for the analysis-bank filters are as shown in Figure 6.2, for a set of filters h^0, h^1, h^2 with lengths 7, 7, 5 respectively. The h^0 filter is a lowpass filter, while the h^1 and h^2 filters are highpass filters with similar frequency magnitude

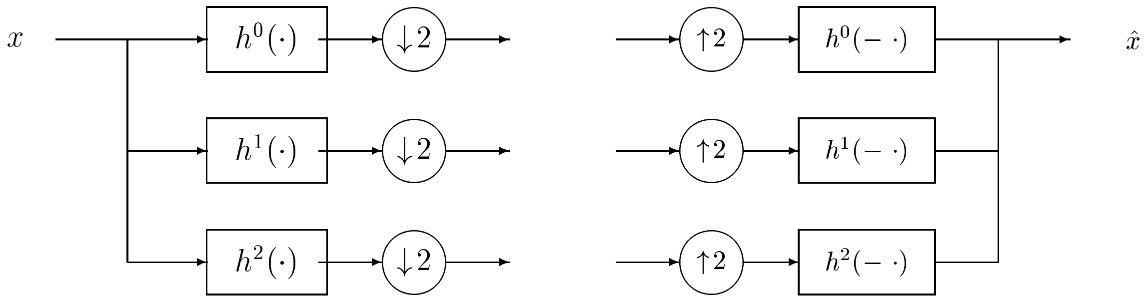


Figure 6.1 : The DDWT analysis and synthesis filter bank.

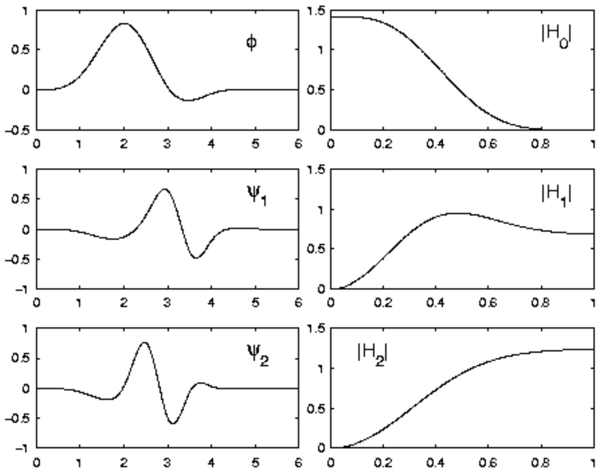


Figure 6.2 : Left column: basis functions. Right column: frequency responses.

responses. An iterated oversampled filter bank is created by the usual iteration on the lowpass branches of the analysis and synthesis banks. Selesnick et al. show that the above DDWT implementation corresponds to a multiresolution analysis with a single scaling function ϕ and two wavelets ψ^1, ψ^2 as shown in Figure 6.2. Observe that the wavelets approximately satisfy the "half-delay" property $\psi^1(t) = \psi^2(t - \frac{1}{2})$.

The DDWT is extended to 2D by iterating the 1D oversampled filter bank on the rows and then on the columns of an image, as is usually done for separable 2D

DWTs. At a given level in the iterated filter bank, this separable extension produces nine 2D subbands. To indicate the filters used along the row and column dimensions to create the nine subbands, we label the subbands $h_y^i h_x^j, \quad i,j \in \{0,1,2\}$. The subscript x indicates filtering along the row dimension, while subscript y indicates filtering along the column dimension. The superscripts $0,1,2$ indicate the particular filter h^0, h^1, h^2 used to filter along a specified dimension to create the subband. Next, we introduce an interleaving process that operates on four disjoint subsets of the nine DDWT subbands producing four new DDWT subbands $HH_{DD}, HL_{DD}, LH_{DD}, LL_{DD}$ with frequency domain supports that are comparable to those of the HH, HL, LH, LL subbands of the 2D DWT. Subsequent processing of the interleaved DDWT subbands is then quite similar to the processing of DWT subbands.

Since h^0 is a lowpass filter while both h^1 and h^2 are highpass filters, the $h_y^2 h_x^2, h_y^2 h_x^1, h_y^1 h_x^2, h_y^1 h_x^1$ subbands each have a frequency-domain support comparable to that of the HH subband in a DWT. Although these subbands all have the same frequency domain support, the half-delay property implies that the spatial-domain supports of the wavelets corresponding to the coefficients in the columns of the $h_y^2 h_x^1$ subband lie a half-integer to the right of those in the columns of the $h_y^2 h_x^2$ subband. Therefore we may interleave the wavelet-coefficient columns in the $h_y^2 h_x^1$ subband inbetween and to the right of the wavelet-coefficient columns in the $h_y^2 h_x^2$ subband. We may also interleave the wavelet-coefficient rows in the $h_y^1 h_x^2$ subband inbetween and below the wavelet-coefficient rows in the $h_y^2 h_x^2$ subband. We complete the interleaving process by interleaving the wavelet-coefficient rows in the $h_y^1 h_x^1$ subband inbetween and below the wavelet-coefficient rows in the $h_y^2 h_x^1$ subband. We represent this interleaving process concisely in Matlab-compatible notation by

$$HH_{DD}(m,n) = \begin{cases} h_y^2 h_x^2\left(\frac{m}{2}, \frac{n}{2}\right), & m \text{ even}, \quad n \text{ even}, \\ h_y^2 h_x^1\left(\frac{m}{2}, \frac{n-1}{2}\right), & m \text{ even}, \quad n \text{ odd}, \\ h_y^1 h_x^2\left(\frac{m-1}{2}, \frac{n}{2}\right), & m \text{ odd}, \quad n \text{ even}, \\ h_y^1 h_x^1\left(\frac{m-1}{2}, \frac{n-1}{2}\right), & m \text{ odd}, \quad n \text{ odd}. \end{cases}$$

HH_{DD} is a new subband with four times as many coefficients but the same frequency-domain support as the HH subband at the same level in a DWT. A similar interleaving scheme creates the HL_{DD} (LH_{DD}) subband with the same frequency-domain support as the corresponding HL (LH) subband of the DWT, but with twice as many coefficients. We describe this interleaving scheme by

$$HL_{DD}(m,n) = \begin{cases} h_y^2 h_x^0 \left(\frac{m}{2}, n\right), & m \text{ even}, \\ h_y^1 h_x^0 \left(\frac{m-1}{2}, n\right), & m \text{ odd}, \end{cases}$$

$$LH_{DD}(m,n) = \begin{cases} h_y^0 h_x^2 \left(m, \frac{n}{2}\right), & n \text{ even}, \\ h_y^0 h_x^1 \left(m, \frac{n-1}{2}\right), & n \text{ odd}. \end{cases}$$

Finally, note that there is only one subband $h_y^0 h_x^0$ with the same frequency domain support as the LL subband in a DWT. To maintain consistency, we label this $h_y^0 h_x^0$ subband as LL_{DD}.

6.2 Enhancing DDWT Directional Selectivity

In this section, we demonstrate how separable 2D complex-coefficient filters can project the DDWT coefficients to obtain the CDDWT Fourier-plane partitioning

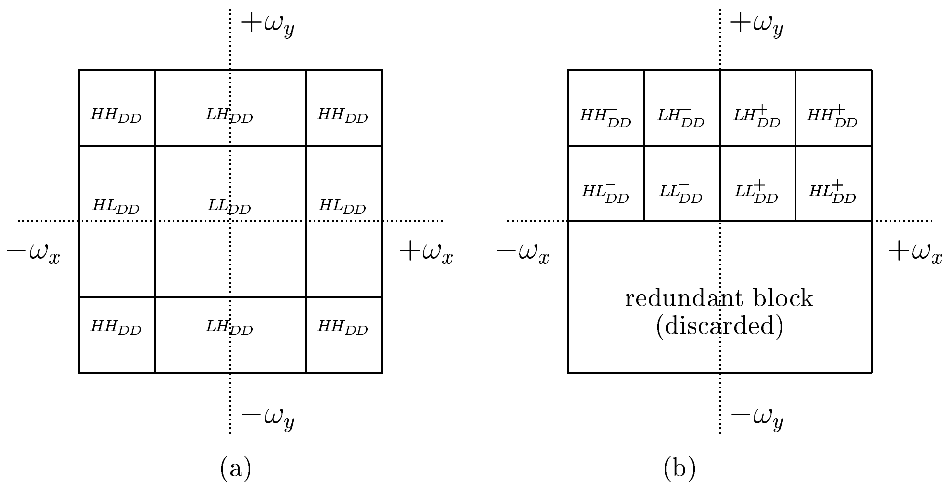


Figure 6.3 : **(a)** DDWT Fourier-plane partitioning, **(b)** CDDWT Fourier-plane partitioning.

