19980309 014

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

The Rational Resolution Analysis: A Generalization of Multiresolution Analyses with Application to the Specific Emitter Identification Problem

DISSERTATION

Bruce P. Anderson
Captain, USAF

AFIT/DS/ENC/97D-2

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/DS/ENC/97D-2

The Rational Resolution Analysis: A Generalization of Multiresolution Analyses with Application to the Specific Emitter Identification Problem

DISSERTATION

Bruce P. Anderson
Captain, USAF

AFIT/DS/ENC/97D-2

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

AFIT/DS/ENC/97D-2

The Rational Resolution Analysis: A Generalization of Multiresolution Analyses with Application to the Specific Emitter Identification Problem

Bruce P. Anderson, B.S.E.E., M.S., M.S.E.E.

Captain, USAF

Approved:

Research Advisor, Dr. Gregory T. Warhola	24 Nov 97
Mark E. Oxley Committee Member, Dr. Mark E. Oxley	24 Nov 97
Joseph J. Sacchini Committee Member, Dr. Joseph J. Sacchini	24 Nov 97
Dean's Representative, Dr. Milton E. Franke	8 Dec 97

Robert A. Calico, Jr.

Dean

AFIT/DS/ENC/97D-2

The Rational Resolution Analysis: A Generalization of Multiresolution Analyses with Application to the Specific Emitter Identification Problem

DISSERTATION

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Bruce P. Anderson, B.S.E.E., M.S., M.S.E.E.
Captain, USAF

December 16, 1997

Acknowledgements

The rational resolution analysis is an area I began researching as a Masters Degree student at AFIT in 1991. The rational resolution analysis is a small part of the revolutionary mathematical theory of wavelets. It provides a general framework of resolution analyses which contains the well-known multiresolution analyis of Stephane Mallat and Yves Meyer as a special case. It builds on the work of Ingrid Daubechies and Pascal Auscher along with a list of others too long to list here. I hope this dissertation proves to be a lasting tribute to their work. There are some others that need to be mentioned as well.

I would like to thank my research advisor Dr. Greg Warhola for his guidance throughout the last six years. He has always demanded my best work and I thank him for setting the bar so high. I thank Dr. Mark Oxley for his timely guidance as well, especially in the last few months of the dissertation. I thank Dr. Joe Sacchini for providing me a solid background in signal processing and being a good friend and ally in recent years. I would also like to acknowledge Dr. Larry Merkle and Laura Suzuki for many thought-provoking conversations. I will miss that part of the research the most.

I would like to thank the National Reconaissance Office for its corporate support of advanced degrees and for providing me time and other resources necessary to bring the dissertation to fruition over this last year.

Finally, I acknowledge an immense debt of gratitude to my wife for her patience, support, understanding, and love, especially when the dissertation kept me away from home for long hours and preoccupied me when I was there. Thank you for letting me pursue my dream at the expense of your own. It is a gift I can never repay. I love you very much.

Bruce P. Anderson

Table of Contents

Page

Page

List of Figures

List of Tables

AFIT/DS/ENC/97D-2

Abstract

The rational resolution analysis (RRA) is introduced and developed as a generalization of the integer-dilation multiresolution analyses (MRA) developed by Mallat and Meyer. Rational dilation factors are achieved by relaxing the condition on MRAs that successive approximation spaces be embedded. Conditions for perfect reconstruction are discussed and it is shown that perfect reconstruction is possible with specific constraints on the scaling function: the scaling filter must have its roots on the unit circle. Furthermore, the required arrangement of the roots indicate the scaling function must be derived from a B-spline of some degree. It is proven the only compactly supported scaling function which satisfies these constraints is the Haar basis. Generation of wavelets for arbitrary integer dilation factors is an important part of RRA perfect reconstruction and an algorithmic approach to constructing these wavelets is presented. The frame properties of perfect reconstruction RRAs is presented and the notion of adjoint wavelets are made rigorous. It is shown the adjoint wavelets form a frame for V_0 and that the corresponding decomposition is both stable and unique. Finally, the redundant representation of the detail coefficients is exploited as a solution to the specific emitter problem. Results demonstrate the RRA is far superior to the traditional MRA and wavepacket approaches when used as a feature extractor in Bayesian classification schemes.

The Rational Resolution Analysis: A Generalization of Multiresolution Analyses with Application to the Specific Emitter Identification Problem

I. *Introduction*

1.1 Introduction

Wavelet theory has emerged in the last decade and has provided a fresh perspective on many significant problems across a wide range of disciplines. Pioneered by the French geophysicist Jean Morlet as a tool for oil prospecting, the theory was brought to practical fruition by Stephane Mallat in his famous paper[27]. Mallat introduced the concept of a *multiresolution analysis* whereby a signal is decomposed by iteratively applying an approximation operation which resulted in a representation of the signal "seen" at sucessively lower resolutions. By keeping the parts of the signal which is lost at each level of approximation, the "details," Mallat showed that the original signal could be exactly reconstructed starting with the lowest level approximation and successively adding more and more of the details.

There were two remarkable things about Mallat's multiresolution analysis. First, the algorithms which implemented it were extremely fast; the approximation and detail operations could be performed as discrete filter operations for which extremely fast dedicated hardware exists. The impact was similar to that of the introduction of the fast Fourier transform (FFT); it made the calculations practical for digital computers and led to many real-world applications.

Secondly, in the paper which introduced the multiresolution analysis, Mallat provided a unifying framework for most, if not all, of the wavelet-like processing being done at the time. Mallat showed that the pyramid algorithms of image processing[8], the subband coding and quadrature-mirror schemes of signal processing[32, 36, 37] were all essentially based on wavelet theory. The multiresolution analysis became the vehicle for the application of wavelets and wavelet theory across an even wider range of disciplines.

Wavelet analysis and Fourier analysis are inextricably linked in the sense that much of the development of wavelet theory uses Fourier theory as building blocks. While the FFT had far-reaching impacts on society, some argued it was a victim of its own success. According to Yves Meyer, the French mathematician widely credited with helping Mallat write his monumental paper, "Because the FFT is very effective, people have used it in problems where it is not useful–the way Americans use a car to go half a block. Cars are very useful, but that's a waste of a car.[20] " Fourier analysis does not work equally well for all kinds of signals or for all kinds of problems.

The main problem with Fourier analysis is that information about time is concealed. Fourier analysis is underpinned by sines and cosines which exist for all time. It assumes the frequency content of a signal is a homogeneous property of the signal and ignores transient behavior or changing frequencies. As Barbara Hubbard puts it, "[the Fourier transform] pretends, so to speak, that every instant of signal is identical to every other, even if the signal is complex as a Mozart symphony or changes dramatically as the electrocardiogram of a fatal heart attack.[20]" The Fourier transform takes a transient of a signal, a local property of a signal, and makes it to a global property of the transform. Wavelet analysis allows representation of a signal's transients since bonafide wavelets are localized in time having sufficient decay for a finite energy property.

This dissertation mostly deals with multiresolution analyses of one form or another. Where Mallat's original multiresolution analysis was based on a dilation factor of two (each level of approximation has half the resolution of the previous), this research introduces the rational multiresolution analysis, or simply the rational resolution analysis (RRA), whereby the dilation factors taken from the set of rational numbers are considered. The properties of this new analysis are investigated and identified. As an example of its applicability, the rational resolution analysis is used to create feature vectors in the specific emitter identification problem. The following sections detail the specific motivation for this dissertation, provide its specific objectives, describe the notation used throughout the document, and preview the remainder of the document.

1.2 Motivation

The multiresolution analyses á la Mallat with integer dilation factors provide a representation of $L_2(\mathbb{R})$ which is "tight" in that it is complete without being redundant. In the Fourier domain, the dilation factor determines the ratio of the bandwidth of the approximation filter at one dilation level to that of the next. This dissertation is primarily motivated by an apparent lack of knowledge in the general body concerning multiresolution-like analyses based on rational dilation factors. Rational dilation factors may be more appropriate in applications where a logarithmic segmentation of the Fourier spectrum is desired, but greater freedom is needed in defining the bandwidths associated with the approximation spaces.

Previous work by Auscher[4, 5] has discovered that no useful wavelet bases exist based on rational dilation factors when the wavelets are orthogonal from one decomposition level to the next. The resulting wavelets from such attempts were proven to exist, but lacked exponential decay and hence, were not bonafide wavelets in the sense that they lacked localization in time. By relaxing the condition that the wavelets be orthogonal across dilation levels, one can construct multiresolution analyses whereby detail spaces overlap in the Fourier domain. This overlap translates into a redundancy in the decomposition which can be exploited in certain applications. As a notional example, consider the case where a frequency of interest falls near the boundary between the Fourier representations of two adjacent detail spaces. Energy at that frequency has a divided representation; part of the energy is represented by wavelets in one detail space and part by the wavelets in the other. There is no clear winner. With a redundant representation, successive wavelet representations would provide a more accurate picture of how the energy is distributed throughout a particular signal.

A secondary motivation for this dissertation involves the mathematical theory surrounding the development of a rational-dilation multiresolution analysis. Questions that quickly arise in this research are whether there exist functions that satisfy a dual system of two-scale equations and what are the characteristics of such functions. A function ϕ which satisfies a two-scale equation satisfies

$$\phi(t) = \sum_n h_n \phi(pt - n)$$

where the h_n are the scaling coefficients and p is the dilation factor. A function which satisfies a dual system of two-scale equations is one for which a set of scaling coefficients exist for two different dilation factors.

In the course of pursuing the motivations described above, several theoretical problems arise which have become ancillary motivations for this dissertation. They represent unique challenges which, when resolved, contribute to the overall state-of-the-art in multiresolution-like analyses and associated mathematical foundations.

1.3 Scope

This dissertation focuses on several specific objectives which clearly delineate its scope. Each of the objectives roughly translates to a chapter. The objectives are not independent since the achieving of one objective usually raises another issue, the resolution of which becomes a subsequent objective. The remainder of this section briefly describes these objectives and their interrelationship.

1. The first objective involves the the two-scale equation introduced in the previous section. Characterizing solutions to a dual system of two-scale equations, where the two dilation factors p and q are positive integers, is extremely important in studying the rational resolution analysis which is introduced later. With respect to this objective, certain questions naturally arise which this dissertation attempts to answer. Is there a class of compactly supported functions which solves the dual system of two-scale equations? In the context of the $L_2(\mathbb{R})$ Hilbert space, which solutions are self-orthogonal

with respect to their integer translations? The answers to these questions are a goal of this dissertation.

2. The answer to the previous questions lead naturally into B-spline functions. The second major objective of this research is to provide a rigorous algorithmic approach to constructing B-spline wavelets of arbitrary regularity corresponding to an arbitrary positive integer dilation factor. The relationship of this objective to the previous one will be made clear in the subsequent chapters.

3. The third major objective is the determination of the frame characteristics of the rational resolution analysis. Specifically, that the "adjoint" wavelets associated with a rational resolution analysis form a frame for an arbitrarily fine approximation space of $L_2(\mathbb{R})$ will be proven. Furthermore, useful estimates of the frame bounds will be determined.

4. The final objective is to apply the rational resolution to the specific emitter identification problem and demonstrate its applicability and perhaps its superiority to similar techniques.

These four objectives, along with related secondary objectives, form the corners of a tetrahedron within which the scope of the dissertation is confined. However, in the course of research toward the objectives, questions arise which are related to the objectives but outside the scope of the work. These issues will either be identified in the final chapter as an area for future work or else discussed where they present themselves as an area of interest, but not within the scope. The relative overall importance of the issue will determine under which category it falls.

1.4 Notation and Preliminaries

1.4.1 Notation.
The following notation is used throughout this dissertation.

1.4.1.1 Number Systems and Sets.
Table 1.4.1.1 lists the notation for the various number sets used in this document.

Symbol	Meaning
$\mathbb{N}$	the set of natural numbers: $\{1, 2, 3, \ldots\}$
$\mathbb{C}$	the set of complex numbers
$\mathbb{Z}$	the set of integers
$\mathbb{Z}^+$	the set of non-negative integers
$\mathbb{Z}_p$	the set of integers modulo p where $p \in \mathbb{N}$: $Z_p = \{0, 1, 2, \ldots, p-1\}$.
$\mathbb{R}$	the set of real numbers
$\mathbb{Q}$	the set of rational numbers
$\mathbb{T}$	the unit circle: $\mathbb{T} = \{z \in \mathbb{C} : \lvert z\rvert = 1\}$

Table 1.1 Number Set Notation

1.4.1.2 Vector Spaces.

- $L_1(\mathbb{R})$ for the space of Lebesgue measurable and absolutely integrable complex-valued functions:

$$L_1(\mathbb{R}) = \{f : f \text{ is Lebesgue-measureable and } \int_{\mathbb{R}} |f(x)|dx < \infty\}.$$

- $L_2(\mathbb{R})$ for the space of Lebesgue-measurable, square-integrable complex-valued functions:

$$L_2(\mathbb{R}) = \{f : f \text{ is Lebesgue-measureable and } \int_{\mathbb{R}} |f(x)|^2 dx < \infty\}.$$

  If $f \in L_2(\mathbb{R})$, f is sometimes referred to as a finite-energy function. With the following standard definition for the inner product, $L_2(\mathbb{R})$ is a Hilbert space.

- The standard inner product of $f, g \in L_2(\mathbb{R})$ will be denoted by

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx,$$

  where the overbar represents complex conjugation.

- $l_2(\mathbb{Z})$ for the space of square-summable sequences:

$$l_2(\mathbb{Z}) = \left\{ a = (\ldots, a_{-1}, a_0, a_1, \ldots) : a_k \in \mathbb{C}, \sum_{k=-\infty}^{+\infty} |a_k|^2 < \infty \right\}.$$

With the standard definition of the inner product, $l_2(\mathbb{Z})$ is also a Hilbert space.

1.4.1.3 Transforms and Operations.

- The Fourier transform of a Lebesgue-integrable function f will be denoted by $\hat{f}$. It is defined as

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x} dx \quad f \in L_1$$

For $f \in L_2(\mathbb{R})$, an appropriate limiting definition for the Fourier transform must be used[10]. For f which is $L_2(\mathbb{R})$ but not $L_1(\mathbb{R})$, the Fourier transform can be defined by

$$\hat{f}(\xi) = \lim_{\lambda \to \infty} \int_{\mathbb{R}} e^{|x|/\lambda} f(x) e^{-i\xi x} dx \quad f \in L_2(\mathbb{R}).$$

- The z-transform of a sequence $x = (\ldots, x_{-1}, x_0, x_1, \ldots)$ is defined by the Laurent series

$$X(z) = \sum_{k \in \mathbb{Z}} x_k z^{-k}$$

where X is often called the *symbol* of the sequence x. The z-transform does not converge for all sequences or for all values of z. For any given sequence, the set of values for which the z-transform converges is called the *region of convergence*, abbreviated by ROC.

- The Fourier transform of a sequence x is defined to be the symbol X restricted to the unit circle. This assumes the ROC of x includes the unit circle; the Fourier transform does not converge otherwise. Uniform convergence of the Fourier transform requires the sequence to be absolutely summable. The Fourier transform of sequences which are not absolutely summable, but which have finite energy (square summable) can be defined in a manner analagous to the continuous case with an appropriate limiting operation. For instance, the Fourier transform of the sequence

$$x(n) = \frac{\sin(\omega n)}{\pi n}$$

converges in a mean-square sense to a discontinuous periodic function[31].

- The Kronecker delta function, δ, is defined in two forms depending on the domain. The proper form will be clear from the context. For $\delta : \mathbb{Z} \to \mathbb{R}$:

$$\delta(k) = \delta_k = \delta_{k,0} = \begin{cases} 1, & k = 0 \\ 0, & \text{otherwise.} \end{cases}$$

  It is often seen with two arguments, for which is defined: $\delta(j,k) = \delta_{j,k} = \delta(j-k)$.

- The characteristic (indicator) function $\chi : \mathbb{R} \to \mathbb{R}$ is defined for a subset $A \subset \mathbb{R}$.

$$\chi_A(t) = \begin{cases} 1, & t \in A, \\ 0, & \text{otherwise.} \end{cases}$$

- Given a polynomial $F : \mathbb{C} \to \mathbb{C}$, let $\mathcal{Z}(F)$ be the zero set of F:

$$\mathcal{Z}(F) = \{z \in \mathbb{C} : F(z) = 0\}$$

1.4.2 The Orthonormalization Technique. One technique which will be used throughout this dissertation is the "orthogonalization trick" attributed to Daubechies[15]. It is best described by example. Given a function $g \in L_2(\mathbb{R})$ whose integer translations form a Riesz (unconditional) basis for some Hilbert space $\mathcal{H} \subset L_2(\mathbb{R})$, there is an invertible operator T which maps the sequence

$$G = \{g_n\}_{n \in \mathbb{Z}}$$

where $g_n = g(\cdot - n)$ to an orthonormal basis[15], $\{\gamma_k\}_{k\in\mathbb{Z}}$ of $\mathcal{H}$. An orthonormal basis satisfies

$$\langle \gamma_n, \gamma_m \rangle = \delta_{m-n}$$

which can be written equivalently via the Poisson summation formula

$$\sum_{k\in\mathbb{Z}} |\hat{\gamma}(\xi + 2\pi k)|^2 = 1 \quad \text{a.e.} \xi \in \mathbb{R}.$$

Then we can define T by

$$\hat{\gamma}(\xi) = (\widehat{Tg})(\xi) = \frac{\hat{g}(\xi)}{\left(\sum_{k\in\mathbb{Z}} |\hat{g}(\xi + 2\pi k)|^2\right)^{1/2}}$$

so that γ is an "orthonormalized" version of g. The denominator in the definition does not vanish as a consequence of g forming a Riesz basis for $\mathcal{H}$ which is equivalent to the existence of positive constants A and B such that

$$A\|f\|^2 \leq \sum_n |\langle f, g_n \rangle|^2 \leq B\|f\|^2 \quad \forall f \in \mathcal{H}.$$

This Riesz condition will have some usefulness in the discussion of frames later in the dissertation. Frames differ from Riesz bases in that frames are not necessarily a linearly independent set.

1.4.3 The Downsample Operator. The downsample operator defined below is used periodically throughout this dissertation.

Definition 1.1. Given a sequence x with corresponding symbol X, and a subsample factor $p \in \mathbb{N}$, the subsample (or downsample) operator, D_p, is defined by

$$(D_pX)(z) = \frac{1}{p}\sum_{l=0}^{p-1} X(r_l(z))$$

where $r_l(z)$ is the l-th p-root of z. Specifically, with $z = re^{i\omega}$, r_l is defined for $l \in \mathbb{Z}_p$ by

$$r_l(z) = r^{1/p} e^{i(\omega+2\pi l)/p} \qquad \omega \in [0, 2\pi).$$

It is assumed that z is restricted to the region of convergence of $D_p X$. ∎

For $z = re^{i\omega}$,

$$\begin{aligned}
(D_p X)(re^{i\omega}) &= \frac{1}{p} \sum_{l=0}^{p-1} X(r^{-1/p} e^{-i(\omega+2\pi l)/p}) \\
&= \frac{1}{p} \sum_{l=0}^{p-1} \sum_{n} x(n) r^{n/p} e^{-in(\omega+2\pi l)/p} \\
&= \sum_{n} x(n) r^{-n/p} e^{-in\omega/p} \frac{1}{p} \sum_{l=0}^{p-1} e^{-i2\pi ln/p} \\
&= \sum_{n} x(pn) (re^{i\omega})^{-n} \\
&= \sum_{n} x(pn) z^{-n}.
\end{aligned}$$

The last step is a consequence of

$$\frac{1}{p} \sum_{l=0}^{p-1} e^{-i2\pi ln/p} = \begin{cases} 1, & n \in p\mathbb{Z} \\ 0, & \text{otherwise.} \end{cases}$$

The final expression for $D_p X$ above indicates why it is called the downsample operator; only the p-th elements of x are kept.

1.5 Preview

As was mentioned earlier, this dissertation has specific objectives which form the backbone of the research. Each chapter is primarily dedicated to an objective and the chapters flow according to the logical progression of the objectives.

Chapter II begins the dissertation with a brief review of multiresolution analyses of $L_2(\mathbb{R})$ which includes a discussion of wavelets and scaling functions and the discrete filter operators which form the algorithmic heart of the MRA. Spline-based scaling functions and wavelets are important to the rational resolution analysis and their properties are discussed. The chapter concludes with a formal introduction to the rational resolution analysis (RRA). The associated filter operations are discussed for both analysis and synthesis and the requirements for exact (perfect) reconstruction are derived and presented.

Chapter III is primarily concerned with the existence and characterization of the class of compactly supported RRA scaling functions which result in perfect reconstruction. The dual dilation symbol equation is presented and solutions are characterized in detail with the development of several novel theorems. The chapter concludes with a characterization of the class of scaling functions for which the associated RRA yields perfect reconstruction.

Chapter IV deals with the algorithmic construction of spline-based wavelets which complement the scaling functions introduced and used in Chapters II and III. Semi-orthogonal wavelets are presented as a means of constructing orthogonal ones. The algorithmic development is framed as finding a set of vectors which span the nullspace of a particular operator related to the scaling function. The completeness of the method is proven. The chapter concludes with a three-step pseudo-algorithm to contstruct spline-based wavelets bases with an arbitrary dilation factor and regularity.

Characterizing the extent to which the wavelets associated with an RRA form a frame for $L_2(\mathbb{R})$ is the topic of Chapter V. The analysis focuses on two fundamental questions. First, is the RRA wavelet decomposition unique to a particular $L_2(\mathbb{R})$ function, and second, can every $L_2(\mathbb{R})$ function be represented in a stable way. Stability is equivalent to continuity and implies that if two sets of RRA wavelet coefficients are "close", then the $L_2(\mathbb{R})$ functions reconstructed from these sequences are "close" as well. In this context, the metric induced by the inner product norm is used to define closeness. The uniqueness and stability of the RRA decomposition can be characterized by what are known as the frame bounds. This chapter provides estimates for these frame bounds which are functions of regularity and, of course, dilation factor.

Chapter VI provides the details of an application of an RRA decomposition a real-world problem. The specific emitter identification problem is one in which a particular radar system must by identified by one pulse of its collected signal as being one from a set of which are intended to be identical. That is, a radar system must be identified at the serial number level. This chapter develops a feature extraction scheme using the RRA and entropy calculations. Classification performance using these features is compared to similar non-RRA techniques with remarkable results.

Chapter VII is the final chapter in the dissertation and provides a summary of the research. Conclusions are made and recommendations for further work outside the scope of this dissertation are presented.

II. *Mathematical Foundations of the Rational Resolution Analysis*

2.1 Introduction

This chapter introduces the rational resolution analysis (RRA). The theoretical foundations of the RRA are found in the standard multiresolution analyses with dilation factor $p = 2$ and its natural extension to integer dilation factors $p > 2$. Scaling functions and wavelets are discussed as they pertain to both multiresolution and rational resolution analyses. Approximation and detail spaces are reviewed along with the associated projection operators which form the heart of the analysis.

The B-spline scaling functions are given special treatment. They have unique properties which are both useful and necessary in the RRA. In fact, it is shown later in this work that the spline-based scaling functions and wavelets are the only such class of functions which satisfy the unique requirements for perfect reconstruction with the RRA.

The final section of this chapter provides a formal definition of the RRA. It describes the analysis (or decomposition) operators in terms of discrete filter operations. It also presents the scheme for reconstruction which highlights the necessary and sufficient conditions for perfect reconstruction.

2.2 Integer Dilation Multiresolution Analysis

This section reviews the standard multiresolution analysis of Mallat[27, 28] and Meyer[29] and its generalization to arbitrary integer dilation factors[1, 33].

2.2.1 Definitions and Preliminaries.

Definition 2.1. A multiresolution analysis (MRA) of $L_2(\mathbb{R})$ (with dilation factor $p \in \mathbb{N}$) is a sequence $\{V_k\}_{k\in\mathbb{Z}}$ of closed subspaces of $L_2(\mathbb{R})$ such that the following hold:

1. The approximation spaces are embedded: $V_k \subset V_{k-1},\ k \in \mathbb{Z}$.

2. The approximation spaces satisfy

$$\bigcap_{k\in\mathbf{Z}} V_k = \{0\} \quad \text{and} \quad \overline{\bigcup_{k\in\mathbf{Z}} V_k} = L_2(\mathbb{R}),$$

3. $$f \in V_k \Longleftrightarrow f(p\cdot) \in V_{k-1} \quad \forall\, k \in \mathbb{Z},$$

4. $$f \in V_k \Longrightarrow f(\cdot - p^{-k}k) \in V_k \quad \forall\, k \in \mathbb{Z}.$$

Furthermore, there exists a scaling function $\phi \in V_0$ such that $\{\phi_{k,n}\}_{n\in\mathbf{Z}}$ forms a Riesz basis basis for V_k so that

$$V_k = \overline{\text{span}\{\phi_{k,n}\}_{n\in\mathbf{Z}}},$$

where

$$\phi_{k,n} = p^{-k/2}\phi(p^{-k}\cdot - n).$$

■

The integer translations of ϕ are not necessarily orthogonal, but with Daubechies'[14] "orthonormalization trick", we can find an orthogonal basis with no loss of generality. It is described briefly in the first chapter. In what follows, we assume $\{\phi_{0,n}\}_{n\in\mathbf{Z}}$ is an orthonormal set.

The detail space, W_k, is defined as the orthogonal complement of V_k in V_{k-1}. The existence of W_k for integer dilation factors has been proven in [33] and [27] among others. This means

$$W_k \perp V_k$$
$$W_k \subset V_{k-1},$$

and

$$V_k \oplus W_k = V_{k-1}.$$

In general, the detail space is spanned by the integer translations of $p-1$ linearly independent wavelets, $\psi^{(j)}$ for $j = 1, \ldots, p-1$. That is,

$$W_k^{(j)} = \overline{\text{span}\{\psi_{k,n}^{(j)}\}_{n \in \mathbf{Z}}}$$

and

$$W_k = \sum_{j=1}^{p-1} W_k^{(j)},$$

where the summation is used to mean the direct sum of the spaces and

$$\psi_{k,n}^{(j)} = p^{-k/2} \psi^{(j)}(p^{-k}x - n).$$

In the case where the $\psi^{(j)}$ are all orthonormal with respect to all dilations and translations, i.e. $\langle \psi_{k,n}^{(j)}, \psi_{k',n'}^{(j')} \rangle = \delta_{j,j'} \delta_{k,k'} \delta_{n,n'}$, the expression for W_k above can be written as an orthogonal sum instead of a direct sum.

An element $f \in L_2(\mathbb{R})$ can be represented as a linear combination of these integer-indexed wavelets $\psi_{k,n}^{(j)}$ in the following manner. The approximation of f at the k-th resolution level is the orthogonal projection of f onto V_k. We write

$$P_k f \in V_k \Longrightarrow (P_k f)(t) = \sum_n c_{k,n} \phi_{k,n}(t),$$

where P_k is the orthogonal projection operator, $P_k : L_2(\mathbb{R}) \longrightarrow V_k$, and

$$c_{k,n} = \langle f, \phi_{k,n} \rangle$$

are known as the approximation coefficients. The approximation at level k is entirely characterized by the sequence $c_{k,n}$ for $n \in \mathbb{Z}$. We can define a similar projection operator to project f onto W_k:

$$Q_k f = \sum_{j=1}^{p-1} Q_k^{(j)} f \in W_k$$

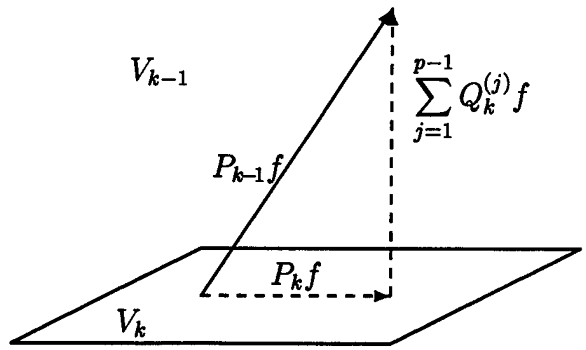

Figure 2.1 Approximation Spaces and Projection Operations

with

$$(Q_k^{(j)} f)(t) = \sum_n d_{k,n}^{(j)} \psi_{k,n}^{(j)}(t) \quad j \in \{1, 2, \ldots, p-1\},$$

where the $d_{k,n}^{(j)}$ are known as detail coefficients and are defined similarly to the $c_{k,n}$. From these two projections, we can reconstruct the approximation of f at the $m-1$ level by

$$P_{k-1} f = P_k f + \sum_{j=1}^{p-1} Q_k^{(j)} f. \tag{2.1}$$

This is illustrated graphically in Figure 2.1. It is shown in [28] for dilation factor $p = 2$, and later in [33] for aribitrary $p \in \mathbf{N}$ that

$$\bigoplus_{k \in \mathbf{Z}} W_k = L_2(\mathbb{R}),$$

which implies that all p dilations and translations of the mother wavelet ψ form an orthonormal basis for $L_2(\mathbb{R})$. The multiresolution framework provides an efficient and elegant algorithm to find the approximation and detail coefficients at each resolution level by using discrete filters on the approximation coefficients of the next higher level. Each successive set of approximation and detail coefficients can be calculated from the previous set of approximation coefficients and this calculation is independent of level.

To see this, consider the scaling function $\phi_{1,0}$. Because we have $V_1 \subset V_0$, we can express $\phi_{1,0}$ as a linear combination of $\phi_{0,k}$

$$\phi_{1,n} = p^{-1/2}\phi(\cdot p^{-1} - n) = \sum_k h_p(k)\phi(\cdot - k), \tag{2.2}$$

where

$$h_p(n) = \langle \phi_{1,0}, \phi_{0,n} \rangle. \tag{2.3}$$

Taking the Fourier transform of both sides of Equation 2.2 for $n = 0$ yields

$$\hat{\phi}(p\xi) = p^{-1/2} H_p(e^{i\xi})\hat{\phi}(\xi), \tag{2.4}$$

where H_p is the Fourier transform of the sequence h_p:

$$H_p(e^{i\xi}) = \sum_k h_p(k)e^{-i\xi k}.$$

The function H_p acts as a filter in the sense that the Fourier transform of the lower resolution scaling function, $\widehat{\phi_{1,0}}$, is found by filtering $\widehat{\phi_{0,0}}$ by H_p. This filter has a great deal of significance in the multiresolution analysis. Mallat[28] has demonstrated that approximation coefficients $c_{k,n}$ at one level can be easily calculated from the approximation coefficients of the previous level by using the coefficients of the filter H_p for the case $p = 2$. The extension for arbitrary $p \in \mathrm{N}$ is

$$c_{k,l} = \sum_n h_p(n - pl)c_{k-1,n}. \tag{2.5}$$

Notice that the approximation coefficients $c_{k,n}$, which completely characterize the approximation, can be calculated without having to calculate the inner product of f with $\phi_{k,n}$ directly except for some initial approximation level. This is the elegance of Mallat's algorithm.

The detail coefficients $d_{k,n}^{(j)}$ can also be calculated from the $c_{k-1,n}$ in a similar way. If we define

$$g_p^{(j)}(n) = \langle \psi_{1,0}^{(j)}, \phi_{0,n} \rangle \quad j \in \{1, 2, \ldots, p-1\}$$

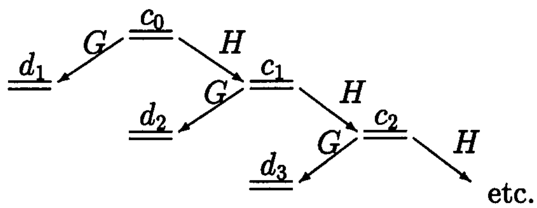

Figure 2.2 Illustration of the p-Dilation Multiresolution Analysis

then, taking Fourier transforms, we get

$$p^{1/2}\hat{\psi^{(j)}}(p\xi) = G_p^{(j)}(e^{i\xi})\hat{\phi}(\xi), \tag{2.6}$$

with $G_p^{(j)}(e^{i\xi})$ defined to be the Fourier transform of $g_p^{(j)}$. Now, at each resolution level k in a multiresolution decomposition, we get a set of detail coefficients $\{d_{k,n}^{(j)}\}_{n\in\mathbf{Z}}$ for $j = 1, 2, \ldots, p-1$. With $G = \sum_{j=1}^{p-1} G^{(j)}$, this is illustrated in Figure 2.2. As the number of decompositions grows without bound, the function will be represented by the sets of detail coefficients

$$d_{k,n}^{(j)} = \langle f, \psi_{k,n}^{(j)} \rangle.$$

Using the approximation and detail coefficients at a particular resolution level k, we can reconstruct the approximation at the $k-1$ level by using the same filter coefficients

$$c_{k-1,n} = \sum_l h_p(n-pl)c_{k,l} + \sum_{j=1}^{p-1}\sum_l g_p^{(j)}(n-pl)d_{k,l}^{(j)}.$$

This is a consequence of Equation 2.1.

2.2.2 Filter Operations. The filters H_p and $G_p^{(j)}$ are very important. However, as Daubechies points out in [14], Mallat's algorithm deals only with sequences; the underlying multiresolution analysis is only used in the computation of the filters H_p and $G_p^{(j)}$. Daubechies studied the filters and determined they had certain properties which allowed Mallat's discrete algorithm to be "weaned" from its multiresolution parent.

By defining $\mathbf{H}_p$ and $\mathbf{G}_p^{(j)}$ as bounded operators from $l^2(\mathbb{Z})$ to itself:

$$(\mathbf{H}a)_k = \sum_n h(n-pk)a_n,$$
$$(\mathbf{G}^{(j)}a)_k = \sum_n g^{(j)}(n-pk)a_n, \tag{2.7}$$

Daubechies showed that necessary conditions for Mallat's algorithm to work could be expressed as

$$\sum_n |h_p(n)| < \infty$$
$$\sum_n |g_p^{(j)}(n)| < \infty \quad j \in \{1,2,\ldots,p-1\}. \tag{2.8}$$

which is equivalent to requiring the filter sequences h_p and $g_p^{(j)}$ for $j \in \{1,2,\ldots,p-1\}$ to elements of $l_1(\mathbb{Z})$, absolutely summable sequences. In the signal processing community, they are known as a stability requirements[31]. We want to perfectly reconstruct a sequence from its decomposition, so with the adjoint operators $\mathbf{H}_p^*$ and $\mathbf{G}_p^{(j)*}$ given by

$$(\mathbf{H}^*a)_n = \sum_k h_p(n-pk)a_k,$$
$$(\mathbf{G}^{(j)*}a)_n = \sum_k g_p^{(j)}(n-pk)a_k, \quad j\{1,2,\ldots,p-1\}, \tag{2.9}$$

we also require a resolution of the identity by

$$\mathbf{HH}^* + \sum_{j=1}^{p-1} \mathbf{G}^{(j)}\mathbf{G}^{(j)*} = \mathbf{1}, \tag{2.10}$$

where $\mathbf{1}$ is the identity operator on $l_1(\mathbb{Z})$. Similar to the requirement that the detail and approximation spaces be orthogonal, we require

$$\mathbf{G}^{(j)}\mathbf{G}^{(i)*} = \mathbf{0} \quad \forall i, j \in \{1, 2, \ldots, p-1\}, i \neq j$$
$$\mathbf{H}\mathbf{G}^{(j)*} = \mathbf{0} \quad \forall j \in \{1, 2, \ldots, p-1\}.$$

Finally, approximation and detail roles are assigned to the operators by requiring

$$\sum_n h_p(n) = \sqrt{p},$$
$$\sum_n g_p^{(j)}(n) = 0, \quad \forall j \in \{1, 2, \ldots, p-1\}. \tag{2.11}$$

The properties described above are identified by Daubechies[14] as the essence of Mallat's algorithm. They are the conditions which allow you to separate the algorithm from the multiresolution analysis. This is important since it allows the transition from an elegant theory to a practical implementation.

2.2.3 Scaling Functions and Wavelets. The scaling functions and wavelets are related to the approximation and detail filters H and $G^{(j)}$, respectively. Conditions on these filters were presented which allow Mallat's algorithm to work. This section reviews how these conditions affect the scaling functions and wavelets. The following definition is useful for that discussion.

Definition 2.2. The regularity polynomial, S_p, with degree $p \in \mathbb{N}$ is defined by

$$S_p(z) = \frac{1}{p}\sum_{n=0}^{p-1} z^{-n} \quad z \in \mathbb{C} \tag{2.12}$$

■

Equations 2.4 and 2.6 express the relationship between the filters and the corresponding basis functions. Iterating Equation 2.4 and implicitly redefining H_p to include the scaling factor $p^{-1/2}$ yields

$$\hat{\phi}(\xi) = \hat{\phi}(0) \prod_{k=1}^{\infty} H_p(e^{i\xi p^{-k}}), \tag{2.13}$$

provided the infinite product converges. Thus, the scaling function is determined by the iterated product of the approximation filter. Since we assume ϕ exists, we assume this infinite product converges. In addition, one necessary property of the scaling filter is that S_p must divide H_p denoted by

$$S_p \mid H_p.$$

where S_p is the p-th order regularity polynomial defined above. That is, the scaling filter must have at least one regularity spectral factor[9, 15]. This characteristic is discussed in the next paragraph. The corresponding wavelets can be found, provided one has $G_p^{(j)}$, by substituting the expression for $\hat{\phi}$ above into Equation 2.6 and changing variables to get

$$\hat{\psi}^{(j)}(\xi) = \hat{\phi}(0) G_p^{(j)}(\xi/p) \prod_{k=2}^{\infty} H_p(e^{i\xi p^{-k}}). \tag{2.14}$$

We mentioned earlier that we want the scaling functions and wavelets to look relatively "nice" (continuous, differentiable, etc.) which implies they are somewhat regular. Daubechies[14] has shown that a necessary condition for the iterated scaling function in Equation 2.13 to converge to a regular function is that the filter $H_p(e^{i\xi})$ have zeros of sufficiently high degree at $\xi = 2\pi l/p$ for $l \in \{1, 2, \ldots, p-1\}$. This causes the zeroes of one dilation of H_p in Equation 2.13 to attenuate the peaks of the previous dilation. As H_p is successively dilated and multiplied in the infinite product, the high frequency peaks will be attenuated. The greater the degree of the zero, the more attenuation occurs in the high frequencies and the time-domain function is more regular. Consequently, for a regular scaling function, we must have $H_p(e^{i\xi})$ of the form

$$H_p(e^{i\xi}) = (S_p(e^{i\xi}))^m \mathcal{L}(e^{i\xi}) \quad m \geq 1 \tag{2.15}$$

where S_p is the regularity polynomial defined earlier. With that definition, the sequence $h_p(n)$ is formed by m-fold discrete convolutions of the sequence

$$s_p(n) = \begin{cases} p^{-1} & n \in \mathbb{Z}_p \\ 0 & \text{otherwise} \end{cases} \tag{2.16}$$

with the sequence l, where

$$\mathcal{L}(e^{i\xi}) = \sum_n l(n) e^{-i\xi n}. \tag{2.17}$$

The most significant contribution of Daubechies' work in [14] is the development for $p = 2$ of a set of scaling functions (and thus wavelets via Equation 2.14) which are compactly supported and have an arbitrary degree of regularity. Equivalently, an approximation filter is found which has the form of Equation 2.15 and also satisfies the conditions of an MRA. If the sequence l is finite, then $h_p(n)$ will be finite and the scaling function generated by iterating $H_p(e^{i\xi})$ will be compactly supported[14].

2.3 Spline-Based Scaling Functions

This section describes a class of scaling functions whose importance will become evident in subsequent chapters. The m-th order cardinal B-spline, N_m, is defined recursively by:

$$N_1 = \chi_{[0,1)}$$
$$N_m = N_{m-1} * N_1 = \int_0^1 N_{m-1}(\cdot - t)dt.$$

It is the m-fold convolution of N_1, the characteristic function on $[0, 1)$. The cardinal B-splines have many useful properties which facilitate their manipulation (see [11]). There are two that are the most important in the context of this dissertation.

First, N_m satisfies a dilation equation for every $p \in \mathbb{N}$. That is, given a dilation factor p, we have

$$N_m(t) = \sum_{k=0}^{m(p-1)} \mu_p(k) N_m(tp - k),$$

where μ_p is the sequence constructed by the m-fold discrete convolution of s_p, the regularity sequence defined earlier in Equation 2.16 . Note that, with the exception of N_1, the integer translates of N_m do not form an orthogonal set.

The second property is related to this lack of orthogonality. In order to "orthogonalize", it is necessary that the Euler-Frobenius polynomial, E_{N_m}, generated by N_m,

$$E_{N_m}(e^{i\xi}) = \sum_{k \in \mathbf{Z}} |\widehat{N_m}(\xi + 2\pi k)|^2 = \sum_{k=1-m}^{m-1} N_{2m}(m+k)e^{ik\xi}$$

not vanish for $\xi \in [0, 2\pi)$. Chui[11] proves that, indeed, this is the case for the cardinal B-splines and provides the following bounds:

$$0 < \sum_{k \in \mathbf{Z}} |\widehat{N_m}(\pi + 2\pi k)|^2 \leq \sum_{k \in \mathbf{Z}} |\widehat{N_m}(\xi + 2\pi k)|^2 \leq 1.$$

This implies the cardinal B-splines can always be "orthogonalized" which will become important later.

While orthogonalizing the B-spline is important, it is more important algorithmically to have an expression for the corresponding orthonormal scaling filter H_p. Let ϕ be the orthonormalized version of N_m for an arbitrary $m \in \mathbf{N}$:

$$\hat{\phi}(\xi) = \frac{\widehat{N_m}(\xi)}{\sum_k |\widehat{N_m}(\xi + 2\pi k)|^2}.$$

The scaling filter H_p must satisfy

$$H_p(e^{i\xi}) = \frac{p^{1/2}\hat{\phi}(p\xi)}{\hat{\phi}(\xi)}.$$

Substituting for $\hat{\phi}$ in terms of the $\widehat{N_m}$,

$$H_p(e^{i\xi}) = \frac{p^{1/2}\widehat{N_m}(p\xi)}{\widehat{N_m}(\xi)} \left(\frac{\sum_k |\widehat{N_m}(\xi + 2\pi k)|^2}{\sum_k |\widehat{N_m}(p\xi + 2\pi k)|^2} \right)^{1/2}$$

$$= M(e^{i\xi}) \left(\frac{E_{N_m}(e^{i\xi})}{E_{N_m}(e^{ip\xi})} \right)^{1/2},$$

where $M(e^{i\xi}) = \sum_k \mu_p(k) e^{-ik\xi}$. In the more general case where

$$p^{1/2} \hat{f}(p\xi) = F(e^{i\xi}) \widehat{N_m},$$

the orthonormalized version of F is given by

$$F^o(e^{i\xi}) = F(e^{i\xi}) \left(\frac{E_{N_m}(e^{i\xi})}{(D_p(|F|^2 E_{N_m})(e^{ip\xi})} \right)^{1/2},$$

where D_p is the downsample operator discussed earlier. The more general form will be useful in calculating orthogonal detail filters which will be discussed later.

2.4 The Rational Resolution Analysis

The rational resolution analysis (RRA), the main topic of this dissertation, is a relatively new concept and has little or no literature currently available[1, 2]. However, there are two notable items which are related to the research at hand and it is appropriate to mention them here.

The first is the work of Pascal Auscher in both [4] and [5]. Auscher has shown that multiresolution analyses with rational dilation factors are possible and that the corresponding wavelets form an orthonormal basis for $L_2(\mathbb{R})$. However, the wavelets are neither compactly supported nor do they have exponential decay. Non-compactly supported wavelets are still useful, but without sufficient (exponential) decay, their usefulness diminishes. Sufficient decay is necessary for good localization in time (or position). Without good localization, the wavelets' usefulness in analyzing non-stationary signals is limited. This is the main reason for the development of wavelets.

The second item is not as important as the first but has some relevance. Kovačević and Vetterli have done some work[22, 23, 24] on perfect reconstruction filter banks which have rational sampling rate changes. Their work is significant in that the RRA developed in this section involves rational sampling rate changes, but their work does not deal with additional constraints imposed by multiresolution/rational-resolution framework.

2.4.1 Definitions and Preliminaries. The definition of the rational resolution analysis (RRA) is similar to MRA with the exception that the requirement for embedded approximation spaces is relaxed. Relaxing other requirements of MRA definition lead to other constructions. For instance, if the integer translates of the scaling function ϕ only form a frame[38] for the approximation space V_0, Benedetto and Li[7] define the frame multiresolution analysis (FMRA).

Throughout the remainder of this dissertation, let r be a rational number defined by $r = p/q$ where $p, q \in \mathrm{N}$ with $p > q$. Furthermore, assume that p and q are relatively prime denoted by $\gcd(p, q) = 1$.

Definition 2.3. The rational resolution analysis (RRA) of $L_2(\mathbb{R})$ (with dilation factor r, described above) is a sequence $\{V_k\}_{k\in\mathbf{Z}}$ of closed subspaces of $L_2(\mathbb{R})$ such that the following hold:

1. The spaces are ordered: $V_k < V_{k-1}$ $k \in \mathbb{Z}$ where the ordering is defined by

$$V_{k_1} < V_{k_2} \Longleftrightarrow k_1 > k_2.$$

2. The approximation spaces satisfy

$$\bigcap_{k\in\mathbf{Z}} V_k = \{0\} \quad \text{and} \quad \overline{\bigcup_{k\in\mathbf{Z}} V_k} = L_2(\mathbb{R}).$$

3. $$f \in V_k \Longleftrightarrow f(r\cdot) \in V_{k-1} \quad \forall\, k \in \mathbb{Z}$$

4. $$f \in V_k \Longleftrightarrow f(\cdot - r^{-k}n) \in V_k \quad \forall\, k, n \in \mathbb{Z}$$

Furthermore, there exists a scaling function $\phi \in L_2(\mathbb{R})$ such that $\{\phi_{k,n}\}_{n\in\mathbb{Z}}$ forms an unconditional (Riesz) basis for V_k where $\phi_{k,n}(t) = r^{-k/2}\phi(r^{-k}t - n)$. ■

The definition is very similar to the MRA. As mentioned above, the significant difference is that the approximation spaces have a different ordering from the ordering ased upon reversed set containment. The approximation spaces of an RRA are sometimes embedded, however. Since the integers are a subset of the rational numbers, the RRA becomes an MRA when $q = 1$ and the approximation spaces are embedded. In this sense, the RRA generalizes the MRA.

2.4.2 Rational Resolution Approximation Operators. Let f_{k-1} be an arbitrary element of V_{k-1}. The orthogonal projection of f_{k-1} onto V_k is given by

$$\begin{aligned} P_{V_k} f_{k-1} &= \sum_{n\in\mathbf{Z}} \langle f_{k-1}, \phi_{k,n} \rangle \phi_{k,n} \\ &= \sum_{n,l\in\mathbf{Z}}\sum c_{k-1,l} \langle \phi_{k-1,l}, \phi_{k,n} \rangle \phi_{k,n} \end{aligned}$$

where the last step involves writing $f_{k-1} = \sum_n c_{k-1,n}\phi_{k-1,n}$ since $f_{k-1} \in V_{k-1}$. To analyze the inner product in the preceeding expression, consider the following derivation of its conjugate

$$\begin{aligned} \overline{\langle \phi_{k-1,n}, \phi_{k,l} \rangle} &= \langle \phi_{k,l}, \phi_{k-1,n} \rangle \\ &= \langle \phi_{1,l}, \phi_{0,n} \rangle \\ &= \frac{1}{2\pi} \langle \widehat{\phi_{1,l}}, \widehat{\phi_{0,n}} \rangle \\ &= \frac{r^{1/2}}{2\pi} \int_{\mathbf{R}} \hat{\phi}(r\xi) \overline{\hat{\phi}(\xi)} e^{i\xi(n-rn)} d\xi \\ &= \frac{(pq)^{1/2}}{2\pi} \int_{\mathbf{R}} \hat{\phi}(p\omega) \overline{\hat{\phi}(q\omega)} e^{i\omega(qn-pl)} d\omega \\ &= (\mathcal{F}^{-1} H_r)(qn - pl) \\ &= h_r(qn - pl) \end{aligned} \tag{2.18}$$

where $H_r(\omega) = \sqrt{pq}\hat{\phi}(p\omega)\overline{\hat{\phi}(q\omega)}$. At this point, we know nothing of the periodicity of H_r. In the following section, further requiring the scaling function to satisfy integer dilation equations will impose a great deal of structure to H_r and h_r. The end result is that

$$\langle \phi_{k-1,n}, \phi_{k,l} \rangle = \overline{h_r(qn - pl)}$$

and

$$P_{V_{k+1}} f_k = \sum_{n \in \mathbf{Z}} \sum_{l \in \mathbf{Z}} c_{k,n} \overline{h_r(qn - pl)} \phi_{k+1,l}.$$

Thus, similar to Equation (2.5) the projection from one approximation space to the next can be written as a discrete filter operation on the approximation coefficients

$$c_{k,l} = \sum_n \overline{h_r(qn - pl)} c_{k-1,n}.$$

If $\{c_{k-1,n}\}_{n \in \mathbf{Z}}$ is considered a discrete-time signal, then the operation described above is effectively a rational sampling rate change[31]. The approximation coefficients are upsampled by a factor of q, filtered via h_r, and downsampled by p. The role of the filter is to prevent aliasing.

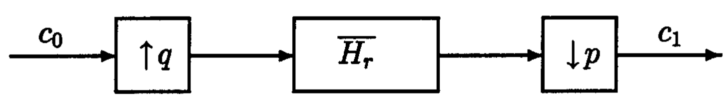

Figure 2.3 Processing Diagram for Rational Resolution Approximations

It is important at this point to mention one of the fundamental differences between the MRA and RRA. With the MRA, the projection from one approximation space to the next is equivalent to the projection from $L_2(\mathbb{R})$ to the particular approximation space. That is, given an arbitrary $f \in L_2(\mathbb{R})$ and V_k a subspace of an MRA

$$f_k = P_{V_k} f.$$

This is a consequence of $V_k \subset V_{k'}$ for $k > k'$. With the RRA, we have the recursive definition

$$f_k = P_{V_k} f_{k-1}.$$

This definition relies on property (2) of the RRA definition. Given an arbitrary $f \in L_2(\mathbb{R})$, there exists a sequence of functions $f_k \in V_k$ which converges to f. That is, there exists a $K(\epsilon) \in \mathbb{Z}$ sufficiently negative such that

$$\|f - P_{V_K} f\| < \epsilon$$

for any arbitrarily small ϵ. This fact is given a more thorough treatment in §5.3.2. of [15]. Daubechies' proof does not depend on the dilation factor and only requires the scaling function satisfy a Riesz condition. Thus, since we assumed the scaling function provides an orthonormal basis so that the Riesz conditions are satisfied, her analysis is directly extendable to the RRA and provides justification to the Equation 2.4.2 above.

This difference between the RRA and MRA is generally not a problem in practice since the discrete-time filter implementation of the RRA projection is defined between adjacent approximation spaces just like the MRA. However, in order to have $f_k = P_{V_k} f$ at each level for the RRA, we must abandon the fast filter implementation of the projection operator and calculate the inner products analytically for each f. This is because the approximation spaces are not embedded and hence $P_{V_k} P_{V_{k-1}} f \neq P_{V_k} f$ in general. (This characteristic causes some difficulties when looking at the frame properties of the RRA which will be discussed in Chapter V.) However, in practical uses for which f is assume to live in one of the V_k, fast algorithm exist which are analogous to those in the MRA case.

2.4.3 Rational Resolution Reconstruction. Reconstruction in the RRA is more complicated due to the approximation spaces generally not being embedded. Recall, with the MRA, an element of a particular approximation space can be written as an orthogonal decomposition of its projection onto the next lower approximation space and the corresponding detail space. This is generally not possible with the RRA. Consider two adjacent approximation spaces V_0 and V_1 and assume the set $D = V_1 \backslash V_0$ is non-empty. Then there is no orthogonal detail space that would sum with V_1 to form V_0. The elements of D could not be "cancelled" by any element in an orthogonal detail space. This suggests a processing scheme on V_1 which eliminates the elements in D and combines with some orthogonal detail space to form V_0. This is the approach taken in this research.

Define the sub-approximation spaces V_k' by

$$\begin{aligned} V_k' &= \overline{\operatorname{span}\{p^{-1/2}\phi_{k-1,n}(\cdot/p)\}_{n\in\mathbf{Z}}} \\ &= \overline{\operatorname{span}\{q^{-1/2}\phi_{k,n}(\cdot/q)\}_{n\in\mathbf{Z}}}. \end{aligned}$$

The normalization constants above are not necessary for the definition of V_k', but are used to normalize a Riesz basis for the space. If ϕ satisfies a dilation equation based on p, then we have $V_k' \subset V_{k-1}$ in exact analogy to the MRA. Furthermore, we have a detail space which the orthogonal complement of V_k' in V_{k-1}. Thus, perfect reconstruction is suggested via MRA techniques described earlier. This is illustrated in Figure 2.4.3.

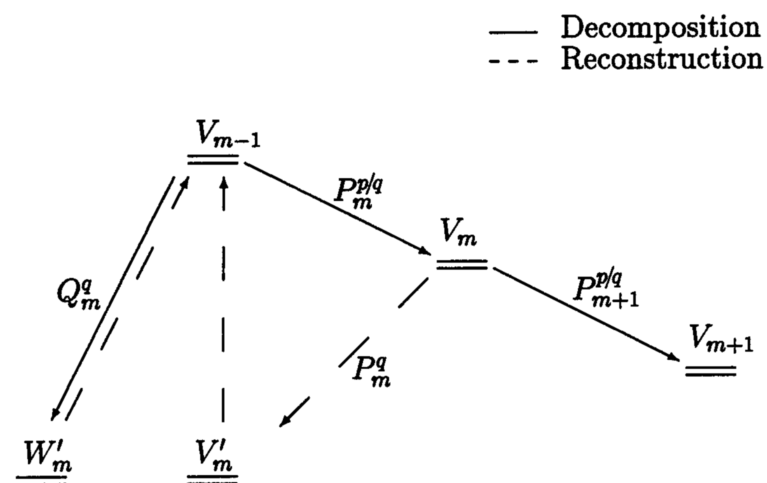

Figure 2.4 Illustration of the p/q-Dilation Rational Resolution Analysis

In order to ensure this scheme will work, it is necessary to have

$$f_k' = P_{V_k'} f_{k-1} = P_{V_k'} P_{V_k} f_{k-1}, \quad \forall\, f_{k-1} \in V_{k-1}. \tag{2.19}$$

Some implications of this condition are presented in the following theorem.

Theorem 2.4. *Let V_k and V'_k be as defined previously for some rational dilation factor $r = p/q$. Then*

$$P_{V'_k} f_{k-1} = P_{V'_k} P_{V_k} f_{k-1}, \quad \forall\, f_{k-1} \in V_{k-1}$$

is equivalent to $V'_k \subset V_k$.

Proof. Let f_{k-1} be chosen arbitrarily in V_{k-1}. Suppose Equation 2.19 is valid. Without loss of generality, let $k = 1$. When $\{\phi_{k,n}\}_{n\in\mathbf{Z}}$ and $\{\phi'_{k,n}\}_{n\in\mathbf{Z}}$ are Riesz bases for V_k and V'_k, respectively, the following analysis holds:

$$\begin{aligned}
P_{V'_1} P_{V_1} f_0 &= P_{V'_1} \sum_n \langle f_0, \phi_{1,n} \rangle \phi_{1,n} \\
&= \sum_l \sum_n \langle f_0, \phi_{1,n} \rangle \langle \phi_{1,n}, \phi'_{1,l} \rangle \phi'_{1,l} \\
&= \sum_l \sum_n \left\langle f_0, \overline{\langle \phi_{1,n}, \phi'_{1,l} \rangle} \phi_{1,n} \right\rangle \phi'_{1,l} \\
&= \sum_l \sum_n \left\langle f_0, \langle \phi'_{1,l}, \phi_{1,n} \rangle \phi_{1,n} \right\rangle \phi'_{1,l} \\
&= \sum_l \left\langle f_0, \sum_n \langle \phi'_{1,l}, \phi_{1,n} \rangle \phi_{1,n} \right\rangle \phi'_{1,l}.
\end{aligned}$$

However,

$$P_{V'_1} f_0 = \sum_l \langle f_0, \phi'_{1,l} \rangle \phi'_{1,l}$$

which implies

$$\phi'_{1,l} = \sum_n \langle \phi'_{1,l}, \phi_{1,n} \rangle \phi_{1,n}$$

since $f_0 \in V_0$ arbitrarily. From this equation $\phi'_{1,l}$ is a linear combination of the $\phi_{1,n}$. It follows immediately that this is equivalent to having $V'_1 \subset V_1$.

Assuming $V'_1 \subset V_1$, and hence

$$\phi'_{1,l} = \sum_n \langle \phi'_{1,l}, \phi_{1,n} \rangle \phi_{1,n}, \tag{2.20}$$

the derivation can be reversed to show Equation 2.19 holds.

■

The perfect reconstruction scheme assumes the scaling function satisfies a p dilation equation and Theorem 2.4. indicates that it must also satisfy a q dilation equation (via Equation 2.20). Thus, for the remainder of this chapter, we will concern ourselves only with those scaling functions which satisfy the dual dilation system

$$\phi(\cdot - k) = \begin{cases} \sqrt{p}\sum_n h_p(n-pk)\phi(p\cdot -n) \\ \sqrt{q}\sum_n h_q(n-qk)\phi(q\cdot -n) \end{cases}, \tag{2.21}$$

where the h_p and h_q are defined below.

With such a scaling function, the reconstruction scheme is as follows. Given an arbitrary function $f_{k-1} \in V_{k-1}$, then

$$f_{k-1} = P_{V_k'} f_k + \sum_{j=1}^{p-1} Q_k^{(j)} f_{k-1} \tag{2.22}$$

where the $Q^{(j)}$ are the orthogonal projectors onto the $p-1$ detail spaces associated with a p-dilation MRA partition of V_{k-1}.

This scaling function assumption allows the orthogonal projection P_{V_k} to be written in more detail. As with the MRA, define the following sequences:

$$h_p(n) = \langle \phi'_{k,0}, \phi_{k-1,n}\rangle \quad n \in \mathbb{Z}$$
$$h_q(n) = \langle \phi'_{k,0}, \phi_{k,n}\rangle \quad n \in \mathbb{Z}.$$

Then, h_r defined in (2.18) can be explicitly written in terms of h_p and h_q via the following analysis. Define V''_{k-1} by

$$\begin{aligned} V''_{k-1} &= \overline{\operatorname{span}\{q^{1/2}\phi_{k-1,n}(q\cdot)\}_{n\in\mathbb{Z}}} \\ &= \overline{\operatorname{span}\{p^{1/2}\phi_{k,n}(p\cdot)\}_{n\in\mathbb{Z}}}. \end{aligned}$$

Since ϕ satisfies the pair of dilation equations, we have $V_{k-1} \subset V_k''$ and $V_k \subset V_{k-1}''$. Let ϕ_{k-1}'' denote the orthonormal scaling function associated with V_{k-1}''. Then,

$$\begin{aligned}
\langle \phi_{k,l}, \phi_{k-1,n} \rangle &= \left\langle \sum_j h_p(j-pl)\phi_{0,j}'', \sum_i h_q(i-qn)\phi_{0,i}'' \right\rangle \\
&= \sum_j \sum_i h_p(j-pl)\overline{h_q(i-qn)}\langle \phi_{0,j}'', \phi_{0,i}'' \rangle \\
&= \sum_j \sum_i h_p(j-pl)\overline{h_q(i-qn)}\delta_{j,i} \\
&= \sum_j h_p(j-pl)\overline{h_q(j-qn)} \\
&= h_r(qn-pl).
\end{aligned}$$

Recall that we defined h_r previously in the same way. With the requirement that ϕ satisfy dilation equations based on p and q, the composition of h_r is made explicit. In the Fourier domain, we use Equation 2.18 to write

$$\begin{aligned}
h_r(qn-pl) &= \frac{(pq)^{1/2}}{2\pi} \int_{\mathbb{R}} \hat{\phi}(p\xi)\overline{\hat{\phi}(q\xi)} e^{i\xi(qn-pl)} d\xi \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} H_p(e^{i\xi})\overline{H_q(e^{i\xi})}|\hat{\phi}(\xi)|^2 e^{i\xi(qn-pl)} d\xi \\
&= \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} \int_0^{2\pi} H_p(e^{i\xi})\overline{H_q(e^{i\xi})}|\hat{\phi}(\xi+2\pi k)|^2 e^{i\xi(qn-pl)} d\xi \\
&= \frac{1}{2\pi} \int_0^{2\pi} H_p(e^{i\xi})\overline{H_q(e^{i\xi})} \sum_{k\in\mathbb{Z}} |\hat{\phi}(\xi+2\pi k)|^2 e^{i\xi(qn-pl)} d\xi \\
&= \frac{1}{2\pi} \int_0^{2\pi} H_p(e^{i\xi})\overline{H_q(e^{i\xi})} e^{i\xi(qn-pl)} d\xi
\end{aligned}$$

which implies $H_r = H_p\overline{H_q}$ is the discrete Fourier transform of h_r. The last step in the analysis is due to the orthogonality of the integer translates of ϕ

$$\langle \phi_{k,n}, \phi_{k,m} \rangle = \delta_{n,m} \Longleftrightarrow \sum_l |\hat{\phi}(\xi+2\pi l)|^2 = 1 \quad \text{a.e. } \xi \in \mathbb{R}$$

This definition of H_r is a slight abuse of notation since we have previously defined $H_r(\omega) = \sqrt{pq}\hat{\phi}(p\omega)\overline{\hat{\phi}(q\omega)}$. Via the analysis above, this current definition of H_r can be seen as the

2π-periodization of the previous definition. The two definitions are not inconsistent with one another in the sense that we evaluate the inverse Fourier transform of H_r only at the integers to get h_r. In the previous case h_r is a continuous function sampled only at the integers and in the current case, h_r is a sequence whose domain is the integers. The integer values are the same. The difference comes about from our assumption in the latter case that ϕ satisfies 2-scale equations for both dilations p and q so that we have scaling sequences h_p and h_q. The original development made no such assumption.

2.5 Summary

This chapter has provided a mathematical launching point for the following chapters. The RRA was rigorously defined and the various projection operators in terms of discrete filters were described. Necessary and sufficient conditions for perfect reconstruction within the RRA framework were presented. All this description was couched within the notation of existing multiresolution paradigms. In fact, the multiresolution analysis fits within the RRA framework as a special case. This is expected since the integers are a subset of the rational numbers.

The next chapter studies scaling functions and wavelets associated with a perfect-reconstruction RRA. It further examines the existence and characterization of such scaling functions and wavelets which are compactly supported. The results are surprising.

III. Compactly Supported Orthonormal RRA Scaling Functions

3.1 Dual Dilation Symbol Requirement

This chapter presents an analysis of the existence and characterization of compactly supported, orthonormal scaling functions which yield perfect recontruction in an RRA. The analysis centers on the dual dilation requirement described briefly in the previous chapter and further discussed below.

The synthesis scheme presented in the last chapter relies on the scaling function satisfying a pair of dilation equations

$$\phi(t) = \begin{cases} p^{1/2} \sum_n h_p(n)\phi(pt-n) \\ q^{1/2} \sum_n h_q(n)\phi(qt-n) \end{cases} \qquad t \in \mathbb{R}, \tag{3.1a}$$

or in the Fourier domain,

$$\hat{\phi}(pq\xi) = \begin{cases} p^{-1/2} H_p(e^{iq\xi})\hat{\phi}(q\xi) \\ q^{-1/2} H_q(e^{ip\xi})\hat{\phi}(p\xi) \end{cases} \qquad \xi \in \mathbb{R}, \tag{3.1b}$$

where $H_p(z) = \sum_n h_p(n)z^{-n}$. The two Fourier domain expressions can be substituted into each other to yield

$$\hat{\phi}(pq\xi) = \begin{cases} (pq)^{-1/2} H_p(e^{iq\xi})H_q(e^{i\xi})\hat{\phi}(\xi) \\ (pq)^{-1/2} H_q(e^{ip\xi})H_p(e^{i\xi})\hat{\phi}(\xi) \end{cases} \qquad \xi \in \mathbb{R}.$$

which implies

$$H_p(z^q)H_q(z) = H_q(z^p)H_p(z) \qquad |z| = 1. \tag{3.2}$$

This dual symbol dilation requirement implies several interesting properties of the associated scaling function. This remaining sections of this chapter are primarily focused on the requirement and how it affects the existence of orthogonal and/or compactly supported scaling and wavelet functions.

3.2 Root Loci

In a separable Hilbert space, two bases $\{\varphi_j\}_{j\in\mathbb{Z}}$ and $\{\tilde{\varphi}_k\}_{k\in\mathbb{Z}}$ are duals if they satisfy

$$\langle \varphi_j, \tilde{\varphi}_k \rangle = \delta_{j,k} \quad \forall\, j,k \in \mathbb{Z}. \tag{3.3}$$

Orthonormal scaling functions and wavelets are desirable in resolution analyses since they are self-dual[11] (or self-reciprocal[21]) in the sense that that $\{\varphi_j\} = \{\tilde{\varphi}_k\}$ so that $\langle \varphi_j, \varphi_k \rangle = \delta_{j,k}$. This means the analysis and synthesis kernels are identical so that one filter works for both operations. There has been significant work in the area of biorthogonal wavelets[13, 34] in which the analyzing and synthesizing are not identical but merely duals of one another, i.e. one filter is used for analysis, and one for synthesis. For the most part, this research is concerned solely with orthonormal wavelets, but biorthogonal scaling functions and wavelets, and duality in general, will be mentioned briefly in later chapters.

Scaling functions and wavelets with compact support are also desirable in resolution analyses since they correspond to finite scaling and detail dilation sequences (h_p and $g_p^{(j)}$) as well[14]. In signal processing terms, this is equivalent to filters with finite impulse response (FIR) which allow for fast, efficient implementations.

In this section, two theorems are presented which provide a good foundation for the characterization of compactly supported scaling functions which allow perfect reconstruction with the RRA. The first theorem demonstrates that any two polynomials which satisfy a dual dilation condition (3.5) must necessarily have their roots on the unit circle. The second theorem builds upon the first by explicitly characterizing the arrangement of the roots of such polynomials. These two theorems have some specific implications in regard to compactly supported RRA scaling functions. Since the scaling symbols H_p and H_q completely characterize the scaling functions, the requirement stated in Equation 3.2 will be used as the basis for the analysis.

Given a scaling function, ϕ, with scaling filter H_p, orthonormality of the scaling function translates into the following condition on the scaling coefficients, h_p

$$\sum_k h_p(k)\overline{h_p(k+pn)} = \delta_{n,0} \quad n \in \mathbb{Z} \tag{3.4a}$$

which, when written in the Fourier domain, becomes

$$(D_p(|H_p|^2))(e^{i\xi}) = 1 \quad \text{a.a } \xi \in \mathbb{R}. \tag{3.4b}$$

The following definitions will be useful in subsequent lemmas and theorems.

Definition 3.1. An *R-set* of z with respect to p, denoted $R_p(z)$ is the set of the p-th roots of z. Writing $z = re^{i\theta}$,

$$R_p(z) = \{r^{1/p}e^{i(\theta+2\pi l)/p}\}_{l\in\mathbb{Z}_p}.$$

The term R-set will also be used to describe any set of complex numbers which share the same magnitude and have arguments uniformly spaced in $\mathbb{R}$ (modulo 2π). ■

Define A_z by

$$A_z = R_{pq}(z^{pq}).$$

With a slight abuse of notation, let

$$A_z^p = R_q(z^{pq})$$
$$A_z^q = R_p(z^{pq}).$$

Define $B_z = A_z \cup A_z^p \cup A_z^q$. The set B_z can be given practical meaning by recognizing it as the collection of all p-th, q-th, and pq-th roots of the complex number z^{pq}. It is used in the following lemma.

Lemma 3.2. *Given two polynomials P and Q which satisfy*

$$P(z^q)Q(z) = P(z)Q(z^p) \quad \forall\, z \in \mathbb{C}, \tag{3.5}$$

define $f(z) = P(z)Q(z^p) = P(z^q)Q(z)$. Define $R \subset \mathcal{Z}(f)$ by

$$R = \{z \in \mathcal{Z}(f) : |z| > 1\}.$$

If R is non-empty, and a minimum root α is well-defined such that $|z| \geq |\alpha|$, for all $z \in R$, then $f(B_\alpha) = \{0\}$.

Proof. For notational convenience, define the polynomials P^q and Q^p by

$$P^q(z) = P(z^q) \quad \forall\, z \in \mathbb{C}$$
$$Q^p(z) = Q(z^p) \quad \forall\, z \in \mathbb{C}.$$

Then $f = P^qQ = PQ^p$. First of all, $f(\alpha) = 0 \Longrightarrow P(\alpha)Q(\alpha^p) = 0$. This implies either $P(\alpha) = 0$ or $Q(\alpha^p) = 0$. Consider the first possibility:

$$\begin{aligned} P(\alpha) = 0 &\Longrightarrow P^q(r_l^q(\alpha)) = 0 \quad l \in \mathbb{Z}_q \\ &\Longrightarrow \{r_l^q(\alpha)\}_{l\in\mathbb{Z}_q} \subset R. \end{aligned}$$

This implies $\alpha^{1/q}$ is also a root of f which contradicts the assumption that α was the smallest root in R since $1 < |\alpha^{1/q}| < |\alpha|$. Thus, $Q^p(\alpha) = 0$ and by a similar argument using the other representation of f, $P^q(\alpha) = 0$.

With $P(\alpha^q) = Q(\alpha^p) = 0$, the following holds:

$$\begin{aligned} P^q(\alpha) = 0 &\Longrightarrow P^q(\alpha e^{i2\pi l/q}) = 0 \quad \forall\, l \in \mathbb{Z}_q, \\ &\Longrightarrow \{\alpha e^{i2\pi l/q}\}_{l\in\mathbb{Z}_q} \subset R. \end{aligned}$$

Now, each root in this R-set of roots is also a root of Q^p by the preceeding argument. Hence,

$$\begin{aligned}
P^q(\alpha e^{i2\pi l/q}) = 0 &\Longrightarrow Q^p(\alpha e^{i2\pi l/q}) = 0, \quad l \in \mathbb{Z}_q. \\
&\Longrightarrow Q^p(\alpha e^{i2\pi l/q} e^{i2\pi k/p}) = 0, \quad l \in \mathbb{Z}_q, k \in \mathbb{Z}_p. \\
&\Longrightarrow Q^p(\alpha e^{\frac{i2\pi(pl+qk)}{pq}}) = 0, \quad l \in \mathbb{Z}_q, k \in \mathbb{Z}_p. \\
&\Longrightarrow Q^p(\alpha e^{\frac{i2\pi l}{pq}}) = 0, \quad l \in \mathbb{Z}_{pq}. \\
&\Longrightarrow \{\alpha e^{\frac{i2\pi l}{pq}}\}_{l \in \mathbb{Z}_{pq}} \subset R. \\
&\Longrightarrow A_\alpha \subset R. \\
&\Longrightarrow f(A_\alpha) = \{0\}
\end{aligned}$$

The third and fourth above steps use the fact that p and q are relatively prime and an application of Lemma A.1. Continuing,

$$\begin{aligned}
P^q(A_\alpha) = \{0\} &\Longrightarrow P(A_\alpha^q) = \{0\} \\
&\Longrightarrow A_\alpha^q \subset R
\end{aligned}$$

and

$$\begin{aligned}
Q^p(A_\alpha) = \{0\} &\Longrightarrow Q(A_\alpha^p) = \{0\} \\
&\Longrightarrow A_\alpha^p \subset R
\end{aligned}$$

which implies $B_\alpha \subset R$ since $B_\alpha = A_\alpha \cup A_\alpha^p \cup A_\alpha^q$. This leads to $f(B_\alpha) = \{0\}$

■

Corollary 3.3. *If* $R = \{z \in \mathcal{Z}(f) : |z| < 1\}$ *and* R *is non-empty with* α *as a maximum root, then* $f(B_\alpha) = \{0\}$.

Proof. By switching the inequalities, the previous proof holds.

■

Theorem 3.4. *If two polynomials P and Q satisfy (3.5), and the polynomial f is defined as in Lemma 3.2., then $\mathcal{Z}(f) \subset \mathbb{T} \cup \{0\}$.*

Proof. If two polynomials P and Q satisfy (3.5), then $\tilde{P}(z) = z^k P(z)$ and $\tilde{Q}(z) = z^j Q(z)$ satisfy (3.5) as well for $k, j \in \mathbf{N}$ such that

$$\frac{k}{j} = \frac{p-1}{q-1}.$$

This accounts for any roots at the origin. Hence, assume P and Q have non-zero constant terms. To show the roots of the individual polynomials P and Q lie on the unit circle, it is necessary and sufficient to show the roots of f lie on the unit circle. The approach is to assume there exists some finite number of roots outside the unit circle and show this leads to a contradiction. With slight modifications, the same approach is then used to contradict the existence of a finite set of non-zero roots inside the unit circle.

Assume the set R as defined in Lemma 3.2. is non-empty. There are a finite number of roots in R since both P and Q (and hence f) are polynomials. Choose $\alpha \in R$ to satisfy

$$1 < |\alpha| \leq |z| \quad \forall z \in R.$$

That is, choose α so that no other root in R has a smaller magnitude.

The implication of Lemma 3.2. is that α necessarily exists as one of a set of pq roots which share a common magnitude and are equally spaced in phase by $2\pi/pq$. Furthermore, the sets A_α, A_α^p, and A_α^q are disjoint since they contain roots of different magnitudes (when $|\alpha| \neq 1$). Figure 3.1 illustrates a possible set B_α with $p = 3$ and $q = 2$.

Now choose α' to be a root with the smallest magnitude in $R \setminus B_\alpha$. Such a root exists since it can be shown that $B_\alpha \neq R$ as follows.

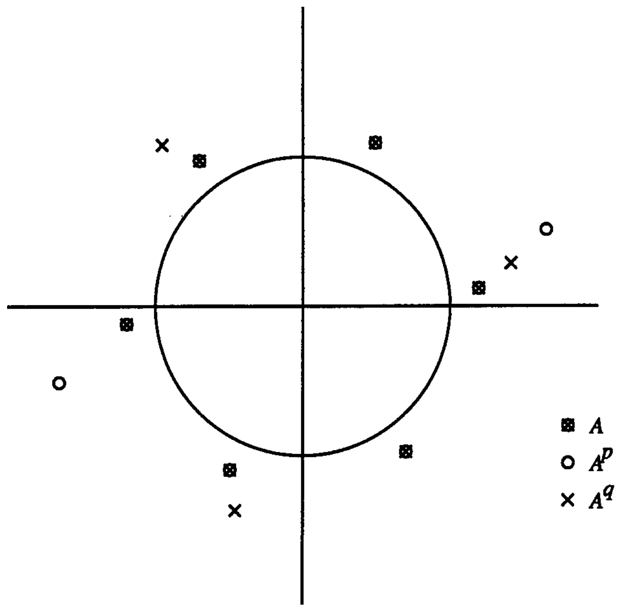

Figure 3.1 A Possible B_α for $p = 3$, $q = 2$, and $\alpha = 1.2\,e^{i(0.1)}$

Existence of α'.

Suppose the opposite. That is, suppose $R = B_\alpha$. Choose $\beta \in A^q_\alpha$ arbitrarily. We have

$$Q(\beta) = 0 \Longrightarrow P(\beta)Q^p(\beta) = 0$$

so that either $P(\beta) = 0$ or $Q^p(\beta) = 0$. First, $P(\beta) = 0$ implies the existence of roots $r^q_l(\beta)$ whose magnitudes are $|\alpha|^{\frac{p}{q}}$ (since $|\beta| = |\alpha|^p$). This is not a possible magnitude for A_α, A^q_α, or A^p_α so that $P(\beta) = 0$ implies there exists another root of f which is not an element of B_α.

The other alternative is that $Q^p(\beta) = 0$ which implies β exists within a set of p roots phase-spaced by $2\pi/p$. But the roots in A^q_α are phase-spaced by $\frac{2\pi}{q}$ which implies there are roots of f which have the same magnitude as those in A^q_α but with different phases. The result in either of these cases is that we contradict our original assumption which implies $R \neq B_\alpha$.