with enhanced directional selectivity. As explained in Section 6.1, the Fourier-plane partitioning associated with the $HH_{DD}, HL_{DD}, LH_{DD}, LL_{DD}$ subbands of the DDWT is comparable to that of the separable real-valued DWT. Consequently, the Fourier-plane partitioning of the DDWT may be represented as in Figure 6.3(a). The HH_{DD} subband is associated with diagonally-oriented spatial features and concentrates its energy in the four blocks labeled HH_{DD}. The HH_{DD} blocks in the upper-left and lower-right corners indicate features with gradients at -45 degrees, while those in the upper-right and lower-left corners indicate features with gradients at $+45$ degrees. Since all four blocks are associated with one interleaved subband in the DDWT, we cannot differentiate between these two orientations. Similarly, features with gradients at $+15(+75)$ degrees and $-15(-75)$ degrees are indistinguishable from their wavelet coefficients since they are both associated with energy in the $HL_{DD}(LH_{DD})$ subband. Therefore, decoupling the positively- and negatively-oriented blocks associated with each subband of the DDWT improves directional selectivity. The CDDWT analysis filter bank in Figure 6.4 shows how this may be done using our non-redundant post-projection technique of Chapter 5 to increase the directionality of the DDWT while preserving its low transform-domain redundancy of 2.67. The complex-coefficient filters in the block labeled "projection" have $+(-)$ superscripts to indicate that they retain positive (negative) frequencies and suppress negative (positive) frequencies. These $+(-)$ filters are 1D projection filters that project real signals onto subspaces consisting of predominantly positive (negative) frequencies in $[-\pi, \pi]$. As explained below, these projection filters enhance the directional selectivity of the DDWT transform preceding them by decoupling the DDWT subband blocks in Figure 6.3(a) to obtain the CDDWT subbands in Figure 6.3(b).

Consider the subband corresponding to the two blocks labeled LH_{DD} in Figure 6.3(a). If we filter this subband in the y direction (along columns) with a filter h_y^+ as shown in Figure 6.4, we obtain a new subband containing the two blocks labeled LH_{DD}^- and LH_{DD}^+ in Figure 6.3(b). We then obtain a decoupled subband LH_{DD}^+ by filtering

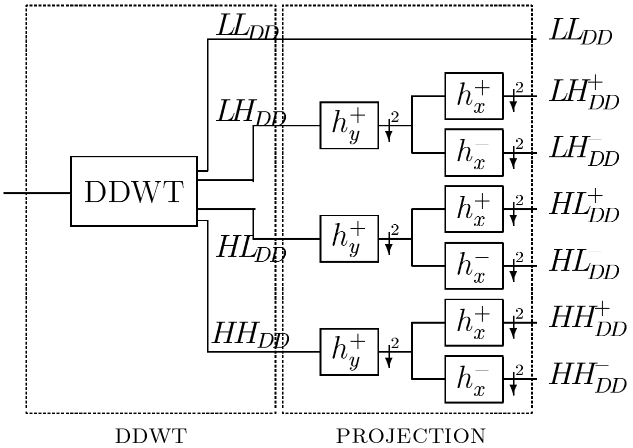
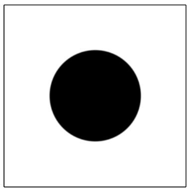
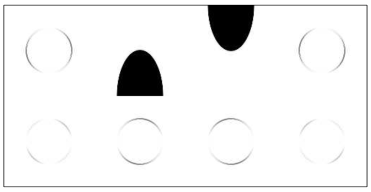


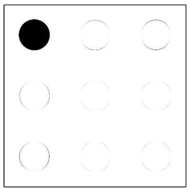
Figure 6.4 : CDDWT analysis filter-bank. The y (x) subscripts indicate filtering along the columns (rows) of an image. The $+$ ($-$) superscripts indicate projection filters that attenuate negative (positive) frequencies.



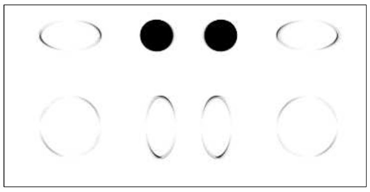
(a) Disc



(b) DTWT



(c) DDWT



(d) CDDWT

Figure 6.5 : **(a)** Disc Image **(b)** CDDWT Clockwise from upper-rightmost block: -75, +45, -15, +15, -45, +75 degree highpass subband, upper half of lowpass subband, lower half of lowpass subband. **(c)** DDWT: From top to bottom, left to right: $h_y^0 h_x^0, h_y^1 h_x^0, h_y^2 h_x^0, h_y^0 h_x^1, h_y^1 h_x^1, h_y^1 h_x^2, h_y^0 h_x^2, h_y^2 h_x^1, h_y^2 h_x^2$ **(d)** CDDWT: Clockwise from upper-rightmost block: -75, +45, -15, +15, -45, +75 degree highpass subbands, -45, +45 degree lowpass subbands.

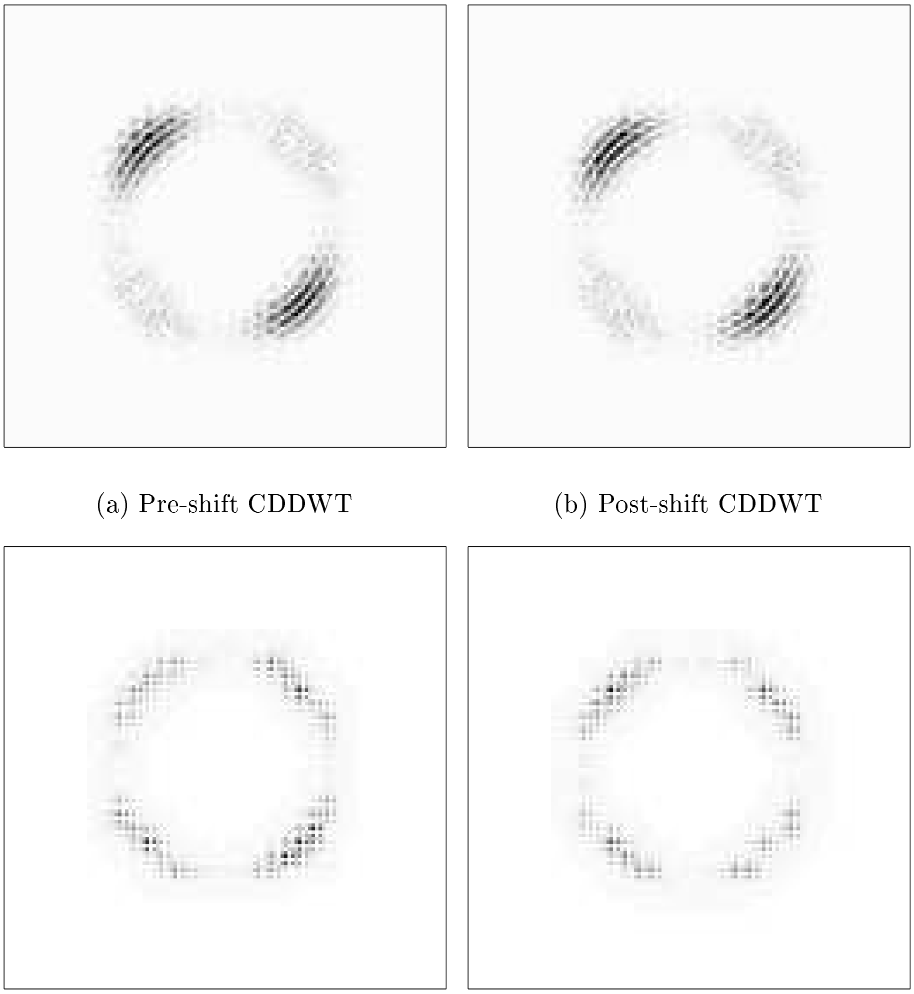
this new subband in the x direction (along rows) with h_x^+. The other decoupled subband LH_{DD}^- is similarly obtained by row filtering with h_x^-. The LH_{DD}^- subband reveals features with gradients at -75 degrees, the LH^{1+}_{DD} subband reveals features with gradients at +75 degrees. Likewise, the subbands labeled $HL^{+}_{DD}, HL^{-}_{DD}, HH^{+}_{DD}, HH^{-}_{DD}$ indicate features with gradients at +15, -15, +45, -45 degrees respectively, giving a total of 6 directional subbands at each scale.

Obtaining perfect reconstruction for the CDDWT requires the design of a synthesis filter bank that inverts the analysis filter bank in Figure 6.4. In Section 5.6 we discussed the inversion of the projection block to recover the 2D DWT subbands. Performing that processing on the CDDWT subbands will reconstruct the DDWT subbands, in this case. The DDWT is then easily inverted by a synthesis filter bank that performs the inverse DDWT [15].

6.3 CDDWT Properties: Shift-Invariance and Improved Directionality

To compare the directional selectivity of various different transforms, we used the disc image displayed in Figure 6.5(a). Figure 6.5(c) shows the nine subbands $h^i_y h^j_x,\ \ i, j \in \{0, 1, 2\}$ from the single-level DDWT of the disc image. Observe that directional information is not readily available in this transform. Figure 6.5 (b) shows the single-level DTWT of the disc image. Due to its large size, the lowpass subband has been displayed in halves. Figure 6.5 (d) presents the single-level CDDWT obtained using the filter-bank structure of Figure 6.4. Note that the interleaving process creates rectangular $LH^{+}_{DD}, LH^{-}_{DD}, HL^{+}_{DD}, HL^{-}_{DD}$ subbands in the CDDWT; although the disc appears elliptical in these subbands, the CDDWT directionality is unaffected. While the DTWT and the CDDWT each provide six highpass directional subbands at each scale, the DTWT transform-domain redundancy is 4.0, whereas that of the CDDWT is only 2.67.

Next, we used the following experiment to compare the shift-invariance properties of the CDDWT with that of other multiscale tranforms. We generated 2-level DWT, DDWT, DTWT and CDDWT transforms of the disc image (Figure 6.5(a)) and then



(a) Pre-shift CDDWT (b) Post-shift CDDWT

(c) Pre-shift DWT (d) Post-shift DWT

Figure 6.6 : Single-subband disc reconstructions comparing shift-sensitivity.

we zeroed out all subbands besides the lowest scale $HH, HH_{DD}, HH^+, HH^+_{DD}$ subbands respectively. Figures 6.6(a) and 6.6(c) show the reconstruction of the disc image from the lowest scale HH^+_{DD} and HH subbands of the 2-level CDDWT and DWT respectively. The input disc image was then shifted one pixel rightwards and downwards and the same reconstruction was performed using each of the above transforms. Figures 6.6(b) and 6.6(d) show the reconstruction of the shifted-disc image from the level-2 CDDWT HH^+_{DD} subband and the level-2 DWT HH subband respectively. Observe that the pre-shift and post-shift DWT reconstructions appear significantly different, whereas there are very minor differences in the CDDWT reconstructions. This indicates that the HH_{DD}^{+} CDDWT subband is much less shift sensitive than the HH DWT subband. For each transform, we then spatially translated the post-shifted reconstruction so that its center coincided with that of the pre-shifted reconstruction. Next we computed the root mean square error between the translated post-shifted reconstruction and the pre-shifted reconstruction. This was repeated for shifts ranging from one to eight pixels. The bar graph in the upper-half of Figure 6.7 summarizes the results. The DWT HH subband is the most shift-sensitive, while the DDWT HH_{DD} and DTWT HH^{+} subbands are the least shift-sensitive. The CDDWT HH_{DD}^{+} subband is much less shift-sensitive than the DWT HH subband and also enjoys both the low-redundancy (2.67) of the DDWT as well as the enhanced directionality of the DTWT. The bar graph in the lower half of Figure 6.7 shows the results of the experiment when performed with HL subbands instead of HH subbands; comparable results are obtained for these subbands as well.

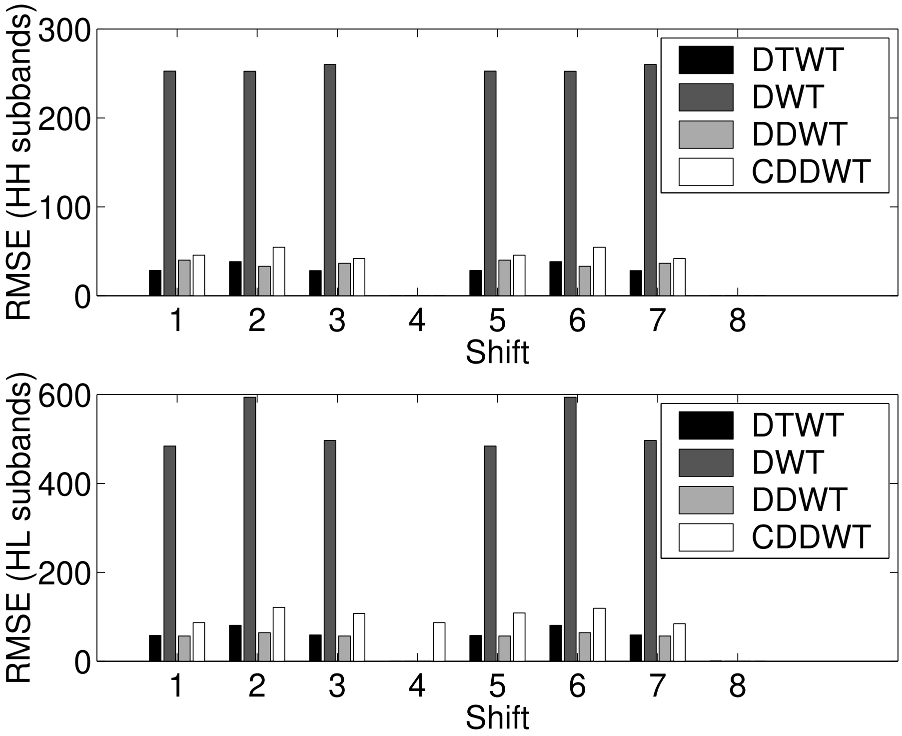


Figure 6.7 : Bar graphs comparing RMSE\shift for various transforms.

Chapter 7

Seismic Signal Processing With Complex Wavelet Transforms

Seismic imagery of the earth's subsurface plays a critical role in all aspects of oil and gas exploration and production — from the location of reserves to their appraisal and subsequent monitoring. In oil and gas exploration, seismic cross-sections are scrutinized by interpreters who search for features that indicate possible hydrocarbon reservoirs. Previously, interpreters visually examined 2D seismic cross-sections to detect faults and extract other stratigraphic information about rock properties and sedimentary environments [44]. They must now use computers to analyze 3D volumes comprising gigabytes of high quality, densely sampled data. Computer programs process the seismic data volumes and output signal attributes that aid interpretation of the data, elucidating its salient characteristics.

Since seismic data is highly non-stationary, traditional Fourier-transform based spectral analysis techniques yield inaccurate attributes. Steeghs [54] proposed the use of time-frequency representations for the derivation of local signal attributes that accurately describe underlying geological features. Subsequently, van Spaendonck and Baraniuk [44] suggested attribute derivation using the complex steerable pyramid, a transform that provides a scale-orientation representation. They argued that seismic attributes must describe the amplitude, frequency and geometry of the geological features. Since the earth's amplitude and frequency response appear through different scales of resolution and the geometry manifests itself as spatial orientations, the scale-orientation representation is a convenient formulation for the computation of local signal attributes. Unfortunately, as explained in Chapter 1, the complex steerable pyramid is overwhelmingly redundant and its spatial-domain implementation does not have perfect reconstruction. To overcome these drawbacks, we used the CWT, the non-redundant CWT and the CDDWT as parsimonious scale-orientation representations for attribute extraction [55, 56, 57, 58, 59, 60]. Our projection-based transforms provide a scale-orientation representation since they are multiscale and directional; however, unlike the complex steerable pyramid they have low redundancy factors and efficient spatial-domain implementations with perfect reconstruction. In the following sections, we shall explain the use of our projection-based transforms for the computation of seismic attributes from 2D and 3D data.