Generation of $B_{\alpha'}$

We show that α' "generates" the sets $A_{\alpha'}$, $A^q_{\alpha'}$, and $A^p_{\alpha'}$ in the same way as α. As before, we start by showing $P^q(\alpha') = Q^p(\alpha') = 0$. This is not as easy as before because we need to consider the possibility of $P(\alpha') = 0$ or $Q(\alpha') = 0$ as well. Recall we concluded $P^q(\alpha) = Q^p(\alpha) = 0$ by showing that $Q(\alpha) = 0$ or $P(\alpha) = 0$ implies the existence of roots whose magnitudes were less than α which was a contradiction. In this case, we know there may exist roots whose magnitudes are less than α' since $\alpha' \in R \setminus B_\alpha$. These roots would be in B_α.

Suppose we have $Q(\alpha') = 0$. Then, as before, we also have $\{r^p_l(\alpha')\}_{l\in\mathbb{Z}^p} \subset R$ which are the roots of Q^p. These roots have magnitudes less than $|\alpha'|$ which is an apparent contradiction unless we have $\{r^p_l(\alpha')\}_{l\in\mathbb{Z}_p} \subset B_\alpha$. We need to show this is not possible either.

First of all, since A, A^q, and A^p are disjoint, we can consider them independently. We can immediately rule out the possibility that $\{r^p_l(\alpha')\}_{l\in\mathbb{Z}_p} = A^q$ since the sets have inconsistent phase-spacings ($2\pi/p$ versus $2\pi/q$) and different cardinalities (p vs. q).

So, consider first the possibility that $\{r^p_l(\alpha')\}_{l\in\mathbb{Z}_p} = A^p$. Then we have $P(r^p_l(\alpha')) = 0$ for $l \in \mathbb{Z}_p$. Looking at the other side of the equation, it follows:

$$P^q(r^p_l(\alpha'))Q(r^p_l(\alpha')) = 0$$

for $l \in \mathbb{Z}^p$. But $Q(r^p_l(\alpha')) \neq 0$ since that would imply roots of Q^p with magnitude $|\alpha'|^{\frac{1}{p^2}} < |\alpha'|$ and these roots are clearly not in B_α since they have different magnitudes. This contradicts the assumption that α' was chosen with the smallest magnitude in $R \setminus B$.

Now consider the possibility that $P^q(r^p_l(\alpha')) = 0$.

$$P^q(r^p_l(\alpha')) = 0 \Longrightarrow P^q(r^p_l(\alpha')e^{i2\pi k/q}) = 0 \quad l \in \mathbb{Z}_p, k \in \mathbb{Z}_q \tag{3.6}$$

$$\Longrightarrow \{|\alpha'|^{\frac{1}{p}} e^{\frac{i\arg(\theta')}{p}} e^{\frac{i2\pi l}{pq}}\}_{l\in\mathbb{Z}_{pq}} \subset R \tag{3.7}$$

Notice we have pq roots with magnitude $|\alpha'|^{\frac{1}{p}}$. Some of these roots will not be in A^p since they have phase spacings of $\frac{2\pi}{q}$. This also contradicts the assumption that α' was chosen with the smallest magnitude in $R \setminus B$. Hence, we conclude

$$P^q(r_l^p(\alpha'))Q(r_l^p(\alpha')) \neq 0 \quad l \in \mathbb{Z}_p$$

which implies

$$P(r_l^p(\alpha')) \neq 0 \quad l \in \mathbb{Z}_p.$$

The result is $\{r_l^p(\alpha')\}_{l\in\mathbb{Z}_p} \neq A^p$. We also have $\{r_l^p(\alpha')\}_{l\in\mathbb{Z}_p} \not\subset A$ since that would imply that $\alpha' \in A^p$, a contradiction in the assumption of how it was chosen.

The end result is that $Q(\alpha') \neq 0$ which forces $P^q(\alpha') = 0$. Using a nearly identical argument (switching the roles of p and q) we also get $Q^p(\alpha') = 0$. The generation of $A_{\alpha'}$, $A^p_{\alpha'}$, and $A^q_{\alpha'}$ directly follow from these two results.

Comparing B_α and B'_α.

To show the number of roots "generated" by $\alpha, \alpha', \alpha'', \ldots$ increases without bound, consider the intersection between B_α and $B_{\alpha'}$. There are three cases to consider.

Case 1: $|\alpha| = |\alpha'|$

Since α' is chosen from $R \setminus B_\alpha$, we can assume without loss of generality that

$$\alpha' = \alpha e^{i\epsilon}$$

for some $\epsilon \in (0, \frac{2\pi}{pq})$. There will be no intersection between A and A' since one set is an ϵ-phase shifted version of the other and $\epsilon < \frac{2\pi}{pq}$. Furthermore:

$$\begin{aligned} A_{\alpha'} &= \{\alpha' e^{\frac{i2\pi l}{pq}}\}_{l\in\mathbb{Z}_{pq}} \\ &= \{\alpha e^{i(\frac{2\pi l}{pq}+\epsilon)}\}_{l\in\mathbb{Z}_{pq}} \end{aligned}$$

which implies

$$A^p_{\alpha'} = \{\alpha^p e^{i(\frac{i2\pi l}{q}+p\epsilon)}\}_{l\in \mathbf{Z}_q}$$
$$A^q_{\alpha'} = \{\alpha^q e^{i(\frac{2\pi l}{p}+q\epsilon)}\}_{l\in \mathbf{Z}_p}.$$

A non-null intersection occurs between A^q_α and $A^q_{\alpha'}$ only when $q\epsilon = 2\pi/p$ or $\epsilon = 2\pi/(pq)$. This is a contradiction since $\epsilon < 2\pi/(pq)$. The same is true for the intersection between A^p_α and $A^p_{\alpha'}$. Hence, for $|\alpha| = |\alpha'|$, B_α and $B_{\alpha'}$ are disjoint.

Case 2: $|\alpha'| = |\alpha|^p$ **or** $|\alpha'| = |\alpha|^q$

Consider the case where $|\alpha|^p = |\alpha'|$. There can be at most q intersections between A^q_α and $A_{\alpha'}$ since A^q_α only has q elements and they are phase-consistent with elements in $A_{\alpha'}$. In the other case, $|\alpha|^q = |\alpha'|$, there will be at most p intersections between A^p_α and $A_{\alpha'}$ for the same reasons. Consequently, $B_{\alpha'}$ will contain either $pq+p$ or $pq+q$ roots distinct from B_α.

Case 3: $|\alpha|^p = |\alpha'|^q$

For this case, the intersection between A^q and $A^p_{\alpha'}$ can contain at most one element since the phase spacings of the two sets are not equal and, more importantly, the spacings are relatively prime. In other words, the only way for there to be more than one intersection would be for p and q to not be relatively prime. $|A^p| = |A^p_{\alpha'}|$ and $|A^q| = |A^q_{\alpha'}|$ but the pairs of sets will be disjoint because of the initial phase difference between $|\alpha| = |\alpha'|$. The other possibility is that $|\alpha| \neq |\alpha'|$. This immediately leads to the conclusion that B_α and $B_{\alpha'}$ are disjoint. Notice that the case $|\alpha|^q = |\alpha'|^p$ is not considered since, for $p > q$, it immediately implies $|\alpha'| < |\alpha|$. This violates the assumption that α has the smallest magnitude.

Conclusion.

With the assumption $p > q$, each set $B_{\alpha'}, B_{\alpha''}, B_{\alpha'''}, \ldots$ contributes at least $pq+q-1$ unique roots of f. Thus the total number of unique roots grows without bound. But, we know the number of roots is finite, hence the original assumption that there exists a nonempty set R containing roots of f which lie outside the unit circle must be false.

It is now easy to show there are no roots inside the unit circle by choosing α to have the largest magnitude of any possible roots within the unit circle. The argument is nearly identical to the one above except that the the magnitude of sets $B_\alpha, B_{\alpha'}, B_{\alpha''}, \ldots$ approaches zero instead of infinity as with the roots outside the unit circle.

■

The previous theorem indicates that the scaling filters associated with scaling functions satisfying (3.1) must have all their zeros on the unit circle. The arrangement of the roots on the unit circle is addressed in the following theorem.

Definition 3.5. The *modulated regularity function*, $S_p(\cdot;\gamma)$ is defined for $p \in \mathbf{N}$ and $\gamma \in \mathbb{C}$ by

$$S_p(z;\gamma) = \frac{1}{p} \sum_{n \in \mathbf{Z}_p} (z\gamma)^{-n}$$

■

Lemma 3.6. *If* $f = P^q Q = PQ^p = S_{pq}(\cdot;\gamma)$, *then*

$$P = S_p(\cdot;\gamma)$$
$$Q = S_q(\cdot;\gamma)$$
$$\gamma \in \{e^{i2\pi k/d}\}_{k \in \mathbf{Z}_d}$$

where $d = \gcd(p-1, q-1)$.

Proof. Consider $S_{pq}(\cdot;\gamma)$ and the arrangement of its roots as in Figure 3.2. The phase space between rows is $2\pi/q$ and between columns, $2\pi/p$. When the roots are partitioned according to PQ^p, there are $q-1$ R_p-sets of roots, which account for the first $q-1$ rows in Figure 3.2. The remaining $p-1$ roots (the last row) are phase spaced by $2\pi/p$, which includes the phase spacing between the actual roots and the deleted root, which corresponds to $\bar{\gamma}$. Hence, $P = S_p(\cdot\,;\gamma)$. By a similar argument, $Q = S_q(\cdot\,;\gamma)$.

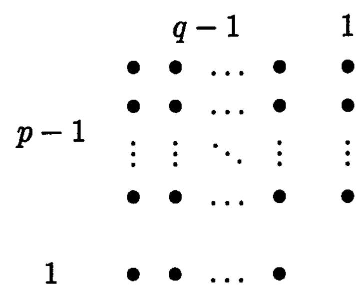

Figure 3.2 Roots of f Arranged on a Grid Showing Phase-Spacings

We need to characterize the possible values of γ. We have

$$pS_p(z;\gamma) = \frac{(z\gamma)^p - 1}{(z\gamma) - 1}$$

which can be substituted into the expression for f to yield:

$$\frac{((z^q\gamma)^p - 1)((z\gamma)^q - 1)}{(z^q\gamma - 1)(z\gamma - 1)} = \frac{((z^p\gamma)^q - 1)((z\gamma)^p - 1)}{(z^p\gamma - 1)(z\gamma - 1)}.$$

After expanding each side of the equation, multiplying to remove the denominator terms, and equating the coefficients of the powers of z, a dual condition arises, namely,

$$\gamma^p = \gamma^q = \gamma.$$

This implies

$$\gamma \in \{e^{i2\pi k/(p-1)}\}_{k \in \mathbb{Z}_{p-1}} \cap \{e^{i2\pi l/(q-1)}\}_{l \in \mathbb{Z}_{q-1}}$$

which leads to

$$\gamma \in \{e^{i2\pi j/d}\}_{j \in \mathbb{Z}_d}$$

where $d = \gcd(p-1, q-1)$.

∎

Theorem 3.7. *With $f = P^q Q = PQ^p$ as before, one (or both) of the following conditions must hold:*

$$\exists\, z \in \mathbb{T} : A_z \subset \mathcal{Z}(f),$$

i.e., $\mathcal{Z}(f)$ contains an R_{pq}-set, or f can be written as

$$f(z) = \prod_{j=1}^{M} S_{pq}(z;\gamma_j) \qquad \gamma_j \in \{2\pi k/d\}_{k\in\mathbb{Z}_d}$$

for some $M \in \mathbf{N}$ with $d = \gcd(p-1, q-1)$.

Proof. From the previous theorem, $\mathcal{Z}(f) \subset \mathbb{T}$. Furthermore, the number of roots of f, denoted $|\mathcal{Z}(f)|$, must satisfy

$$|\mathcal{Z}(f)| = N_p + pN_q = N_q + qN_p$$

where N_p and N_q are the degrees of P and Q, respectively. This implies the degrees of P and Q must satisfy

$$\frac{N_p}{N_q} = \frac{(p-1)}{(q-1)}.$$

Consider the following rational function

$$R(z) = \frac{P(z)}{Q(z)} = \frac{P(z^q)}{Q(z^p)} \quad \forall z \in \mathbb{C}. \tag{3.8}$$

Since there are at most N_p zeros in the above expression, this implies there are at least $N_p(q-1) = N_q(p-1)$ roots common between P^q and Q^p. Define the multiset C by

$$C = \mathcal{Z}(P^q) \cap \mathcal{Z}(Q^p)$$

where the term *multiset* is used to mean a set whose repeated elements are treated as distinct. Thus C consists of the roots which are in $\mathcal{Z}(P^q)$ and $\mathcal{Z}(Q^p)$ simultaneously, along with all their multiplicities. Thus,

$$|C| \geq N_p(q-1) = N_q(p-1).$$

We now consider two cases: $|C| > N_p(q-1)$ and $|C| = N_p(q-1)$.

Case 1: $|C| > N_p(q-1)$

Suppose $|C| > N_p(q-1)$. The roots of P^q are arranged into N_p R_q-sets. That is, each root of P "generates" an R_q-set of roots for P^q. Similarly, the roots of Q^p can be arranged into N_q R_p-sets.

There exists an R_p-set and an R_q-set of roots in C. To see this, define the set $\widetilde{\mathcal{Z}(P^q)} = \mathcal{Z}(P^q) \setminus C$. In order to avoid having an R_p-set of roots in C, $\widetilde{\mathcal{Z}(P^q)}$ must contain at least one root from each of the N_p R_q-sets in $\mathcal{Z}(P^q)$. This implies

$$|\widetilde{\mathcal{Z}(P^q)}| \geq N_p$$

which contradicts our assumption about the cardinality of C since

$$|C| + |\widetilde{\mathcal{Z}(P^q)}| = |\mathcal{Z}(P^q)| = qN_p$$

must hold. That is, $|\widetilde{\mathcal{Z}(P^q)}| \geq N_p$ implies $|C| \leq (q-1)N_p$, a contradiction. Thus there must be an R_q-set of roots in C. By a nearly identical argument, it is easy to show there is a R_p-set of roots in C as well.

The R_q-set of roots in C can be written as

$$\{e^{i(\omega+2\pi l/q)}\}_{l\in\mathbb{Z}_q}$$

where ω represents the phase orientation of the set. Each root in the above set is also a member of $\mathcal{Z}(Q^p)$, which implies it is a member of some R_p-set. The union of the q R_p-sets associated with the above set can be written as

$$\bigcup_{l\in\mathbb{Z}_q}\{(\omega+2\pi l/q+2\pi k/p) \bmod 2\pi\}_{k\in\mathbb{Z}_p} = \{(\omega+2\pi(qk+pl)/pq) \bmod 2\pi\}_{l\in\mathbb{Z}_q,k\in\mathbb{Z}_p}$$
$$= \{(\omega+2\pi n/pq) \bmod 2\pi\}_{n\in\mathbb{Z}_{pq}}$$

which corresponds to an R_{pq}-set of roots. The last step is a consequence of Lemma A.1. Thus, if $|C| > N_p(q-1)$, there exists a R_{pq}-set of roots in $\mathcal{Z}(f)$.

Case 2: $|C| = N_p(q-1) = N_q(p-1)$

For this case, we assume C contains no R_p-set or R_q-set since by the argument of the previous section, we would be done. The only way for C to avoid having a R_p-set or R_q-set is for the "completion" of each of the sets to exist in either $\widetilde{\mathcal{Z}(P^q)}$ or $\widetilde{\mathcal{Z}(Q^p)}$. Hence

$$|\widetilde{\mathcal{Z}(P^q)}| = N_p$$
$$|\widetilde{\mathcal{Z}(Q^p)}| = N_q.$$

Note that $\widetilde{\mathcal{Z}(P^q)}$ and $\widetilde{\mathcal{Z}(Q^p)}$ have no intersection since otherwise this would imply another root in C. Furthermore, the following holds

$$\widetilde{\mathcal{Z}(P^q)} = \mathcal{Z}(P)$$
$$\widetilde{\mathcal{Z}(Q^p)} = \mathcal{Z}(Q).$$

To see this, choose $z_0 \in \mathcal{Z}(P)$. This implies $z_0 \in \mathcal{Z}(P^q)$ or $z_0 \in \mathcal{Z}(Q)$. But $z_0 \in \mathcal{Z}(Q)$ implies a common root between P and Q which forces $|\mathcal{Z}(R)| < N_p$ and consequently implies there is one more root common to $\mathcal{Z}(P^q)$ and $\mathcal{Z}(Q^p)$, which, for this case, forces $|C| > (q-1)N_p$, a contradiction.

Choosing $z_0 \in \widetilde{\mathcal{Z}(P^q)}$ implies $z_0 \notin \mathcal{Z}(Q^p)$ so that $z_0 \in \mathcal{Z}(P)$ must hold. A similar argument holds for $\mathcal{Z}(Q)$ and $\widetilde{\mathcal{Z}(Q^p)}$. The end result is the "completion" of each of the N_p

R_q-sets of roots of P^q must be contained in $\mathcal{Z}(P)$. Conversely, the "completion" of each of the N_q R_p-sets of roots of Q^p must be contained in $\mathcal{Z}(Q)$.

The focus of the proof at this point is to reconcile the two arrangements of roots implied by the left and right sides of $PQ^p = QP^q$. The key observation is that the left side points to roots which are organized into N_q R_p-sets and the right side implies N_p R_q-sets of roots.

Consider the sets $\mathcal{Z}(P^q)$ and $\mathcal{Z}(Q^p)$ arranged as shown in Figure 3.2. On the left, each

$$\overset{\mathcal{Z}(P^q)}{\left\{\begin{array}{ccccc} r_{1,1} & r_{1,2} & \cdots & r_{1,q-1} & r_{1,q} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,q-1} & r_{2,q} \\ r_{3,1} & r_{3,2} & \cdots & r_{3,q-1} & r_{3,q} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{N_p,1} & r_{N_p,2} & \cdots & r_{N_p,q-1} & r_{N_p,q} \end{array}\right\}} \qquad \overset{\mathcal{Z}(Q^p)}{\left\{\begin{array}{cccccc} s_{1,1} & s_{1,2} & s_{1,3} & \cdots & s_{1,p-1} & s_{1,p} \\ s_{2,1} & s_{2,2} & s_{2,3} & \cdots & s_{2,p-1} & s_{2,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{N_q,1} & s_{N_q,2} & s_{N_q,3} & \cdots & s_{N_q,p-1} & s_{N_q,p} \end{array}\right\}}$$

Figure 3.3 Graphical Representation of $\mathcal{Z}(P^q)$ and $\mathcal{Z}(Q^p)$

row is an R_q-set and on the right, each row is an R_p set of roots. Without loss of generality, assume first $q-1$ columns on the left and the first $p-1$ columns on the right form the intersection set C. Specifically, assume

$$C = \{r_{k,l}\}_{k\in\mathbb{N}_{N_p},l\in\mathbb{N}_{q-1}} = \{s_{i,j}\}_{i\in\mathbb{N}_{N_q},j\in\mathbb{N}_{p-1}}$$

where for $p \in \mathbb{N}$, the notation $\mathbb{N}_p$ is used to mean $\mathbb{N}_p = \{1, 2, \ldots, p\}$. On the left, assume the "completion" of each of theR_q sets is represented in the last column and likewise for the R_p sets on the right.

Define the invertible mapping $M : \mathbb{N}_{N_q} \times \mathbb{N}_{p-1} \rightarrow \mathbb{N}_{N_p} \times \mathbb{N}_{q-1}$ to be the index mapping between the two representations of C:

$$s_{k,l} = r_{i,j} \Longleftrightarrow M(i,j) = (k,l).$$

It is an invertible mapping since there we assume unique indices and the cardinalities of the two representations of C are equal.

Now consider the set $\{s_{1,j}\}_{j\in \mathrm{N}_{p-1}}$. This is a partial R_p set so the phase of each element differs from the others by an integer multiple of $2\pi/p$. We say these roots are *phase-spaced* by $2\pi/p$, which is a term used loosely previously but made more rigorous here. Since these roots can be found among the $r_{k,l}$ which are organized as N_p partial R_q sets, the next observation is key. The roots on the left corresponding to the $\{s_{1,j}\}_{j\in \mathrm{N}_{p-1}}$ must each be located in a different row since two roots cannot be simultaneously phase-spaced by $2\pi/q$ and $2\pi/p$ unless they are equal. Rigorously, with $(k_1, l_1) = M(i, j_1)$ and $(k_2, l_2) = M(i, j_2)$,

$$k_1 = k_2 \Longrightarrow j_1 = j_2.$$

Since M is invertible, the converse is also true; any two distinct roots on the left in the same row must map to different rows on the right.

Without loss of generality, we can arrange the roots on the left so that the $p-1$ rows into which the $\{s_{1,j}\}_{j\in \mathrm{N}_{p-1}}$ map are the first $p-1$ rows. Furthermore, we can assume

$$s_{1,j} = r_{j,1} \quad j \in \mathrm{N}_{p-1}.$$

Now, the first partial row on the left $\{r_{1,l}\}_{l\in \mathrm{N}_{q-1}}$ maps onto different rows on the right by the same reasoning as above. As above, we can assume these rows are the first $q-1$. That is, we can assign $r_{1,j} = s_{j,1}$ which results in the first row and column on the left identically mapping to the the first colunm and row, respectively, on the right, and vice-versa.

At this point, we want to show

$$\{r_{k,l}\}_{k\in \mathrm{N}_{p-1}, l\in \mathrm{N}_{q-1}} = \{s_{i,j}\}_{i\in \mathrm{N}_{q-1}, j\in \mathrm{N}_{p-1}}$$

To do this, we show subset relationships in both directions. Choose $r_{k,l}$ arbitrarily. We want to show there exists a $(j,n) \in \mathbf{N}_{q-1} \times \mathbf{N}_{p-1}$ such that $r_{k,l} = s_{j,n}$. To do this we use the following facts.

The k-th row on the left is an R_q-set of roots which means the phase spacing between $r_{k,l}$ and $r_{k,1}$ is $2\pi/q$. Rigorously, there exists a $n_l \in \mathbf{N}_q$ such that

$$r_{k,l} = e^{i2\pi n_l/q} r_{k,1}.$$

Furthermore, since the first column is phase-spaced by $2\pi/p$, there exist $n_k \in \mathbf{N}_p$ such that

$$r_{k,1} = e^{i2\pi n_k/p} r_{1,1}.$$

Substituting $s_{1,1} = r_{1,1}$ and combining the previous two expressions, we get

$$r_{k,l} = e^{i2\pi k_2/p} e^{i2\pi k_1/q} s_{1,1}.$$

Letting $s_{j,1} = e^{i2\pi k_1/q} s_{1,1}$ and $s_{j,n} = e^{i2\pi k_2/p} s_{j,1}$, we need to ensure two conditions. We need to have $j \in \mathbf{N}_{q-1}$ and $n \in \mathbf{N}_{p-1}$. Clearly $j \neq q$ since the set $\{s_{i,1}\}_{i \in \mathbf{N}_q}$ would form an R_q-set of roots in C which is a contradiction as mentioned before. That is, $j = q$ would imply $s_{q,1}$ is the "completion" of an R_q-set of roots in C since we assumed the set $\{s_{i,1}\}_{i \in \mathbf{N}_{q-1}} \subset C$ was phase-spaced by $2\pi/q$.

Showing $n \in \mathbf{N}_{p-1}$ is easier once we know $j \in \mathbf{N}_{q-1}$. Since $r_{k,l}$ and $s_{j,1}$ are phase-spaced by $2\pi/p$, $r_{k,l}$ exists in row j on the right side of Figure 3.2. We know $n \in \mathbf{N}_{p-1}$ since $n = p$ implies $r_{k,l} = s_{j,p}$ which is a contradiction since $r_{k,l} \in C$ and $s_{j,p} \notin C$.

Thus we have shown that given $(k,l) \in \mathbf{N}_{p-1} \times \mathbf{N}_{q-1}$ there exists a $(j,n) \in \mathbf{N}_{q-1} \times \mathbf{N}_{p-1}$for every (k,l) there exists (j,n) such that $r_{k,l} = s_{j,n}$. The reasoning can be applied from right to left to show the converse is true as well. Thus we have shown

$$\{r_{k,l}\}_{k \in \mathbf{N}_{p-1}, l \in \mathbf{N}_{q-1}} = \{s_{i,j}\}_{i \in \mathbf{N}_{q-1}, j \in \mathbf{N}_{p-1}}.$$

The union of the first $p-1$ rows on the left of Figure 3.2 with the last column on the right is a set of $pq-1$ roots which have a phase-spacing of $2\pi/pq$ due to the fact that p and q are relatively prime and an application of Lemma A.1. This set of roots correspond to those of modulated regularity function $S_{pq}(\cdot,\gamma)$, and form a self-consistent set with respect to the structure imposed by the dual dilation equation. This is to say that $S_{pq}(\cdot,\gamma)$ can be factored from the polynomial $f = Q^pP = P^qQ$ so that $S_p(\cdot,\gamma)$ divides P and $S_q(\cdot,\gamma)$ divides Q. We can continue factoring $S_{pq}(\cdot,\gamma)$ from f as above to show that f can only be factored as a product of some finite number, M, of these modulated regularity functions which is what we set out to prove.

■

Lemma 3.6. then indicates the following forms for P and Q

$$\left.\begin{array}{rl} P(z) = & \prod_{j=1}^{M} S_p(z;\gamma_j) \\ Q(z) = & \prod_{j=1}^{M} S_q(z;\gamma_j) \end{array}\right\} \quad \gamma_j \in \{2\pi k/d\}_{k\in\mathbf{Z}_d}.$$

where $d = \gcd(p-1, q-1)$.

3.3 Rational Resolution Scaling Functions

Before applying the two theorems of the previous section to the discussion of RRA scaling functions, it is necessary to discuss duality in the context of approximation spaces. If $\{\phi_{0,j}\}_{j\in\mathbf{Z}}$ is a Riesz basis for the Hilbert space V_0, then it has a unique dual basis $\{\tilde{\phi}_{0,j}\}_{j\in\mathbf{Z}}$ which is also a Riesz basis for V_0 and satisfies

$$\langle\tilde{\phi}(\cdot - j), \phi(\cdot - k)\rangle = \delta_{k,j}.$$

The dual scaling function $\tilde{\phi}$ can be derived from this requirement via the following analysis. Writing the inner product in terms of the Fourier transform and expanding it in terms of its integral definition yields

$$\begin{aligned}
\delta_{k,j} &= \frac{1}{2\pi}\int_{\xi\in\mathbb{R}} \hat{\tilde{\phi}}(\xi)\overline{\hat{\phi}(\xi)}e^{i\xi(k-j)}d\xi \\
&= \frac{1}{2\pi}\int_{\xi\in\mathbb{R}} \hat{\tilde{\phi}}(\xi)\overline{\hat{\phi}(\xi)}e^{i\xi(k-j)}d\xi \\
&= \frac{1}{2\pi}\sum_{l\in\mathbf{Z}}\int_0^{2\pi} \hat{\tilde{\phi}}(\xi+2\pi l)\overline{\hat{\phi}(\xi+2\pi l)}e^{i\xi(k-j)}d\xi \\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{i\xi(k-j)}\sum_{l\in\mathbf{Z}} \hat{\tilde{\phi}}(\xi+2\pi l)\overline{\hat{\phi}(\xi+2\pi l)}d\xi
\end{aligned}$$

where the interchange of summation and integration above is justified by an application of the Levi theorem for series of Lebesgue-integrable functions[3]§10.8. If the final expression above is considered as an inverse Fourier transform, we have

$$\sum_{l\in\mathbf{Z}} \hat{\tilde{\phi}}(\xi+2\pi l)\overline{\hat{\phi}(\xi+2\pi l)} = 1 \quad \text{a.e. } \xi\in\mathbb{R}$$

from which an expression for $\hat{\tilde{\phi}}$ can be written as

$$\hat{\tilde{\phi}}(\xi) = \frac{\hat{\phi}(\xi)}{\sum_l |\hat{\phi}(\xi+2\pi l)|^2}. \tag{3.9}$$

The denominator in the above is known as the Euler-Frobenius function[11] generated by ϕ, denoted by E_ϕ. Expressing $\hat{\phi}$ in terms of (3.1b) and substituting into E_ϕ yields

$$\begin{aligned}
E_\phi(e^{i\xi}) &= \sum_l |\hat{\phi}(\xi+2\pi l)|^2 \\
&= \frac{1}{p}\sum_l |H_p(e^{i(\xi+2\pi l)/p})|^2|\hat{\phi}((\xi+2\pi l)/p)|^2 \\
&= \frac{1}{p}\sum_{k=0}^{p-1} |H_p(e^{i(\xi+2\pi k)/p})|^2\sum_l |\hat{\phi}((\xi+2\pi k)/p+2\pi l)|^2.
\end{aligned}$$

This expression will be useful in the following lemmas and theorems.

Lemma 3.8. *If ϕ simultaneously satisfies dilation equations for p and q with scaling filters H_p and H_q, then ϕ satisfies dilation equations for dilation factor $p^k q^l$ for $k, l \in \mathbb{Z}^+$. Furthermore, the corresponding scaling filter is given by*

$$H_{p^k q^l}(e^{i\xi}) = p^{-k/2} q^{-l/2} \prod_{n=0}^{k} \prod_{m=0}^{l} H_p(e^{ip^n\xi}) H_q(e^{iq^m\xi})$$

Proof. From the two scale relationships we have

$$\hat{\phi}(p\xi) = p^{-1/2} H_p(e^{i\xi})\hat{\phi}(\xi)$$

and similarly for the q-dilation. By substituting this equation into itself, we can write

$$\hat{\phi}(p^k\xi) = p^{-k/2} \prod_{n=0} H_p(e^{ip^n\xi})\hat{\phi}(\xi).$$

Consequently, we have

$$\begin{aligned}\hat{\phi}(p^k q^l \xi) &= p^{-k/2} \prod_{n=0} H_p(e^{ip^n\xi})\hat{\phi}(q^l\xi) \\ &= p^{-k/2} q^{-l/2} \prod_{n=0}^{k} \prod_{m=0}^{l} H_p(e^{ip^n\xi}) H_q(e^{iq^m\xi})\hat{\phi}(p\xi)\end{aligned}$$

which is what we set out to prove.

■

Lemma 3.9. *If H_p be the scaling symbol (filter) associated with the scaling function ϕ of a p-dilation MRA, then $\mathcal{Z}(H_p)$ cannot contain an R_p-set of roots. That is, there exists no $\alpha \in [0, 2\pi)$ such that $\{e^{i(\alpha+2\pi k)/p}\}_{k\in\mathbb{Z}_p} \subset \mathcal{Z}(H_p)$.*

Proof. Assume such a set of roots exists. Then

$$E_\phi(e^{i\xi}) = \sum_{k\in\mathbb{Z}} |\hat{\phi}(\xi + 2\pi k)|^2$$

$$= \frac{1}{p}\sum_{k=0}^{p-1} |H_p(e^{i(\xi+2\pi k)/p})|^2 \sum_{l\in\mathbb{Z}} |\hat{\phi}((\xi+2\pi k)/p+2\pi l)|^2$$

which vanishes when evaluated at $\xi = \alpha$. Since we assumed ϕ was a scaling function associated with an MRA, it forms a Riesz basis for the approximation space and thus satisfies the Riesz condition. That is, there exist positive constants A and B such that the following condition[12] holds

$$A \leq \sum_{k\in\mathbb{Z}} |\hat{\phi}(\xi+2\pi k)|^2 \leq B, \qquad \text{a.e. } \xi \in \mathbb{R}.$$

Now, since $h_p \in l_1$, (see Equation 2.8 and following discussion) its Fourier transform, H_p, is continuous. Thus $|H_p(e^{i(\xi+2\pi k)/p})|^2$ is continuous and the Euler-Frobenius polynomial will go to zero in a continuous fashion forcing a violation of the lower Riesz bound A. This contradicts the assumption that ϕ formed a Riesz basis for approximation space of the MRA. Hence, no such R_p-set can exist.

■

The implication of this theorem and Theorem 3.7. is that while there exists many classes of polynomials which satisfy (3.5), the only ones which need further consideration are the ones associated with regularity factors:

$$\begin{aligned} H_p(z) &= a_p \textstyle\prod_{j=1}^{m} S_p(z;\gamma_j) \\ H_q(z) &= a_q \textstyle\prod_{j=1}^{m} S_q(z;\gamma_j) \end{aligned} \qquad \gamma_j \in \{e^{i2\pi l/d}\}_{l\in\mathbb{Z}_d}. \tag{3.10}$$

Before characterizing the allowable values of γ_j, the following lemma will be useful.

Lemma 3.10. *If $H_p = \sqrt{p}S_p^m$ corresponds to an orthonormal scaling function, then $m = 1$.*

Proof. For $m = 1$, we have the scaling filter for the well-known Haar scaling function which is orthonormal. We assume $m > 1$ and show the corresponding scaling filter violates a validity condition which is necessary for an orthonormal scaling function.

Since $|S_p|^2$ is positive and bounded above by 1, we have

$$|S_p^m(e^{i\xi})|^2 \leq |S_p(e^{i\xi})|^2 \quad \xi \in [0, 2\pi)$$

with equality only at $\xi = 0$. Hence

$$\frac{1}{2\pi}\int_0^{2\pi} |H_p(e^{i\xi})|^2 d\xi = \frac{1}{2\pi}\int_0^{2\pi} p|S_p^m(e^{i\xi})|^2 d\xi < \frac{1}{2\pi}\int_0^{2\pi} p|S_p(e^{i\xi})|^2 d\xi = 1 \quad \forall\, m > 1,$$

which violates the requirement that a valid scaling filter have unit energy. It is closely related to the requirement that $\|h_p\|_{l_2} = 1$. Hence, $m = 1$ is the only allowable value.

■

The following theorem characterizes the allowable values of γ_j. and proves that any scaling filter of the form given in Equation 3.10 and corresponding to an orthonormal scaling function is the single regularity function $S_p(\cdot; 1)$ or simply, S_p. As background to this result, recall the Fourier-domain relationship btween the scaling filter H_p and the scaling function $\hat{\phi}$:

$$\hat{\phi}(\xi) = \prod_{k=1}^{\infty} H_p(e^{i\xi p^{-k}}) \tag{3.11}$$

where the H_p have been renormalized so that $H_p(1) = 1$ instead of $H_p(1) = p^{1/2}$ as was originally defined.

Theorem 3.11. *If $H_p(z) = a_p \prod_{j=1}^{m} S_p(z;\gamma_j)$ with $\gamma_j \in \{e^{i2\pi l/d}\}_{l\in \mathbb{Z}_d}$ corresponds to an orthonormal scaling function, then $m = 1$, $\gamma_1 = 1$, and $a_p = p^{1/2}$. More precisely, $H_p = p^{1/2} S_p$.*

Proof. Assume H_p has $S_p(\cdot;\gamma_j)$ as a factor with $\gamma_j \neq 1$. This factor contributes zeroes to $\hat{\phi}$ via the expansion given in (3.11) above. That is, the zeros of the product

$$\prod_{k=1}^{\infty} S_p(e^{i\xi p^{-k}};\gamma_j)$$

are also zeroes of $\hat{\phi}$. The approach of this part of the proof is to demonstrate that the zeroes contributed by this infinite product are uniformly spaced by 2π and thus violate the orthonormality condition for $\hat{\phi}$.

To see this, first consider the zeros contributed by the $k = 1$ term of the infinite product. These zeros are given by ξ such that

$$\xi/p - 2\pi j/d = 2\pi l/p \quad l \in \mathbb{Z}, l \bmod p \neq 0.$$

Defining ω as the normalized frequency $\xi = 2\pi\omega$ and solving the equation above in terms of ω, we have

$$\omega = l + pj/d \quad l \in \mathbb{Z}, l \bmod p \neq 0.$$

In terms of ω, the unity points (defined to the points for which $S_p = 1$) for $k = 1$ are

$$\begin{aligned} \omega &= l + pj/d & & l \in \mathbb{Z}, l \bmod p = 0, \\ &= pl + pj/d & & l \in \mathbb{Z}, \\ &= pl + nj + j/d & & l \in \mathbb{Z}. \end{aligned}$$

The last step above highlights the relationship between p and d and is crucial in this analysis. Recall that $d = (p-1, q-1)$ which implies that d divides $p - 1$. Define n by $nd = p - 1$ so that

$$\begin{aligned} pj/d &= (p-1)j/d + j/d \\ &= nj + j/d. \end{aligned}$$

The importance of this observation will become apparent.

The zeros of the infinite product associated with an arbitrary k are given by

$$\begin{aligned} \omega &= p^{k-1}l + p^k j/d & & l \in \mathbb{Z}, l \bmod p \neq 0, \\ &= p^{k-1} + nj(p^{k-1} + p^{k-2} + \cdots + p) + nj + j/d & & l \in Z, l \bmod p \neq 0. \end{aligned}$$

Notice that the zeros for an arbitrary level k are a subset of the $k = 1$ unity points. Furthermore, the zeros at one value of k are disjoint from those of another value. This can be seen by observing that the zeros at level $k+1$ are determined by multiplying the zeros at level k by p which violates the restriction that the iteration index not be a multiple of p. The result is that, in the limit, all the $k = 1$ unity points are cancelled by the zeros associated with subsequent values of k. Thus, $\hat{\phi}(2\pi(j/d + l)) = 0$ identically for $l \in \mathbb{Z}$.

For orthonormality, recall the requirement on $\hat{\phi}$:

$$\sum_{k\in\mathbb{Z}} |\hat{\phi}(\xi + 2\pi k)|^2 = 1 \quad a.e. \xi \in \mathbb{R}.$$

By choosing $\xi = 2\pi j/d$, we have a violation which implies ϕ is not an orthonormal scaling function. Thus, H_p cannot contain any factor $S_p(\cdot\,;\gamma_j)$ when $\gamma_j \neq 1$.

With this intermediate result, we are left to consider H_p of the form $H_p = S_p^m$. According to Lemma 3.10., we must have $m = 1$. The value $a_p = p^{1/2}$ follows from the required normalization of H_p. Thus, scaling filters of the form given in Equation 3.10 which correspond to orthonormal scaling functions can be written as

$$H_p = p^{1/2} S_p,$$
$$H_q = q^{1/2} S_q.$$

■

Theorem 3.11. provides an interesting insight into spline-based scaling functions. Without Lemma 3.10., the theorem proves that any scaling filter containing a modulated regularity factor $S_p(\cdot;\gamma)$ with $\gamma \neq 1$ cannot be orthonormal. Furthermore, it cannot be transformed into an orthonormal scaling function via the "orthonormalization" trick since there exist points ξ for which

$$\sum_{k\in\mathbb{Z}} |\hat{\phi}(\xi + 2\pi k)|^2 = 0.$$

Recall the left side of the expression above is required in the denominator of the trick which implies division by zero.