7.1 Computing Seismic Attributes from 2D Images

A particularly useful attribute is the local angle (dip) of the reflecting surface. Angle representations enable 3D interpretation of *structures* using seismic depth-slices. Channels and faults appear as angle variations, but are often barely visible in the 2D seismic images.

We now outline a two-stage method to compute local angle representations for 2D seismic images using our projection-based transforms. In the first stage, we apply a projection-based transform to the seismic image. For comparison purposes, we used the CWT, the non-redundant CWT and the CDDWT to generate six directional subbands at each scale. In the second stage, we perform six separate reconstructions to get six reconstructed directional images. In each reconstruction, we preserve, over all scales, the directional subbands corresponding to one of the six directions; all the other directional subbands are zeroed out. Therefore, each reconstructed image reflects the extent to which the seismic cross-section is oriented in one of the six directions. At each pixel we finally estimate the local angle as the weighted average of the six corresponding pixels in the six reconstructed images.

Figure 7.1 compares the angle representations generated for a seismic cross-section. The geometry in the section (Figure 7.1(a)) shows a structure with substantial angle variations and several smaller scale events such as faults. In Figures 7.1(b)-(e), white shading represents an uphill-dipping reflection with an angle of +90 degrees, black shading indicates a downward-dipping reflection with an angle of +90 degrees; gray shades represent intermediate angles. The CSP angle-representation in Figure 7.1(b) provides an accurate and smooth representation of the local angles in the seismic section. The non-redundant 2D post-projection CWT generates a noisier and less smooth angle representation with a finer resolution [Figure 7.1(c)]. This transform however requires *14 times* less redundancy than the CSP which has a redundancy factor of 14. The angle-detection result from the CWT in Figure 7.1(d) appears equally informative as the result from the CSP [Figure 7.1(b)], but it has a higher spatial resolution and is obtained with a transform redundancy of four. Finally in Figure 7.1(e) we present the angle-representation from the CDDWT. This result is comparable to the CSP angle-representation in Figure 7.1(e). However, the CDDWT transform redundancy of 2.67 is much lower than that of the CSP which has a redundancy factor of 14. This lower redundancy factor results in modest storage and computational requirements.

7.2 Computing Seismic Attributes from 3D Volumes

In the preceding section, we demonstrated the effectiveness of our projection-based transforms for 2D seismic-attribute extraction from 2D slices of 3D seismic volumes. We explained that these attributes are indispensable information sources for seismic interpreters. Nevertheless, because seismic phenomenon occur in three dimensions, 3D seismic-attributes computed directly from the 3D volume characterize the underlying geological phenomena more accurately than 2D attributes derived from 2D slices of the 3D volume. Although it is possible to infer 3D attributes from 2D attributes, this process is inefficient and so the direct calculation of 3D attributes is preferable. Previously, 3D seismic-attribute extraction from large 3D volumes was computationally unattractive due to the overwhelming redundancy of 3D complex, multiscale

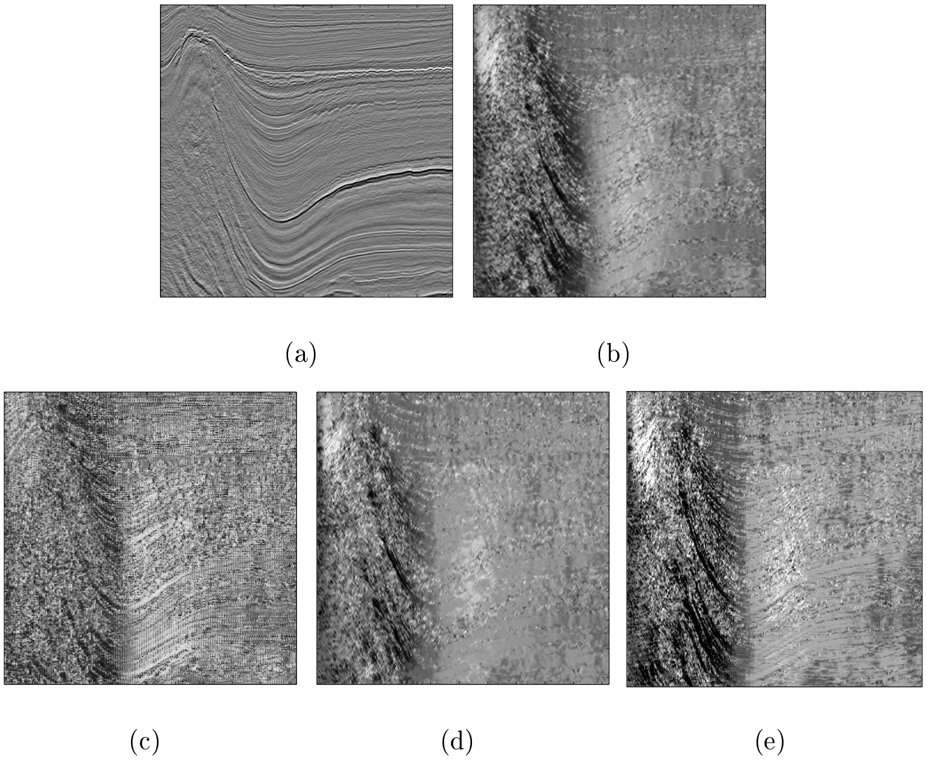


Figure 7.1 : (a) Seismic section showing substantial angle variations, (b) Local angle using complex steerable pyramid, (c) Local angle using the non-redundant post-projection CWT, (d) Local angle using the CWT, (e) Local angle using the CDDWT . In (b),(c), (d) and (e) intensity represents angle, black (-90 degrees) through white (+90 degrees).

transforms such as the complex steerable pyramid and the dual-tree wavelet transform. However, projection-based transforms such as the CWT, non-redundant CWT and CDDWT are either non-redundant or only slightly oversampled. In our earlier paper [59], we demonstrate that the CWT may be used efficiently for 3D seismic-attribute extraction from large 3D seismic volumes. In the following sections, we discuss the extension of the CWT to three dimensions and show how this transform can be used to extract frequency-related and geometry-related attributes. We use a seismic volume from the F09 block in the North Sea to illustrate attribute extraction.

7.2.1 The 3D CWT

In Section 5.2, we explained that the 2D extension of the CWT is obtained by eliminating redundant information in the negative vertical-frequency half-plane , decoupling the positive horizontal frequencies from the negative horizontal frequencies and then performing 2D DWTs on the decoupled images as illustrated in Fig. 5.2 of Chapter 5. We define the 3D extension of the CWT analogously, as shown in Fig. 7.2. The h_z^+ Softy-space projection filter operates along the z-dimension to eliminate the negative z-frequency half-cube. The h_y^+ and h_y^- projection filters decouple positive vertical frequencies from negative vertical frequencies and the h_x^+ and h_x^- projection filters decouple positive horizontal frequencies from negative horizontal frequencies. We have now decoupled the positive z-frequencies into four cubes. We obtain the 3D CWT by computing the 3D DWT of each of these cubes. The right half of Figure 7.2 depicts the two-level 3D CWT tiling. The redundancy of the 3D CWT is much lower than that of the complex steerable pyramid as explained below.

To derive the redundancy factor for the d-dimensional extension of the CWT, let us consider a d-dimensional, real-valued hypercube of length N in each dimension. This hypercube occupies N^d memory units, each memory unit being large enough to store a real number. We now perform a d-dimensional projection of the hypercube to eliminate redundant negative frequencies and decouple the remaining positive frequencies from negative frequencies. This projection yields 2^{d-1} complex-valued, d-dimensional hypercubes of length N in each dimension. Each hypercube occupies $2 \times N^d$ memory units, since a complex number occupies two memory units. Because we use critically sampled DWTs, the d-dimensional DWT of each hypercube will also occupy $2 \times N^d$ memory units. Since we compute 2^{d-1} DWTs, the redundancy factor is

$$\frac{\text{transform storage requirement}}{\text{input storage requirement}} = \frac{2^{d-1} \times 2 \times N^d}{N^d} = 2^d.$$

Hence in three dimensions, the CWT has a redundancy factor of eight. In contrast, van Spaendonck [60] shows that the 3D complex steerable pyramid with six directional subbands is 39 times redundant! This prohibits its application to 3D seismic-attribute extraction. On the other hand, the d-dimensional Dual-Tree Wavelet Transform (DTWT) has the same redundancy as the CWT: 2^d. Therefore, the CWT and the DTWT are practical alternatives to the complex steerable pyramid for 3D seismic-attribute extraction. Nevertheless, the continuing depletion of existing oil reserves will accelerate the search for new reservoirs. Larger volumes of 3D seismic data will need to be processed and then, even the factor-eight redundancy of the CWT and the DTWT will be unacceptable. The DTWT cannot reduce its redundancy and retain directionality. On the other hand, the d-dimensional CDDWT has a redundancy factor of

$$\frac{3^d - 1}{2^d - 1},$$

or 3.67 in three dimensions. If we are willing to sacrifice shiftability, we can use the d-dimensional extension of the non-redundant CWT. This transform has the lowest possible redundancy factor of 1, $\forall d \in \mathbb{N}$. Our projection-based transforms, the CWT, non-redundant CWT and CDDWT, clearly provide the most flexible tools for 3D seismic-attribute extraction: they provide directionality and phase information along with a trade-off between shiftability and redundancy. In the next section we demonstrate the applicability of the CWT to 3D seismic-attribute extraction.