3.4 Conclusion

The previous theorem proves that the only compactly supported scaling function associated with a perfect-reconstruction RRA is the well-known characteristic function or the Haar scaling function. Furthermore, the required form of the scaling filters demonstrates the only scaling functions which yield perfect reconstruction are those which are spline-based. That is, they are scaling functions which are constructed from successive convolutions of the regularity function S_p.

IV. Spline-Based Wavelets

4.1 Introduction and Motivation

The results of the last chapter indicate that the only useful RRA scaling functions are those which are based on the B-splines. According to the scheme proposed in Chapter II, wavelets which complement these types of scaling functions are necessary. These spline-based wavelets are the topic of this chapter.

There are two major portions of this chapter. First, some initial analysis is performed to quantify the requirements of spline-based wavelets. The approach is to define an operator which represents the projection of $\psi \in V_{-1}$ onto V_0 and show the nullspace of this operator and the detail space W_0 are isomorphic. The approach is to start with compactly supported functions which are orthogonal to the approximation space, then show the integer shifts of these functions form a basis for the detail space and are, therefore, semi-orthonormal wavelets. Furthermore, they can be "orthonormalized" to form an orthogonal basis for the detail space.

Then, the theory above is put into practice via an algorithmic description of the technique to construct the B-spline wavelets. The description will include such issues as symmetry/anti-symmetry, redundancy, orthogonalization, and numerical accuracy.

4.2 Analysis and Problem Description

4.2.1 Conditions for Semi-orthogonal Wavelets. Suppose we have a p-dilation MRA with the (non-orthonormal) B-spline scaling function $\phi = N_m$ for some $m \in \mathbf{N}$. The integer translations of ϕ form a basis for an approximation space V_0. The orthogonal complement of V_0 in V_{-1} is the detail space, W_0, which is spanned by $p-1$ wavelets.

Definition 4.1. In the context of an MRA, a *semi-orthogonal wavelet* $\psi \in V_{-1}$ is a function which satisfies the following:

$$\langle \phi_{j,k}, \psi_{j,m} \rangle = 0 \quad j, k, m \in \mathbb{Z}$$
$$\langle \psi_{j,k}, \psi_{l,m} \rangle = 0 \quad j \neq l,\ j, k, l, m \in \mathbb{Z}$$

which is to say that ψ is orthogonal to the corresponding approximation space and self-orthogonal across dilation levels, but generally not self-orthogonal with respect to its own integer translations within the same dilation level. ∎

Semi-orthogonal wavelets are important in that they can form Riesz bases for detail spaces in the same way $\phi = N_m$ forms a basis for the approximation space. In this sense, ϕ can generally be considered as a semi-orthogonal scaling function. Semi-orthogonal basis functions are usually constructed with compact support; their associated symbols are polynomials which facilitates their manipulation.

We want to find semi-orthonormal wavelets, ψ, which satisfy

$$\langle \phi(\cdot - n), \psi \rangle = \int_{-\infty}^{\infty} \phi(t-n)\psi(t)dt = 0, \quad \forall\, n \in \mathbb{Z}. \tag{4.1}$$

Expressing the scaling function and the wavelet in terms of their respective two-scale equations

$$\phi(t) = \sum_n h_p(n)\phi(pt-n)$$
$$\psi(t) = \sum_n g_p(n)\phi(pt-n),$$

so that the orthogonality expressed in (4.1) can be written as:

$$\begin{aligned}\langle \phi(\cdot - n), \psi \rangle &= \sum_l \sum_k h_p(l) g_p(k) \int_{-\infty}^{\infty} \phi(pt - pn - l)\phi(pt-k)dt \\ &= \frac{1}{p}\sum_l \sum_k h_p(l) g_p(k) \int_{-\infty}^{\infty} \phi(t - pn - l + k)\phi(t)dt\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{p} \sum_l \sum_k h_p(l) g_p(k) \int_0^m N_m(t - pn - l + k) N_m(t) dt \\
&= \frac{1}{p} \sum_l \sum_k h_p(l) g_p(k) \int_0^m N_m(m + pn + l - k - t) N_m(t) dt \\
&= \frac{1}{p} \sum_l \sum_k h_p(l) g_p(k) N_{2m}(m + pn + l - k) \\
&= 0 \quad n \in \mathbb{Z},
\end{aligned}
$$

where the following identities have been used: $N_m(t) = N_m(m-t)$, and $N_m * N_m = N_{2m}$ where $*$ denotes continuous convolution. We can write this previous expression in matrix form:

$$\mathbf{Ag} = 0$$

where

$$[\mathbf{A}]_{n,k} = \frac{1}{p} \sum_l h_p(l) N_{2m}(m + pn + l - k)$$

and $[\mathbf{g}]_k = g_p(k)$. It is clear that any solution, $\mathbf{g}$, must lie in the nullspace of $\mathbf{A}$. If we are seeking solutions which are compactly supported, the problem can be reduced to finding nullspace vectors of a truncated version of $\mathbf{A}$; the columns of $\mathbf{A}$ corresponding to the zero elements of any solution need not be considered. Furthermore, truncating columns allows us to truncate rows as well since $\mathbf{A}$ is banded in the sense that each row is finite and is a p-shifted version of a single fundamental row. Thus, there are only a finite number of rows with non-zero elements in the column-truncated $\mathbf{A}$. The desired support of the nullspace vector will determine the truncation of $\mathbf{A}$. Hence, some criteria must be applied to determine the support of possible solutions.

There is a great deal of structure to $\mathbf{A}$. In particular, the rows of $\mathbf{A}$ are p-shifted versions of the fundamental row

$$\mathbf{a}_l = [\mathbf{A}]_{0,l} = \frac{1}{p} \sum_k h_p(k) N_{2m}(m + l - k).$$

This row is necessarily compactly supported since both h_p and N_{2m} are compactly supported. Let $\mathbf{A}'$ denote the truncated version of $\mathbf{A}$ which contains the fundamental row (plus additional columns if necessary to make the matrix rank deficient). This definition effectively truncates the columns of $\mathbf{A}$ for which the $n = 0$ row vanishes.

The reason for this choice of truncation is simple. If $\mathbf{A}'$ contains a complete row, then any vector in $\mathcal{N}(\mathbf{A}')$ will also be in $\mathcal{N}(\mathbf{A})$ (with appropriate zero padding). More importantly, any p-shift of the vector will be in $\mathcal{N}(\mathbf{A})$ as well. Considering any $\mathbf{A}'$ with fewer columns will not necessarily preserve this property.[1] In general the size of $\mathbf{A}$ will depend upon the dilation factor p and regularity factor m. The lengths of h_p and N_{2m} are $m(p-1)+1$ and $2m-1$, respectively, which implies the length of $\mathbf{a}$ is $L = mp+m-1$. The number of rows can be calculated by

$$\text{rows} = 1 + 2\left\lfloor \frac{L-1}{p} \right\rfloor .$$

However, in the case where p divides $L-1$ evenly, the number of rows can be reduced by 2 since this corresponds to the case where the top and bottom rows are all zero exept for the first and last element, respectively. Eliminating these rows in this case does not change the nullspace of the matrix since it simply forces the first and last elements of any nullspace vector to be zero. Consequently, the first and last colums of the matrix can also be eliminated. This will be more fully explained in the example which follows.

Then end result is that $\mathbf{A}'$ will have more rows than columns, which is as expected since we assumed a nontrivial nullspace. It would be ideal to have $\dim\mathcal{N}(\mathbf{A}') = p-1$ so that there are as many nullspace vectors as there are wavelets. This is generally not the case. In fact we generally find $\dim\mathcal{N}(\mathbf{A}') > p-1$ so that there are more nullspace vectors than desired wavelets. This will be addressed in the next section. This section is concluded with two simple examples.

[1] The case where $p = 2$ as in Example 4.2. is a special case. Including an entire fundamental row in the formulation of $\mathbf{A}'$ precludes any nontrival nullspace vectors since the matix is full rank for all values of m. This is handled by shifting the support of the desired solutions by 1.

Example 4.2. Consider the case where $m = 2$ and $p = 2$. This example is consistent with the derivation above. We have $h_p = \{1,2,1\}$, $N_{2m} = \{0,1,4,1,0\}$, and $h_p * N_{2m} = 0,1,6,10,1,0$. Then for g_p supported on $\{0,1,\ldots,3m-2\}$ (as given in Chui[11]), the system matrix $\mathbf{A}'$ is given by:

$$\mathbf{A}' = \begin{bmatrix} 6 & 1 & 0 & 0 & 0 \\ 6 & 10 & 6 & 1 & 0 \\ 0 & 1 & 6 & 10 & 6 \\ 0 & 0 & 0 & 1 & 6 \end{bmatrix}$$

which corresponds to $k \in \{0,1,\ldots,4\}$ and $n \in \{0,1,2,3\}$. The range of n is determined explicitly by the range of l. Also, the entries have been normalized so they are all integer values. This can always be accomplished since both h_p and N_{2m} are spline-based.

We want to solve the system $\mathbf{A}'\mathbf{g} = 0$. As noted before, $\mathbf{g}$ must be in the null space of $\mathbf{A}'$. The system is rank-deficient so there is a non-trivial nullspace. Calculating the spanning vector in this nullspace leads to:

$$\mathbf{q}^T = \begin{bmatrix} 1 & -6 & 10 & -6 & 1 \end{bmatrix}$$

which agrees exactly with Chui's coefficients for the minimally supported semi-orthonormal wavelets corresponding to the B-spline N_2[11]. ■

Example 4.3. Now consider the case where $m = 2$ (linear splines) and $p = 3$. The two-scale symbol H_p can be calculated as $H_p = (pS_p)^m$, where S is the regularity symbol defined and used previously. Assume that we are looking for semi-orthonormal wavelets with supports in $[0, m]$ which implies $k \in \{0,1,\ldots,4\}$ and $n \in \{-1,0,1\}$. With $h_p = \{1,2,3,2,1\}$, $N_{2m} = \{0,1,4,1,0\}$, and $h_p * N_{2m} = \{0,1,6,12,16,12,6,1,0\}$, the system matrix $\mathbf{A}'$ is given by:

$$\mathbf{A}' = \begin{bmatrix} 12 & 6 & 1 & 0 & 0 \\ 6 & 12 & 16 & 12 & 6 \\ 0 & 0 & 1 & 6 & 12 \end{bmatrix}$$

from which we can calculate a null space spanned by two vectors:

$$(\mathbf{g}^{(1)})^T = \begin{bmatrix} 2 & -5 & 6 & -5 & 2 \end{bmatrix}$$

and

$$(\mathbf{g}^{(2)})^T = \begin{bmatrix} 1 & -2 & 0 & 2 & -1 \end{bmatrix}.$$

The implementation details of this example are provided in Appendix D. ■

4.2.2 Completeness. The question which must be answered in light of the construction technique outlined above is how the nullspace vectors of $\mathbf{A}'$ and the detail space, W_0, are related. Specifically, it would be desirable to show the nullspace vectors correspond to semi-orthogonal wavelets which span the detail space of the p-dilation MRA.

Lemma 4.4. *Let the approximation space V_0 be defined by the scaling function $\phi = N_m$ for some $m \in \mathbf{N}$. As usual, let the detail space W_0 be the orthogonal complement of V_0 in V_{-1}. Let $\mathbf{A}$ be defined as above. Then $\mathcal{N}(\mathbf{A})$ and W_0 are isomorphic.*

Proof. Consider the transformation $\mathbf{Q} : V_{-1} \to l_2$ which "extracts" approximation coefficients. That is, for any $f \in V_{-1}$, we can express f by:

$$f = \sum_{k \in \mathbf{Z}} c_k N_m(p \cdot -k),$$

where $c \in l_2$. Define $\mathbf{Q}$ by $\mathbf{Q}f = c$. Since the appoximation coefficients uniquely define the functions, $\mathbf{Q}$ is invertible:

$$\mathbf{Q}^{-1}c = \sum_{k \in \mathbf{Z}} c(k) N_m(p \cdot -k) = f.$$

and both $\mathbf{Q}$ and $\mathbf{Q}^{-1}$ are linear. To show $\mathbf{Q}$ is an isomorphism between W_0 and $\mathcal{N}(\mathbf{A})$, it is sufficient to show $\mathbf{Q}(W_0) = \mathcal{N}(\mathbf{A})$ which implies $\mathbf{Q}$ is one-to-one and onto.

To show $\mathbf{Q}$ is one-to-one, it is sufficient (and necessary) to show $\mathcal{N}(\mathbf{Q}) = \{0\}$. This holds since

$$\mathbf{Q}f = c = 0 \Longrightarrow \sum_{k\in\mathbf{Z}} c(k)N_m(p\cdot -k) = f = 0.$$

To show "onto", it is sufficient to show that $\mathbf{Q}^{-1}$ is one-to-one, or that $\mathcal{N}(\mathbf{Q}^{-1}) = \{0\}$. This also holds since

$$\begin{aligned}\mathbf{Q}^{-1}c = f = 0 &\Longrightarrow \sum_{k\in\mathbf{Z}} c(k)N_m(p\cdot -k) = 0\\ &\Longrightarrow c = 0\end{aligned}$$

The last implication is a consequence of the linear independence of the translates of N_m. Thus, $\mathbf{Q}$ is an invertible, linear, one-to-one and onto transformation between $\mathcal{N}(\mathbf{A})$ and W_0 which implies these spaces are isomorphic.

■

As was discussed earlier, one way to describe the structure of the infinite matrix $\mathbf{A}$ is given in Figure 4.1. The matrix $\mathbf{A}'$ is repeated along the main diagonal. We would like to show that

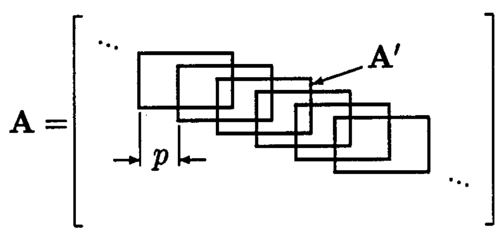

Figure 4.1 Matrix Structure of operator $\mathbf{A}$

the nullspace of the matrix $\mathbf{A}'$, discussed earlier, is related to the nullspace of the operator $\mathbf{A}$ by showing that all integer translations of these (finite) nullspace vectors span the nullspace of $\mathbf{A}$. Formally, we want to show:

$$\mathcal{N}(\mathbf{A}) = \bigcup_{i=1}^{N} G_i$$

where

$$G_i = \overline{\text{span}\{g_i(\cdot - pk)\}_{k\in\mathbb{Z}}}$$

and g_i is a nullspace vector of $\mathbf{A}'$. Note there is a slight abuse of notation here in that the nullspace vectors of $\mathbf{A}'$ are elements of $\mathbb{R}^L$, yet they are treated as if they are in l_2. The inconsistency can be overlooked with the assumption that g_i vanishes outside this region of interest or that we can append and prepend zeros to the various vectors without changing the nature of the problem. Here, we use L to denote the the length of the nullspace vector; it is also the number of columns in $\mathbf{A}'$.

Definition 4.5. The truncated matrix $\mathbf{A}^{(k)}$ is defined to be the columnwise truncation of $\mathbf{A}$ to include rows corresponding to $n \in \{-k, -k+1, \ldots, 0, \ldots, k-1, k\}$. The non-zero portion of $\mathbf{A}^{(k)}$ contains $2k+1$ cascaded and overlapping copies of the fundamental matrix $\mathbf{A}'$ (see Figure 4.1). ■

Lemma 4.6. *The sequence $\{\mathbf{A}^{(k)}\}$ converges strongly (pointwise) to $\mathbf{A}$:*

$$\lim_{k\to\infty} \|(\mathbf{A} - \mathbf{A}^{(k)})(x)\|_{l_2} = 0 \quad \forall x \in l_2.$$

Proof. Fix x arbitrarily in l_2. Let $M_k = \{1-m-pk, \ldots, L-m+pk\}$ be the set of indices corresponding to the non-zero entries in $\mathbf{A}^{(k)}$. Choose an $\epsilon > 0$. We need to find a $K \in \mathbf{N}$ such that

$$\|(\mathbf{A} - \mathbf{A}^{(k)})(x)\|^2 < \epsilon \quad \forall k > K$$

Let $y = \mathbf{A}x$. This represents the discrete convolution of x with the fundamental row $\mathbf{a}$, followed by a p-downsample. Since a has compact support and is bounded (by 1), then $y \in l_2$. Furthermore,

$$((\mathbf{A} - \mathbf{A}^{(k)})x)(n) = \begin{cases} 0 & |n| \le k \\ e^{(k)}(n) & k < |n| \le \lfloor \frac{L-1}{p} \rfloor \\ y(n) & \text{otherwise} \end{cases}$$

where

$$e^{(k)}(n) = \sum_{l \notin M_k} a(l - pn)x(l).$$

Since $y \in l_2$, there exists a K_1 such that

$$\sum_{|n|>K_1} |y(n)|^2 < \epsilon/2.$$

But we need to contend with the "error" terms, $e^{(k)}$. Define the finite-range max operator $M_L : l_2 \to l_2$ by

$$(M_L x)(n) = \max_{j \in \mathbb{Z}_L}\{x(n + j)\}.$$

Since the fundamental row **a** is bounded by 1 and has finite support, L, the following holds for $n > 0$

$$\begin{aligned}
|e^{(k)}(n)| &= \left| \sum_{l=L-m+1+pk}^{L-m+pn} a(l - pn)x(l) \right| \\
&\leq \sum_{l=L-m+1+pk}^{L-m+pn} a(l - pn)|x(l)| \\
&\leq \sum_{l=L-m+1+pk}^{L-m+pn} |x(l)| \\
&\leq L|(M_L x)(L - m + 1 + pk)|
\end{aligned}$$

and for $n < 0$

$$\begin{aligned}
|e^{(k)}(n)| &= \left| \sum_{l=1-m+pn}^{-m-pk} a(l - pn)x(l) \right| \\
&\leq \sum_{l=1-m+pn}^{-m-pk} a(l - pn)|x(l)| \\
&\leq \sum_{l=1-m+pn}^{-m-pk} |x(l)| \\
&\leq L|(M_L x)(-m - pk - L)|.
\end{aligned}$$

Since $M_L x \in l_2$ (see Lemma A.2), there exist K_2 such that

$$\sum_{|n|>K_2} |e^{(k)}(n)|^2 < \epsilon/2.$$

Choosing $K = \max(K_1, K_2)$,

$$\sum_{|n|>K} |e^{(k)}(n)|^2 + |y(n)|^2 < \epsilon$$
$$||(\mathbf{A} - \mathbf{A}^{(K)})x||^2 < \epsilon$$

which is sufficient to satisfy the hypothesis of the lemma.

■

Remark. We do not have convergence in the uniform topology since it is always possible to choose an $x_k \in l_2$ such that $||x_k|| = 1$ and $x_k(l) = 0$ for $l \in M_k$. Hence, $(\mathbf{A} - \mathbf{A}^{(k)})x_k = \mathbf{A}(x_k)$.

The ultimate goal is to show the nullspace of the infinite matrix operator $\mathbf{A}$ is spanned by the p-translates of the nullspace vectors of $\mathbf{A}^{(0)} = \mathbf{A}'$. Let Q be a set of vectors which span the nullspace of $\mathbf{A}^{(0)}$ given by

$$Q = \{q_j\}_{j=1}^{N}$$

so that $\mathcal{N}\left(\mathbf{A}^{(0)}\right) = \text{span}(Q)$.

Lemma 4.7. *For all $k \in \mathbb{Z}^+$, the nullspace of $\mathbf{A}^{(k)}$ is spanned by the p-translates of the nullspace vectors of $\mathbf{A}^{(0)}$.*

Proof. This proof is done by induction. First, note that the q_i span $\mathcal{N}(\mathbf{A}^{(0)})$ by definition. Now, assume

$$\mathcal{N}(\mathbf{A}^{(k-1)}) = \text{span}\{q_i(\cdot - pn)\}_{i=1}^{N}$$

for $|n| \leq k-1$. We want to show the $\mathcal{N}\left(\mathbf{A}^{(k)}\right)$ also spanned by the q_i as well.

Fix q arbitrarily in the nullspace of $\mathbf{A}^{(k)}$. The first L columns are identical to those in $\mathbf{A}^{(0)}$ so the first L elements of q can be represented as some element in the nullspace of $\mathbf{A}^{(0)}$, say $q_-^{(0)}(\cdot + pk)$. Likewise, the last L elements of q, taken as a vector, is a nullspace vector of $\mathbf{A}^{(0)}$, say $q_+^{(0)}(\cdot - pk)$.

Now both $q_-^{(0)}(\cdot + pk)$ and $q_+^{(0)}(\cdot - pk)$ are nullspace vectors of $\mathbf{A}^{(k)}$ and can be chosen so that the first and last p elements of $\tilde{q}$ are zero so that

$$\tilde{q} = q - q_-^{(0)}(\cdot + pk) - q_+^{(0)}(\cdot - pk) \in \mathcal{N}(\mathbf{A}^{(k)}).$$

But the first and last p elements of $\tilde{q}$ are zero, which implies $\tilde{q} \in \mathcal{N}(\mathbf{A}^{(k-1)})$. Thus we can write

$$\tilde{q} = \sum_{i=1}^{N} \sum_{n=1-k}^{k-1} \alpha^{(k-1)} q_i(\cdot - pn)$$

or substituting the definition of $\tilde{q}$:

$$q = \sum_{i=1}^{N} \sum_{n=1-k}^{k-1} \alpha^{(k-1)} q_i(\cdot - pn) + q_+^{(0)}(\cdot - pk) + q_-^{(0)}(\cdot + pk).$$

Since q was chosen arbitrarily, the previous analysis must hold for all such q which gives us the desired result for the nullspace of $\mathbf{A}^{(k)}$. Since it is clear the nullspace of $\mathbf{A}^{(0)}$ is spanned by the q_i, the induction on k is complete and so is the proof.

■

Theorem 4.8. *With* $\mathbf{A}$ *and* $\mathbf{A}^{(k)}$ *defined above:*

$$\mathcal{N}(\mathbf{A}) = \overline{\bigcup_{k \in \mathbb{Z}^+} \mathcal{N}(\mathbf{A}^{(k)})}.$$

Proof. Let $\{q_i\}_{i=1}^N$ be a basis for $\mathcal{N}\left(\mathbf{A}^{(0)}\right)$. Fix $q \in \mathcal{N}(\mathbf{A}) \cap l_2$ arbitrarily. We want to show there exists a sequence $\{q^{(k)}\}$, such that $q^{(k)} \in \mathcal{N}\left(\mathbf{A}^{(k)}\right)$ which converges to q. Define

$q^{(k)}$ by

$$q^{(k)}(l) = \begin{cases} q(l) & l \in M_k \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathbf{A}^{(k)} q^{(k)} = 0$, we can write $q^{(k)}$ as some linear combination of the shifted nullspace vectors of $\mathbf{A}^{(0)}$ as a result of the previous lemma. Now, since

$$\lim_{k \to \infty} q^{(k)} = q,$$

if we choose an ϵ arbitrarily small, there exists a $K \in \mathbf{N}$ such that

$$\|q - \sum_{i=1}^{N} \sum_{j=-k}^{k} \alpha_{i,j} q_i(\cdot - pj)\| < \epsilon \quad k > K,$$

since $q \in l_2$. This implies

$$\mathcal{N}(\mathbf{A}) \subset \overline{\bigcup_{k \geq 0} \mathcal{N}(\mathbf{A}^{(k)})}.$$

We also have subset inclusion in the other direction if $\mathbf{A}$ is a continuous operator[30]. It is sufficient to show $\mathbf{A}$ is linear and bounded. Linearity is a direct consequence of the definition of the operator. For boundedness, observe that $\mathbf{A}x$ is a discrete convolution of x with the fundamental row a, which is compactly supported and thus bounded. Thus

$$(\mathbf{A}x)(n) = \sum_{l} \mathbf{a}(l) x(l + pn)$$

which implies via the triangle inequality

$$\begin{aligned} \|\mathbf{A}x\| &\leq \sum_{l=1-m}^{L-m} \mathbf{a}(l) \|x\| \\ &\leq L\|x\| \end{aligned}$$

which satisfies the boundedness requirement. Hence, $\mathbf{A}$ is a continuous operator and so the subset inclusion in the other direction can be shown:

$$\begin{aligned}\mathbf{A}(\lim_{k\to\infty} q^{(k)}) &= \lim_{k\to\infty} \mathbf{A}q^{(k)} \\ &= \lim_{k\to\infty} (0) \\ &= 0.\end{aligned}$$

■

Since we have the isomorphism between W_0 and $\mathcal{N}(\mathbf{A})$, this theorem prescribes a technique for finding wavelets which span the detail space W_0, and by dilation, all the W_k for $k \in \mathbb{Z}$. In fact, the problem of finding the wavelets is reduced to a numeric calculation of the null space of an integer-valued matrix $\mathbf{A}'$. In [25], Lian proves the completeness of such a technique for a dilation factor $p = 3$ by constructing a system matrix which is non-singular. The technique relies on being able to symbolically calculate the determinant of a 3×3 matrix, which becomes nearly impossible for larger dilation factors. This result proves completeness without the cumbersome calculations.

There are many computer codes to calculate a set of nullspace vectors of finite matrics. The result is usually an orthogonal set of unit-norm vectors whose linear combinations span the nullspace. While this is useful in practice, this approach generally sacrifices symmetry or anti-symmetry in the wavelets, which corresponds to linear phase in the corresponding detail filters.

However, symmetric and anti-symmetric wavelet vectors can be constructed as follows. Since the fundamental row used to construct $\mathbf{A}'$ is symmetric, the reversed or "flipped" version of any nullspace vector q_i is also in the nullspace. Consequently, the sum and difference are symmetric and anti-symmetric nullspace vectors. This new set of nullspace vectors are not necessarily linearly independent though.

There are only $p-1$ wavelets whose integer translations span the detail space. In general, the nullspace of $\mathbf{A}'$ has a dimensionality greater than $p-1$ so that the wavelets corresponding to these nullspace vectors overspecify the detail space. To form an orthonormal basis for W_0, a technique such as Gramm-Schmidt can be used to orthogonalize the $p-1$ wavelets with respect to each other and eliminate the redundant nullspace vectors.

4.2.3 Construction of Orthogonal Wavelets. In Example 4.3., two semi-orthogonal wavelets were calculated. These two wavelets are not mutually orthogonal and so the sub-detail spaces they span, respectively, are not orthogonal to one another, although they collectively span the entire detail space W_0. An MRA or RRA analysis using these two wavelets would lead to two sets of detail coefficients which are correlated. This can be a problem in certain applications where the desired representation of the signal needs to be as efficient as possible. Furthermore, correlated signals during analysis imply a need for an implicit decorrelation during synthesis, should synthesis be necessary for the application. In order to decorrelate the two wavelets of the previous example, a minimum-norm approach can be taken. In implementation it is similar to the Gramm-Schmitt procedure.

For the sake of illustration, assume a dilation factor of $p = 3$. Suppose we start with two semi-orthogonal wavelets ψ_1 and ψ_2 which are orthogonal to the corresponding scaling function ϕ and orthogonal to themselves across dilation levels. We want to fix one wavelet, say ψ_1, and calculate another wavelet ψ_2' which is orthogonal to it.

We want to define ψ_2' by

$$\psi_2' = \psi_2 - f_\alpha \tag{4.2}$$

where $f_\alpha = \sum_k \alpha_k \psi_{1,k}$ and

$$\alpha = \arg\min_{\alpha} ||\psi_2 - f_\alpha||^2.$$

This effectively describes the projection of ψ_2 onto the space spanned by the integer translates of ψ_1. We then have $\psi_2' \perp \psi_1(\cdot - k)$ for all $k \in \mathbb{Z}$. This last statement is a consequence of working within a Hilbert space[26]. Notice there is a non-standard use of notation. The subscript denotes the particular wavelet and not the dilation level as used previously. A second subscript is used to denote an integer translation.

For a particular example, let $m = 2$ (corresponding to the linear splines, functions which are piecewise linear with knots at the integers) and let the dilation factor $p = 3$. The two semi-orthonormal wavelets are shown in Figure 4.2 below. They are the same ones calculated by Lian in [25]. The values of these wavelets are normalized versions of thos

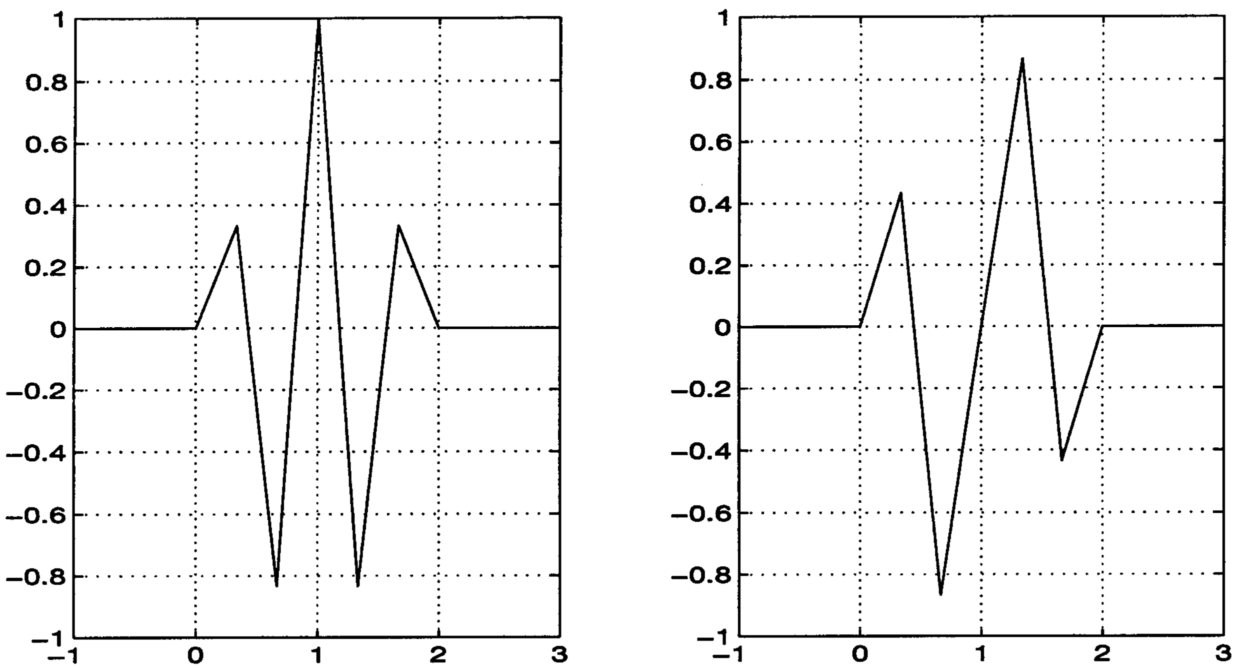

Figure 4.2 Linear Spline Semiorthonormal Wavelets for $p = 3$

calculated in Example 4.3. Details on their calculation are contained in Appendix D.

The minimization function is

$$\begin{aligned}
||\psi_2 - f_\alpha||^2 &= ||\psi_2 - \sum_k \alpha_k \psi_{1,k}||^2 \\
&= \langle \psi_2 - \sum_k \alpha_k \psi_{1,k},\ \psi_2 - \sum_l \alpha_l \psi_{1,l} \rangle \\
&= \langle \psi_2, \psi_2 \rangle + \langle \sum_k \alpha_k \psi_{1,k}, \sum_l \alpha_l \psi_{1,l} \rangle - 2 \langle \psi_2, \sum_k \alpha_k \psi_{1,k} \rangle
\end{aligned}$$

where the last step uses the fact that the wavelets are real-valued. The minimum of the function will occur at the point were the gradient vanishes. Define

$$J_\alpha = \frac{1}{2}\langle \sum_k \alpha_k \psi_{1,k}, \sum_l \alpha_l \psi_{1,l}\rangle - \langle \psi_2, \sum_k \alpha_k \psi_{1,k}\rangle.$$

We want to solve for α such that

$$[\nabla J_\alpha]_i = \frac{\partial J_\alpha}{\partial \alpha_i} = 0 \quad \forall i \in \mathbb{Z}.$$

The gradient of J_α is given by

$$\frac{\partial J(\alpha)}{\partial \alpha_i} = \sum_k \alpha_k \langle \psi_{1,k}, \psi_{1,i}\rangle - \langle \psi_2, \psi_{1,i}\rangle.$$

The minimizer of J_α satisfies

$$\mathbf{C}\alpha = \mathbf{b}$$

where

$$[\mathbf{C}]_{i,k} = \langle \psi_{1,k}, \psi_{1,i}\rangle$$

and

$$[\mathbf{b}]_i = \langle \psi_2, \psi_{1,i}\rangle.$$

So, to find this minimizer, we need to find an inverse to the matrix $\mathbf{C}$ (if possible) so that

$$\alpha_{\min} = \mathbf{C}^{-1}\mathbf{b}.$$

There is a great deal of structure to the matrix $\mathbf{C}$. Specifically, since we are dealing with ψ_1 which is compactly supported, the elements of $\mathbf{C}$ will be non-zero only in a localized region about the main diagonal. Furthermore, $\mathbf{C}$ will be a symmetric Toeplitz matrix since one row will be a shifted version of some fundamental row, similar to the matrix $\mathbf{A}$ discussed previously. Thus, the operation of $\mathbf{C}$ on any vector is equivalent to a discrete convolution of the fundamental row and the vector followed by downsampling in some cases. Thus the problem can be formulated as an inverse filter problem.

Let $\mathbf{c}$ be the fundamental column of $\mathbf{C}$ defined by

$$[\mathbf{c}]_k = [\mathbf{C}]_{k,0} = \langle \psi_{1,0}, \psi_{1,k} \rangle.$$

Then

$$[\mathbf{C}\alpha]_k = (\mathbf{c} * \alpha)(k)$$

where $*$ denotes discrete convolution and in the Fourier domain:

$$\sum_k (\mathbf{c} * \alpha)(k) e^{-i\omega k} = C(e^{i\omega}) A(e^{i\omega})$$

where C and A are the discrete Fourier transforms of $\mathbf{c}$ and α, respectively. Consequently, we want to solve for

$$A(e^{i\omega}) = B(e^{i\omega}) / C(e^{i\omega}) \quad \omega \in \mathbb{R}.$$

Looking again at the definition of C, we can make an interesting observation.

$$\begin{aligned}
[\mathbf{c}]_k &= \langle \psi_{1,0}, \psi_{1,k} \rangle \\
&= \frac{1}{2\pi} \langle \widehat{\psi_{1,0}}, \widehat{\psi_{1,k}} \rangle \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{\psi}_1(\xi)|^2 e^{ik\xi} \, d\xi \\
&= \frac{1}{2\pi} \sum_l \int_0^{2\pi} |\hat{\psi}_1(\xi + 2\pi l)|^2 e^{ik(\xi + 2\pi l)} \, d\xi \\
&= \frac{1}{2\pi} \int_0^{2\pi} \sum_l |\hat{\psi}_1(\xi + 2\pi l)|^2 e^{ik\xi} \, d\xi
\end{aligned}$$

which implies

$$C(e^{i\xi}) = \sum_l |\hat{\psi}_1(\xi + 2\pi l)|^2$$

$$= E_{\psi_1}(e^{i\xi})$$

With this in mind, we can express the second wavelet (orthogonal to the first) as

$$\begin{aligned}\hat{\psi}_2'(\xi) &= \hat{\psi}_2(\xi) - B(e^{i\xi})\frac{\hat{\psi}_1(\xi)}{E_{\psi_1}(e^{i\xi})} \\ &= \frac{E_{\psi_1}(e^{i\xi})\hat{\psi}_2(\xi) - B(e^{i\xi})\hat{\psi}_1(\xi)}{E_{\psi_1}(e^{i\xi})}\end{aligned} \tag{4.3}$$

The connection between this approach and a typical Gramm-Schmidt procedure is evident in the expression above. In Equation 4.3, the portion of $\hat{\psi}_1$ being subtracted from $\hat{\psi}_2$ is exactly the orthogonal projection of ψ_2 onto the space spanned by the integer translations of ψ_1. To see this, write the dual of $\hat{\psi}_1$ as

$$\hat{\tilde{\psi}}_1(\xi) = \frac{\hat{\psi}_1(\xi)}{E_{\psi_1}(e^{i\xi})}.$$

Then

$$\begin{aligned}\widehat{P^{\perp}_{\psi_1}\psi_2}(\xi) &= \mathcal{F}\left\{\sum_{k\in\mathbf{Z}}\langle\psi_2, \psi_1(\cdot - k)\rangle\tilde{\psi}_1(\cdot - k)\right\}(\xi) \\ &= \sum_{k\in\mathbf{Z}}\langle\psi_2, \psi_1(\cdot - k)\rangle\hat{\tilde{\psi}}_1(\xi)e^{-ik\xi} \\ &= \frac{B(e^{i\xi})}{E_{\psi_1}(e^{i\xi})}\hat{\psi}_1(\xi).\end{aligned}$$

Writing the orthogonalized version of ψ_2' via the "orthogonalization trick", we have:

$$\hat{\psi}_2^{o\prime}(\xi) = \frac{E_{\psi_1}(e^{i\xi})\hat{\psi}_2(\xi) - B(e^{i\xi})\hat{\psi}_1(\xi)}{\left(\sum_k \left|E_{\psi_1}(e^{i\xi})\hat{\psi}_2(\xi + 2\pi k) - B(e^{i\xi})\hat{\psi}_1(\xi + 2\pi k)\right|^2\right)^{\frac{1}{2}}}.$$

Define R by

$$R(e^{i\xi}) = E_{\psi_1}(e^{ip\xi})Q_2(e^{i\xi}) - B(e^{ip\xi})Q_1(e^{i\xi}).$$

where Q_1 and Q_2 correspond to the sequences which define ψ_1 and ψ_2 in terms of N_m. Rewriting the expression for $\hat{\psi}_2^{o\prime}$ as

$$\hat{\psi}_2^{o\prime}(\xi) = \frac{R(e^{i\xi/p})\hat{N}_m(\xi/p)}{(p\,(D_p|R|^2 E_{N_m})(e^{i\xi}))^{\frac{1}{2}}}$$

implies that R corresponds to a sequence which defines a semi-orthogonal wavelet. That is, the wavelet defined by

$$\psi_2'(t) = \sum_k r(k) N_m(pt - k)$$

satisfies

$$\langle \psi_2'(\cdot - n), \psi_1 \rangle \quad \forall n \in \mathbb{Z}$$

Notice that the definition for R involves sequences which are compactly supported so that the resulting semi-orthonormal wavelet ψ_2' is compactly supported as well. The wavelet corresponding to the particular example is shown in Figure 4.3. We are assured this new

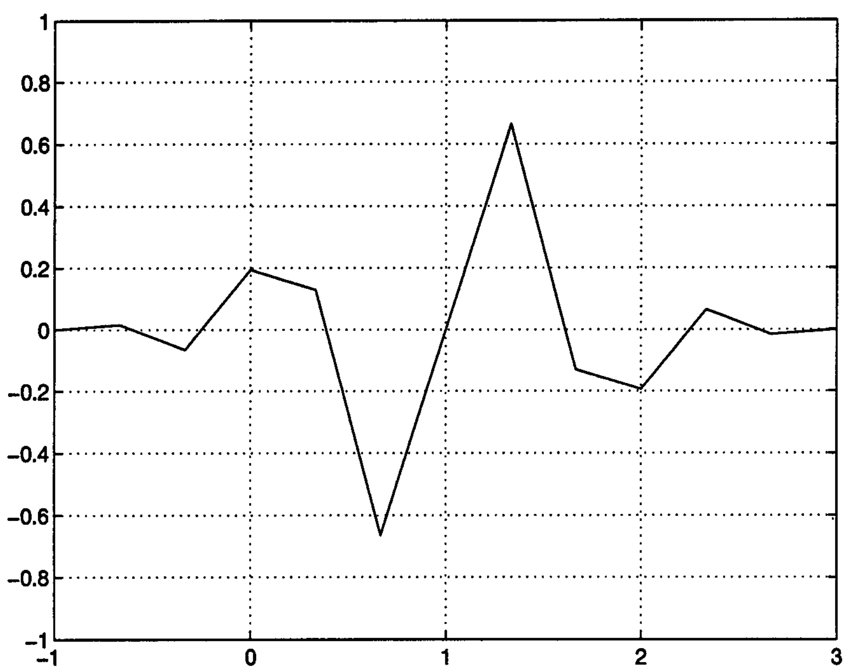

Figure 4.3 Linear Spline Semiorthonormal Wavelet for $p = 3$

semi-orthonormal wavelet in conjuction with the first span the same space as the original two since Equation 4.3 is a Fourier domain expression for the Gramm-Schmidt procedure.