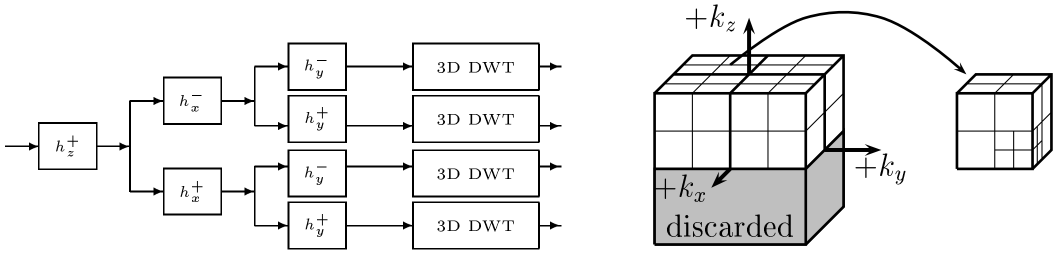


Figure 7.2 : **Left:** The 3D CWT analysis filter bank. **Right:** Fourier-plane partitioning of the 3D CWT.

7.2.2 3D Seismic-Attribute Extraction with the CWT

We now outline a method for the extraction of local 3D-scale and local 3D-angle (dip) attributes from seismic volumes. First we apply the 3D-CWT transform to the seismic volume. In this particular experiment, we used only three directional subbands: horizontal, vertical and diagonal. For angle extraction, we perform three separate reconstructions from horizontal, vertical and diagonal orientations in isolation. The diagonal features are located in all subbands along the corners of the Fourier volume as indicated by the gray colored subbands in Fig. 7.3. The vertical features are located in Fourier domain in the subbands adjacent to the $k_x - k_y$ axes. The horizontal features are localized in the upper bands adjacent to the k_z-axis. Each reconstructed volume reflects the extent to which the seismic cross-section is oriented in the corresponding direction. We estimate the local angle as the weighted average of the three reconstructions..

For the scale extraction, we compute the local scale as the weighted average from separate reconstructions, each reconstruction created from volume information at a particular scale. The subbands contributing to a particular scale are radially oriented around the origin. For example, the gray region in Fig. 7.3 shows the subbands that are used for the reconstruction from the finest-scale volume information.

To test the 3D attribute-extraction process, we used the 3D F09 block from the Southern North Sea. This seismic volume was surveyed and time migrated in 1994

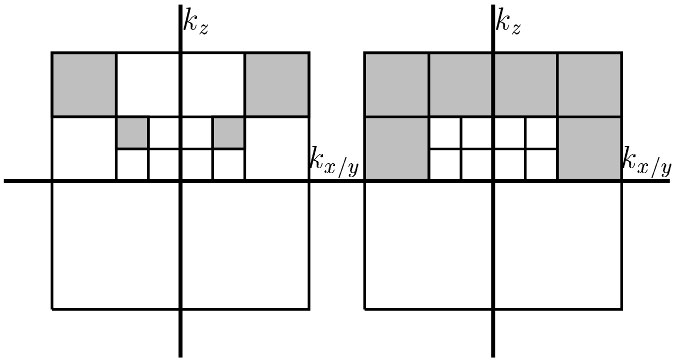
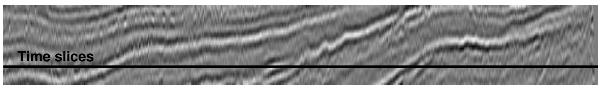


Figure 7.3 : Attribute extraction. **Left:** Subbands used for extraction of diagonal features. **Right:** Subbands used for extraction of finest scale.

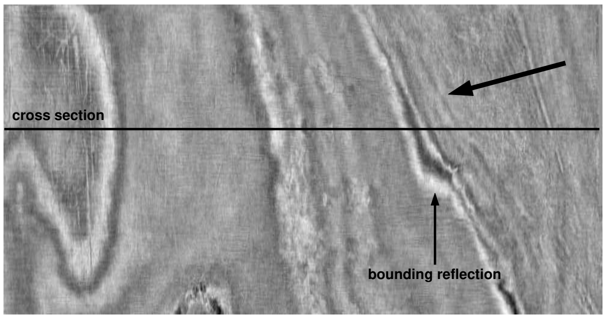
by TransCanada. The survey covers 28.2 by 16 km. and consists of 640 in-lines and 1128 cross-lines with 25-meter spacing. The focus of this interpretation is an enormous delta system that expanded into the North Sea basin during the Late Cenozoic era [61].

For attribute extraction, we used the 552-680 ms, 6.2 km$\times$24.8 km sub-volume of the F09 dataset. Figures 7.4 (a) and (b) show a section and an amplitude slice from the sub-volume. The striations in these figures illustrate the layering of seismic reflectors in the sub-volume. An expert visual interpreter would observe that the layering-sequences in these figures are indicative of an expanding delta system. He/she would label the reflector that bounds the delta as shown in Fig. 7.4 (b). Oil reservoirs are frequently located around such bounding reflectors. Hence it is essential to develop interpretation programs that rapidly examine seismic volumes and detect bounding reflectors.

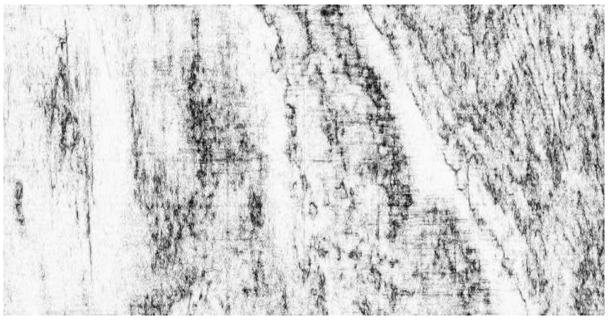
As a basis for machine interpretation of the delta system in the sub-volume, we consider the 3D local-angle and scale attributes in Figures 7.4 (c) and (d) respectively. The local angle strikingly portrays the reflector that bounds the delta. The scale attribute shows alternating zones of fine and coarse scale, indicating alternating frequency content of reflections. Furthermore there is a distinct difference between the scale pattern in the (steeply dipping) delta lobe and the subsequent layer. We conclude that the local-angle and scale attributes provide valuable information that can be exploited for the detection of oil reservoirs.



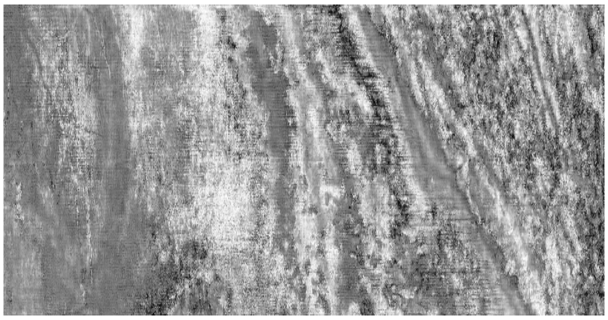
(a) section



(b) amplitude



(c) dip



(d) scale

Figure 7.4 : (a) Seismic section. (b) Seismic amplitude timeslice showing direction of deltaic progradation. (c) Angle attribute (black=vert., white=horz.). (d) Scale attribute (black=fine, white=coarse).