A Fourier domain representation of the associated orthonormalized scaling and detail filters is useful in visualizing the results of this technique. Figure 4.2.3 is the magnitude-squared frequency response of filters associated with the the wavelets of the previous example. Performing the same type of analysis with $m = 4$, we can get the filters shown in Figure 4.2.3.

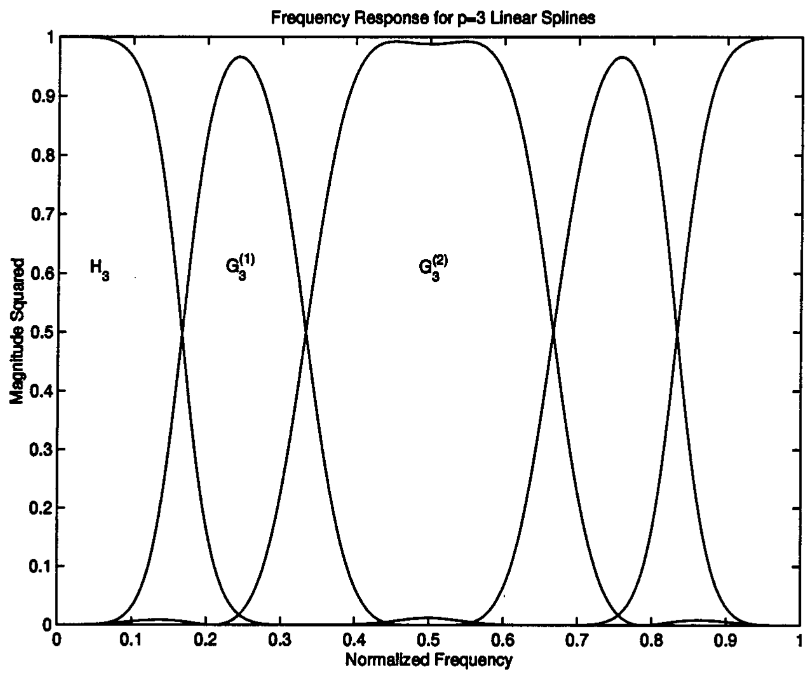

Figure 4.4 Frequency Response of $p = 3$ Linear Spline ($m = 2$) Approximation and Detail Filters

For larger dilation factors, the technique outlined in the previous example is essientially the same. The Gramm-Schmidt procedure is performed iteratively so that at step k, a new semi-orthogonal wavelet is constructed which is orthogonal to the previous $k - 1$ semi-orthogonal wavelets. The k-th semi-orthogonal wavelet can be written as

$$\hat{\psi}'_k(\xi) = \hat{\psi}_2(\xi) - \sum_{l=1}^{k-1} \frac{B_l^{(k)}(e^{i\xi})}{E_{\psi'_l}(e^{i\xi})} \hat{\psi}'_l(\xi) \tag{4.4}$$

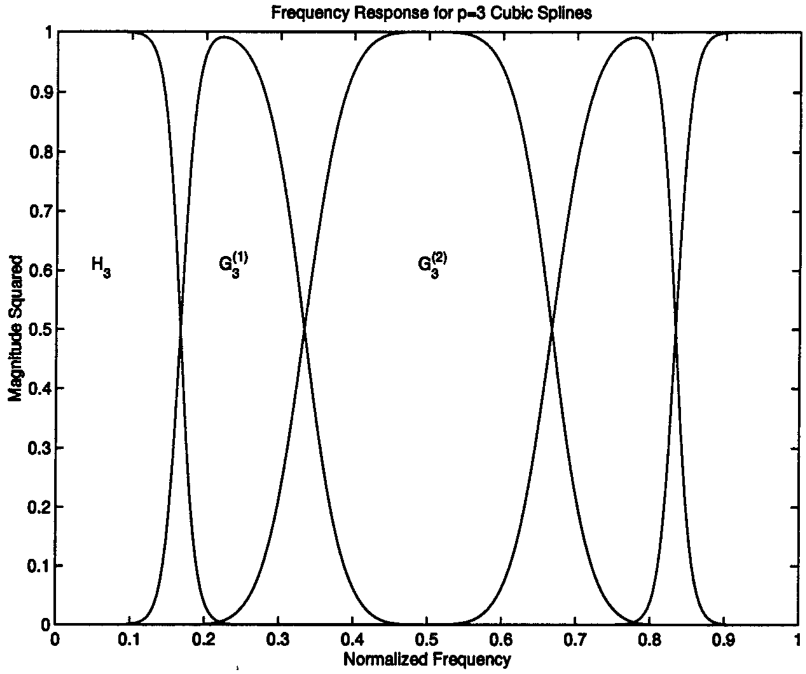

Figure 4.5 Frequency Response of $p = 3$ Cubic Spline ($m = 4$) Approximation and Detail Filters

where, as expected,

$$B_l^{(k)}(e^{i\xi}) = \sum_n \langle \psi_k, \psi_l \rangle e^{i\xi n}$$

and $\psi_1' = \psi_1$. By expressing $\hat{\psi}_k'$ in (4.4) as a rational function via a common denominator, the compactly supported filter coefficients can be extracted in a manner similar to the example above.

As an example, Figure 4.2.3 is the magnitude-squared frequency response of a set of linear spline approximation and detail filters when $p = 4$.

From a practical standpoint, there are several things which must be considered. As with any Gramm-Schmidt orthogonalization procedure, the order in which the vectors are "orthogonalized" has a great deal of effect on the final set of vectors. Furthermore, Gramm-Schmidt can have notoriously poor numerical properties[18], especially when the angle between two vectors is small.

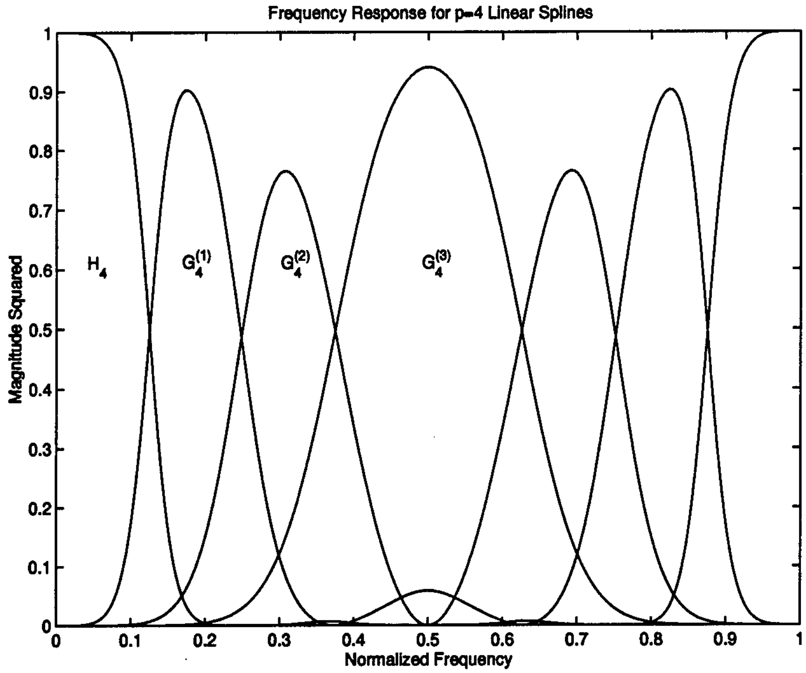

Figure 4.6 Frequency Response of $p = 4$ Linear Spline ($m = 2$) Approximation and Detail Filters

Since it is generally desirable to have detail filters which have well-localized spectra, some design criteria must be applied to the construction of the semi-orthogonal wavelets. The technique used in this research is as follows.

1. The nullspace vectors of $\mathbf{A}'$ are numerically calculated. From these vectors, a symmetric and anti-symmetric set are formed from the sum and difference of the original nullspace vectors and their "flipped" versions. If possible, these vectors are renormalized so they have integer values. This is generally possible since it is assumed that elements of $\mathbf{A}'$ have a similar integer-valued normalization.

2. The discrete Fourier transforms are calculated for each of the vectors. For each of the discrete frequency bins $\{[\pi k/p, \pi(k+1)/p]\}_{k=1}^{p-1}$, a nullspace vector is chosen which has its Fourier transform generally well-localized in that bin. This is a generally subjective choice since for some bins, especially those corresponding to the lower frequencies, no

one spectrum is localized there. In this case, the spectrum with the largest contribution is chosen.

3. These $p-1$ nullspace vectors are ordered according to the frequency bin with which they were associated. Then, starting with the higest frequency and progressing to the lowest, the vectors are "orthogonalized" via the Gramm-Schmidt procedure discussed above. If the vectors were chosen to have good frequency independence, the procedure will generally produce a set of $p-1$ vectors corresponding to semi-orthogonal wavelets.

A short remark about the last item above is in order. There are two cases where the technique fails to produce the desired results. First, if one of the selected nullspace vectors is a linear combination of the integer translations of the others, the Gramm-Schmidt procedure will produce the zero vector for that step in the iteration. Substituting another nullspace vector with similar frequency characteristics usually solves this problem. The second failure occurs when the Euler-Frobenius polynomials for the respective wavelets, E_{ψ_k} has a singular point. That is, $E_{\psi_k}(e^{i\xi}) = 0$ for some $\xi \in [0, 2\pi)$. It is fairly easy to avoid this situation by judicious choice of the wavelets. Consider the following:

$$\begin{aligned}
E_{\psi_k}(e^{i\xi}) &= \sum_l |\hat{\psi}_k(\xi + 2\pi l)|^2 \\
&= \frac{1}{p}\sum_l |Q_k(e^{i(\xi+2\pi l)/p})\widehat{N_m}((\xi + 2\pi l)/p)|^2 \\
&= \frac{1}{p}\sum_{j\in\mathbb{Z}_p} |Q_k(e^{i(\xi+2\pi j)/p})\sum_l \widehat{N_m}((\xi + 2\pi j)/p) + 2\pi l)|^2 \\
&= D_p\{|Q_k|^2 E_{N_m}\}(\xi).
\end{aligned}$$

Since E_{N_m} has no zeros on the unit circle, E_{ψ_k} will vanish only if $|Q_k|^2$ has a p ring of roots. This occurrence can be avoided by inspection of the spectra of the nullspace vectors.

4.3 Summary

This chapter presented an algorithmic approach to the construction of spline-based wavelets for an MRA with an arbitrary dilation factor. One of the major results of this chapter is the analysis of the relationship between the nullspace vectors of the integer-valued truncated system matrix $\mathbf{A}'$ and W_0, the $L_2(\mathbb{R})$ detail space associated with an MRA. Specifically, there is an isomorphism between the two spaces. Constructing spline wavelets can be cast as calculating nullspace vectors of special approximation operators.

The $p-1$ wavelets associated with a p-dilation MRA are not unique. There are degrees of freedom which can be exploited to construct wavelets which are optimized to criteria determined by the specific application. In this chapter, the wavelets which were constructed in the various examples were roughly optimized with localized bandwidths in mind. But the wavelets can be optimized with other criteria in mind simply be choosing the initial vectors in another fashion.

V. RRA Frame Characteristics

5.1 Introduction and Problem Statement

Wavelets associated with multiresolution analyses form orthonormal bases for $L_2(\mathbb{R})$. Because of the nice properties of orthonormal bases, the wavelet coefficients, $d_{k,n}^{(j)} = \langle f, \psi_{k,n}^{(j)} \rangle$, uniquely characterize f in the following sense. If $f, g \in L_2(\mathbb{R})$, and $\langle f, \psi_{k,n}^{(j)} \rangle = \langle g, \psi_{k,n}^{(j)} \rangle$ for all $k, n \in \mathbb{Z}$, then $f = g$ a.e. on $\mathbb{R}$ Since the MRA wavelets form an orthonormal basis, the above expression holds as a simple consequence. Furthermore, an algorithm to reconstruct a given function from its detail coefficients always exists via the projection theorem and is stable, specifically,

$$f = \sum_n \sum_k \sum_{j=1}^{p-1} \langle f, \psi_{k,n}^{(j)} \rangle \psi_{k,n}^{(j)}.$$

A stable reconstruction implies that if two detail coefficient sequences are "close" in the l_2 topology, then the corresponding functions are "close" in the $L_2(\mathbb{R})$ Hilbert space topology as well. Stability is a practical issue in the sense that it is not desireable to have small changes in the detail coefficients result in large changes in the reconstructed function. This is important for noise rejection. Since the wavelets associated with a rational resolution analysis are not orthogonal, these properties of unique characterization and stable reconstruction are not automatic and must be studied.

The focus of this chapter is to study the extent to which RRA detail coefficients provide a unique characterization for functions in V_0 and to what extent the reconstruction is stable. That is, we want to study the RRA detail coefficients

$$d_{k,n}^{(j)} = \langle f_{k-1}, \psi_{k,n}^{(j)} \rangle \tag{5.1}$$

for $f_0 \in V_0$. The two issues of unique characterization and stability are actually related and within the context of a Hilbert space, lead naturally to the concept of *frames.*

Definition 5.1. [Duffin and Schaeffer] A set of elements $\{\varphi_n\}_{n\in\mathbb{Z}}$, in a separable Hilbert space, $\mathcal{H}$, is a *frame* provided there exist positive constants A and B such that

$$A\|f\|^2 \le \sum_{n\in\mathbb{Z}} |\langle f, \varphi_n\rangle|^2 \le B\|f\|^2$$

holds for all $f \in \mathcal{H}$[17]. The positive constants A and B are known as lower and upper *frame bounds*, respectively. ∎

If $\{\varphi_n\}$ is a frame for a Hilbert space $\mathcal{H}$, then the *frame operator* $T : \mathcal{H} \to \mathcal{H}$ is defined by

$$Tf = \sum_{n\in\mathbb{Z}} \langle f, \varphi_n\rangle \varphi_n.$$

The following theorem is taken from [6] (Theorem 3.2).

Theorem 5.2. *Let $\{\varphi_n\}$ be a frame for $\mathcal{H}$ and let T be the associated frame operator. Then the following hold:*

1. *$AI \le T \le BI$, where A and B are the frame bounds and I is the identity operator on $\mathcal{H}$ and inequality is in the sense that $AI \le T \Longrightarrow \langle Ax, x\rangle \le \langle Tx, x\langle\ \forall x \in \mathcal{H}$.*
2. *T is invertible and $B^{-1}I \le T^{-1} \le A^{-1}I$.*
3. *$\{T^{-1}\varphi_n\}$ is also a frame with frame bounds B^{-1} and A^{-1}; this frame is known as the* dual frame *of $\{\varphi_n\}$.*
4. *For every $f \in \mathcal{H}$, $f = \sum_n \langle f, T^{-1}\varphi_n\rangle\varphi_n = \sum_n \langle f, \varphi_n\rangle T^{-1}\varphi_n$.*

Proof. See Theorem 3.2 in [6].

∎

The existence of the lower frame bound ensures the frame coefficients uniquely characterize f. That is, a non-zero lower frame bound is necessary for the following analysis to hold. Let $f, g \in \mathcal{H}$.

$$\langle f, \varphi_n\rangle = \langle g, \varphi_n\rangle \Longrightarrow \langle f - g, \varphi_n\rangle = 0 \quad \forall n \in \mathbb{Z}$$

$$\Longrightarrow \sum_n |\langle f-g, \varphi_n\rangle|^2 = 0 \quad \forall n \in \mathbb{Z}$$

which is a contradiction of the frame condition unless $f = g$ since we assumed $A > 0$ existed. Thus, the lower frame bound characterizes in some sense the uniqueness of the representations. The smaller the lower bound, the "less" unique the representation

It is possible, in principle, to reconstruct f from its frame coefficients if T is only an injective operator. However the stability requirement forces some reasonable limit on how much a function can change due to small changes in its detail coefficients. This amount of change is measured in the appropriate topology. In a rigorous sense, we want the frame operator T to be a bounded operator[19]:

$$\|T\| = \sup_{f\in\mathcal{H}} \left\{ \frac{\|Tf\|}{\|f\|} \right\} = C < \infty.$$

The upper frame bound condition is similar to the one above and also characterizes the continuity of the frame operator since it is linear. It can be shown that the existence of the upper frame bound B implies the existence of C. Consider the following analysis:

$$\begin{aligned}
\sum_n |\langle f, \varphi_n\rangle|^2 &= \sum_n \langle f, \varphi_n\rangle \overline{\langle f, \varphi_n\rangle} \\
&= \sum_n \langle f, \varphi_n\rangle \langle \varphi_n, f\rangle \\
&= \left\langle \sum_n \langle f, \varphi_n\rangle \varphi_n, f \right\rangle \\
&= \langle Tf, f\rangle.
\end{aligned}$$

The frame condition is then equivalent to

$$A\|f\|^2 \leq \langle Tf, f\rangle \leq B\|f\|^2 \quad \forall f \in \mathcal{H}$$

where $A, B > 0$ as usual. Now, $\langle Tf, f\rangle = \|Tf\|\|f\| \cos\gamma_f$ where γ_f is the angle between Tf and f for a given f. We assume $\gamma \in [0, \pi/2]$ since $\langle Tf, f\rangle$ is non-negative. Moreover, we can bound $\cos\gamma_f \geq a > 0$ with $a \in (0,1]$ since otherwise would contradict the existence of the lower frame bound A in the following manner:

$$A\|f\|^2 \leq \sum_n |\langle f, \varphi_n\rangle|^2 = \langle Tf, f\rangle \quad \forall\, f \in \mathcal{H}$$

which implies

$$A\|f\| \leq \|Tf\| \cos\gamma_f \quad \forall f \in \mathcal{H}.$$

Thus, there exists no $f \in \mathcal{H}$ such that $\|f\| \neq 0$ for which $\cos\gamma_f = 0$. Therefore, by applying the upper frame bound condition, we have

$$\frac{\|Tf\|}{\|f\|} a \leq B \quad \forall f \in \mathcal{H}$$

so that

$$\frac{\|Tf\|}{\|f\|} \leq \frac{B}{a} = C < \infty \quad \forall f \in \mathcal{H}.$$

From a heuristic standpoint, it is difficult to assign physical meaning to the values of A and B. In some sense, the relative difference between the values gives an indication as to how "uniform" the frame elements are distributed in the Hilbert space. Specifically, the greater the difference between A and B, the less uniform the distribution. The magnitudes of the frame bounds give some indication as to the normalization of the frame elements and the redundancy.

A simple example of these concepts is the Euclidean plane, $\mathbb{R} \times \mathbb{R}$, with two frame elements $\varphi_1 = (r\cos\gamma, r\sin\gamma)$ and $\varphi_2 = (r\cos\gamma, -r\sin\gamma)$ where $\gamma \in [0, \pi)$. This is illustrated in Figure 5.1. Considering elements on the unit circle, $f_\theta = (\cos\theta, \sin\theta)$, the upper and lower frame bounds are $A = 2r^2\sin^2\gamma$ and $B = 2r^2\cos^2\gamma$. Thus, as the angle between the frame elements gets smaller, the difference between A and B gets larger. The case where $\gamma = \pi/4$ or $\gamma = 3\pi/4$ corresponds to an orthogonal basis and $A = B$. Scaling the frame elements simply scales the frame bounds. With $r = 1$, we have $A = B = 1$. The case where

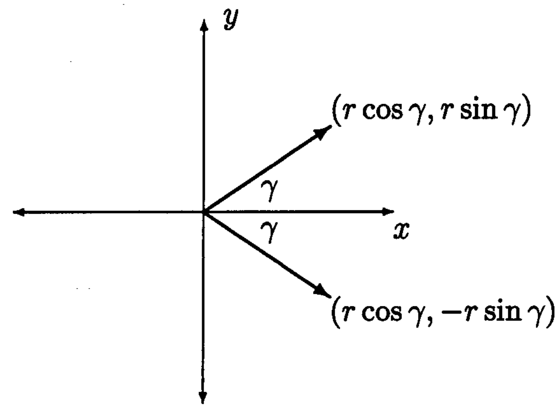

Figure 5.1 Euclidean Plane Frame Illustration with 2 Elements

$\gamma = 0$ does not correspond to a frame since $A = 0$, but it does give some insight into the effect redundancy has on the upper frame bound which is $B = 0$ in this case.

To see the effect of redundancy consider a modification of the previous example where an additional frame vector $\varphi = (r, 0)$ is included as shown in Figure 5.2. In this example, the

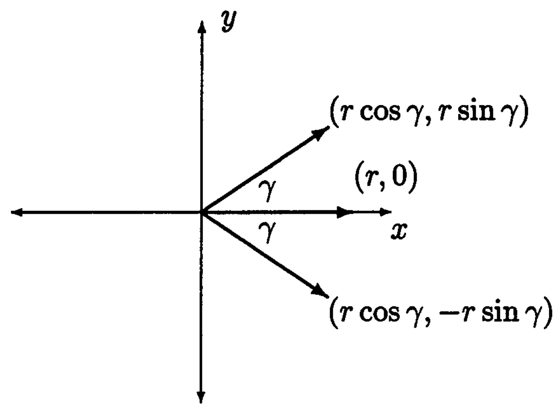

Figure 5.2 Euclidean Plane Frame Illustration with 3 Elements

upper and lower bounds are $A = 2r^2\sin^2\gamma$ and $B = r^2(2\cos^2\gamma + 1)$. Thus, the case where $\gamma = \pi/4$ corresponds to $A = 1$ and $B = 2$ (when $r = 1$). Thus, redundancy affects the upper frame bound. Notice that $A = B = 3/2$ when $\gamma = 2\pi/3$. The value of the frame bound gives an indication of the redundancy; we have three frame elements in a two dimensional space. In that case, the three frame vectors are uniformly spaced around the unit circle. If the frame vectors are normalized, $\|\varphi_n\| = 1$, and $A = B = 1$, then the frame is actually an orthonormal basis as shown in [15].

The remainder of this chapter is organized as follows. First, we discuss the RRA in the context of frame calculations and introduce adjoint wavelets. In the next section we prove the the RRA detail coefficients satisfy the frame conditions and thus prove the adjoint wavelets form a frame for an approximation space of arbitrarily fine resolution. The final section is a discussion of the frame bound estimation for the RRA and associated issues.

5.2 Adjoint Wavelets

5.2.1 Introduction and Preliminaries. The RRA has certain peculiarities which make calculating frame bounds more difficult than the MRA. It was mentioned in Chapter II that one major difference between the two was the manner in which successive approximations were performed. An elegant feature of the MRA is that the projection of an arbitrary element $f \in L_2(\mathbb{R})$ onto a particular approximation space is calculated by using the projection of f from the previous approximation space. Thus, for the MRA, the following is an equivalent statement of the frame condition:

$$0 < A\|f\|^2 \leq \sum_{j=1}^{p-1} \sum_{k\in\mathbf{Z}} \sum_{n\in\mathbf{Z}} |d_{k,n}^{(j)}|^2 \leq B\|f\|^2 < \infty. \tag{5.2}$$

With the RRA, this is not the case since the projection of f onto a particular approximation space is generally not equivalent to the projection from the previous approximation space. That is, given $f \in L_2(\mathbb{R})$, we have $P_{V_k} f \neq P_{V_k} P_{V_{k-1}} f$ in general.

This requires us to modify the RRA frame calculations in two ways. First, we assume we want to investigate the frame characteristics of the wavelets for an approximation space V_k, where k is chosen sufficiently negative so that V_k is arbitrarily "close" to $L_2(\mathbb{R})$. Without loss of generality, we can assume $k = 0$. Making this assumption gives us a "starting point" in the frame analysis. This assumption is justified in any practical application of the RRA since the "starting point" is usually defined to correspond to the sampling rate associated with the signal in question.

The second modification involves the detail coefficients defined in (5.1). Suppose we choose $f_0 \in V_0$ arbitrarily and calculate the corresponding detail coefficients for a particular RRA. The coefficients at a particular dilation level, $k > 0$, are calculated as the the inner product between the particular wavelet and f_{k-1}, not f_0. Except for $k = 1$, the detail coefficients do not represent the inner product of the original function with the dilated wavelets. Thus, it is difficult to correlate a representation such as (5.2) with the frame definition. To compensate for this, *adjoint wavelets* are introduced.

5.2.2 Definition of Adjoint Wavelets. An adjoint wavelet is defined implicitly as follows:

Definition 5.3. Within the context of an RRA, the *adjoint wavelets*, $\tilde{\psi}^{(j)}_{k,n}$, are defined by

$$\tilde{d}^{(j)}_{k,n} = \langle f_{k-1}, \psi^{(j)}_{k,n} \rangle = \langle f_0, \tilde{\psi}^{(j)}_{k,n} \rangle \quad k \in \mathbf{N} \tag{5.3}$$

■

The reason for the use of the term *adjoint* is now clear from this definition. Now, the study of the RRA detail coefficients can be posed as whether or not the adjoint wavelets form a frame for V_0.

The adjoint wavelets are easily calculated in terms of the various filter coefficients. Consider the following example.

Example 5.4. We want to calculate $\tilde{\psi}^{(j)}_{2,l}$ so that $d^{(j)}_{2,l} = \langle f_0, \tilde{\psi}^{(j)}_{2,l} \rangle$.

$$\begin{aligned}
\tilde{d}^{(j)}_{2,l} &= \langle f_1, \psi^{(j)}_{2,l} \rangle \\
&= \sum_n c_1(n) g^{(j)}(n - pl) \\
&= \sum_n \sum_k c_{0,n} h_r(qk - pn)\, g^{(j)}(n - pl)
\end{aligned}$$

$$= \sum_n c_{0,n} \sum_k h_r(qn - pk)\, g^{(j)}(k - pl)$$
$$= \left\langle f_0, \sum_k h_r(qn - pk)\, g^{(j)}(k - pl)\phi_{0,n} \right\rangle$$

which implies

$$\tilde{\psi}^{(j)}_{2,l} = \sum_n \sum_k h_r(qn - pk)\, g^{(j)}(k - pl)\phi_{0,n}.$$

■

Generalizing the result of this example, we can explicitly write an expression for the adjoint wavelets in terms of the approximation and detail filter coefficients:

$$\tilde{\psi}^{(j)}_{k,l_k} = \sum_{l_0,l_1,\ldots,l_{k-1}\in\mathbf{Z}} \sum \cdots \sum \phi_{0,l_0} h_r(ql_0 - pl_1) h_r(ql_1 - pl_2) \cdots h_r(ql_{k-2} - pl_{k-1}) g^{(j)}(l_{k-1} - pl_k)$$

Recognizing this expression as the sequential projection of $\psi^{(j)}_{k,l_k}$ from one approximation space to the previous, we can succinctly define the adjoint wavelets by

$$\tilde{\psi}^{(j)}_{k,l_k} = P^*_{W_k} \psi^{(j)}_{k,l_k}$$

where

$$P^*_{W_k} = P_{V_0} P_{V_1} P_{V_2} \cdots P_{V_k}.$$

Notice that $P^*_{W_k}$ is simply the adjoint of the sequence of orthonormal projection operators which take f to W_k via an RRA.

5.2.3 Characteristics of Adjoint Wavelets. As a side effect of the RRA, the adjoint wavelets have several unfortunate characteristics. The major disadvantage of the adjoint wavelet approach is translational dependency. Consider Example 5.4. again. Ideally we would like to have the subscripts on the adjoint wavelets represent integer translations of the "mother" adjoint wavelet. This is not the case as seen from the following analysis.

With the RRA, we desire $\tilde{\psi}^{(j)}_{k,l}(t) = \tilde{\psi}^{(j)}_{k,0}(t - r^k q l)$ similar to the MRA where we desire $\psi^{(j)}_{k,l}(t) = \psi^{(j)}_{k,0}(t - p^k l)$. Consider the case where $k = 2$. Let $l \in \mathbb{Z}$ be fixed arbitrarily. Define integers u and m by $l = qu + m$ where $m \in \mathbb{Z}_q$. That is, let $u = l$ div q and $m = l \bmod q$.

$$\begin{aligned}
\tilde{\psi}^{(j)}_{2,qu+m}(t) &= \sum_n \sum_i h_r(qn - pi) g^{(j)}(i - pqu - pm) \phi_{0,n}(t) \\
&= \sum_n \sum_i h_r(q(n - p^2 u) - pi) g^{(j)}(i - pm) \phi_{0,n}(t) \\
&= \sum_n \sum_i h_r(qn - pi) g^{(j)}(i - pm) \phi_{0,n+p^2 u}(t) \\
&= \sum_n \sum_i h_r(qn - pi) g^{(j)}(i - pm) \phi_{0,n}(t - p^2 u) \\
&= \tilde{\psi}^{(j)}_{2,m}(t - p^2 u).
\end{aligned}$$

The analysis above is easily generalized to other dilation levels k:

$$\tilde{\psi}^{(j)}_{k,q^{k-1}u+m}(t) = \tilde{\psi}^{(j)}_{k,m}(t - p^k u). \tag{5.4}$$

This indicates that at a given dilation level, k, the adjoint wavelets are translated versions of a set of q^{k-1} adjoint wavlets. That the set of q^{k-1} adjoint wavelets are unique is evident by observing the translation of a single adjoint wavelet by $r^k q$ forces a misalignment of the knots of the representation in V_0. As k increases, number of unique wavelets grows exponentially.

While this is generally discouraging from an analysis point of view, in practice, the actual wavelets are never used in their continuous forms. Only the corresponding filters are used and we can determine the frame characteristics directly from the filters and never concern ourselves with the wavelets. However, it is important to realize the role the adjoint wavelets play in the frame analysis since it is they that form the frame for an arbitrary V_0 and not the wavelets used at each level of the decomposition. The majority of the attention from this point forward will be paid to the analytical and numerical existence and estimation of the RRA frame bounds.

5.3 *Existence and Estimation of RRA Frame Bounds*

5.3.1 Preliminaries. Given a spline-based RRA with an arbitrary dilation and regularity, the adjoint wavelets associated with a particular approximation space V_k form a frame for that space. The proof of the existence of the upper and lower frame bounds are treated separately. The following lemmas will be useful in both cases.

Lemma 5.5. *If V_{k-1} and V_k are adjacent approximation spaces in a spline-based RRA, then*

$$V_k' = V_{k-1} \cap V_k$$

where V_k' is the sub-approximation space defined earlier.

Proof. As usual, let p/q be the rational dilation factor and m be the regularity factor. Without loss of generality, assume $k = 1$. We need to show subset inclusion in both directions. From the definition and the fact we are dealing with spline-based scaling functions and wavelets, we have $V_1' \subseteq V_0 \cap V_1$. Thus we need to show $V_0 \cap V_1 \subseteq V_1'$.

Choosing $f \in V_0 \cap V_1$ arbitrarily, there exist l_2 sequences c_0 and c_1 such that

$$\begin{aligned} f(t) &= \sum_k c_0(k) N_m(t-k) \\ &= \sum_k c_1(k) N_m(qt/p - k), \end{aligned}$$

where the above expressions represent expansions in the standard non-orthogonal B-spline bases for V_0 and V_1. Expressing these relationships in the Fourier domain yields

$$\begin{aligned} \hat{f}(\xi) &= C_0(e^{i\xi}) \widehat{N_m}(\xi) \\ &= p/q\, C_1(e^{ip\xi/q}) \widehat{N_m}(p\xi/q), \end{aligned}$$

from which we have

$$q\hat{f}(q\xi) = q\, C_0(e^{iq\xi}) \widehat{N_m}(q\xi)$$

$$= p\, C_1(e^{ip\xi})\widehat{N_m}(p\xi).$$

From the definition of m-th order cardinal B-spline, we have

$$|\widehat{N_m}(\xi)| = \left|\frac{1-e^{i\xi}}{i\xi}\right|^m, \quad \xi \in \mathbb{R}.$$

For the interval $[0, 2\pi)$, this implies the zero sets of $\widehat{N_m}(q\cdot)$ and $\widehat{N_m}(p\cdot)$ are $\{2\pi l/q\}_{l=1}^{q-1}$ and $\{2\pi l/p\}_{l=1}^{p-1}$, respectively. Furthermore, since the two expressions for $q\hat{f}(q\xi)$ above are equivalent and are continuous, their zero sets must be equal. This has useful implications for C_0 and C_1. Specifically, the zero sets of $\widehat{N_m}(p\xi)$ and $\widehat{N_m}(q\xi)$ intersect only at $\xi = 2\pi l$ for $l \in \mathbb{Z}$ and $l \neq 0$ since p and q are relatively prime.

To see this, suppose there was a point r in $(0, 2\pi)$ which was a zero point of both $\widehat{N_m}(p\xi)$ and $\widehat{N_m}(q\xi)$. Then there exist $k, l \in \mathbf{N}$ with $k < p$ and $l < q$ such that $pr = 2\pi k$ and $qr = 2\pi l$. Equating r in both of these expressions leads to $p/q = k/l$. Since $k < p$ and $l < q$, this implies there exists an integer $j > 1$ such that $p = kj$ and $q = lj$, or that j is a factor common to both p and q. This contradicts that p and q are relatively prime. Hence, no such r can exist.

The result is that $C_0(e^{iq\xi})$ must vanish for $\xi = 2\pi l/p$ for $l = 1, 2, \ldots, p-1$ and, conversely, $C_1(e^{ip\xi})$ must vanish for $\xi = 2\pi l/q$ for $l = 1, 2, \ldots, q-1$. Looking at C_0 for now, the following holds due to the 2π-periodicity of C_0 and the relative primeness of p and q

$$\left\{C_0(e^{i2\pi ql/p})\right\}_{l=1}^{p-1} = \left\{C_0(e^{i2\pi l/p})\right\}_{l=1}^{p-1} = \{0\}_{l=1}^{p-1}.$$

The validity of this last statement holds since $\{ql \bmod p\}_{l\in\mathbb{Z}_p} = \mathbb{Z}_p$ which is a restatement of Lemma A.1.

Thus, it follows that $S_p(e^{i\xi})$ divides $C_0(e^{i\xi})$. More accurately, $S_p^m(e^{i\xi})$ divides $C_0(e^{i\xi})$ since the degree of the zeros is m. Thus, with $\tilde{C}_0$ defined by $\tilde{C}_0 S_p^m = C_0$,

$$\hat{f}(q\xi) = \tilde{C}_0(e^{iq\xi}) S_p^m(e^{iq\xi}) \widehat{N_m}(q\xi),$$

or

$$\hat{f}(\xi) = \tilde{C}_0(e^{i\xi}) S_p^m(e^{i\xi}) \widehat{N_m}(\xi)$$

The last two factors in the expression above combine via the cardinal B-spline identity:

$$\widehat{N_m}(p\xi) = S_p^m(e^{i\xi}) \widehat{N_m}(\xi)$$

so that

$$\hat{f}(\xi) = \tilde{C}_0(e^{i\xi}) \widehat{N_m}(p\xi)$$

which is to say that f can be written as

$$f = \sum_n \tilde{c}_0(n) N_m(\cdot/p - n)$$

which implies $f \in V_1'$. Since f was chosen arbitrarily in the intersection of V_0 and V_1, the analysis holds for all such elements so that $V_0 \cap V_1 \subseteq V_1'$ which completes the proof.

■

Corollary 5.6. *Let V_{k-1} and V_k be adjacent approximation spaces in a spline-based RRA. Let V_k' be the associated sub-approximation space as in the previous lemma and let W_k be the usual detail space which is the orthogonal complement of V_k' in V_{k-1}. Define W_k' to be the orthgonal complement of V_k' in V_k. That is, $V_k' \oplus W_k' = V_k$. Then*

$$W_k \cap W_k' = \{0\}.$$

Proof. We need to show subset inclusion in both directions. Without loss of generality, assume $k = 1$ as in the previous lemma. We have $\{0\} \subseteq W_k \cap W_k'$ by definition. Now, choose $g \in W_k \cap W_k'$ arbitrarily. This implies $g \in V_{k-1} \cap V_k$, which implies via Lemma 5.5. that $g \in V_k'$. Since $g \in W_k'$ which is orthogonal to V_k', we must have $g = 0$. Similarly, $g \in W_k \cap V_1'$ imples $g = 0$. Since g was chosen arbitrarily, the analysis holds for all such g which implies $W_k \cap W_k' = \{0\}$ and completes the proof.

■

5.3.2 Existence of an RRA Lower Frame Bound. A non-zero lower bound indicates the RRA frame operator has a bounded inverse and that the detail coefficients for a particular signal is a unique characterization of that signal. The proof of a non-zero lower bound is provided in the following Theorem.

Theorem 5.7. *For an RRA, the adjoint wavelets satisfy the lower frame bound condition for V_0 with $A = 1$.*

Proof. Assume f_0 is an arbitrary unit-norm function in V_0 and that $f_k = P_{V_k} f_{k-1}$ as usual. At each iteration of an RRA decomposition, the following holds:

$$||f_{k-1}||^2 = ||P_{V_k'} f_{k-1}||^2 + ||P_{W_k} f_{k-1}||^2$$

since V_k' and W_k form an orthogonal decomposition of V_{k-1}. Furthermore,

$$||P_{V_k} f_{k-1}||^2 = ||f_k||^2 = ||P_{V_k'} f_k||^2 + ||P_{W_k'} f_k||^2 \tag{5.5}$$

since since V_k' and W_k' form an orthogonal decomposition of V_k. Substituting $P_{V_k'} f_{k-1} = P_{V_k'} f_k$ in the previous expression leads to

$$||P_{V_k} f_{k-1}||^2 \geq ||P_{V_k'} f_{k-1}||^2$$

with equality if and only if $P_{W_k'} f_k = 0$. Substituting this expression into (5.5) and iterating on the index k yields

$$\sum_{k=1}^{N} \|P_{W_k} f_{k-1}\|^2 + \|f_N\|^2 \geq \|f_0\|^2 = 1.$$

Taking the limit as $N \to \infty$ and using that

$$\lim_{N\to\infty} \|f_N\| = 0,$$

which is a consequence of the fact that $\cap_{j\in\mathbf{Z}} V_j = \{0\}$ which was proven by Daubechies[15] independent of dilation factor and the embedding of approximation spaces. This gives us the following result:

$$\lim_{N\to\infty} \sum_{k=1}^{N} \|P_{W_k} f_{k-1}\|^2 = \sum_{k=1}^{\infty} \|P_{W_k} f_{k-1}\|^2 = \sum_{j=1}^{p-1} \sum_{k=1}^{\infty} \sum_{n\in\mathbf{Z}} |d_{k,n}^{(j)}|^2 \geq 1$$

which implies a lower bound $A \geq 1$ since the $d_{k,n}^{(j)}$ represent the inner product of f_0 with the corresponding adjoint wavelet.

■

This result is somewhat intuitive since the wavelets are normalized but there is energy "carry-over" from one detail space to the next since they are not orthogonal and are correlated in some sense.

One question which naturally arises is whether or not there exists some function $f \in V_0$ for which the lower frame bound is acheived. The answer to this is not clear. For the Haar case ($m = 1$) a class of functions which acheive the lower frame bound is simple to describe and contruct. However, for more regular RRA's the numerical evidence supports the conjecture that the best you can do to achieve the lower bound is to construct a sequence of functions whose lower bounds approach unity. Furthermore, it is not clear what effect the choice of dilation factors has in this result.

As indicated in the previous section, the existence of a non-zero lower bound implies the frame coefficients provide a unique representation of V_0 and that the frame mapping T

has no null space. This has the further implication that any signal can be represented via the frame coefficients. That is not to say the adjoint wavelets span the approximation space V_0. We would have to have a closed-form synthesis formula whereas now we only have a reconstruction algorithm which is iterative and the existence of a closed-form equivalent is questionable.

5.3.3 Existence of an RRA Upper Frame Bound. The proof and estimation of an upper frame bound B follows closely from the analysis of the lower frame bound. The following lemma will be useful in proving the existence of an upper frame bound for the spline-based RRA.

Lemma 5.8. *Let W_k and W_k' be detail spaces of a spline RRA as defined as in Corollary 5.6.. Then there exists an $a < 1$ such that*

$$\sup_{\substack{g \in W_k \\ g \neq 0}} \frac{\|P_{W_k'} g\|^2}{\|g\|^2} = a < 1.$$

Proof. Without loss of generality, assume $k = 1$. The first step in this proof is to show the supremum is achieved for some $g \in W_1$. Since W_k is a Hilbert space, the closed unit sphere is a weak* compact subset of W_k. Since $\|P_{W_k'} \cdot \|^2$ is a continuous real-valued functional defined on a weak* compact subset of W_k, it is bounded and achieves its maximum. For details on this result and on weak* convergence, see Luenberger[26], section 5.10.

Now suppose the maximum $a = 1$. Since W_1 is closed, there exists a $g \in W_1$ such that $\|g\| = 1$ and $\|P_{W_1'} g\| = 1$. This implies that $g \in W_1'$ since $P_{W_1'}$ is an orthogonal projector. This is a contradiction in light of Corollary 5.6. and the assumption $g \neq 0$. Hence $a < 1$ exists.

■

Theorem 5.9. *For an RRA, there exists a finite upper frame bound B for the adjoint wavelets associated with V_0.*

Proof. The analysis of the upper frame bound is ultimately concerned with the energy in the successive detail spaces. At each level in a decomposition, the energy in a particular approximation space, V_{k-1}, is partitioned into the primed approximation space, V_k', and the detail space, W_k. The energy in V_k' is directly and entirely "transferred" into V_k since there is a subset relationship between the two. The maximum fraction of the energy from the detail space W_k which is effectively "transferred" via projection from V_{k-1} to V_k is a defined in the lemma above.

Hence, if $x_k \in [0,1]$ is defined to be the fraction of the energy of V_{k-1} contained in W_k, then $1 - x_k$ is the fraction contained in V_k' (and thus V_k) and the maximum proportion transferred from the detail space will ax_k. Letting v_k denote the energy of a particular signal in V_k, we have

$$v_k \leq v_{k-1}(1 - x_k(1-a)), \tag{5.6}$$

and the total energy associated with the RRA representation of the particular signal is

$$E = \sum_{k=1}^{\infty} x_k v_{k-1}$$

and the upper frame bound is the supremum of this expression taken over all signals in V_0 such that $v_0 = 1$ and all such x_k. This supremum corresponds to the case when $x_k = 1$ identically for all $k \in \mathbf{N}$ as seen by the following analysis.

If we assume equality in (5.6) and let $b = 1 - a$ then

$$v_k = \prod_{j=1}^{k}(1 - bx_j)$$

and the upper frame bound satisfies

$$\begin{aligned} B &\leq \sup_x \sum_{k=1}^{\infty} x_k v_{k-1} \\ &\leq \sup_x \sum_{k=1}^{\infty} x_k \left[\prod_{j=1}^{k-1} 1 - bx_j\right]. \end{aligned}$$

and

$$\begin{aligned}
\sum_{k=1}^{\infty} x_k \left[\prod_{j=1}^{k-1}(1-bx_j)\right] &= \lim_{N\to\infty} \sum_{k=1}^{N} x_k \left[\prod_{j=1}^{k-1}(1-bx_j)\right] \\
&= \lim_{N\to\infty} \frac{1-\prod_{j=1}^{N}(1-bx_j)}{b} \\
&= \frac{1-\lim_{N\to\infty}\prod_{j=1}^{N}(1-bx_j)}{b} \\
&= \frac{1-\prod_{j=1}^{\infty}(1-bx_j)}{b} \\
&= \frac{1}{b} - \prod_{j=1}^{\infty} \frac{1-bx_j}{b}.
\end{aligned}$$

The second step in the derivation above is a result of an identity which can be easily proven by induction as follows.

It is clear that for $N = 1$ we have

$$b\sum_{k=1}^{N} x_k \prod_{j=1}^{k-1}(1-bx_j) = 1 - \prod_{j=1}^{N}(1-bx_j)$$

Assuming the previous statement holds for an arbitrary N, it is easy to show it holds for $N+1$ as follows:

$$\begin{aligned}
b\sum_{k=1}^{N+1} x_k \prod_{j=1}^{k-1}(1-bx_j) &= b\sum_{k=1}^{N} x_k \prod_{j=1}^{k-1}(1-bx_j) + bx_{N+1}\prod_{j=1}^{N}(1-bx_j) \\
&= 1 - \prod_{j=1}^{N}(1-bx_j) + bx_{N+1}\prod_{j=1}^{N}(1-bx_j) \\
&= 1 - \prod_{j=1}^{N}(1-bx_j)\left[1-bx_{N+1}\right] \\
&= 1 - \prod_{j=1}^{N+1}(1-bx_j)
\end{aligned}$$

Thus by induction, the statement holds for an arbitrary $N \in \mathbb{N}$. The end result is

$$\begin{aligned} B &\leq \sup_x \{\frac{1}{b} - \frac{1}{b}\prod_{j=1}^{\infty}(1 - bx_j)\} \\ &\leq \frac{1}{b} - \inf_x \{\frac{1}{b}\prod_{j=1}^{\infty}(1 - bx_j)\} \\ &\leq \frac{1}{b} = \frac{1}{1-a}. \end{aligned}$$

■

In the last step, the infinum is achieved for $x_j = 1$ identically. With the assumption $v_0 = 1$, then $B < \infty$ for all choices of x_k since $a < 1$ as seen in Lemma 5.8. This result, along with that of the previous section, indicates that the spline-based RRA satisfies the requirements for a frame. Furthermore, the result is independent of dilation factor and/or regularity.

5.3.4 Estimation of the RRA Upper Frame Bound. To actually calculate the value of a (or b) for an arbitrary dilation factor and regularity, the analysis quickly becomes more complicated. Consider the following optimization problem. Given V_1 and W_1, find a such that

$$a = \sup_{\substack{g \in W_1 \\ \|g\|=1}} \|P_{V_1} g\|^2.$$

With g expanded in terms of the orthonormal basis for W_1:

$$g = \sum_{j=1}^{p-1} \sum_n d^{(j)}(n) \psi_{1,n}^{(j)}, \quad d \in l_2^{p-1},$$

the objective function can be expressed as the inner product of $P_{V_1} g$ with itself:

$$\begin{aligned} J(d) &= \|P_{V_1} g\|^2 \\ &= \langle P_{V_1} g, P_{V_1} g \rangle \end{aligned}$$

$$= \left\langle \sum_k \sum_{j=1}^{p-1} \sum_n d^{(j)}(n) \langle \psi_{1,n}^{(j)}, \phi_{1,k} \rangle \phi_{1,k}, \sum_l \sum_{i=1}^{p-1} \sum_m d^{(i)}(m) \langle \psi_{1,m}^{(i)}, \phi_{1,l} \rangle \phi_{1,l} \right\rangle$$

$$= \sum_{j=1}^{p-1} \sum_n \sum_{i=1}^{p-1} \sum_m d^{(j)}(n) d^{(i)}(m) \left\langle \sum_k \langle \psi_{1,n}^{(j)}, \phi_{1,k} \rangle \phi_{1,k}, \sum_l \langle \psi_{1,m}^{(i)}, \phi_{1,l} \rangle \phi_{1,l} \right\rangle$$

$$= \sum_{j=1}^{p-1} \sum_n \sum_{i=1}^{p-1} \sum_m d^{(j)}(n) d^{(i)}(m) \langle P_{V_1} \psi_{1,n}^{(j)}, P_{V_1} \psi_{1,m}^{(i)} \rangle$$

$$= \sum_{j=1}^{p-1} \sum_n \sum_{i=1}^{p-1} \sum_m d^{(j)}(n) d^{(i)}(m) \langle \varphi_{1,n}^{(j)}, \varphi_{1,m}^{(i)} \rangle$$

where $\varphi_{1,n}^{(j)} = P_{V_1} \psi_{1,n}^{(j)}$. This objective function can be written succinctly in tensor notation:

$$J(d) = d_{jn} A_{jnil} d_{il}$$

where $A_{jnil} = \langle \varphi_{1,n}^{(j)}, \varphi_{1,m}^{(i)} \rangle$ is a Hermitian-symmetric fourth-order tensor. As with matrices, the Hermitian symmetry ensures the eigenvalues will be non-negative. The maximum value is the largest eigenvalue and the maximizer is the corresponding eigenvector.

This theoretical result is all well and good, but calculating a in this way is not practical because the values of the tensor are calculated by taking the inner product of one adjoint wavelet with a translated version of another. Since the adjoint wavelets are infinitely supported (albeit with exponential decay) the tensor is effectively infinite as well which makes the calculation of its eigenvectors and eigenvalues very difficult.

One approach is to use the exponential decay of the adjoint wavelets to form an approximation of the tensor which is finite. The tensor is truncated at the point where the magnitude of its values fall below some threshold (the machine precision, for instance) and the corresponding eigenvalues and eigenvectors can be calculated since the tensor is now effectively finite.

A simple example of this is to consider the $m = 1$ RRA with $p = 3$ and $q = 2$. Since the adjoint wavelets are compactly supported, we would expect the associated tensor to be finite as well. The following example illustrates this.

Example 5.10. Assume we have an RRA with $p = 3$, $q = 2$, and $m = 1$ (the piecewise constant splines). We want to calculate the constant a defined above. Without loss of generality, assume $k = 1$ as usual. Furthermore, assume we have a signal $g_1 \in W_1$ and that $\|g_1\| = 1$. We can represent g_1 as

$$g_1 = \sum_{j=1}^{2} \sum_{n} d^{(j)}(n) \psi_{1,n}^{(j)}.$$

The projection onto V_1 is given by

$$\begin{aligned} P_{V_1} g_1 &= \sum_{l} \sum_{j=1}^{2} \sum_{n} d^{(j)}(n) \langle \psi_{1,n}^{(j)}, \phi_{1,l} \rangle \phi_{1,l} \\ &= \sum_{j=1}^{2} \sum_{n} d^{(j)}(n) P_{V_k} \psi_{1,n}^{(j)}. \end{aligned}$$

With the two well-used "standard" wavelets

$$\begin{aligned} \psi_{1,0}^{(1)} &= -\sqrt{1/6}\phi_{0,0} + \sqrt{2/3}\phi_{0,1} - \sqrt{1/6}\phi_{0,2}) \\ \psi_{1,0}^{(2)} &= \sqrt{1/2}\phi_{0,0} - \sqrt{1/2}\phi_{0,2}, \end{aligned}$$

we have

$$\begin{aligned} P_{V_1} \psi_{1,0}^{(1)} &= 0 \\ P_{V_1} \psi_{1,0}^{(2)} &= (1/3)^{-1/2} (\phi_{1,0} - \phi_{1,1}) \end{aligned}$$

so that

$$P_{V_1} g_1 = (1/3)^{1/2} \sum_{n} d^{(2)}(n) (\phi_{1,2n} - \phi_{1,2n+1})$$

and

$$\|P_{V_1} g_1\|^2 = 1/3 \sum_{n} |d^{(2)}(n)|^2 \|\phi_{1,2n} - \phi_{1,2n+1}\|^2$$

$$= 2/3 \sum_n |d^{(2)}(n)|^2$$
$$\leq 2/3$$

where the last step is a consequence of the assumption that $||g_1|| = 1$. Thus we have calculated a for the piecewise consant spline RRA with dilation factor 3/2. From this we conclude $B \leq 3$. ■

The calculation of an estimate of the upper frame bound in the way described above has a severe drawback. The value of a represents the maximum fraction of energy which is "transferred" between the detail space W_k and the next approximation space V_k. The estimation of the upper frame bound is based on the worst-case scenario by which we could construct a function which could transfer that maximum fraction of energy, a, at each and every level of the RRA.

This scenario is unrealistic because it does not reflect that each level of decomposition in an RRA (or MRA for that matter) corresponds to a bandpass filtering of the original signal. The location of the passband corresponding to a particular level of decomposition relative to that of an ajacent level depends upon the dilation factor. For MRAs, adjacent passbands do not "overlap" in the sense that their corresponding detail spaces are orthogonal. For RRAs, the passbands do intersect and a particular frequency might be "covered" by the passbands corresponding to a series of ajacent levels of decomposition. The important similarity in both cases is that the center frequencies of the passbands approach zero as successive levels of decomposition are performed. Thus, a function constructed to have a maximum energy transfer for all levels of a particular RRA would necessarily have a spectrum which is not well-localized. Considering a unit-energy function, the maximum fraction of energy might be transferred at a particular level, but the amount of that energy would have to significantly less than unity which automatically ensures the estimate of the upper bound developed earlier will not be achieved.

This observation suggests another way to get a handle on the upper frame bound estimation. Suppose we assume the detail filters have a perfect passband. This corresponds to taking the regularity m to infinity. Thus, in the frequency domain we can assume the various detail filters are piecewise constant and well-localized. For sake of illustration, consider the case where the filters corresponding to particular level of an RRA decomposition occupy the frequencies from $1/p$ to 1. The next detail space will have filters which occupy the frequencies between q/p^2 and q/p. Further multiplication by q/p yields the bandwidth limits of subsequent decomposition levels.

We can estimate the upper frame bound by calculating the maximum number of adjacent detail spaces which have an intersection. Formulated in terms of the "perfect" bandwith limits in the previous paragraph, we want to find the maximum n such that

$$\left(\frac{q}{p}\right)^n \geq \frac{1}{p}.$$

Solving for n, we have

$$n = \left\lceil \frac{1}{1 - \log_p q} \right\rceil. \tag{5.7}$$

Note that for the case where $q = 1$, the MRA case, this formula for the estimate of the upper bound yields $n = 1$ which is as expected.

It is assumed the accuracy of this upper bound for the frame bound B increases monotonically with the regularity. This assumption is justified by the observation that the width of the transition regions of the detail filters becomes more narrow with the regularity and that the frame bounds are greatest when the passband is maximally flat.

5.4 Summary

The most significant result of this chapter is that the spline-based RRA is a stable and unique decomposition of the initial approximation space V_0. The supporting analysis for this result focuses on the theory of frames and, specifically, the proof of finite frame bounds. In order to apply the concept of frames, it was necessary to introduce the concept of adjoint wavelets and discuss their properties.

It was also necessary to explore the relationship between adjacent approximation spaces in terms of the projection operator induced by the RRA. It was especially necessary to characterize the projection from one to the next those functions which were specifically not in the intersection. This characterization let to an important lemma which stated the intersection between two adjacent approximation spaces, V_{k-1} and V_k, was solely the primed approximation space, V_k'. The corollary to this lemma stated that the intersection of the detail space, W_k and the primed detail space W_k' was solely the null vector. Furthermore, it was shown in a subsequent lemma that the orthogonal projector was suboptimal in preserving energy in the sense that the projected function always had less energy than the original function. This observation formed the basis for the proof of the existence of the upper frame bound.

Estimating the upper frame bound led to two different formulations, one more useful than the other, though both were illuminating. The first focused on the calculation of the frame bound as a worst-case scenario used a fourth-order tensor eigenvalue calculation. This was of limited practical use since the manipulation of the tensor was difficult and necessarily involved an approximation since certain infinite sequences had to truncated.

The second formulation dealt the limiting case where the regularity of the RRA approached infinity and the relative bandwidths of the filters associated with adjacent detail spaces as defined by the dilation factor. As regularity increased, the bandwidths of the various filters became more localized with sharper cutoffs. The estimate simply relied on counting the maximum number of overlapping detail filter bandwidths which reduced to a simple function of the dilation factor.

Further work needs to be done in estimating the upper bound, however. An analysis similar to that by Daubechies in [15] would be the most useful for developing a tighter upper frame bound and, in fact, is easily extendable for integer dilation factors greater than 2. However, that analysis depends upon the frame elements being translated and dilated versions of a set of mother wavelets which is not the case with adjoint wavelets. It is not clear how to apply such an analysis to the case of the RRA and so it remains an open research area.

The primary issue in further characterization of the RRA is the analysis of the effect of the upsample and downsample operations in the implementation of the projection operation. It is difficult to visualize, in the frequency domain, the dilation of a signal by q followed by the "compression" (via downsampling) by p. This remains the primary focus for subsequent research in this area.

VI. Application of the RRA to the Specific Emitter Problem

6.1 Introduction and Background

This chapter demonstrates a simple application of the RRA to a real-world, practical problem. The goal is to illustrate the advantage the RRA has over the MRA measured by the ability to create salient features in the specific emitter identification (SEI) problem. The application follows along the lines of that developed by Devaney[16] at Northeastern University where the entropy of the detail and approximation coefficients of an MRA and wavepacket decomposition were used to form feature vectors from radar pulses. These feature vectors were then subsequently used for training and identification of the radar pulses. The remainder of this section provides some background on the SEI problem, the use of resolution analyses (MRA, RRA, and wavepackets) and the entropy metric to generate feature vectors from sampled radar pulses, and development of a simple Bayesian-like classifier.

6.1.1 Specific Emitter Identification. In a tactical intelligence gathering scenario, many radar pulses are received from various locations of interest. Often these radar signals are differentiated by obvious characteristics such as pulse repetition frequency, pulse width, and carrier frequency. These differences are often enough to distinguish one emitter from another, such as a search radar from a tracking radar. There are times, however, when two radars are nearly identical and share nearly the same characteristics mentioned above and there is a need, from an intelligence standpoint, to be able to determine which specific radar generated a particular intercepted signal. This type of information is useful in tracking the "when" and "where" of specfic air defense radars in tactical environments such as the Persian Gulf or Bosnia.

Assigning a particular signal to a particular radar emitter in this case is a difficult problem because of the reason mentioned above: the radars are designed and manufactured to be identical aside from their serial numbers. It follows that any solution to the problem must rely on the differences between the radars which are not intended, such as minor variations in the tolerances of critical parts during the manufacturing process. Examples of some critical parts are the travelling wave tubes, power supplies, antennas, and switching hardware.

When we speak of radar signals, we mean the radar pulse. Radar pulses are characterized by what is known as the amplitude and frequency *videos.* A video is simply the baseband information coded into the pulse. The following figure illustrates this. The frequency video

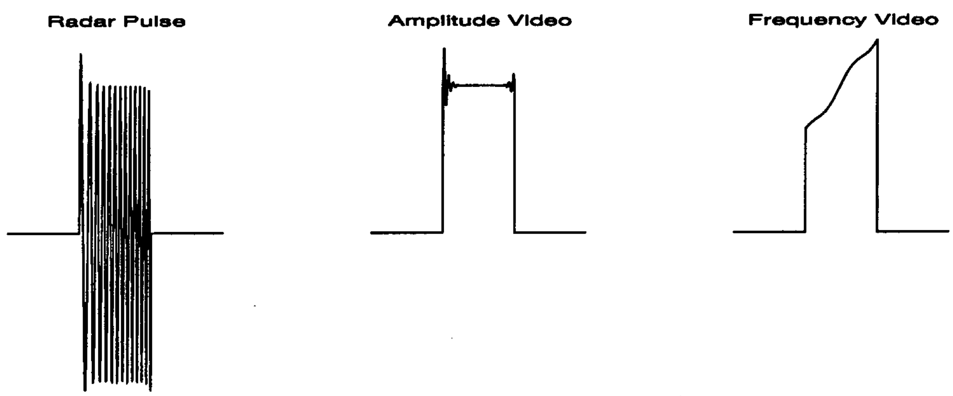

Figure 6.1 Illustration of Radar Pulse Videos

illustrates a nearly linear increase in frequency throughout the duration of the pulse which is characteristic of a linearly chirped signal. Both amplitude and frequency videos illustrate the irregularities due to manufacturing which serve to identify the specific emitter. While this suggests a processing scheme which first extracts the videos from the particular pulse, this is not necessarily the only way to design an algorithm. In fact, the schemes presented in the remainder of this introduction are based on processing the (sampled) pulse as an entire unit rather than deal with the videos as entities unto themselves. This is discussed next.

6.2 Feature Vectors from Radar Pulses

The processing we perform on a the sampled radar pulse data results in a feature vector which is then used for classification and/or training. In pattern recognition algorithms, feature vectors are designed to highlight the differences between signals so that classification can be performed. Furthermore, processing signals to yield feature vectors generally reduces the amount of data involved with the classification. A feature is *salient* when it is especially important in distinguishing between the patterns. Given a processing scheme which generates salient features, classification is often almost trivial. However, knowing a priori the saliency of features generated via various feature extraction algorithms is sometimes difficult. This complicates the design of those algorithms.

The approach taken in this chapter is to use the RRA, MRA, and wavepacket decomposition coefficients (both detail and approximation) as the basis for the construction of a feature vector. The idea is relatively simple. We assume the sampled radar signal represents a set of approximation coefficients of some signal which exists in some approximation space V_0. Of course, this approximation space is defined in terms of the particular analysis we are performing, the choice of basis functions, and the dilation factor. The signal is then decomposed and the sets of detail and approximation coefficients are then used to create the feature vector. An specific example follows shortly.

Resolution analyses do not really provide any data reduction in the sense that the resulting number of approximation and detail coefficients are less than the number of samples in the original signal. In fact, due the redundancy in representation, the RRA produces more coefficients than the original signal. Hence, this information must somehow be reduced to form the final feature vector. A logical choice, and the method chosen by Devaney, is to reduce each set of approximation and detail coefficients to a single number via the entropy metric. Entropy is a rough measurement of the spread of a set of numbers or a distribution. For continuous probability distributions, the entropy, $\mathcal{E}$, is is defined to be

$$\mathcal{E} = -E\{\ln f_X\} = -\int_x f_X(x) \ln f_X(x) dx$$

where f_X is the probability distribution of a random variable X[35] and E is the expected value operator. It is assumed the distribution is unit normalized since a measurement of spread (or localization) should be independent of relative magnitude.

For a discrete set of numbers, entropy can be expressed as

$$\mathcal{E} = -\sum_n |x_n|^2 \ln\left(|x_n|^2\right) \tag{6.1}$$

where it must be assumed that x is unit norm (l_2 norm). Looking at these two entropy expressions, the greater the entropy, the more evenly distributed the function (or set) is.

As a discrete example, consider the case were we have two unit-norm l_2 signals, say x and y. Further suppose these signals are compactly supported on $[0, N-1]$ and that

$$x(n) = \delta(n)$$
$$y(n) = 1/\sqrt{N}$$

The energy in x is perfectly localized so that $\mathcal{E}_x = 0$ while the energy in y is uniformly distributed so that $\mathcal{E}_y = \ln N$.

Used in conjuntion with resolution analyses, entropy measurement can be very useful for determining the significance (and thus, saliency) of a particular set of decomposition coefficients. Since entropy is not dependent upon the amount of energy in the signal, the resulting feature is not dependent upon the gain of the system used to collect the signal.

Throughout the previous chapters and especially throughout the development of the RRA, the implicit goal was to use the RRA in the front-end of a feature extraction scheme. The goal of this chapter is present a classification experiment which compares and contrasts the results obtained with the RRA with that of the MRA and wavelet packets. The exact description of the classification trials will be described in the next section, but a justification of why the RRA should outperform the MRA and/or wavelet packets follows here.

From a frequency spectrum point of view, the RRA detail coefficients provide a redundant representation of the signal since the bandpass filters corresponding to the various adjoint wavelets are not "orthogonal" to one another as the MRA and wavelet packet filters are. More specifically, the RRA filters overlap from one decomposition level to the next which increases the liklihood that an important piece of spectral information will be highlighted instead of slipping through the MRA and wavelet packet "gaps".

6.2.1 Classification. Given a particular radar signal, classification is the problem of correctly determining the specific radar which generated the signal. In a more general sense, we are trying to classify a pattern from its associated feature vector. In our case, the feature vectors are composed of the entropies of several levels of detail coefficients. The feature vectors for each class have some probabilty distribution which is generally unknown to us or we assume knowledge of.

Classifying an unknown feature vector becomes a simple matter. An unknown feature vector is assigned to the class for which the probability distribution is maximum. Another way of thinking about the classification is to imagine the each class as an ellipsoid in multi-dimensional space. The center of the ellipsoid is characterized by the mean and the shape and/or orientation is determined by the covariance matrix. Given a feature vector, we choose the class for which the distance to the associated ellipsoid is minimized. This distance is known as the Mahalanobis distance[35]. This distance from a feature vector $\mathbf{x}$ to the ellipsoid of class j can be written as

$$r_j(\mathbf{x}) = (\mathbf{x} - \mathbf{m}_j)\mathbf{C}_j^{-1}(\mathbf{x} - \mathbf{m}_j)$$

where $\mathbf{m}_j$ and $\mathbf{C}_j$ represent the mean vector and covariance matrix, respectively. Thus, the assigned class k is defined to be

$$k = \arg\min_j \{r_j(\mathbf{x})\}.$$

6.2.2 Example Feature Vector Formation. To illustrate the resolution analyses which go into the processing to form feature vectors, consider the RRA case where $p = 3$ and $q = 2$ and we will use 5 levels of decomposition to form the feature vectors. This is the actual processing for one of the RRA cases described later in this chapter. For this example,

Set	Entropy
$d_1(1)$	3.7970
$d_2(1)$	3.6702
$d_1(2)$	3.5811
$d_2(2)$	3.3897
$d_1(3)$	2.9680
$d_2(3)$	3.0509
$d_1(4)$	2.6487
$d_2(4)$	2.6180
$d_1(5)$	2.2103
$d_2(5)$	2.2437

Table 6.1 Entropy Values for Example Feature Vector Formation

we use linear spline scaling functions and wavelets ($m = 2$). The processing is shown in Figure 6.2.2. The pulse in this example is number 17 taken from the second Granny emitter set. This pulse is also part of the associated culled data set. This nomenclature is more thoroughly defined in the following section. For our purposes it sufficient at this point to understand this data is actual radar pulse data.

We assume the sampled pulse is a sequence of approximation coefficients (c_0)associated with some function living in V_0 which is defined by the scaling function and dilation factor. Since we have $p = 3$, there are $p - 1 = 2$ wavelets and sets of detail coefficients at each level. The left column of graphs in Figure 6.2.2 represent the successive sets of approximation coefficients for 4 levels of decomposition (c_1, c_2, c_3, c_4). To the right of each approximation coefficient graph are the two graphs of the corresponding detail coefficients.

The entropy for each set of detail coefficients is given in Table 6.2.2. The values of the approximation coefficients shown in Figure 6.2.2 have been normalized for the entropy calculation and the c_0 approximation coefficients (the actual sampled pulse data) has been normalized prior to performing the RRA decomposition. The feature vectors for the other types of processing (MRA and WP) are formed in a similar manner.

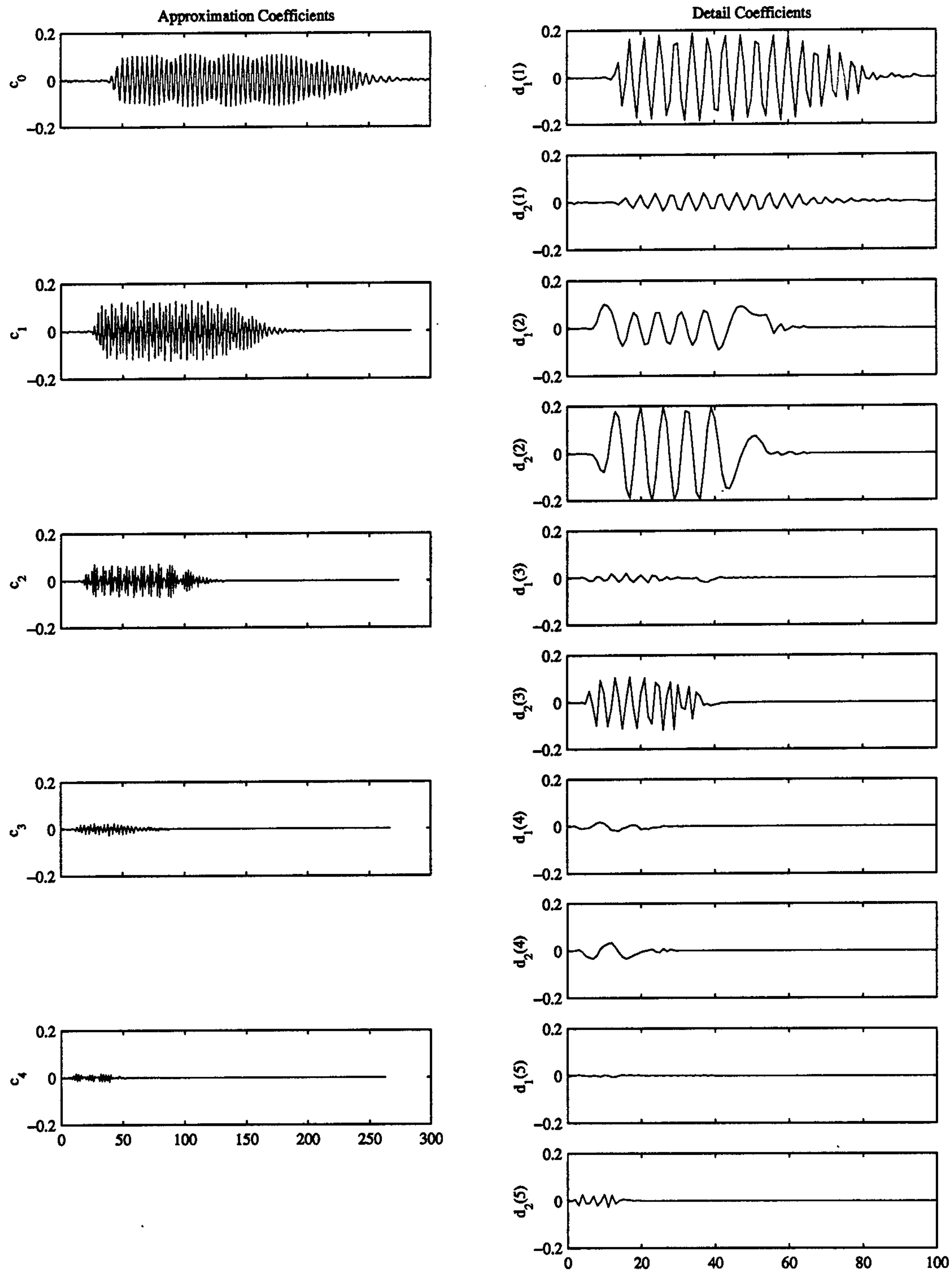

Figure 6.2 Illustration of RRA Processing for Feature Vector Formation with $p = 3$, $q = 2$, $m = 2$ and Actual Sampled Pulse Data

6.3 Data and Processing Description

6.3.1 Radar Data. The data used for this application consists of sampled radar pulses of four emitters. For reference, the entire data set is called the "apples " and the emitters are labelled "golden," "granny," "ida," and "mac." For each emitter, there are three data sets and each radar pulse in the data set is represented by 300 samples. From data set, a subset has been set aside. This subset, which is called the "culled" data set, represents a portion of the sampled radar pulses which are considered "clean". For instance, these pulses all have values which did not exceed the dynamic range of the collection equipment so that no clipping is evident. Furthermore, the pulse transitions are clearly evident. These data points will be used to "train" the classifier in certain experiments since they are potentially the most representative of the specific emitter. The total number of data points is shown in Table 6.2.

	Data Set					
Emitter	1		2		3	
	unculled	culled	unculled	culled	unculled	culled
Golden	141	98	141	110	141	111
Granny	141	113	141	115	141	125
Ida	177	93	65	38	65	32
Mac	196	167	50	30	157	89

Table 6.2 Number of Sampled Radar Pulses per Class

6.3.2 Wavepacket Decomposition. There are three basic types of processing which are used in this chapter: the RRA decomposition, the MRA decomposition, and the binary wavepackets decomposition. Wavepacket decompositions are considered in this research for comparison with the work of Devaney cited earlier. They require some explanation since both the RRA and MRA have been descibed in previous chapters.

In simplest terms, a wavepacket decomposition can be though of as a recursive application of the MRA. For the dyadic case, the approximation coefficients c_0 are processed into approximation coefficients c_1 and detail coefficients d_1. The difference between the MRA and the wavepacket decompostion is that for the MRA, only c_1 is further decomposed, while for the wavepacket decomposition, both c_1 and d_1 are further decomposed. At each level, decomposition is performed on both the "approximation" and "detail" coeffcients. For example, in a two-level wavepacket decomposition, seven sets of coefficients are generated: c_0, c_{00}, c_{01}, c_{000}, c_{001}, c_{010}, and c_{011}. A binary labelling scheme has been used. The original set of approximation coefficients can be reconstructed by considering the sets c_{000}, c_{001}, c_{010}, and c_{011}. In contrast, the MRA would require c_2, d_2, and d_1. The numbers of total coefficients in either case would be roughly equivalent.

From a frequency standpoint, a wavepacket decomposition covers the spectrum with frequency bins whose widths are constant and therefore independent of the the center frequency of particular band. In contrast, the MRA covers the spectrum with frequency bins whose widths are proportional to the center frequency. The effect of a wavepacket decomposition is a partitioning of the spectrum which is in some sense similar to that of a discrete Fourier transform. One difference is that the wavepacket decomposition allows the flexibility to consider groups of frequency bins as a single bin (by dealing with c_{01} instead of both c_{010} and c_{011}, for example). More will be said on the wavepacket decomposition later.

6.3.3 Feature Vector Construction. The feature vectors for the various decompositions are constructed in roughly similar ways. The radar pulse is first normalized (l_2) and then decomposed via the RRA, MRA, or wavepacket (WP) analysis. For certain sets of decomposition coefficients, the entropy is calculated and the result becomes an entry in the corresponding feature vector. It should be noted at this point that the decomposition coefficients are normalized prior to calculating the entropy so that all levels of decomposition coefficients are treated equally. The length of the feature vectors in all three cases are determined so they are roughly equivalent. This allows a more meaningful comparison between the decompositions.

Feature vectors are used in one of two ways. The first way is for use in "training" the classifier. A set of feature vectors corresponding to a particular emitter are used to calculate a mean and variance vector. These vectors are used to construct a Bayesian-like classifier whereby the distribution of the feature vectors belonging to a particular class is assumed to have a normal distribution characterized by the mean and variance vectors. The other use of the feature vectors is for use in classification. Once the mean and variance vectors have been calculated, a specific feature can be "compared" to the various distributions to determine which is the most likely class.