Chapter 8

Conclusions and Future Work

The DWT provides a rapid, local, sparse, multiscale decomposition for a large class of signals. These important properties contribute to make the DWT a powerful tool for signal- and image-processing. Nevertheless, this popular transform is seriously disadvantaged in three ways. First, DWT coefficients are shift sensitive because input-signal shifts result in unpredictable changes in DWT coefficients. Shift sensitivity compromises transform robustness to input-perturbations, and prevents discrimination between input-signal shifts. Second, the separable, multi-dimensional DWT suffers from poor directionality because DWT coefficients reveal only three spatial orientations. Since natural images are characterized by smooth regions punctuated with edges at several different spatial orientations, the DWT fails to represent natural images optimally. High transform directionality is essential for the efficient representation of natural images. Third, DWT analysis with real-valued wavelets provides no phase information. Transform-coefficient phase is important because it enables the amplitude-phase signal representation that reveals local energy, phase and orientation. These quantities accurately describe non-stationary signal behavior.

In Chapter 1 we investigated several different techniques for the mitigation of DWT disadvantages. Most methods address only one of the three drawbacks described above. Notable exceptions are the Complex Steerable Pyramid (CSP) and the Dual Tree Wavelet Transform (DTWT); both transforms overcome all three shortcomings: shift sensitivity, poor directionality and absence of phase information. The CSP uses a non-separable, highly redundant implementation to mitigate DWT disadvantages. However, the overwhelming 3D CSP redundancy of 39 prohibits its usage in data-intensive applications such as seismic signal processing. Moreover, the spatial-domain CSP implementation lacks perfect reconstruction, an important requirement for most signal processing applications. On the other hand, the separable m-dimensional DTWT has a lower redundancy of 2^m, is perfectly reconstructing and also overcomes all DWT shortcomings. This transform has been used in various applications with excellent results. To obtain perfect reconstruction, the DTWT uses real-coefficient filters in two DWT trees. Conceptually however, the DTWT coefficients may be obtained using a DWT with quasi-analytic filters. This quasi-analytic filter decomposition enables the mitigation of all three DWT disadvantages.

The central concept in this thesis is that the analytic-filter decomposition of a real signal is equivalent to the real-filter decomposition of its Hardy-space projection. Therefore, DTWT-like properties are attainable using *any* DWT preceded by a projection onto Hardy space. Our projection-based approach has two advantages over the DTWT: first, the flexibility to use an arbitrary DWT decomposition; and second, the ability to control transform redundancy.

The flexibility to use an arbitrary DWT in our projection-based approach arises from the decoupling in its implementation. Whereas the DTWT implementation requires the joint design of a quadrature-pair of wavelets, our projection-based approach consists simply of a projection stage followed by a DWT stage. This decoupling into two stages allows for the flexibility to use any arbitrary DWT decomposition. We showed how this flexibility can be exploited, for example, to create the Complex Double-density DWT (CDDWT): a shift-insensitive, directional, complex wavelet transform with a low-redundancy of $\frac{3^m-1}{2^m-1}$ in m dimensions. To the best of our knowledge, no other transform achieves all these properties at a redundancy that is lower than $\frac{3^m-1}{2^m-1}$.

The ability to control the transform redundancy of our projection-based approach arises from the controllability of the redundancy in the projection stage. The decoupling in the implementation of our projection-based approach implies that transform redundancy is determined by the projection step. Consequently, by creating a non-redundant projection stage, we obtain a directional, non-redundant, complex wavelet transform: the non-redundant CWT. To the best of our knowledge, no other complex wavelet transform is simultaneously directional and non-redundant. This transform has significant potential benefits for applications such as image coding, in which parsimonious signal representations are critical.

We have shown that our decoupled projection stage mitigates DWT shortcomings and provides important advantages that engender new complex wavelet transforms with unprecedented properties. In theory, the projection stage must project a real input signal onto the Hardy space $H^2(\mathbb{R} \to \mathbb{C})$; in practice, such a projection is unrealizable. To circumvent this obstacle, we defined a new function space, the Softy space S^+. We accomplish the projection onto Softy space by using a digital projection filter h^+. We proved that the Softy space S^+ is isomorphic and approximately equal to the Hardy space $H^2(\mathbb{R} \to \mathbb{C})$. To create the non-redundant CWT, we defined the non-redundant projection that consists of a projection filter followed by a dowsampler. This projection stage introduces no redundancy into the transform. We proved that the non-redundant projection of a real signal approximates its Softy-space projection, provided that the signal has a lowpass characteristic.

In summary, the definition of the Softy space enables the implementation of our projection-based complex wavelet transforms. Like the complex steerable pyramid and the dual-tree wavelet transform, our projection-based transforms simultaneously mitigate all DWT shortcomings. However, in contrast to these other transforms, our projection-based approach has additional benefits that allow for the creation of new complex wavelet transforms with significant potential advantages for a variety of applications. As an example, we demonstrated that our projection-based complex wavelet transforms achieve state-of-the-art results in 2D and 3D seismic attribute-extraction.

8.1 Future Work

We conclude this chapter by suggesting promising directions for future research. First, we consider projection-filter design constraints that will improve the properties of the complex wavelet transforms. Next, we consider potential applications for our complex wavelet transforms. Finally, we shall indicate possible generalizations of important concepts that we discussed earlier.

8.1.1 Projection-Filter Design Constraints

In Chapters 2 and 4 we derived projection-filter design constraints that ensured unitary maps between $L^2(\mathbb{R} \to \mathbb{R})$ functions and their projections. Recall that this unitarity is important because it ensures that the projection preserves inner products and hence the structure of $L^2(\mathbb{R} \to \mathbb{R})$. We now consider other important projection-filter design constraints.

Constraints for Filter Symmetry and Linear-Phase Response

In image-processing applications, filter symmetry and linear-phase response are important filter-design constraints. Symmetric filters enable boundary conditions to be handled with minimal distortion by using symmetric extensions at image boundaries [46]. Symmetric synthesis-bank filters also minimize perceptual distortions in low bit-rate image-coding systems. Linear-phase response is an important filter property because it eliminates phase distortion in filtered signals. For real-coefficient filters, filter symmetry and linear-phase response are equivalent. However, Zhang et al. [37] proved that this equivalence does not hold for filters with complex coefficients. In our earlier paper [56], we derived conditions for the design of projection filters that allow symmetric extension to be used at image boundaries. Since these conditions do not necessarily imply linear-phase response or unitarity, an important avenue for further research is to consider the design of unitary, symmetric projection filters with linear-phase response.

Constraints for Shift Invariance

In Chapter 6 we demonstrated that the non-redundant post-projection may be used to create the Complex Double-density DWT (CDDWT) from the Double-density Wavelet Transform (DDWT). The CDDWT introduces directionality and phase information into the DDWT without increasing transform redundancy. We also showed experimentally that the CDDWT preserves the shift insensitivity of the DDWT, although the non-redundant projection was not explicitly designed for this purpose. Therefore, another important consideration in the projection-filter design is the minimization of a quantitative measure of the shift variance properties of the non-redundant projection. This minimization will create the optimally shift-invariant CDDWT.

8.1.2 Applications for Complex Wavelet Transforms

In Chapter 7 we proved that the m-dimensional CWT and the m-dimensional Dual-Tree Wavelet Transform (DTWT) have the same transform redundancy of 2^m. These transforms have similar properties since both are shiftable, directional and provide phase information. Consequently, we expect that the CWT may be used for the same applications in which the DTWT has yielded excellent results. We listed some of these applications in Chapter 1.

In Chapter 7, we demonstrated the effectiveness of the CWT for 3D seismic-attribute extraction. In that chapter, we also showed that the CDDWT can be used for the accurate extraction of 3D seismic attributes. Since the m-dimensional CDDWT has similar properties to the m-dimensional DTWT, but attains these properties at a lower redundancy of $\frac{3^m-1}{2^m-1}$, it is important to compare the performance of the CDDWT and the DTWT in signal-processing applications. As an indicator of CDDWT potential in signal-processing applications, note that Selesnick and Sendur [15] used the DDWT, from which the CDDWT is derived, to obtain excellent denoising results. These results were achieved without exploiting directionality and phase-information, properties that are imparted to the DDWT by the non-redundant projection that generates the CDDWT from the DDWT. It is therefore realistic to expect significant signal-processing benefits from the projection-based CDDWT.