6.3.4 Pulse Classification Experiments. As was roughly described above, a pulse is classified based on the comparison of its feature vector with the distributions of feature vectors associated with the various emitters. An explicit assumption in this process is that pulses from a specific emitter, or rather, the associated feature vectors, are normally distributed. That is, the distibution is Gaussian and is parameterized entirely by the mean and variance. There are two primary reasons for this assumption. The first has to do with the Central Limit Theorem which states the distribution of an infinite sequence of functions of random variables is a normal distribution regardless of the distribution of the random variable. The second reason is more practical than theoretical. The sample mean and variance, used as estimates of the actual mean and variance, are easy to calculate. Thus, without knowledge of the distribution of the sampled radar data and the effect the various resolution analyses have on that distribution, the assumption of a normal distribution is warranted.

For each of the apples there are three unculled data sets and each of these has an associated culled data set. Each classification experiment uses one or more of the three data sets to "train" the classifier as was described above. Then each of the radar pulses in the remaining data sets are compared to the distributions established by the training sets and assigned a class (golden, granny, ida, or mac) based on that result. As a measure of saliency, the training sets can also be classified. This gives a qualitative estimate of how well the feature vectors uniquely characterize a particular class of signals.

The results of the various classification experiments are presented in confusion matrices. A confusion matrix shows the performance of the classifier by showing how the pulses were classified versus the class they actually belong to. An example confusion matrix is shown below in Table 6.3.

	Golden	Granny	Ida	Mac
Golden	133	1	7	0
Granny	3	132	4	2
Ida	3	14	143	17
Mac	1	0	5	190

Table 6.3 Example Confusion Matrix

For example, the entry in the third column and second row is the number of the pulses in the second class (granny, in our case) which the classifier incorrectly assigned to the third class (ida). The sum of any row is the total number of pulses of a particular class used in the experiment and the distribution of values within a row gives an indication of how well the classifier performs on the class associated with that row. A perfect classifier would have a strictly diagonal matrix. The overall accuracy of the classifier (feature vectors and classifier) is the trace of the confusion matrix divided by the total number of pulses used in the experiment.

6.4 Classification Results

6.4.1 Description. Each processing method (RRA, WP, and MRA) was used, each with various sets of parameters. For the RRA, the 3/2 dilation factor was used with two regularity factors ($m = 2$ and $m = 4$). For the RRA case, 5 levels of decomposition were performed and entropies calculated for the resulting 10 sets of detail coefficients. The 10 entropy values were concatenated to form a feature vector. An entropy was not calculated for the final approximation coefficients. The number of decomposition levels was chosen to minimize the size of the resulting feature vectors and to correspond to a bandwidth coverage roughly equivalent to the WP processing. More will be said on this later.

For the WP processing, Daubechies wavelets and scaling functions were used. In particular, the 12-tap filters (Db-6) and the 6-tap filters (Db-3) were used since they had regularities which corresponded most closely with $m = 4$ and $m = 2$, respectively. In each case, 3 decomposition levels were calculated and entropies were calculated for each set of approximation and detail coefficients (15 entropies). As with the RRA case, the 15 entropy values were concatenated to form a feature vector. The number of decomposition levels was chosen to correspond to a particular bandwith coverage discussed below.

In order to equate the RRA and WP processing in terms of information extraction, the equivalent bandwidth of the processing filters must be approximately equivalent. If we consider a starting point of unity bandwith and assume the coverage of a particular decomposition level is determined be the center point of the bandwidth of the final level of decomposition, the five levels of the RRA decomposition for $p/q = 3/2$ cover down to $(2/3)^5 = 0.13168$ and the 3 levels of the WP decomposition covers to $(1/2)^3 = 0.125$. Thus, although the resulting feature vectors have unequal lengths, the bandwidth coverages are roughly equivalent. It was mentioned above that the approximation coefficients associated with the final level of RRA decomposition were not included. For the $p/q = 3/2$ case, the coverage of this part of the processing is centered at $2^4/3^5$ which is exactly the lower edge of the coverage of the detail spaces. The decision to do without these coefficients stems from this observation and a desire to use as few coefficients as possible to cover the same bandwidth.

It is convenient at this point to adopt a nomenclature which will make it easier to refer to the various processing techniques and their associated parameters. For the RRA, the notation RRA(p, q, m, l) will be used to mean RRA processing with dialation factor p/q, regularity m, and l levels of decomposition. For the wavepacket processing, the notation WP(*Db6*,l) will be used to mean an l-level wavepacket decomposition (in this example, using the 12-tap Daubechies scaling function and wavelet). The dilation factor 2 is implicit in the wavepacket processing to be consistent with Devaney. For the MRA, the notation MRA(*type*,p, l) is used to mean an l-level MRA decomposition using wavelets and scaling functions associated with *type* and dilation factor p. In both the MRA and RRA processing schemes, the entropy of the final set of approximation coefficients was used in constructing the feature vector unless explicitly specified.

6.4.2 Discussion of Results. The results of classification processing is summarized in Tables 6.4 and 6.5.

Case	Train	Test	RRA m=4	RRA m=2	WP Db6	WP Db4	MRA 3,4,2	MRA 3,2,2
1	123	123	0.8985	0.8901	0.8335	0.8335	0.7584	0.7841
2	123	123c	0.9233	0.9153	0.8849	0.8956	0.7886	0.7975
3	123c	123	0.8560	0.8837	0.7815	0.7770	0.7237	0.7873
4	123c	123c	0.9224	0.9376	0.8796	0.8894	0.8037	0.8260
5	12	12	0.9078	0.9040	0.8527	0.8565	0.7757	0.7899
6	12	12c	0.9123	0.9188	0.8861	0.9084	0.7801	0.7932
7	12	3	0.8849	0.8631	0.7996	0.7817	0.7242	0.7917
8	12	3c	0.9440	0.9216	0.8627	0.8711	0.7787	0.8291
9	12c	12	0.8622	0.9030	0.8099	0.8118	0.7567	0.8156
10	12c	12c	0.9241	0.9476	0.8927	0.9097	0.8207	0.8390
11	12c	3	0.8333	0.8274	0.7480	0.7222	0.7004	0.7698
12	12c	3c	0.9216	0.9048	0.8739	0.8655	0.7843	0.8151
13	13	13	0.8999	0.8921	0.8516	0.8447	0.7765	0.7808
14	13	13c	0.9324	0.9203	0.9118	0.9034	0.8104	0.7911
15	13	2	0.8992	0.8615	0.7935	0.7909	0.7154	0.7557
16	13	2c	0.9044	0.8840	0.8532	0.8464	0.7474	0.7747
17	13c	13	0.8447	0.8904	0.8050	0.7852	0.7343	0.7947
18	13c	13c	0.9312	0.9457	0.9106	0.9022	0.8321	0.8418
19	13c	2	0.8363	0.8589	0.7305	0.7506	0.6650	0.7582
20	13c	2c	0.8874	0.8976	0.8294	0.8396	0.7406	0.7952
21	23	23	0.8923	0.8824	0.8080	0.8002	0.7170	0.7647
22	23	23c	0.9246	0.9185	0.8677	0.8677	0.7692	0.7954
23	23	1	0.8947	0.8962	0.8092	0.8473	0.8198	0.7191
24	23	1c	0.9087	0.9172	0.8195	0.8832	0.8280	0.6921
25	23c	23	0.8446	0.8801	0.7403	0.7203	0.6837	0.7492
26	23c	23c	0.9292	0.9385	0.8523	0.8385	0.7692	0.8000

Table 6.4 Overall Classification Results

The table contains the overall classification rate (the number of successful classifications divided by the total number of attempts). In terms of the confusion matrix, it is simply the trace divided by the sum of the column sums.

The discussion of the results can be naturally divided into two areas: the feature saliency and classification accuracy. Feature saliency deals with the classifications where the training and test dat are identical. Hopefully, a classifier should perform well on the same

Case	Train	Test	RRA m=4	RRA m=2	WP Db6	WP Db3	MRA 3,4,2	MRA 3,2,2
27	23c	1	0.8443	0.8733	0.7557	0.7496	0.7481	0.7496
28	23c	1c	0.8981	0.9257	0.8429	0.8514	0.8174	0.8004
29	1	1	0.9160	0.9130	0.8809	0.8794	0.8198	0.8046
30	1	1c	0.9278	0.9278	0.9151	0.9257	0.8217	0.8068
31	1	23	0.8868	0.8535	0.7980	0.7869	0.7214	0.7669
32	1	23c	0.9231	0.8923	0.8754	0.8662	0.7692	0.7938
33	1c	1	0.8519	0.9115	0.8534	0.8412	0.7756	0.8443
34	1c	1c	0.9342	0.9575	0.9321	0.9236	0.8705	0.8726
35	1c	23	0.8257	0.8313	0.7625	0.7214	0.6848	0.7547
36	1c	23c	0.9077	0.8954	0.8815	0.8308	0.7738	0.8169
37	2	2	0.8967	0.8992	0.8262	0.8161	0.7229	0.7683
38	2	2c	0.8976	0.9078	0.8703	0.8669	0.7440	0.7713
39	2	13	0.8809	0.8887	0.7222	0.7368	0.7575	0.8067
40	2	13c	0.9082	0.9251	0.7560	0.7995	0.7778	0.8225
41	2c	2	0.8816	0.8791	0.8060	0.7733	0.7053	0.7708
42	2c	2c	0.9113	0.9283	0.8874	0.8805	0.7611	0.7986
43	2c	13	0.8352	0.8628	0.7282	0.6747	0.7187	0.7532
44	2c	13c	0.8949	0.9372	0.8080	0.7935	0.7778	0.7947
45	3	3	0.8909	0.8750	0.8075	0.8016	0.7282	0.7599
46	3	3c	0.9412	0.9272	0.8992	0.8768	0.7871	0.7843
47	3	12	0.8821	0.8650	0.8213	0.7928	0.7662	0.7300
48	3	12c	0.9018	0.8874	0.8665	0.8364	0.7945	0.7186
49	3c	3	0.8333	0.8631	0.7421	0.7242	0.6607	0.7560
50	3c	3c	0.9468	0.9440	0.8683	0.8431	0.7759	0.8235
51	3c	12	0.7966	0.8441	0.7262	0.7120	0.6635	0.7139
52	3c	12c	0.8704	0.9005	0.8285	0.8024	0.7683	0.7683

Table 6.5 Overall Classification Results (continued)

data which was used to train it. If this is not the case, it generally signifies that the features used to train the classifier are not salient. That is, the features do not provide "handles" to determine which class a particular test vector belongs to. Determining feature saliency also depends upon the particular classifier.

Classification accuracy are associated with the test cases where the training set and test sets do not intersect. It is simply a measure of how well a classifer can do with a given feature set. It is representative of the results one would expect if the classifier was operationally fielded.

6.4.2.1 Entropy Vector Feature Saliency. For convenience, the classifications where the training and test sets are the same are sifted from Tables 6.4 and 6.5 and shown in Table 6.6. Comparing the RRA and wavepacket techniques independent of filter regularity,

Case	Train	Test	RRA m=4	RRA m=2	WP Db6	WP Db4	MRA 3,4,2	MRA 3,2,2
1	123	123	0.8985	0.8901	0.8335	0.8335	0.7584	0.7841
4	123c	123c	0.9224	0.9376	0.8796	0.8894	0.8037	0.8260
5	12	12	0.9078	0.9040	0.8527	0.8565	0.7757	0.7899
10	12c	12c	0.9241	0.9476	0.8927	0.9097	0.8207	0.8390
13	13	13	0.8999	0.8921	0.8516	0.8447	0.7765	0.7808
18	13c	13c	0.9312	0.9457	0.9106	0.9022	0.8321	0.8418
21	23	23	0.8923	0.8824	0.8080	0.8002	0.7170	0.7647
26	23c	23c	0.9292	0.9385	0.8523	0.8385	0.7692	0.8000
29	1	1	0.9160	0.9130	0.8809	0.8794	0.8198	0.8046
34	1c	1c	0.9342	0.9575	0.9321	0.9236	0.8705	0.8726
37	2	2	0.8967	0.8992	0.8262	0.8161	0.7229	0.7683
42	2c	2c	0.9113	0.9283	0.8874	0.8805	0.7611	0.7986
45	3	3	0.8909	0.8750	0.8075	0.8016	0.7282	0.7599
50	3c	3c	0.9468	0.9440	0.8683	0.8431	0.7759	0.8235

Table 6.6 Overall Saliency Results

the RRA has equivalent or better performance in all of the 14 cases (100%). When comparing similar regularities, the RRA performed better in all of the 14 cases (100%) for the $m = 4$ and "db6" filters. The RRA performed equivalently or better in all of the 14 cases (100%) for the $m = 2$ and "db3" filters as well.

Looking at Table 6.6, the overall classification rates for these saliency cases are generally at 89%. The conclusion is that although the RRA performs better than the wavepackets in terms of feature saliency, the margin of improvement is relatively small. While this statement may appear to trivialize any advantage the RRA has over wavepacket processing, the size of the feature vectors needs to be considered. The wavepacket entropy vector has 15 elements compared to 10 for RRA. From a purely theoretical standpoint, one would expect better performance from the wavepacket feature vector simply because it has the potential to contain so much more information. But in this case, the relative equivalence of the performances supports a conclusion that the RRA feature vectors are significantly more salient than those from wavepacket processing.

6.4.2.2 Classifier Accuracy. This part of the results focuses on the classification trials where the training set and test set are different. As was mentioned earlier, each data set has a culled subset which contains those pulses which are cleanly sampled with no aliasing or clipping. Both the unculled and culled data sets were used to train the classifier. Theoretically, training with culled data would let the classifier better adapt to a particular emmitter since the associated exemplars are "better."

This theory does not hold in this case however. When comparing training with culled and unculled datasets on the same test sets, it is clear that, in fact, using culled data to train the classifier generally hurts the overall accuracy. For the 26 pairs of classifications, training with unculled data resulted in significantly better overall results as shown in Table 6.7.

	RRA m=4	Db6 WP	RRA m=2	Db3 WP
Ratio	21/26	16/26	14/26	24/26
Percentage	80.77%	61.54%	53.85%	92.31%

Table 6.7 Percentage of Classifications where Unculled Training Data Resulted in Greater Overall Accuracy

While it is difficult to determine the exact cause of this result, one speculation might be that forming feature vectors via entropy calculation nullifies any potential advantage from training with clean data. That is, by "cleaning up" the data, information useful to this particular classification scheme is lost. Since the culling criteria is unknown, it is not clear what type of classification scheme could benefit from culling the data, but for this scheme, there is a general penalty for using culled data. With this result in mind, only the classifications with unculled training data will be discussed in the remainder of this section. For completeness, Tables 6.4 and 6.5 contains all the classifications however.

The subset of Table 6.4 and 6.5 for which training was performed on unculled data is shown in Table 6.8. The table shows that in classifications where training was performed

Case	Train	Test	RRA m=4	RRA m=2	WP Db6	WP Db4	MRA 3,4,2	MRA 3,2,2
1	123	123	0.8985	0.8901	0.8335	0.8335	0.7584	0.7841
2	123	123c	0.9233	0.9153	0.8849	0.8956	0.7886	0.7975
5	12	12	0.9078	0.9040	0.8527	0.8565	0.7757	0.7899
6	12	12c	0.9123	0.9188	0.8861	0.9084	0.7801	0.7932
7	12	3	0.8849	0.8631	0.7996	0.7817	0.7242	0.7917
8	12	3c	0.9440	0.9216	0.8627	0.8711	0.7787	0.8291
13	13	13	0.8999	0.8921	0.8516	0.8447	0.7765	0.7808
14	13	13c	0.9324	0.9203	0.9118	0.9034	0.8104	0.7911
15	13	2	0.8992	0.8615	0.7935	0.7909	0.7154	0.7557
16	13	2c	0.9044	0.8840	0.8532	0.8464	0.7474	0.7747
21	23	23	0.8923	0.8824	0.8080	0.8002	0.7170	0.7647
22	23	23c	0.9246	0.9185	0.8677	0.8677	0.7692	0.7954
23	23	1	0.8947	0.8962	0.8092	0.8473	0.8198	0.7191
24	23	1c	0.9087	0.9172	0.8195	0.8832	0.8280	0.6921
29	1	1	0.9160	0.9130	0.8809	0.8794	0.8198	0.8046
30	1	1c	0.9278	0.9278	0.9151	0.9257	0.8217	0.8068
31	1	23	0.8868	0.8535	0.7980	0.7869	0.7214	0.7669
32	1	23c	0.9231	0.8923	0.8754	0.8662	0.7692	0.7938
37	2	2	0.8967	0.8992	0.8262	0.8161	0.7229	0.7683
38	2	2c	0.8976	0.9078	0.8703	0.8669	0.7440	0.7713
39	2	13	0.8809	0.8887	0.7222	0.7368	0.7575	0.8067
40	2	13c	0.9082	0.9251	0.7560	0.7995	0.7778	0.8225
45	3	3	0.8909	0.8750	0.8075	0.8016	0.7282	0.7599
46	3	3c	0.9412	0.9272	0.8992	0.8768	0.7871	0.7843
47	3	12	0.8821	0.8650	0.8213	0.7928	0.7662	0.7300
48	3	12c	0.9018	0.8874	0.8665	0.8364	0.7945	0.7186

Table 6.8 Overall Classification Results for Cases with Unculled Training Data Sets

with unculled data, the RRA out-performed wavepacket processing by a substantial margin. Of the 26 classifications, 24 (92.31%) were associated with one of the two RRA processing schemes. The actual breakout of the numbers is shown in Table 6.9.

6.5 *Modified Wavepacket Processing*

6.5.1 Introduction. As was mentioned in the introduction to this chapter, a radar pulse can be characterized by amplitude and frequency videos. From a strictly Fourier point of view, the most significant aspect of a radar pulse is the frequency chirp or carrier frequency. One might hope to gain better classification accuracy by focusing on these frequencies. Unfortunately, the wavepacket processing described earlier in this chapter subdivide the frequency spectrum in a symetric, tree-structured fashion.

However, an advantage wavepacket processing has over the RRA and MRA is the ease with which different frequency bands can be emphsized. Consider Figure 6.3 which shows a binary-tree representation of the frequency division which occurs during a dyadic wavepackets decomposition. The wavepacket processing earlier in this chapter used coefficients of 3 levels

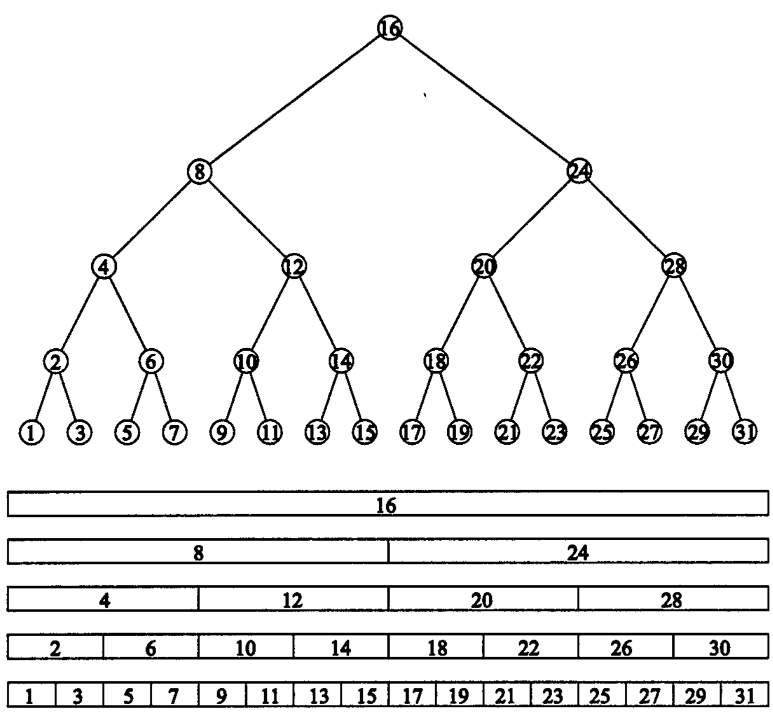

Figure 6.3 Binary Tree Representation of Dyadic Wavepacket Processing

to form the feature vector. If each node in Figure 6.3 represents a set of coefficients and an entry in the feature vector via the entropy calculation, the earlier feature vectors were formed using nodes in the top three levels. In Figure 6.3, this corresponds to even-numbered nodes for a total of 15 nodes.

6.5.2 Description of Modified WP Feature Vectors. The frequency video for a representative pulse is focused mainly in the frequency bins associated with nodes 13 through 19. This suggests a wavepacket feature vector which emphasizes these nodes may offer better classification performance. To test this hypothesis, classification tests using a modified wavepacket feature vector were performed.

In this second set of classifications, feature vectors are constructed according to Figure 6.4 below. This new feature vector has 13 entries, 2 less than the symetric WP decom-

Position	1	2	3	4	5	6	7	8	9	10	11	12	13
Node	16	8	12	10	14	13	15	24	20	18	17	19	22

Figure 6.4 Modified Feature Vector Diagram

position earlier. It de-emphasizes the lower frequencies (the nodes attached below node 8) in favor of the mid to high frequencies bands which are associated with the nodes attached below node 20. With the inclusion of information in nodes 17, 19, 21, and 23, note that this decomposition is also carried one level lower than the symetric decomposition earlier in this chapter. For comparision with the earlier wavepacket results, both the Daubechies 6-tap and 12-tap filters (db3 and db6) were used.

6.5.3 Classification Results with Modified Wavepacket Feature Vectors. Despite emphasizing the frequency video, the classification results using the modified WP feature vectors were only remarkable in that they showed no significant improvement over the symetric WP feature vectors. The overall results are tabulated in Table 6.10. In fact, when comparing the results in Table 6.10 with those in Tables 6.4 and 6.5, the symetric feature vectors outperform the modified feature vectors with no exceptions. This result is not intuitive and suggests the salient features generated from the processing are not caused by the frequency video, which might otherwise be considered a prominent feature.

6.6 *Summary*

In this chapter, the RRA was used to aid in the classification of radar pulses in a specific emitter identification problem. The RRA was compared to a previously published wavepacket entropy technique and was shown to provide significantly better overall results. The "sliding" redundancy of the RRA allows for more salient features by nearly a four to one margin compared to the "subdivision" redundancy of the wavepackets. In terms of accuracy, the results of this chapter indicate that in practical applications of the classifier the RRA-type processing would out-perform the MRA-based wavepackets processing.

While RRA-based processing performed better than wavepackets processing, the results were dependent upon the regularity of the filters used for the processing. In general, the $m = 2$ (linear) and $m = 4$ (cubic) RRA techniques competed for the best performance. This would indicate that the overall results are dependent upon the regularity of the filters and may have something to do with matching the regularity of a particular radar's pulse. Further studies are needed to determine the extent of this observation. One suggestion would be to try the RRA with other regularities.

The hypothesis that the RRA would compare favorably to the MRA (and its derivatives) in feature extraction situations is based on the observation that an orthogonal decomposition allows "gaps" in the frequency coverage. That is not to say there are frequencies which are not covered by an MRA. However, certain frequencies will be covered by two adjacent bands equally. Thus the full impact of these frequencies will be spread among two sets of decomposition coefficients.

Process	RRA m=4	Db6 WP	RRA m=2	Db3 WP
Ratio	19.5/26	0/26	6.5/26	0/26
Percentage	75%	0%	25%	0%

Table 6.9 Percentages of Unculled Trained Classifications by Best Performing Technique

Case	Train	Test	WP Db6	WP Db3	Case	Train	Test	WP Db6	WP Db3
1	123	123	0.7384	0.6947	27	23c	1	0.6840	0.6672
2	123	123	0.7690	0.7386	28	23c	1c	0.7580	0.7686
3	123c	123	0.7159	0.6902	29	1	1	0.7802	0.7359
4	123c	123	0.7814	0.7984	30	1	1c	0.7813	0.7643
5	12	12	0.7519	0.7224	31	1	23	0.6837	0.6382
6	12	12c	0.7618	0.7565	32	1	23c	0.7354	0.6969
7	12	3	0.7024	0.6766	33	1c	1	0.7634	0.7282
8	12	3c	0.7703	0.7339	34	1c	1c	0.8068	0.8195
9	12c	12	0.7500	0.7129	35	1c	23	0.6737	0.6193
10	12c	12c	0.7984	0.8024	36	1c	23c	0.7462	0.7185
11	12c	3	0.6806	0.6409	37	2	2	0.7632	0.7128
12	12c	3c	0.7759	0.7619	38	2	2c	0.7747	0.7406
13	13	13	0.7343	0.6937	39	2	13	0.6903	0.6928
14	13	13c	0.7693	0.7476	40	2	13c	0.7077	0.7403
15	13	2	0.7305	0.6675	41	2c	2	0.7280	0.6977
16	13	2c	0.7611	0.7065	42	2c	2c	0.7816	0.7952
17	13c	13	0.7170	0.6842	43	2c	13	0.6687	0.6324
18	13c	13c	0.7959	0.7995	44	2c	13c	0.7126	0.7331
19	13c	2	0.7003	0.6549	45	3	3	0.7242	0.6528
20	13c	2c	0.7577	0.7406	46	3	3c	0.8067	0.7563
21	23	23	0.7436	0.6792	47	3	12	0.7319	0.6863
22	23	23c	0.7923	0.7477	48	3	12c	0.7539	0.7291
23	23	1	0.7359	0.7099	49	3c	3	0.6786	0.6528
24	23	1c	0.7431	0.7431	50	3c	3c	0.7815	0.7759
25	23c	23	0.6926	0.6681	51	3c	12	0.6540	0.6483
26	23c	23c	0.7677	0.7785	52	3c	12c	0.7225	0.7474

Table 6.10 Overall Classification Results Using Modified WP Feature Vectors

VII. Conclusion

7.1 Introduction

This chapter concludes the dissertation. The first section contains a summary of the major accomplishments. It also compares these accomplishments with the research goals which were set at the outset. These goals were presented in Chapter I.

This research, like any theoretical research, has raised questions beyond those originally envisioned. The second section deals with some of those questions in the form of recommendations for additional research. While these recommendations are related to the issues raised in the research, they are either beyond the scope of the research or deal with a refinement of some result which was not pursued.

The final section draws some conclusions from the research in the context of the larger mathematical and engineering environments. Connections are made to several other research areas and the significance of the results of this dissertation are stated.

7.2 Summary and Evaluation of Objectives

The organization of this research has been focused on answering several fundamental questions which were identified at the outset of the work. The RRA as presented and described in this document was crudely introduced in a previous work[1]. That introduction actually posed more questions than it answered although it did set the stage for this larger and more complete treatment of the subject.

Chapter II contained the theoretical foundations for the RRA. The nomenclature and structure of the RRA is based on the integer-dilation MRA and it is presented in that context. While the wavelets for the MRA form an orthonormal basis for $L_2(\mathbb{R})$, the wavelets associated with the RRA do not. Furthermore, it is not clear if they form a frame for $L_2(\mathbb{R})$. This is a question which is not answered until a later chapter. The conditions on the RRA scaling functions and wavelets required for perfect reconstruction are the most significant point of this first chapter. For perfect reconstruction, it was shown that the scaling function must satisfy a dual system two-scale equation based on two relatively prime positive integers. This was required so that the primed approximation space V_k' was a subset of the higher-resolution approximation spaces V_{k-1} and V_k simultaneously.

This requirement on the scaling function for perfect reconstruction is a segue to Chapter III where the existence of compactly supported scaling functions satisfying the dual two-scale equations is investigated. Here, the implications of that requirement are analyzed. It is shown through a series of Lemmas and Theorems that the only orthonormal compactly supported scaling functions are the well-known characteristic functions which give rise to Haar wavelets. As a related result, it is shown that even when the orthonormality requirement is relaxed, the only functions which simultaneously satisfy the dual two-scale equations are B-splines or their derivatives. The roots of all the characteristic two-scale expression must lie on the unit circle.

Having determined the scaling functions which are allow perfect reconstruction are spline-based, the next question that arises is how to calculate (numerically or symbolically) the corresponding wavelets. The approach taken in Chapter IV is to cast the problem as determining a set of vectors which span the nullspace of a finite approximation projection which is determined by the scaling function. This set of nullspace vectors are then related to nullspace of the larger, infinite approximation projection. Through an isomorphism between this nullspace and the detail space of the associated MRA, it is proven these nullspace vectors can be transformed into the desired orthonormal spline wavelets.

Chapter V deals with the concept of frames and the question of whether the RRA wavelets form a frame for $L_2(\mathbb{R})$. The problem is one of determining how redundant the wavelets are in representing an arbitrary element in $L_2(\mathbb{R})$ and how stable that representation is. A stable representation is one where closeness between two sets of detail coefficients translates into some degree of closeness between the functions which generated the coefficients. It is shown in this chapter that the RRA wavelets do indeed form a frame for $L_2(\mathbb{R})$

and that the lower frame bound is one and the upper frame bound is a function of the regularity and dilation factor.

Chapter VI presents an application of the RRA to a real-world problem: specific emitter identification. The RRA is compared to other techniques such as wavepacket and MRA decompositions with respect to the ability to construct salient feature vectors from a sampled radar pulse. These features are fed into a classifer and the results were compared. It was determined the RRA-formed feature vectors resulted in classifications which were significantly more accurate than either the wavepackets or MRA. The overlapping structure of the RRA feature vector caused features which were more salient than the ultra-redundant wavepacket feature vectors, even when the wavepacket feature vectors were contructed in such a way as to emphasize the (apparent) significant features such as frequency video.

7.3 Recommendations for Future Work

As with any research project, this dissertation raises additional questions which are out of the scope of the original effort. These questions and their associated analysis become recommendations for further research and are the subject of this section. The discussion deals with significant issues which became apparent during the course of the research but is not meant to be an exhaustive list since other issues will likely occur to the reader. This section also discusses some issues with the research itself in terms of how the analysis leading to certain results could be refined.

The characterization of compactly supported scaling functions yielding perfect reconstruction in the RRA is based upon a proof which deals with the zero locations of the corresponding approximation filters. The proof is somewhat inelegant in that it brute forces the conclusions about where the roots can lie by process of elimination. A recommendation in this area would be to research ways in which that part of the proof could be made simple and straight-forward. This recommendation falls in the category of research refinement.

One of the first questions which arose during the course of the research is the requirement for scaling functions and wavelets which lead to perfect reconstruction in the RRA. The research has demonstrated the B-splines are the only family of functions in this class. If one has no interest in synthesis, then perhaps the analysis can be simplified by using wavelets and scaling functions which do not satisfy the dual dilation equation system. After all, the application presented in Chapter VI does not depend upon perfect reconstruction, nor does it depend upon the wavelets forming a frame for $L_2(\mathbb{R})$. Thus, the question of necessity for perfect reconstruction in some applications is open. Aside from the elegance of the theory, perfect reconstruction seems useful only in those applications where analysis and synthesis are both present. Thus, research may be warranted in areas of biorthogonal wavelets and scaling functions. Compactly supported biorthogonal wavelets and scaling functions are certainly possible.

The orthonormalization technique for constructing the final scaling functions and wavelets results in functions which do not have compact support (though they have exponential decay). The corresponding approximation and detail filters are not FIR which complicates the numerical implementation of the various projections. Currently, the filters are truncated so that their values are identically zero after some finite extent. The decay on the filters allow control of the resulting errors, but there is still a tradeoff between error and filter length. More detailed analysis should be performed to understand this trade space.

There are two potential ways to deal with this situation. The first is to observe that the infinite extent of the filters may give rise to a pole-zero representation. That is, the filters may be expressed as a ratio of trigonometric polynomials which migh yield a fast implementation as a finite difference equation. The non-causality of the filters would have to be dealt with however.

The second idea is to allow for bi-orthogonal scaling functions and wavelets. With bi-orthogonal functions, the analyzing and synthesizing wavelets (and scaling functions) are duals of one another and are generally not translational orthogonal. This allows some freedom in designing these functions for specific criteria. For example, the analyzing functions can be designed with compact support while the synthesizing functions are infinitely supported. Alternatively, the analyzing functions can be designed with infinite support but satisfying specific bandwidth criteria.

With respect to the specific emitter identification problem, the application of the RRA is simplistic in the sense that it does not rely on the perfect reconstruction property of the B-splines in the context of the RRA. The use of a Gaussian classifier is also simplistic and may not be the most appropriate type to use in this regard. For instance, the covariance matrix used in the classifier is assumed to be diagonal since the theoretical correlation between successive detail coefficient sets is not apparent. More research should be done to determine a model with greater fidelity in this regard.

7.4 Conclusions

This dissertation provides the theoretical framework for all multiresolution analyses with rational dilation factors. While non-integer dilation factors lead to resolution analyses in which the successive approximation spaces are not embedded, this research has shown the embedded analyses with integer dilation factors are a subset of this larger analysis, albeit with special nice properties.

In the context of the discrete-time and digital signal processing realm, the RRA is recognized to contain the classic rational sampling rate change where the lowpass filter is intrinsicly defined in terms of the scaling function and dilation factor. This connection is satisfying in the sense that it affirms the belief that a significant technical accomplishment has many manifestations. This is similar to Mallat's observation that many people were doing the same thing in area of resolution analyses, each with there own different vernacular. They might have realized their similarity had there areas of research ever intersected; an often unlikely occurrence.

As a final word to this document, the state of the art in resolution analyses has been expanded by this dissertation. Certain results such as the theorems dealing with the dual dilated polynomials are general enough to transcend their limited application to the development of RRA theory. They are applicable to a wider audience than those focusing on wavelets and multiresolution analysis.

Appendix A. Supporting Lemmas

Lemma A.1 *Let $p, q \in \mathbb{N}$ with $(p,q) = 1$. The map $\rho : \mathbb{Z}_q \times \mathbb{Z}_p \to \mathbb{Z}_{pq}$ defined by*

$$\rho(l,k) = (lp + kq) \bmod pq.$$

is an isomorphism.

Proof. We want to show ρ is a bijective mapping. It is sufficient to show ρ is injective (one-to-one) since the range and domain of the map have the same finite cardinality. Thus injective will imply surjective (onto). We need to show:

$$\rho(l_1, k_1) = \rho(l_2, k_2) \Longrightarrow l_1 = l_2, k_1 = k_2.$$

Suppose $\rho(l_1, k_1) = \rho(l_2, k_2)$. Then

$$\begin{aligned}(l_1p + k_1q) \bmod pq = (l_2p + k_2q) \bmod pq &\Longrightarrow (l_1p + k_1q) \bmod pq - (l_2p + k_2q) \bmod pq = 0 \\ &\Longrightarrow ((l_1p + k_1q) - (l_2p + k_2q)) \bmod pq = 0 \\ &\Longrightarrow (l_0p + k_0q) \bmod pq = 0\end{aligned}$$

where $l_0 = l_1 - l_2$ and $k_0 = k_1 - k_2$. Continuing:

$$\begin{aligned}(l_0p + k_0q) \bmod pq = 0 &\Longrightarrow l_0p + k_0q = npq, \qquad n \in \mathbb{Z} \\ &\Longrightarrow l_0 \in q\mathbb{Z} \text{ and } k_0 \in p\mathbb{Z}.\end{aligned}$$

This last implication comes from the observation:

$$ap + bq = npq \Longrightarrow \begin{cases} ap/q = np - b \in \mathbb{Z} \\ bq/p = nq - a \in \mathbb{Z} \end{cases}$$

$$\Longrightarrow \begin{cases} a \in q\mathbb{Z} \\ b \in p\mathbb{Z} \end{cases}$$

since p and q are relatively prime. This justifies the last step in the line of reasoning above. The logical conclusion is that

$$l_1 = l_2 \text{ and } k_1 = k_2$$

which comes from the fact $|l_0| < q$ and $|k_0| < p$ so that $l_0 \in q\mathbb{Z}$ implies $l_0 = 0$ or $l_1 - l_2 = 0$ and similarly for $k_1 - k_2$. This shows that ρ is an invertible map between $\mathbb{Z}_p \times \mathbb{Z}_q$ and $\mathbb{Z}_{pq}$ so that

$$\{(lp + kq) \bmod pq\}_{l\in\mathbb{Z}_q, k\in\mathbb{Z}_p} = \mathbb{Z}_{pq}.$$

which we set out to prove

■

Lemma A.2 *Define the max operator $M_N : l \to l$ by*

$$(M_N x)(n) = \max\{x(n + j)\}_{j\in\mathbb{Z}_N}.$$

If $x \in l^p$, then $y = M_N x \in l^p$ as well where $1 \leq p \leq \infty$.

Proof. Let x and p be fixed arbitrarily. We need to show $||y|| < \infty$. Consider the sequence z defined by

$$z(n) = x(\lfloor n/N \rfloor) \quad n \in \mathbb{Z}.$$

Clearly $||y|| \leq ||z|| = N^{1/p}||x|| < \infty$, which holds for all p such that $1 \leq p \leq \infty$.

■

Appendix B. Description of MATLAB *Processing with Sequence Data Structures*

To facilitate the manipulation of sequences throughout this research, a sequence data structure was developed. The sequence data structure provides an explicit reference to the zeroth position in a vector. MATLAB handles vectors, specifically row vectors, simply as an unindex series of numbers. In the MATLAB indexing scheme, the first element of a vector is indexed as 1, the second as 2, and so on. Thus, prepending a zero to the sequence effectively shifts it one index to the right. The need for this data structure becomes apparent when considering the processing involved with downsampling a vector. Without a zero reference, downsampling cannot be accomplished.

The sequence data structure solves this problem by fixing the first position of a sequence to contain the index of the zeroth position of the remaining data. This is best illustrated by example. Suppose the following symetric sequence is centered at the origin: `x = [1 2 3 2 1]`. To represent its orientation, the corresponding sequence would be `sq_x = [3 1 2 3 2 1]`. The "sq_" prefix is used throughout the processing to indicate an instance of a sequence data type or a subroutine or function which operates on sequence data types.

The remainder of this section is to describe some of the functions and subroutines which are used on sequence data types. For the most part, the functionality of these routines can be inferred from their names with appropriate modifications for the sequence arguments. The following table provides a list of squence functions and a brief description of their functionality. The functions are listed in their entirety on the following pages.