Besides the CDDWT, the non-redundant CWT is another novel transform generated using the non-redundant projection. As explained in Chapter 4, the non-redundant CWT is the only non-redundant, directional, complex wavelet transform with explicit phase information in its transform coefficients. These properties make the non-redundant CWT particularly attractive for use in image-coding systems. Reeves and Kingsbury [43], Portilla and Simoncelli [45], and Romberg et al. [41], have derived accurate statistical models for complex wavelet coefficients. These models have proven useful in various different signal-processing applications. However, none of these models may be used for image coding, due to the redundancy of the transforms on which the models are based. On the other hand, because it can exploit directionality and statistical complex-wavelet models in a non-redundant framework, the non-redundant complex-wavelet transform has significant potential benefits for image coding systems.

8.1.3 Beyond Complex Wavelet Transforms

Most signal-processing applications involve a real-valued signal whose Fourier transform exhibits a conjugate symmetry that is responsible for a redundancy in the Fourier transform of the signal. Eliminating this redundancy creates a complex-valued signal whose DWT is directional, approximately shiftable and packed with useful phase information. Recall that the Fourier transform is an expansion of a real-valued signal in a basis with complex vectors. Therefore, this representation of a real-valued signal with complex basis-vectors generates a redundancy that is beneficial when eliminated.

Now the real field may be extended to the complex field by the introduction of the imaginary number j. It is possible to generalize this extension and derive new algebras by introducing additional imaginary numbers. For example, adding 3 imaginary numbers to the real field creates the quaternion system; adding 7 imaginary numbers to the real field creates the system of Cayley numbers [62]. These generalizations of the complex field are called Clifford algebras. Note that Mitrea [63] has derived wavelet transforms that operate in Clifford algebras. This provokes the following open questions:

1. Does the representation of a real-valued signal using a Clifford-algebra basis generate a redundancy that is also beneficial, when eliminated?

2. Would the DWT of the resulting Clifford-algebra signal representation have stronger shift invariance properties, higher directionality, or more revealing phase information than the complex-signal DWT that has been the subject of this thesis?

In this thesis, we have demonstrated that there are significant benefits to wavelet analysis in the field of complex numbers. We hope that positive answers to the above open questions will motivate wavelet analysis in the more esoteric Clifford algebras.

Appendix A

Relating Hardy Space to $L^2(\mathbb{R} \to \mathbb{R})$

The following results deal with the relationship between the Hardy space $H^2(\mathbb{R} \to \mathbb{C})$ and $L^2(\mathbb{R} \to \mathbb{R})$. Both function spaces are interpreted as vector spaces over the real field $\mathbb{R}$.

Proposition 11 *The Hardy space $H^2(\mathbb{R} \to \mathbb{C})$ is isometric to $L^2(\mathbb{R} \to \mathbb{R})$.*

Proof: Let f be an arbitrary $L^2(\mathbb{R} \to \mathbb{R})$ function. To obtain f^H, the Hardy-space image of f, we define a linear, bijective map between $H^2(\mathbb{R} \to \mathbb{C})$ and $L^2(\mathbb{R} \to \mathbb{R})$ as follows. The forward mapping is given by

$$F^H(\omega) \stackrel{\triangle}{=} F(\omega)\chi_{[0,\infty)}(\omega),$$

and the inverse mapping is specified by

$$F(\omega) \stackrel{\triangle}{=} F^H(\omega) + F^H(-\omega)^* - \chi_{[0,0]}(\omega)\left[\frac{F^H(\omega) + F^H(-\omega)^*}{2}\right]$$

Now,

$$\begin{aligned} ||f^H||_2^2 &= \int_{-\infty}^{\infty} \left|F^H(\omega)\right|^2 d\omega, \\ &= \int_0^{\infty} |F(\omega)|^2 d\omega, \\ &= \frac{1}{2}\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega, \text{ due to the conjugate symmetry of } f, \\ &= \frac{1}{2}||f||_2^2. \end{aligned}$$

This shows that the linear, bijective map between $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$ is also distance-preserving (upto a constant).

Hence $H^2(\mathbb{R} \to \mathbb{C})$ is isometric to $L^2(\mathbb{R} \to \mathbb{R})$. $\square$

Proposition 12 *Let $\langle \cdot, \cdot \rangle$ denote the $L^2(\mathbb{R} \to \mathbb{C})$ inner product. Define the Hardy-space inner product $\langle \cdot, \cdot \rangle_H$ by*

$$\langle \cdot, \cdot \rangle_H \triangleq 2\mathcal{R}e \left\{ \langle \cdot, \cdot \rangle \right\}.$$

Then the Hardy space $H^2(\mathbb{R} \to \mathbb{C})$ is isomorphic to $L^2(\mathbb{R} \to \mathbb{R})$, with respect to the Hardy-space inner product.

Proof: Suppose that f, g are arbitrary $L^2(\mathbb{R} \to \mathbb{R})$ functions. With respect to the linear, bijective map of Proposition 11, let f^H, g^H denote the Hardy-space images of f, g respectively. Now,

$$\begin{aligned}
\langle f, g \rangle &= \int_{-\infty}^{\infty} F(\omega) G^*(\omega)\ d\omega, \\
&= \int_{-\infty}^{0} F(\omega) G^*(\omega)\ d\omega + \int_{0}^{\infty} F(\omega) G^*(\omega)\ d\omega, \\
&= \int_{-\infty}^{0} F(\omega) G^*(\omega)\ d\omega + \int_{-\infty}^{\infty} F^H(\omega) G^H(\omega)^*\ d\omega, \\
&= \int_{-\infty}^{0} F(\omega) G^*(\omega)\ d\omega + \langle f^H, g^H \rangle, \\
&= -\int_{\infty}^{0} F(-\xi) G^*(-\xi)\ d\xi + \langle f^H, g^H \rangle, \text{ putting } \xi = -\omega, \\
&= \int_{0}^{\infty} F(-\xi) G^*(-\xi)\ d\xi + \langle f^H, g^H \rangle, \\
&= \int_{0}^{\infty} F^*(\xi) G(\xi)\ d\xi + \langle f^H, g^H \rangle, \text{ using the conjugate symmetry of } F \text{ and } G, \\
&= \int_{-\infty}^{\infty} F^H(\xi)^* G^H(\xi)\ d\xi + \langle f^H, g^H \rangle, \\
&= \langle g^H, f^H \rangle + \langle f^H, g^H \rangle, \\
&= \langle f^H, g^H \rangle^* + \langle f^H, g^H \rangle, \\
&= 2\mathcal{R}e \left\{ \langle f^H, g^H \rangle \right\}, \\
&= \langle f^H, g^H \rangle_H,
\end{aligned}$$

Therefore the linear, bijective map preserves the Hardy-space inner product. Hence $L^2(\mathbb{R} \to \mathbb{R})$ and $H^2(\mathbb{R} \to \mathbb{C})$ are isomorphic with respect to the Hardy-space inner product. $\square$

Bibliography

[1] A. Grossman and J. Morlet. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, 15(4):723–736, 1984.

[2] S. Mallat. Multiresolution approximation and wavelets. *Trans. of American Math. Soc.*, 315:69–88, 1989.

[3] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Applied Math.*, XLI(41):909–996, November 1988.

[4] I. Daubechies. Orthonormal bases of compactly supported wavelets II, variation on a theme. *SIAM Journal of Mathematical Analysis*, 24(2):499–519, March 1993.

[5] G. Strang. Wavelets and dilation equations: A brief introduction. *SIAM Review*, 31(4):614–627, 1989. also, MIT Numerical Analysis Report 89-9, August 1989.

[6] A. Grossman, R. Kronland-Martinet, and J. Morlet. Reading and understanding continuous wavelet transforms. In J. M. Combes, A. Grossman, and Ph. Tchamitchian, editors, *Wavelets Time-Frequency Methods and Phase Space*, pages 2–20, Berlin, 1989. Springer-Verlag. Proceedings of International Colloquium on Wavelets and Applications, Marseille, France, December 1987.

[7] G. H. Granlund and H. Knutsson. *Signal Processing for Computer Vision*. Kluwer Academic Publishers, 1995.

[8] G. Beylkin and B. Torresani. Implementation of operators via filter banks: Hardy wavelets and autocorrelation shell. *Applied and Computational Harmonic Analysis*, 3:164–185, 1996.

[9] S. Mallat. Zero-crossings of a wavelet transform. *IEEE Trans. Inform. Theory*, 37(4), July 1991.

[10] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.*, 29(6):1716–1740, 1992.

[11] H. Guo. Theory and applications of the shift-invariant, time-varying and undecimated wavelet transform. Master's thesis, Rice University, Houston, TX, May 1995.

[12] J. Liang and T. W. Parks. A translation invariant wavelet representation algorithm with applications. *IEEE Trans. SP*, 44(2):225–232, February 1996.

[13] S. A. Benno and J. M. F. Moura. Scaling functions robust to translations. *IEEE Transactions on Signal Processing*, 46(12):3269–3281, December 1998.

[14] F. Bao and N. Erdol. The optimal wavelet transform and translation invariance. In *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, volume 3, pages 13–16, 1994.

[15] I. W. Selesnick and L. Sendur. Iterated oversampled filter banks and wavelet frames. In *Wavelet Applications VII, Proceedings of SPIE*, 2000.

[16] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger. Shiftable multi-scale transforms. *IEEE Trans. Inform. Theory*, 38(2):587–607, March 1992.

[17] Q. Wang and L. Wu. A unified approach to translation-invariance problem. Preprint, 1998.

[18] P. Abry. *Transformées en ondelettes – Analyses multirésolution et signaux de pression en turbulence.* PhD thesis, l'Universite Claude Bernard - Lyon I, 1994.

[19] J. Magarey. *Motion estimation using complex wavelets.* PhD thesis, University of Cambridge, 1997.

[20] N. G. Kingsbury and J. F. A. Magarey. Wavelet transforms in image processing. In *Proc. First European Conference on Signal Analysis and Prediction*, pages 23–34, Prague, June 24-27 1997.

[21] N. Kingsbury. Image processing with complex wavelets. *Phil. Trans. R. Soc. Lond.*, September 1999.

[22] I. Selesnick. The design of Hilbert transform pairs of wavelet bases. Preprint, 2001.

[23] I. W. Selesnick. Hilbert transform pairs of wavelet bases. *IEEE Signal Processing Letters*, 8(6), June 2001.

[24] R. H. Bamberger and M. J. T. Smith. A filter bank for the directional decomposition of images: Theory and design. *IEEE Transactions on Signal Processing*, 40(4):882–893, April 1992.

[25] S-I. Park, M. J. T. Smith, and R. M. Mersereau. A new directional filter bank for image analysis and classification. In *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1999.

[26] M. Do and M. Vetterli. Pyramidal directional filter banks and curvelets. In *IEEE Proc. Int. Conf. Image Processing*, 2001.

[27] D. L. Donoho. Wedgelets: nearly-minimax estimation of edges, August 1997. Preprint.

[28] R. L. Claypoole, Jr. *Adaptive Wavelet Transforms via Lifting.* PhD thesis, Rice University, 1999.

[29] E. J. Candés. *Ridgelets: theory and applications*. PhD thesis, Stanford University, 1998.

[30] D. L. Donoho. Orthonormal ridgelets and linear singularities, 1998. Preprint.

[31] D. L. Donoho and M. R. Duncan. Digital curvelet transform: Strategy, implementation and experiments. In *Proceedings of SPIE, Wavelet Applications VII*, volume 4056, 2000.

[32] E. Candès and D. Donoho. Curvelets. Manuscript, 1999.

[33] A. Watson and J. A. J. Ahumada. Model of human visual-motion sensing. *Journal of Optical Society of America*, 2(2):322–342, 1985.

[34] T. Burns, S. Rogers, D. Ruck, and M. Oxley. Discrete, spatiotemporal, wavelet multiresolution analysis method for computing optical flow. *Optical Engineering*, 33(7):2236–2247, 1994.

[35] W. Lawton. Applications of complex valued wavelet transforms to subband decomposition. *IEEE Transactions on Signal Processing*, 41(12):3566–3568, December 1993.

[36] J.-M. Lina. Image processing with complex Daubechies wavelets. Preprint, 1997.

[37] X-P. Zhang, M. Desai, and Y-N. Peng. Orthogonal complex filter banks and wavelets: Some properties and design. *IEEE Transactions on Signal Processing*, 47(4), April 1999.

[38] A. Cohen. Ondelettes, analyses multirésolutions et filtres miroirs en quadrature. *Ann. Inst. Henri Poincaré, Analyse non linéaire*, 7(5):439–459, 1990.

[39] A. Cohen and R. D. Ryan. *Wavelets and multiscale processing*. Chapman and Hall, 1992.

[40] C. Bernard. *Ondelettes et problèmes mal posés: la mesure de flot optique et l'interpolation irrégulière*. PhD thesis, École Polytechnique, 1999.

[41] J. K. Romberg, H. Choi, R. G. Baraniuk, and N. G. Kingsbury. Hidden Markov tree models for complex wavelet transforms. *IEEE Transactions on Signal Processing*, 2001 (submitted).

[42] P. de Rivaz and N. Kingsbury. Complex wavelet features for fast texture image retrieval. In *IEEE Proc. Int. Conf. Image Processing*, 1999.

[43] T. H. Reeves and N. G. Kingsbury. Prediction of coefficients from coarse to fine scales in the complex wavelet transform. In *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, 2000.

[44] R. van Spaendonck and R. G. Baraniuk. Directional scale analysis for seismic interpretation. In *Annual Meeting Abstracts*, pages 1844–1847. Society Of Exploration Geophysicists, 1999.

[45] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 2000.

[46] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, MA, 1996.

[47] G. G. Walter. Analytic representations with wavelet expansions. *Complex Variables*, 26:235–243, 1994.

[48] N. Young. *An Introduction to Hilbert Space*. Cambridge University Press, 1988.

[49] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA, 1992. Notes from the 1990 CBMS-NSF Conference on Wavelets and Applications at Lowell, MA.

[50] P. P. Vaidyanathan. *Multirate Systems and Filter Banks.* Prentice Hall, Englewood Cliffs, NJ, 1992.

[51] I. W. Selesnick. Lowpass filters realizable as allpass sums: design via a new flat delay filter. *IEEE Trans. on Circuits and Systems Part II*, 46(1):40–50, January 1999.

[52] M. J. T. Smith and S. L. Eddins. Analysis/synthesis techniques for subband image coding. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(8):1446–1456, August 1990.

[53] G. Beylkin. On factored FIR approximation of IIR filters. *Applied and Computational Harmonic Analysis*, 2:293–298, 1995.

[54] T. Steeghs. *Local power spectra and seismic interpretation.* PhD thesis, Delft University of Technology, 1997.

[55] R. van Spaendonck, F. C. A. Fernandes, Mark Coates, and C. S. Burrus. Non-redundant, directionally selective, complex wavelets. In *Proc. Int. Conf. Image Processing, Extended Abstracts II*, pages 379–382. IEEE, 2000.

[56] F. C. A. Fernandes, R. van Spaendonck, M. Coates, and C. S. Burrus. Directional complex-wavelet processing. In *Wavelet Applications VII, Proceedings of SPIE*, 2000.

[57] F. C. A. Fernandes, R. L. C. van Spaendonck, and C. S. Burrus. A directional, shift-insensitive, low-redundancy, wavelet transform. In *IEEE Proc. Int. Conf. Image Processing*, October 2001, to appear.

[58] R. L. C. van Spaendonck, F. M. Hindriks, F. C. A. Fernandes, and G. G. Drijkoningen. Three-dimensional attributes for seismic interpretation. In *Ann. Mtg. Abstracts*, pages 2059–2062. Society of Exploration Geophysicists, 2000.

[59] F. Hindriks, R. L. C. van Spaendonck, G. G. Drijkoningen, and J. T. Fokkema. Directional scale analysis for three-dimensional seismic data. In *Extended Abstracts*, pages A–02. European Association of Geoscientists and Engineers, 2000.

[60] R.L.C. van Spaendonck. *Wavelet-based Attributes for Seismic Interpretation.* PhD thesis, Delft University of Technology, Delft, Netherlands, 2001.

[61] I. Overeem, C. Bishop, G.J. Weltje, and S.B. Kroonenberg. The late Cenozoic Eridanos delta system in the southern North Sea basin: a climate signal in sediment supply? *Basin Research*, 2001. In press.

[62] J.P. Ward. *Quaternions and Cayley Numbers.* Kluwer Academic Publishers, 1997.

[63] M. Mitrea. *Clifford wavelets, singular integrals and Hardy spaces.* Lecture notes in mathematics 1575. Springer-Verlag, 1994.