Table B.1 Summary of Sequence Processing Functions

Sequence Function	Description
sq_absfft.m	Returns the pointwise modulus-squared of the data portion of the argument sequence.
sq_add.m	Returns the index-aligned pointwise sum of two sequences.
sq_ai.m	A simple routine that returns the values of a sequence around the zero point along with the corresponding indices.
sq_conv.m	Returns the discrete convolution of two sequences.
sq_conv2.m	Returns the discrete convolution of a matrix of sequences with a fixed sequence: row-wise convolution with index alignment.
sq_corr.m	Returns the discrete correlation of two argument sequences.
sq_corr2.m	Returns the discrete correlation of a matrix of sequences with a fixed sequence: row-wise correlation with index alignment.
sq_cut.m	Returns a the argument sequence truncated about the zero point.
sq_downs.m	Downsamples the argument sequence.
sq_fft.m	Performs the FFT on the data portion of a sequence with proper phase correction to account for the zero point.
sq_flip.m	Returns the argument sequence flipped about its zero point.
sq_mult.m	Multiplies the argument sequence s by the scalar; the zero point is unchanged.
sq_norm.m	Returns the p-norm of a sequence; p is passed as an argument.
sq_plot.m	A simple routine which plots a sequence with proper zero alignment.
sq_plot_s.m	A modification of the routine above which allows the plotting range to be scaled.
sq_stem.m	A simple routine which plots a sequence using the "stem" style.
sq_strip.m	Described above.
sq_sub.m	Returns the difference between two argument sequences.
sq_thresh.m	Returns a truncated version of the argument sequence where the truncation is determined by an argument threshold.
sq_ups.m	Returns an upsampled version of the argument sequence.

Table B.2 Listing of sq_absfft.m

```
function Sq=sq_absfft(sq,N);
[z,s]=sq_strip(sq);
if nargin < 2,
  N=length(s);
  end
Sq=absfft(s,N);
```

Table B.3 Listing of sq_add.m

```
function x=sq_add(a,b)
% Usage: x=sq_add(a,b).  Adds two sequences pointwise.  Aligns the
% zeropoints for correct values and assumes zeros for points which
% don't overlap.
%
[az,as]=sq_strip(a);
[bz,bs]=sq_strip(b);
%
% Find the proper length of the resulting sequence
xz=max(az,bz);
la=length(as)-az;lb=length(bs)-bz;
xs=zeros(1,xz+max(la,lb));
xs((xz-az+1):xz+la)=as;
xs((xz-bz+1):xz+lb)=xs((xz-bz+1):xz+lb)+bs;
%
x=[xz,xs];
```

Table B.4 Listing of sq_ai.m

```
function f=sq_ai(x)
%
[xz,xs]=sq_strip(x);
a=[[(1:length(xs))-xz]',xs(:)];
f=a(find(abs(a(:,1))<11),:);
```

Table B.5 Listing of sq_conv.m

```
function x=sq_conv(a,b)
%
[az,as]=sq_strip(a);
[bz,bs]=sq_strip(b);
%
xs=conv(as,bs);
xz=az+bz-1;
%
x=[xz,xs];
```

Table B.6 Listing of sq_conv2.m

```
% function x=sq_conv2(a,b)
%
% This function takes a matrix sequence a and convolves it
% with a scalar sequence b: 'a' is a matrix whose rows are
% sequences; 'b' is a standard row sequence.
function x=sq_conv2(a,b)
%
[az,as]=sq_strip(a);
[bz,bs]=sq_strip(b);
%
xs=conv2(bs,as);
xz=az+bz-1;
%
x=[xz,xs];
```

Table B.7 Listing of sq_corr.m

```
function x=sq_corr(a,b)
%
[az,as]=sq_strip(a);
[bz,bs]=sq_strip(b);
%
n=length(as);
i=n:-1:1;
xs=conv(as(i),bs);
xz=n-az+bz;
%
x=[xz,xs];
```

Table B.8 Listing of sq_corr2.m

```
function x=sq_corr2(a,b)
% function x=sq_corr2(a,b)
%
% This function takes a matrix sequence a and correlates  it with a
% scalar sequence b: 'a' is a matrix whose rows are sequences, 'b' is
% a standard row sequence.
%
[az,as]=sq_strip(a); [bz,bs]=sq_strip(b);
%
[m,n]=size(as);
xs=xcorr2(bs,as);
xz=n-az+bz;
%
x=[xz,xs];
```

Table B.9 Listing of sq_cut.m

```
function xc=sq_cut(x,N);
% Usage: xc=sq_cut(x,N);
% This function takes a sequence x, truncates the data sequence
% symmetrically about the zero point, and returns a length N sequence.
[xz,xs]=sq_strip(x);
xcz=ceil(N/2);
o=1:N;
xcs=xs(xz-xcz+o);
xc=[xcz,xcs];
```

Table B.10 Listing of sq_downs.m

```
function xd=sq_downs(x,p)
% Usage: xd=sq_downs(x,p)
%
% x is the sequence row vector; first element is zeropoint
% p is the downsample factor
%
% BUG: for matrix sequences, the downsample assumes all row sequences
% have the same zero point.
%
[z,s]=sq_strip(x);  %separate the data
%
[m,n]=size(s);
%
o=rem(z-1,p)+1;
%
xds=s(:,o:p:n);  % <-- potential problem here
xdz=ceil(z/p);
%
xd=[xdz,xds];
```

Table B.11 Listing of sq_fft.m

```
function X=sq_fft(x,N)
%
[xz,xs]=sq_strip(x);
if nargin <2 N=length(xs);end
%
xp=zeros([1,N]);xp(1:length(xs))=xs;
xp=rshift(xp,xz-1);
Xs=fft(xp);
Xz=1;
X=[Xz,Xs];
```

Table B.12 Listing of sq_flip.m

```
function f=sq_flip(x)
%
[z,s]=sq_strip(x);
n=length(s);
fs=fliplr(s);
fz=n-z+1;
f=[fz,fs];
```

Table B.13 Listing of sq_mult.m

```
function x=sq_mult(a,s)
% Usage: x=sq_mult(a,s).  Multiplies the sequence s by the scalar a.
%
[sz,ss]=sq_strip(s);
%
x=[sz,a*ss];
```

Table B.14 Listing of sq_norm.m

```
function n=sq_norm(sq_x,p)
% Usage: n = sq_norm(sq_x,p)
% This function returns the p-norm of the sequence.
% Without the second argument, p=2 by default.
%
% Modified to sort the array before calculating the norm
if nargin < 2 , p=2; end
[xz,xs]=sq_strip(sq_x);
n=norm(sort(xs),p);
```

Table B.15 Listing of sq_plot.m

```
function f=sq_plot(x,r,c)
%
[xz,xs]=sq_strip(x);
range=[1-xz:length(xs)-xz];
xM=ceil(max(xs));
xm=floor(min(xs));
if nargin <3, c='y'; end
plot(range,xs,c);
if nargin>=2,axis([-r,r,xm,xM]);end
return
```

Table B.16 Listing of sq_plot_s.m

```
function f=sq_plot_s(x,s,r,c)
% A modification of sq_plot to allow a scale parameter
% in the plotting range.
[xz,xs]=sq_strip(x);
range=[1-xz:length(xs)-xz]/s;
xM=ceil(max(xs));
xm=floor(min(xs));
if nargin <4, c='y'; end
plot(range,xs,c);
if nargin>=3,axis([-r,r,xm,xM]);end
return
```

Table B.17 Listing of sq_stem.m

```
function d=sq_stem(x,linetype,dottype)
%
[z,s]=sq_strip(x);
y=[1:length(s)]-zp;
stem(x,y,linetype,dottype);
```

Table B.18 Listing of sq_strip.m

```
function [z,s]=sq_strip(sq_x)
% Usage: [z,s]=sq_strip(sq_x)
% This function separates the data from the zero point
%
% modified 6/18/96 to handle matrix sequences
[m,n]=size(sq_x);
z=sq_x(:,1);
s=sq_x(:,2:n);
```

Table B.19 Listing of sq_sub.m

```
function x=sq_sub(a,b)
% Usage: x=sq_sub(a,b).  Subtracts 'b' from 'a' pointwise. Aligns
% the zeropoints for correct values and assumes zeros for points which
% don't overlap.
%
[az,as]=sq_strip(a);
[bz,bs]=sq_strip(b);
%
% Find the proper length of the resulting sequence
xz=max(az,bz);
la=length(as)-az;lb=length(bs)-bz;
xs=zeros(1,xz+max(la,lb));
xs((xz-az+1):xz+la)=as;
xs((xz-bz+1):xz+lb)=xs((xz-bz+1):xz+lb)-bs;
%
x=[xz,xs];
```

Table B.20 Listing of sq_thresh.m

```
function t = sq_thresh(s,threshold)
% function t = sq_thresh(s,threshold)
%
% This function takes a sequence retains the "middle" (contiguous)
% portion which has element magnitudes above the threshold.
%
z=s(1);
if nargin < 2, threshold = eps; end
l=2;
   while (abs(s(l))< threshold)*(l<=z), l=l+1; end
h=length(s);
   while (abs(s(h))< threshold)*(h>(z+1)), h=h-1; end
%
t=[s(1)-l+2, s(l:h)];
```

Table B.21 Listing of sq_ups.m

```
function xu=sq_ups(x,p)
%
[z,s]=sq_strip(x);
%n=length(s);
[m,n]=size(s);
%xus=zeros(1,p*(n-1)+1);xus(p*[0:n-1]+1)=s;
xus=zeros(m,p*(n-1)+1);xus(:,p*[0:n-1]+1)=s;
xuz=p*(z-1)+1;
%
xu=[xuz,xus];
```

Appendix C. The Orthonormalization of Semiorthogonal B-Spline Scaling and Wavelet Filters

This Appendix contains various MATLAB code listings for the generation of orthonormal spline scaling filters which are used throughout this dissertation. There are several versions of the basic code, each designed for a specific input and to yield a specific output.

The mainstay code is `on_filter.m` which is listed at the end of this section. This routine performs Daubechies' "orthonormalization trick" on a vector of numbers to yield a set of filter coefficients which correspond to the associated orthonormal scaling function or wavelet. The equation which is implemented in the following code was presented in Chapter I. Note that the code does no error checking with respect to division by zero. Hence, the code will fail violently if the discrete Fourier spectrum of the argument sequence has a zero. The spectrum not having a zero in $[0, 2\pi)$ is a condition for orthonormalization which was discussed earlier.

The code takes several arguments: the dilation factor p, the spline degree m, the total number filter coefficients to return N, and an optional argument which allows the user to specify the initial phase on the input sequence. This optional argument is used in some cases where the first element of the sequence to orthonormalized is not indexed to 0.

Example C.1 To calculate the orthonormal B-spline scaling function associated with a p dilation and m regularity, the following code would be used to generate a 1024 tap sequence:

```
matlab> p=3; m=2; hp=[1 2 3 2 1];
matlab> filter=on_filter(hp,p,m,1024);
```

The values for `hp` are determined by the m-fold convolution of the p-length regularity sequence. For calculating wavelets, the implementation is nearly identical except that the semi-orthogonal wavelet sequence is substituted for `hp`. For instance, to orthonormalize semi-orthogonal wavelet sequence, one might use the following:

Table C.1 Listing of `on_filter.m`

```
function q_on = on_filter(q0,p,m,N,ph);
% Usage: q_on = on_filter(q0,p,m,N,ph)
%wave
% q0 is the two-scale sequence,
% p  is the dilation factor (3)
% m  is the degree of the B-spline (2)
% N  is the number of points in the fft and ifft calculations, should
%    be power of 2.  Increasing N yields more accurate filter
%    coefficients. (256)
% ph is an optional argument which represents the zero index of q0. (1)
%
% The resulting filter coefficients, q_on, satisfy the pseudo two-scale
% relationship:
%                \psi_{1,0} = \sum_k q_on(k) \phi_{0,k}
%
% so that the normalization constant is included with the
% coefficients.  That is, sum(q_on)=sqrt(p).
%
if nargin < 5, ph=1;  end
if nargin < 4, N=256; end
if nargin < 3, m=2;   end       %set the defaults
if nargin < 2, p=3;   end
%
L=length(q0);                   %length of the wavelet sequence
Nm=b_spline(2*m,1:2*m-1);
%
E=zeros([1,N]);
E(1:2*m-1)=Nm;
E=fft(rshift(E,m-1));           %the numerator of the filter
%
% set up the unsampled sequence for the denominator
%
x=conv(Nm,corr(q0,q0));
k=1:length(x);
[v,z]=max(x);                   %find the zero point
zid=find(rem(k-z,p)~=0);
x(zid)=zeros(size(zid));        %insert zeros
%
d = zeros([1,N]);               %initializes the denominator sequence
d(k)=x;d=rshift(d,z-1);         %correctly positions the coefficients
D = fft(d);                     %Fourier domain representation of
%denominator
w=exp(i*2*pi*(ph-1)*(0:N-1)/N); %phase adjustment for the zeropoint of q0
Q_on = sqrt(E./D).*fft(q0,N).*w;%pointwise calculate the filter
q_on = fftshift(ifft(Q_on));    %ifft to find the coefficients
if isreal(q0), q_on=real(q_on);end
```

```
matlab> p=3, m=2; gp=[2 -5 6 -5 2];
matlab> w_filter = on_filter(gp, p, m);
```

where the number of taps in the wavelet (detail) filter would default to 256 since it was unspecified. ■

There are several routines which are used in the routine listed above. For completness, they are listed here. The `corr` routine returns the unbiased correlation of two row vector arguments.

Table C.2 Listing of `corr.m`

```
function x=corr(a,b)
% CORR  returns the correlation between a and b. The first argument
% is flipped (lr) and convolved with the second.
% Arguments must be row vectors otherwise a simple
% convolution of the two is returned.
index=length(a):-1:1;
x=conv(a(index),b);
```

The `b_spline` routine is used to calculate B-splines at specific points. The points most often of interest are the integers.

The `rshift` routine returns a right-shifted version of its row-vector argument. Values are wrapped during the shift.

Table C.3 Listing of `b_spline.m`

```
%BSPLINE  This function calculates the value of the m-th order cardinal
% B-spline when evaluated at x:  N_m(x)
%
% N_1(x) is the characteristic function on [0,1), N-m(x) is defined
% recursively by
%
%               N_{m}(x) = \int_0^1 N_{m-1}(x-t) dt,     m > 1.
%
% This function evaluates the spline by using equation 4.2.4 in Chui's
% _A_Introduction_to_Wavelets_.
%
function S=b_spline(m,x)
if m==1,  S = ones(size(x)); i=find(floor(x));  S(i)=zeros(size(i));
else  S = (x.*b_spline(m-1,x)+(m-x).*b_spline(m-1,(x-1)))/(m-1);
end
%
```

Table C.4 Listing of `rshift.m`

```
function y = rshift(x,i)
% returns the periodic version of x right-shifted
% by 'i'
%
n=length(x);
y=x(rem((1:n)+i-1+n,n)+1);
```

Appendix D. Calculation of Semiorthonormal B-Spline Wavelets and Detail Filter

This appendix provides the specific MATLAB code used in the calculation of the semiorthogonal spline wavelets from Chapter IV. The orthogonalized versions of these wavelets are used for the calculation of feature vectors for the $p = 3$ cases in Chapter VI (both MRA and RRA). The theory behind these calculations was discussed in Chapter IV and this appendix will provide the theoretical detail necessary to correlate the code with the theory.

There are two primary parts of the calculation. The first is the calculation of the semiorthogonal wavelets via the nullspace vectors of the approximation projection matrix. Following the example from Chapter IV, consider the following code fragment:

As was mentioned in Chapter IV, the nullspace vectors from the calculation are unit norm, from which two integer-valued nullspace vectors can be formed via adding and subtracting flipped versions of one another. We know the resulting nullspace vectors span the nullspace of the matrix since they are clearly linearly independent and there are two vectors when the dimension of the nullspace is also two. It is interesting to note that we could have created symmetric semiorthogonal wavelet from `v2` and the antisymmetric one from `v1`. However, the resulting vectors would have been identical.

The second phase of the calculation is to transform the semiorthogonal wavelets into orthogonal ones via the orthonormalization trick. The technique to do this is described in detail in Chapter IV, but the actual code which implements the calculation is provided in Table D.2.

The calculation detailed in Table D.2 yields the compactly supported sequence `sq_r` which is plotted in its normalized form in Figure 4.3. The corresponding integer-value coefficients are listed in Table D.3.

The final orthonormal wavelets sq_g1 and sq_g2 were calculated by orthonormalizing sq_q1 and sq_r, respectively. The choice to calculate sq_r from sq_q2 was made arbitrarily. It could have just as well been calculated from sq_q1 with the final orthonormal wavelets being calculated from sq_r and sq_q2. The semiorthogonal wavelet sq_r would generally be different, as would the resulting orthonormal wavelets.

For larger dilation factors, there are more wavelets and more nullspace vectors from which to form semiorthogonal wavelets. Forming symmetric and antisymmetric semiorthogonal wavelets from the nullspace vectors may not necessarily result in a linearly independent set. Furthermore, some vectors may possibly be formed from linear combinations of the kp-shifted versions of the others where k is an arbitrary integer. This means that vector lies in the space spanned by the others, where "space" means the linear span of all p-shifts of the particular wavelet.

These two conditions do not pose any significant problems since they will become apparent during the orthonormalization process. The resulting vector will be zeroed which would indicate the condition. In fact, for larger dilation factors with many more wavelets, it may well be a sound engineering practice to eliminate those wavelets which have values close to zero, i.e. not orthogonal enough to the previous wavelets and the spaces they span. There is a great deal of flexibility in designing the wavelets which widens the engineering trade space for specific applications.

Table D.1 MATLAB Example Illustrating Calculation of Nullspace Vectors from Approximation Projection Matrix

```
matlab>  A=[12 6 1 0 0; 6 12 16 12 6; 0 0 1 6 12]
A =
    12     6     1     0     0
     6    12    16    12     6
     0     0     1     6    12

matlab> q=null(A)
v =
   -0.3776         0
    0.8115   -0.0864
   -0.3381    0.5183
   -0.2480   -0.7775
    0.1522    0.3455

matlab> v1=v(:,1)'; v2=v(:,2)';
matlab> q1=v1+fliplr(v1)  % create the antisymmetric s.o. wavelet sequence
q1 =
   -0.2254  0.5635  -0.6762   0.5635  -0.2254

matlab> q2=v2-fliplr(v2)  % create the antisymmetric s.o. wavelet sequence
q2 =
   -0.3455    0.6911         0   -0.6911    0.3455

matlab> q1=2*q1/q1(1)  % renormalize to get integer values
q1 =
    2.0000   -5.0000    6.0000   -5.0000    2.0000

matlab> q2=q2/q2(1)    % renormalize to get integer values
q2 =
    1.0000   -2.0000         0    2.0000   -1.0000
```

Table D.2 MATLAB Example Illustrating Calculation of Orthonormal Wavelets from a Linearly Independent Set of Semiorthogonal Wavelets

```
%
% Linear spline, p=3
%
N=256;
p=3;m=2;
sq_q1=[1 2 -5  6 -5  2 0];  % set up the s.o. wavelet sequences
sq_q2=[1 1 -2  0  2 -1 0];
%
sq_Nm=[m,b_spline(2*m,1:2*m-1)]; % the autocorr of the lin spline
%
sq_E1=sq_conv(sq_Nm,sq_corr(sq_q1,sq_q1));
sq_B=sq_conv(sq_Nm,sq_corr(sq_q1,sq_q2));
%
sq_E1=sq_ups(sq_downs(sq_E1,p),p);
sq_B=sq_ups(sq_downs(sq_B,p),p);
%
sq_r = sq_sub(sq_conv(sq_E1,sq_q2),sq_conv(sq_B,sq_q1));
%
[g2z,g2]=sq_strip(sq_r);
[g1z,g1]=sq_strip(sq_q1);
%
g1_on=on_filter(g1,p,m,N,g1z);sq_g1_on=[N/2+1,g1_on];
g2_on=on_filter(g2,p,m,N,g2z);sq_g2_on=[N/2+1,g2_on];
```

Table D.3 List of sq_r Values

Index	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
Value	0	3	-12	36	24	-123	0	123	-24	-36	12	-3	0

Appendix E. Calculation of Feature Vectors and SEI Processing

This appendix contains much of the MATLAB code used to implement the processing and classification of the radar pulses for the specific emitter identification problem described in Chapter VI. There are two basic types of processing here: the construction of feature vectors from the sampled radar pulses, and the classification of those radar pulses using the feature vectors. The theory behind the processing is mostly contained in Chapter VI; this appendix augments that information with specific details. It is not intended to substitute.

E.1 Feature Vector Construction

For the purposes of multiple classifications, feature vectors were formed from the raw sampled radar pulses and stored in intermediate files so that the feature vectors would not have to be constructed each time a classification was performed. This was especially important when the training sets were changed from run to run. Each type of processing had a set of MATLAB files which contained the feature vectors associated with that type of processing. The code listing in Table E.1 is a typical example of the code used to construct feature vector files for a specific type of processing. In this instance, feature vector files are contructed for the RRA processing with $p = 3$, $q = 2$, $m = 2$, and 3 levels of decomposition.

Some items to note when evaluating this code. The MATLAB file `rf_3_2_2.mat` contains the actual RRA filters which are made global on the line following the inclusion of the file. The resulting output feature vector files have the form `e_rra_k_j.mat` where `k` and `j` represent the apple (radar) number and data set, respectively. The apple number is 1, 2, 3, or 4 to correspond, respectively, to golden, granny, ida, and mac.

Table E.1 Listing of `e_rra.m` for $p = 3$, $q = 2$, $m = 2$, and 5 decomposition levels

```
% script file to calculate feature vectors
% from the various data sets
```

```
%
apples    = str2mat('gold','gran','ida','mac');
num_apples = 4;
num_datasets = 3;
decomp_levels = 5;
%
% Load the filter for the RRA entropy calculation
%
load rf_3_2_2
global sq_hp_on sq_hq_on sq_g1_on sq_g2_on
%
% Setup the loops for the various data sets
%
for k = 1:num_apples, for j=1:num_datasets
   dataset = [deblank(apples(k,:)),'_',int2str(j)]; disp(dataset)
   eval(['load ',dataset]);
   eval(['c = ',dataset,''';']);
   sq_c = [ones(n_pulses,1),c];
%
% Setup the various data structures
%
   entropy = zeros((p-1)*decomp_levels,n_pulses);
   for l = 1:decomp_levels
      sq_d1 = sq_downs(sq_corr2(sq_g1_on,sq_c),p);
      sq_d2 = sq_downs(sq_corr2(sq_g2_on,sq_c),p);
      sq_c = sq_downs(sq_corr2(...
             sq_hp_on,sq_conv2(sq_hq_on,sq_ups(sq_c,q))),p);

      [Dz,D]=sq_strip(sq_d1); clear Dz
      for ii = 1:n_pulses,
         D(ii,:)=D(ii,:)/norm(D(ii,:));
         % Normalize the detail coefficient sets
      end
      D=D.*D;  % Calculate the pointwise square of the sequences
      [row,col] = find(D==0); % Find the zero elements for handling
      for jj = 1:length(row),
         D(row(jj),col(jj))=1;
      end
      D=D.*log(D); % The actual entropy calculation
      entropy((p-1)*l-1,:) = -sum(D'); % Store in the data structure

      [Dz,D]=sq_strip(sq_d2); clear Dz
      for ii = 1:n_pulses,
         D(ii,:)=D(ii,:)/norm(D(ii,:));
         % Normalize the detail coefficient sets
      end
      D=D.*D; % Calculate the pointwise square of the sequences
      [row,col] = find(D==0); % Find the zero elements for handling
      for jj = 1:length(row),
         D(row(jj),col(jj))=1;
      end
```

```
      D=D.*log(D); % The actual entropy calculation
      entropy((p-1)*l,:) = -sum(D'); % Store in the data structure
   end % 'l' loop

   eval(['save ','e_rra',int2str(k),'_',int2str(j),' entropy ...
        decomp_levels dataset p q m']);
end % 'j' loop
end % 'k' loop
```

E.2 Classification

Classification for a particular type of processing involves selecting a training set and at test set. The following MATLAB script forms the shell of a typical classification run.

Table E.2 Listing of s_cnfrra.m for $p = 3$, $q = 2$, $m = 2$, and 5 decomposition levels

```
% Script file to calculate the confusion matrices
% of the RRA; 5 level decomposition
train_mat=[...      % A text matrix constructing input file names
'123 ';'123 ';'123c';'123c';...
'12  ';'12  ';'12  ';'12  ';...
'12c ';'12c ';'12c ';'12c ';...
'13  ';'13  ';'13  ';'13  ';...
'13c ';'13c ';'13c ';'13c ';...
'23  ';'23  ';'23  ';'23  ';...
'23c ';'23c ';'23c ';'23c ';...
'1   ';'1   ';'1   ';'1   ';...
'1c  ';'1c  ';'1c  ';'1c  ';...
'2   ';'2   ';'2   ';'2   ';...
'2c  ';'2c  ';'2c  ';'2c  ';...
'3   ';'3   ';'3   ';'3   ';...
'3c  ';'3c  ';'3c  ';'3c  '];
test_mat=[...     % A text matrix constructing input file names
'123 ';'123c';'123 ';'123c';...
'12  ';'12c ';'3   ';'3c  ';...
'12  ';'12c ';'3   ';'3c  ';...
'13  ';'13c ';'2   ';'2c  ';...
'13  ';'13c ';'2   ';'2c  ';...
'23  ';'23c ';'1   ';'1c  ';...
'23  ';'23c ';'1   ';'1c  ';...
'1   ';'1c  ';'23  ';'23c ';...
'1   ';'1c  ';'23  ';'23c ';...
'2   ';'2c  ';'13  ';'13c ';...
'2   ';'2c  ';'13  ';'13c ';...
```

```
'3   ';'3c  ';'12  ';'12c ';...
'3   ';'3c  ';'12  ';'12c '];
%
n_apples = 4;  % The number of apples
[n_runs,dummy]=size(test_mat); % Sets up n_runs
o_rate_mat = zeros(n_runs,1); % Initializing the overall rate matrix
rate_mat = zeros(n_runs,n_apples); % Initializing the rate matrix
%
for k=1:n_runs,
  train_string = deblank(train_mat(k,:));
  % Sets the string corresponding to the training files

  test_string = deblank(test_mat(k,:));
  % Sets the string corresponding to the testing files

  [confusion,rate,o_rate] = e_class(train_string,test_string);
  % Calls the routine which performs the classifications and
  % calculates the confusion matrix

  eval(['cnf_',train_string,'_',test_string,' = confusion;']);
  % Sets the cnf_xxx_yyy variable as the confusion matrix

  rate_mat(k,:) = rate;
  o_rate_mat(k,:) = o_rate;
end
```

The majority of the processing is performed in the routine `e_class.m` which is listed in Table E.3. This routine performs all of the classifications regardless of feature vector size or processing source. The only tricky item to note is the use of the `train_with_culled` and `test_with_culled` variables. If these variables are set, only the subset of the feature vectors corresponding to the culled data is used for in the training or testing.

Table E.3 Listing of `e_class.m`

```
function [confusion,rate,overall_rate] = e_class(trainset,testset)
%
% This function set up for RRA classifications
%
datatype = 'rra';
n_apples = 4;
%
apples = str2mat('gold','gran','ida','mac');
train_with_culled = 0;
test_with_culled = 0;
```

```
%
l_trn = length(trainset);
l_tst = length(testset);
%
if lower(trainset(l_trn))=='c',
  train_with_culled = 1;
  l_trn = l_trn - 1;
end
if lower(testset(l_tst)) =='c',
  test_with_culled  = 1;
  l_tst = l_tst - 1;
end
%
for k = 1:n_apples, for j = 1:l_trn,
  e_trn_set = str2mat(e_trn_set,...
    ['e_',datatype,int2str(k),'_',trainset(j)]);
end,end
e_trn_set = e_trn_set(2:j*k+1,:);
%
for k = 1:n_apples, for j = 1:l_tst,
  e_tst_set = str2mat(e_tst_set,...
    ['e_',datatype,int2str(k),'_',testset(j)]);
end,end
e_tst_set = e_tst_set(2:j*k+1,:);

%
confusion=zeros(n_apples,n_apples);
e_m = zeros(10,n_apples);
e_v = zeros(10,n_apples);
%
%  CALCULATING MEANS AND VARIANCES
%
for k = 1:n_apples,
  e = [];
  for j = 1:l_trn,
    tr_file = e_trn_set((k-1)*l_trn+j,:);
    eval(['load ',tr_file]);
    if train_with_culled,
      eval([tr_file,'=',tr_file,'(:,c_index);']);
    end
    eval(['e = [e,',tr_file,'];']);
  end % j loop
  e_m(:,k) = mean(e')';
  e_v(:,k) = std(e')'.^2;
end % k loop
%
% CALCULATING CLASSIFICATIONS
%
confusion = zeros(n_apples,n_apples);
%
for k=1:n_apples
```

```
  E=[];  % Clear the entropy matrix
  for j=1:l_tst,
    tst_file = e_tst_set((k-1)*l_tst+j,:);
    eval(['load ',tst_file]);
    if test_with_culled,
      eval([tst_file,'=',tst_file,'(:,c_index);']);
    end
    eval(['E = [E,',tst_file,'];']);
  end % j loop

  [rs,cs] = size(E); v = ones(1,cs);
  class = zeros(cs,n_apples);

  for l=1:n_apples,
    hh = E-e_m(:,l)*v;
    class(:,l) = 1/sqrt(prod(e_v(:,l))) * ...
      exp(-0.5 * diag(hh' * diag(1./e_v(:,l)) * hh));
  end %  l loop

  [val,indx] = max(class');

  for l=1:n_apples,
    confusion(k,l) = length(find(indx==l));
  end

end % k loop
rate = diag(confusion)'./sum(confusion');
overall_rate = sum(diag(confusion))/sum(sum(confusion));
```

Bibliography

1. Anderson, Bruce P. *Theory and Implementation of Wavelet Analyses in Rational Resolution Decompositions*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433, 1992.

2. Anderson, Bruce P. and Gregory T. Warhola. "Rational Resolution Decompositions for Signal Processing." *Proceedings of the Twenty-Eighth Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, CA: IEEE, October 31–November 2 1994.

3. Apostol, Tom M. *Mathematical Analysis* (2nd Edition). Addison-Wesley, 1974.

4. Auscher, P. "Wavelet Bases for $L_2(\mathbb{R})$ with Rational Dilation Factor." *Wavelets and their Applications* edited by M. B. Ruskai and others, Jones and Barlett, 1992.

5. Auscher, Pascal. *Ondelettes Fractales et Applications*. PhD dissertation, Université de Paris-Dauphine, 1989. Translated from French.

6. Benedetto, J. J. and E. F. Walnut. "Gabor Frames for L^2 and Related Spaces." *Wavelets: Mathematics and Applications* edited by J. J. Benedetto and M. W. Frazier, CRC Press, Inc., 1994.

7. Benedetto, John J. and Shidong Li. *The Theory of Multiresolution Analysis Frames and Applications to Filter Banks*. Technical Report, College Park, MD 20742: Department of Mathematics, Unversity of Maryland, 1994. Preprint.

8. Burt, P. and E. Adelson. "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communications*, *31*:217–236 (1983).

9. Cavaraetta, Alfred S., et al. *Stationary Subdivision*, *93*. Memoirs of the American Mathmatical Society. American Mathematical Society, 1991.

10. Champeney, D. C. *A Handbook of Fourier Transforms*. Cambridge University Press, 1987.

11. Chui, Charles K. *An Introduction to Wavelets*, *1*. Wavelet Analysis and its Applications. San Diego: Academic Press, Inc., 1992.

12. Chui, Charles K., editor. *Wavelets: A Tutorial in Theory and Applications*, *2*. Wavelet Analysis and its Applications. San Diego: Academic Press, Inc., 1992.

13. Cohen, Albert, et al. "Biorthogonal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, *45*:485–560 (1992).

14. Daubechies, Ingrid. "Orthonormal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, *41*:909–996 (November 1988).

15. Daubechies, Ingrid. *Ten Lectures on Wavelets*. Society of Industrial and Applied Mathematics, 1992.

16. Devaney, Anthony J., "Wavelet and Wavepacket Signal Processing." Presentation viewgraphs.

17. Duffin, R. J. and A. C. Schaeffer. "A Class of Nonharmonic Fourier Series," *Transactions of the American Mathematical Society*, *72*:341–366 (March 1952).

18. Golub, G. H. and C. F. Van Loan. *Matrix Computations* (2nd Edition). Baltimore: Johns Hopkins University Press, 1989.

19. Heil, Christopher E. and David F. Walnut. "Continuous and Discrete Wavelet Transforms," *SIAM Review*, *31*(4):628–666 (December 1989).

20. Hubbard, Barbara Burke. *The World According to Wavelets*. A.K. Peters, 1996.

21. Kaiser, Gerald. *A Friendly Guide to Wavelets*. Birkhäuser, 1994.

22. Kovačević, Jelena and Martin Vetterli. "Perfect Reconstruction Filter Banks with Rational Sampling Rates in One and Two Dimensions," *Proc. SPIE Conf. on Visual Communications and Image Processing*, *1199* (November 1989).

23. Kovačević, Jelena and Martin Vetterli. "Perfect Reconstruction Filter Banks with Rational Sampling Rate Changes." *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*. 1785–1788. May 1991.

24. Kovačević, Jelena and Martin Vetterli. "Perfect Reconstruction Filter Banks with Rational Sampling Factors," *IEEE Transactions on Signal Processing*, *41*(6):2047–2066 (June 1993).

25. Lian, J. *Semiorthogonal Wavelets with Dilation Factor a=3*. Technical Report, Texas A&M University, October 1993. CAT Report #318.

26. Luenberger, David G. *Optimization by Vector Space Methods*. John Wiley & Sons, Inc., 1969.

27. Mallat, Stephane G. "Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(R)$," *Transactions of the American Mathematical Society*, *315*(1):69–87 (September 1989).

28. Mallat, Stephane G. "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(7):674–693 (July 1989).

29. Meyer, Yves. "Orthonormal Wavelets." *Wavelets, Time-Frequency Methods, and Phase Space* edited by J. A. Combes and Ph. Tchamichian, 21–37, Berlin: Springer-Verlag, 1989.

30. Naylor, Arch W. and George R. Sell. *Linear Operator Theory in Engineering and Science*, *40*. Applied Mathematical Sciences. New York: Springer-Verlag, 1982.

31. Oppenheim, Alan V. and Ronald W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall, 1989.

32. Smith, Mark J. and Thomas P. Barnwell. "Exact Reconstruction Techniques for Tree-Structured Subband Coders," *assp*, *34*(3):434–441 (June 1986).

33. Steffan, Peter, et al. "Theory of Regular *M*-Band Wavelet Bases," *IEEE Transactions on Signal Processing*, *41*(12):3497–3511 (December 1993).

34. Tchamitchian, Ph. "Biorthogonalité et Théorie des Opérateurs," *Revista Matemática Iberoamericana*, *3*:163–189 (1987).

35. Tou, Julius T. and Rafael C. Gonzalez. *Pattern Recognition Principles*. Number 7 in Applied Mathematics and Computation, Addison-Wesley, 1974.

36. Vaidyanathan, P. P. "Theory and Design of M-Channel Maximally Decimated Quadrature Mirror Filters with Arbitrary M, Having the Perfect-Reconstruction Property," *IEEE Transactions on Acoustics, Speech and Signal Processing*, *35*(4):476–492 (April 1987).

37. Vetterli, Martin and Didier Le Gall. "Perfect Reconstruction FIR Filter Banks: Some Properties and Factorizations," *IEEE Transactions on Acoustics, Speech and Signal Processing*, *37*(7):1057–1071 (July 1989).

38. Young, Robert M. *An Introduction to Nonharmonic Fourier Series*. Academic Press, 1980.

Vita

Captain Bruce Patrick Anderson was born May 30, 1965 in Roswell NM. He attended Woodbridge Senior High School, Woodbridge VA and received his Diploma in 1983. He attended the University of Virginia and graduated with the Bachelor of Science Degree in Elecrical Engineering in 1987 and received an ROTC commission as a Second Lieutenant, United States Air Force.

In 1988, his first assignment was to the 31st Test and Evaluation Squadron (SAC), Edwards AFB, CA where he was responsible for conducting software maintainability evaluations on B1-B operational embedded and support software. He was reassigned in 1989 as an operations engineer where he was responsible for planning, scheduling, and conducting B-1B test missions in support of several national priority test efforts.

In 1991, he was selected to attend the Air Force Institute of Technology, Wright-Patterson AFB, OH. He received the Master of Science in Electrical Engineering as a distinguished graduate and won the Commandant's Award for exceptional thesis research. He continued his studies toward the PhD at AFIT.

Captain Anderson is married to the former Ms. Julie Phipps of Boston, MA. They have one daughter, Sydney Elizabeth, and are expecting their second child in March 1998. Captain Anderson is currently assigned to the National Reconaissance Office where he is a program manager in the Advanced Systems and Technology Directorate.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1995	Ph.D. Dissertation

4. TITLE AND SUBTITLE
The Rational Resolution Analysis: A Generalization of Multiresolution Analyses with Application to Specific Emitter Identification

5. FUNDING NUMBERS

6. AUTHOR(S)
Bruce P. Anderson, Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/DS/ENC/97D-2

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Department of Defense
R571
Ft. Meade, MD 20755
Contract Number: H98230-R5-92-9740

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The rational resolution analysis (RRA) is introduced and developed as a generalization of the integer-dilation multiresolution analyses (MRA) developed by Mallat and Meyer. Rational dilation factors are achieved by relaxing the condition on MRAs that successive approximation spaces be embedded. Conditions for perfect reconstruction are discussed and it is shown that perfect reconstruction is possible with specific constraints on the scaling function: the scaling filter must have its roots on the unit circle. Furthermore, the required arrangement of the roots indicate the scaling function must be derived from a B-spline of some degree. It is proven the only compactly supported scaling function which satisfies these constraints is the Haar basis. An algorithmic approach to constructing p-dilation wavelets is presented. The frame properties along with adjoint wavelets of RRAs are is presented. It is shown the adjoint wavelets form a frame for V_0 and that the corresponding decomposition is both stable and unique. The redundant representation of the detail coefficients is exploited as a solution to the specific emitter problem. Results demonstrate the RRA is far superior to the traditional MRA and wavepacket approaches when used as a feature extractor in Bayesian classification schemes.

14. SUBJECT TERMS
Wavelets, Multiresolution Analysis, Multirate Digital Signal Processing, Rational Sampling Rate Change

15. NUMBER OF PAGES
177

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL