Multidimensional Wavelets

by

Thomas Colthurst

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997

Author ....
Department of Mathematics
May 2, 1997

Certified by ....
W. Gilbert Strang
Professor of Mathematics
Thesis Supervisor

Accepted by ....
Richard B. Melrose
Chairman, Department Committee for Graduate Students

Multidimensional Wavelets

by

Thomas Colthurst

Submitted to the Department of Mathematics
on May 2, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

A multidimensional scaling function $\phi(\vec{x}) \in L^2(\mathbf{R}^n)$ has two fundamental properties: one, it is orthonormal to its translates by an n dimensional lattice Γ, and two, it satisfies a dilation equation $\phi(\vec{x}) = \sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}}\phi(M\vec{x} - \vec{\gamma})$ for some expanding matrix M such that $M\Gamma \subset \Gamma$. If V_j is the space spanned by $\{\phi(M^j\vec{x} - \vec{\gamma})\}_{\vec{\gamma}\in\Gamma}$, then the functions $\psi_1, \ldots, \psi_{|\det M|-1}$ are *wavelets* if they and their Γ translates form a basis for W_0, the orthogonal complement of V_0 in V_1. In this thesis, I first describe the set of $c_{\vec{\gamma}}$'s that determine non-zero compactly supported scaling functions and also how the $c_{\vec{\gamma}}$'s determine the degree of smoothness of ϕ. I then prove that for every compactly supported multidimensional scaling function, there exist $|\det M| - 1$ wavelets, and that in certain special cases these wavelets can be chosen to be compactly supported as well. Finally, I show how to construct, for every acceptable matrix M, a compactly supported scaling function ϕ with compactly supported wavelets ψ_i such that $\int \psi_i(\vec{x})x_1\,d\vec{x} = 0$ for all $i = 1\ldots|\det M| - 1$. The major tool in these constructions is the wavelet system's polyphase matrix.

Thesis Supervisor: W. Gilbert Strang

Title: Professor of Mathematics

Multidimensional Wavelets

by

Thomas Colthurst

Submitted to the Department of Mathematics

on May 2, 1997, in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Abstract

A multidimensional scaling function $\phi(\vec{x}) \in L^2(\mathbf{R}^n)$ has two fundamental properties: one, it is orthonormal to its translates by an n dimensional lattice Γ, and two, it satisfies a dilation equation $\phi(\vec{x}) = \sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}}\phi(M\vec{x} - \vec{\gamma})$ for some expanding matrix M such that $M\Gamma \subset \Gamma$. If V_j is the space spanned by $\{\phi(M^j\vec{x} - \vec{\gamma})\}_{\vec{\gamma}\in\Gamma}$, then the functions $\psi_1, \ldots, \psi_{|\det M|-1}$ are *wavelets* if they and their Γ translates form a basis for W_0, the orthogonal complement of V_0 in V_1. In this thesis, I first describe the set of $c_{\vec{\gamma}}$'s that determine non-zero compactly supported scaling functions and also how the $c_{\vec{\gamma}}$'s determine the degree of smoothness of ϕ. I then prove that for every compactly supported multidimensional scaling function, there exist $|\det M| - 1$ wavelets, and that in certain special cases these wavelets can be chosen to be compactly supported as well. Finally, I show how to construct, for every acceptable matrix M, a compactly supported scaling function ϕ with compactly supported wavelets ψ_i such that $\int \psi_i(\vec{x})x_1\,d\vec{x} = 0$ for all $i = 1\ldots|\det M| - 1$. The major tool in these constructions is the wavelet system's polyphase matrix.

Thesis Supervisor: W. Gilbert Strang

Title: Professor of Mathematics

Acknowledgments

> Like the vague sighings of a wind at even,
> That makes the wavelets of the slumbering sea.
> — Shelley, *Queen Mab* viii. 24

> You only hide it by foam and bubbles, by wavelets and steam-clouds, of ebullient rhetoric.
> — Coleridge, *Literary Remembrances* III 360

First and foremost, I must thank my advisor, Professor Strang, for the many useful suggestions, preprints, and pieces of advice that he has given me over the past three years.

I also acknowledge discussions with Chris Heil, Seth Padowitz, and Vasily Strela which have increased my knowledge of wavelets and mathematics.

The following people proofread versions of this thesis and made valuable comments for which I am indebted: Arun Sannuti, Robert Duvall, Andrew Ross, Manish Butte.

Finally, I must thank my parents, Karen MacArthur, and the Thursday Night Dinner Group, all for their encouragement and understanding.

Contents

List of Figures

Chapter 1

Introduction

1.1 What is a Multiresolution Analysis?

Multiresolution analysis (or multiresolution approximation) was invented by Mallat and Meyer as a way of formalizing the properties of the first wavelets that allowed them to describe functions at finer and finer scales of resolution.

A *rank* 2 *multiresolution analysis* of $L^2(\mathbf{R})$ is a sequence V_j, $j \in \mathbf{Z}$, of subspaces of $L^2(\mathbf{R}^n)$ that satisfy the five conditions of

- **Nesting:** [1] $\cdots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots$,
- **Density:** The closure of $\cup_{j\in\mathbf{Z}} V_j$ is $L^2(\mathbf{R})$,
- **Separation:** $\cap_{j\in\mathbf{Z}} V_j = \{0\}$,
- **Scaling:** $f(x) \in V_j \iff f(2x) \in V_{j+1}$, and
- **Orthonormality:** There exists a *scaling function* $\phi \in V_0$ such that $\{\phi(x - \gamma)\}_{\gamma\in\mathbf{Z}}$, the set of all the $\mathbf{Z}$-translates of ϕ, forms an orthonormal basis for V_0.

The rank 2 part of the name comes from the Scaling property which insures that V_{j+1} is a better resolution approximation to $L^2(\mathbf{R})$ than V_j is. Similarly, one defines

[1] Be warned that there is no agreement yet as to "which way the V_j's go"—Daubechies, for example, has $V_j \subset V_{j-1}$.

a *rank m multiresolution analysis* (with integer $m > 1$) by replacing the Scaling condition with $f(x) \in V_j \iff f(mx) \in V_{j+1}$.

One example of an rank 2 multiresolution analysis of $L^2(\mathbf{R})$ is the Haar basis, for which V_0 is the space of functions constant between integers. V_j is the space of functions constant between values in $2^{-j}\mathbf{Z}$, and the scaling function ϕ is the characteristic function of the unit interval $[0, 1]$. The only nontrivial thing to prove is the Density condition.

Of all the properties that can be derived from the definition of a multiresolution analysis, two stand out as most important. First, it yields a basis for $L^2(\mathbf{R})$. Define W_j to be the orthogonal complement of V_j in V_{j+1}, so that $V_{j+1} = V_j \oplus W_j$ and thus $V_{j+2} = V_{j+1} \oplus W_{j+1} = V_j \oplus W_j \oplus W_{j+1}$. By induction, $V_j = V_0 \oplus \bigoplus_{k=0}^{j} W_k$, and by the Density property, the closure of $V_0 \oplus \bigoplus_{k>=0}^{\infty} W_k$ is $L^2(\mathbf{R})$. We have an orthonormal basis for V_0; assume we have one for W_0 as well. (This basis will later turn out to be generated by the wavelets and their translates. In section 3.1, we will give a recipe for getting wavelets from the scaling function.) But because the W_j spaces have the same scaling property that the V_j spaces do—i.e., that $f \in W_j \iff f(mx) \in W_{j+1}$—having a basis for W_0 makes it easy to get a basis for W_j: just compose all the functions in the basis with $m^j x$ and multiply the function by $m^{(j/2)}$ to preserve orthonormality. An orthonormal basis for V_0 together with orthonormal bases for the W_j means we have an orthonormal basis for $L^2(\mathbf{R})$.

Second, the scaling function satisfies the *dilation equation*:

$$\phi(x) = \sum_{\gamma \in \mathbf{Z}} c_\gamma \phi(mx - \gamma), \quad c_\gamma \in \mathbf{R}.$$

Proof: The combination of the scaling and orthonormality conditions guarantee that $\{\sqrt{m}\phi(mx - \gamma)\}$ forms an orthonormal basis for V_1. But $\phi \in V_0 \subset V_1 \implies \phi \in V_1$. $\square$

Because it (mostly) reduces the behavior of ϕ to the behavior of a concrete set of scalar values (the c_γ), the dilation equation is a powerful tool for the construction and analysis of multiresolution approximations. For the rest of this thesis, we will be concerned with these questions: for which c_γ's is there a solution to the dilation equation? For which c_γ's is ϕ orthogonal to its $\mathbf{Z}$ translates? What conditions can be put on the c_γ's to make ϕ smooth (where smoothness can be anything from being continuous, to having many derivatives, to having many "vanishing moments")? And finally, given that $W_0 \subset V_1$ implies that any function ψ in W_0 also satisfies a dilation equation

$$\psi(x) = \sum_{\gamma \in \mathbf{Z}} d_\gamma \phi(mx - \gamma),$$

how can enough sets of d_γ's be constructed to give a basis of W_0?

Let us now move to the question of multiresolution approximations of $L^2(\mathbf{R}^n)$. Our definition of a mulitresolution analysis over $L^2(\mathbf{R})$ really depended only in two ways on the fact that we were working with functions over $\mathbf{R}$. First, we translated them by integers, and second, we composed them with multiplication by an integer $m > 1$. Both of these are easily generalized: first, to translation by an arbitrary n-dimensional lattice Γ, and second, to multiplication by an expanding (all eigenvalues have magnitude greater than 1) matrix M such that $M\Gamma \subset M$. Such a matrix M is called an *acceptable dilation*.

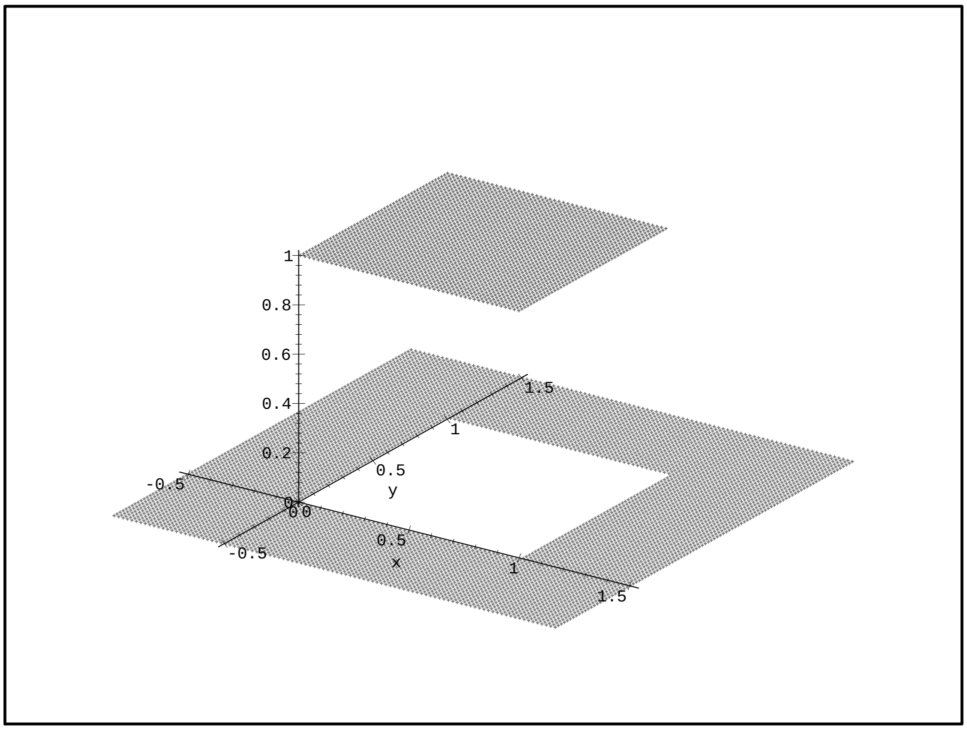
That said, a rank M multiresolution analysis of $L^2(\mathbf{R}^n)$ is the same as a multiresolution analysis of $L^2(\mathbf{R})$, except with the two modified axioms of

- **Scaling:** $f(\vec{x}) \in V_j \iff f(M\vec{x}) \in V_{j+1}$, and
- **Orthonormality:** There exists a scaling function $\phi \in V_0$ such that $\{\phi(\vec{x} - \vec{\gamma})\}_{\vec{\gamma} \in \Gamma}$, the Γ-translates of ϕ, form an orthonormal basis for V_0.

Multiresolution approximations of $L^2(\mathbf{R}^n)$ have most of the same properties as multiresolution approximations of $L^2(\mathbf{R})$. They yield a basis for $L^2(\mathbf{R}^n)$, and the corresponding dilation equation is:

$$\phi(\vec{x}) = \sum_{\vec{\gamma} \in \Gamma} c_{\vec{\gamma}} \phi(M\vec{x} - \vec{\gamma}).$$

Figures 1-1 through 1-5 show the graphs of the scaling functions associated with selected multiresolution approximations of $L^2(\mathbf{R})$ and $L^2(\mathbf{R}^2)$. Note that the mul-

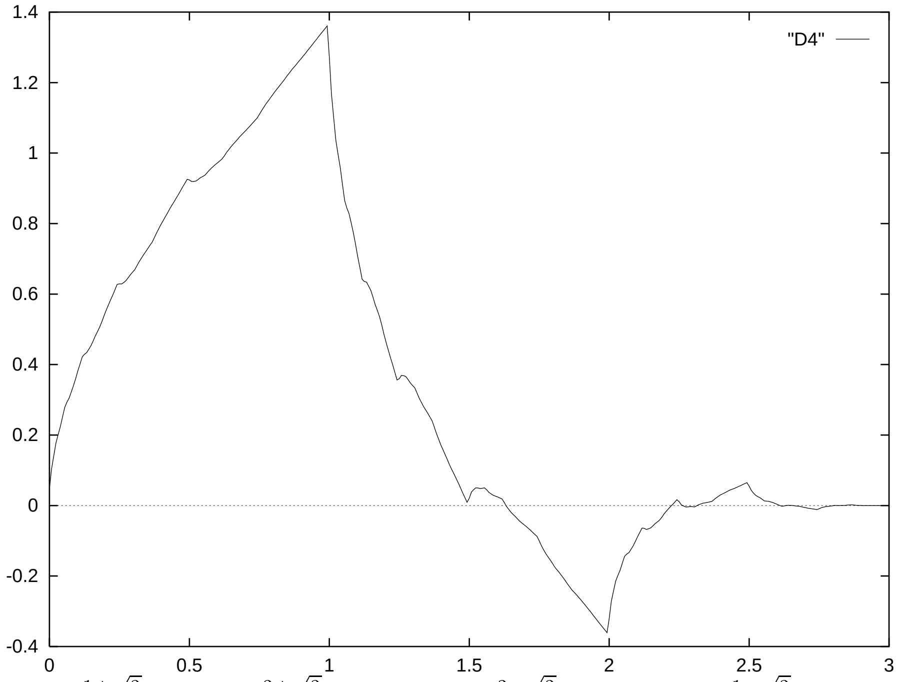


$$\phi(\vec{x}) = \phi(2\vec{x}) + \phi(2\vec{x} + \binom{0}{1}) + \phi(2\vec{x} + \binom{1}{0}) + \phi(2\vec{x} + \binom{1}{1})$$

Figure 1-1: The Haar Rank $2I$ Scaling Function $\phi = \chi_{[0,1]^2}$

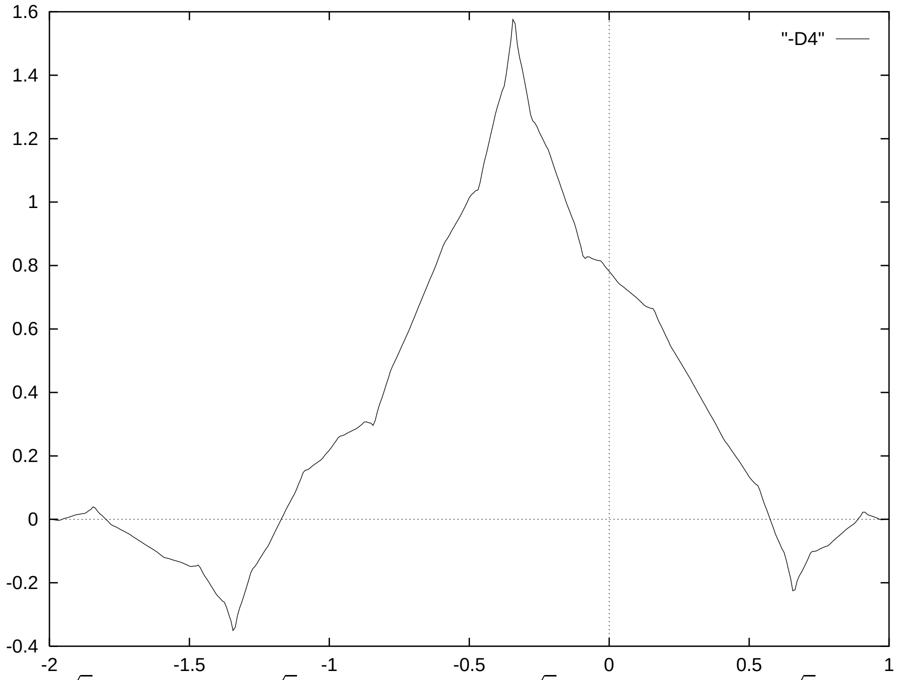
tidimensional multiresolution analysis framework is more general than the normal multiresolution analysis framework, even in the one dimensional case because it allows lattices besides $\mathbf{Z}$ and it allows negative integers $m < -1$ as acceptable dilations. Figures 1-3 and 1-4 show "rank -2" scaling functions; as discussed on pages 256 and 257 of [Daubechies 92a], these are always more symmetric than the associated rank 2 scaling functions.

Figures 1-1 and 1-5 show scaling functions that are the characteristic function of *self-affine tiles*. As discussed in section 3.2, self-affine tiles are tiles of $\mathbf{R}^n$ that satisfy an equation of the form $M\mathcal{T} = \cup_{\vec{k} \in K} \mathcal{T} + \vec{k}$ for some set K of *digits*. Such tiles can always be made into (albeit not very smooth) scaling functions.



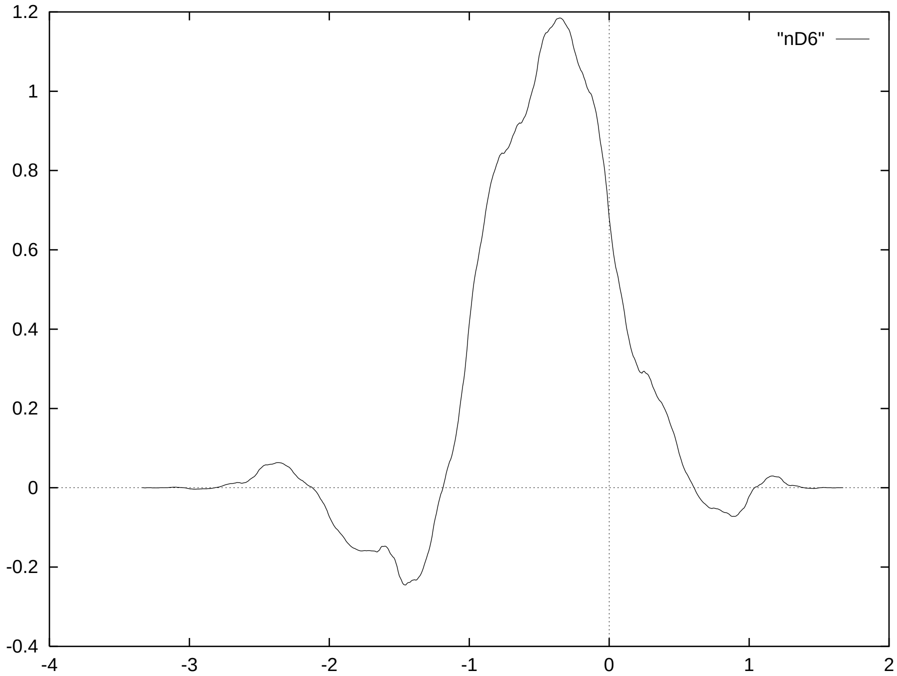
$$\phi(x) = \frac{1+\sqrt{3}}{4}\phi(2x) + \frac{3+\sqrt{3}}{4}\phi(2x-1) + \frac{3-\sqrt{3}}{4}\phi(2x-2) + \frac{1-\sqrt{3}}{4}\phi(2x-3)$$

Figure 1-2: The Daubechies D4 Scaling Function



$$\phi(x) = \frac{1+\sqrt{3}}{4}\phi(-2x) + \frac{3+\sqrt{3}}{4}\phi(-2x-1) + \frac{3-\sqrt{3}}{4}\phi(-2x-2) + \frac{1-\sqrt{3}}{4}\phi(-2x-3)$$

Figure 1-3: "-D4": The Rank -2 Scaling Function with D4 Coefficients

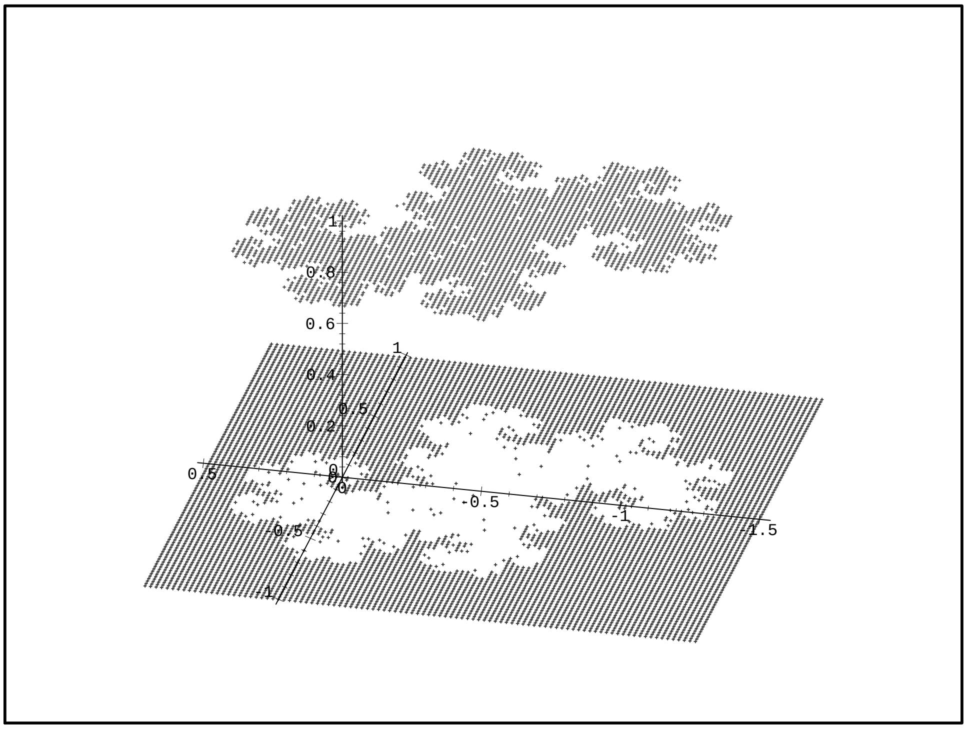


$$\phi(x) = \frac{1}{16}((7-5\sqrt{5+2\sqrt{10}}+3\sqrt{10})\phi(-2x)+(17-9\sqrt{5+2\sqrt{10}}+5\sqrt{10})\phi(-2x-1)$$

$$+(10+2\sqrt{5+2\sqrt{10}}-2\sqrt{10})\phi(-2x-2)+(-2+10\sqrt{5+2\sqrt{10}}-6\sqrt{10})\phi(-2x-3)$$

$$+(-1+3\sqrt{5+2\sqrt{10}}-\sqrt{10})\phi(-2x-4)+(1-\sqrt{5+2\sqrt{10}}+\sqrt{10})\phi(-2x-5))$$

Figure 1-4: "-D6": The Rank -2 Scaling Function with D6 Coefficients



$$\phi(\vec{x}) = \phi(\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\vec{x}) + \phi(\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\vec{x} + \begin{pmatrix} 1 \\ 0 \end{pmatrix})$$

Figure 1-5: Knuth Dragon Scaling Function

1.2 Overview

The remainder of this thesis is divided into two parts. The first part (Chapter 2) describes how to verify whether a given set of $c_{\vec{\gamma}}$ values leads to a multiresolution analysis. The requirements that ϕ be a function in $L^2(\mathbf{R}^n)$, be orthonormal to its Γ translates, and generate a complete multiresolution analysis of $L^2(\mathbf{R}^n)$ are all translated into conditions on the $c_{\vec{\gamma}}$, with the majority of attention going to the case when ϕ is compactly supported. Chapter 2 also discusses wavelets, and reduces their properties to that of the polyphase matrix $A(\vec{x})$. Finally, various measures of the smoothness of a multiresolution analysis (such as its ability to reproduce polynomials or differentiability of its scaling function) are considered, and rules given for their determination.

The second part of this thesis (Chapter 3) describes various methods of constructing multiresolution approximations. By factorizing the polyphase matrix, it is shown how to construct, for every acceptable dilation matrix M, a compactly supported scaling function ϕ with compactly supported wavelets ψ_i such that $\int \psi_i(\vec{x}) x_1 \, d\vec{x} = 0$. (This is the "degree $\vec{e}_1$ vanishing moment" condition, one of the measures of smoothness discussed in Chapter 2.) In addition, the problem of how to construct wavelets given a scaling function is considered, with special consideration paid to the problem of retaining compact support.

1.3 Notation

Symbol	Definition
$\overline{a}$	complex conjugate of a
$\|a\|$	magnitude of a, $\|a\| = \sqrt{a\overline{a}}$
$\mathrm{Re}\, z$	real part of z
M^*	$\overline{M^t}$, conjugate transpose of matrix M
$\vec{x}$	a vector, usually in $\mathbf{R}^n$
Γ	an n dimensional lattice in $\mathbf{R}^n$
$\vec{\gamma}, \vec{\alpha}, \vec{\beta}, \ldots$	vectors in Γ
V_i	subspace of a multiresolution analysis,
	$= \{\sum_{\vec{\gamma}\in\Gamma} a_{\vec{\gamma}}\phi(M^i\vec{x} - \vec{\gamma}) \mid a_{\vec{\gamma}} \in \mathbf{C}, \sum_{\vec{\gamma}\in\Gamma} \|a_{\vec{\gamma}}\|^2 < \infty\}$
$P_i f$	the projection of f onto V_i
W_i	the orthogonal complement of V_j in V_{j+1}
$L^2(\mathbf{R}^n)$	the set of measurable functions $f : \mathbf{R}^n \to \mathbf{C}$ such that $\int_{\mathbf{R}^n} \|f(\vec{x})\|^2\, d\vec{x} < \infty$
$< f, g >$	the inner product in $L^2(\mathbf{R}^n)$, $< f, g > = \int_{\mathbf{R}^n} f(\vec{x})\overline{g(\vec{x})}\, d\vec{x}$
$f = g$	$L^2(\mathbf{R}^n)$ equality: $\int_{\mathbf{R}^n} \|f(\vec{x}) - g(\vec{x})\|^2\, d\vec{x} = 0$
ϕ	a scaling function
ψ_i	a wavelet
M	a dilation matrix
m	$\|\det M\|$
$c_{\vec{\gamma}}$	coefficients of the scaling function in the dilation equation,
	$= < \phi(\vec{x}), \phi(M\vec{x} - \vec{\gamma}) >$
$d_{i,\vec{\gamma}}$	coefficients of the i-th wavelet, $d_{i,\vec{\gamma}} = < \psi_i(\vec{x}), \phi(M\vec{x} - \vec{\gamma}) >$
K	the set of $\vec{\gamma}$ for which $c_{\vec{\gamma}} \neq 0$
$\delta_{\vec{i},\vec{j}}$	Dirac delta function, $= 1$ if $\vec{i} = \vec{j}$ and 0 otherwise
$\vec{v}(i)$	i-th component of the vector $\vec{v}$
$e_{\vec{i}}(\vec{x})$	$= \delta_{\vec{i},\vec{x}}$
$\vec{x}^{\vec{i}}$	$x_1^{i_1} \ldots x_n^{i_n}$
$\|\vec{i}\|$	$i_1 + i_2 + \ldots + i_n$

Symbol	Definition
$\partial^{\vec{i}}$	$(\frac{\partial}{\partial x_1})^{i_1} \dots (\frac{\partial}{\partial x_n})^{i_n}$
$\vec{a} \leq \vec{b}$	$a_i \leq b_i$ for all i
f_0	the starting function in the Cascade algorithm
f_n	$= \sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}} f_{n-1}(M\vec{x} - \vec{\gamma})$, n-th iterate of f_0 in the Cascade algorithm
$\vec{a}_{f,g}(\vec{\gamma})$	$= < f, g(\vec{x} - \vec{\gamma}) >$
T	matrix with $(\vec{\gamma}, \vec{\delta})$ entry $\frac{1}{\lvert\det M\rvert} \sum_{\vec{\beta}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}}$
$\vec{y}_0, \dots, \vec{y}_{m-1}$	a fixed ordering of the elements of $\Gamma/M\Gamma$
$A_{\vec{k}}$	matrix with (i, j)-th entry $d_{i,M\vec{k}+\vec{y}_j}$
$A(\vec{x})$	$= \frac{1}{\sqrt{\lvert\det M\rvert}} \sum_{\vec{k}\in\Gamma} A_{\vec{k}} \vec{x}^{\vec{k}}$, the polyphase matrix
$\vec{k'} \equiv \vec{k}(M)$	$\vec{k'} = \vec{k} + M\vec{\gamma}$ for some $\gamma \in \Gamma$
$\lvert K\rvert$	cardinality of the set K
χ_Q	the characteristic function of the set Q
$\Box$	end of proof

Chapter 2

Verifying a Multiresolution Analysis

This chapter concerns the following questions: Given a set of $c_{\vec{\gamma}}$'s, do they define a scaling function? Does this scaling function generate a multiresolution analysis? How smooth is it?

We start by determining under what circumstances the $c_{\vec{\gamma}}$'s define a non-zero $L^2(\mathbf{R}^n)$ solution to the dilation equation.

2.1 The Cascade Algorithm

In this section, we discuss a simple procedure which under a wide variety of situations allows us to explicitly construct a solution to a dilation equation. This procedure, the *Cascade Algorithm*, starts by choosing a $f_0 \in L^2(\mathbf{R}^n)$, then iterates

$$f_{n+1}(\vec{x}) = \sum_{\vec{\gamma} \in \Gamma} c_{\vec{\gamma}} f_n(M\vec{x} - \vec{\gamma}).$$

If it is successful, then the solution $f = \lim_{n\to\infty} f_n$ will exist and solve the dilation equation $f(\vec{x}) = \sum c_{\vec{\gamma}} f(M\vec{x} - \vec{\gamma})$.

Since we are working in a complete inner product space, the sequence f_n converges if and only if it is a Cauchy sequence. We will thus be very interested in $< f_n, f_m >$

and similar expressions. Luckily, there is a simple recurrence relationship for these. If we are given $f, g \in L^2(\mathbf{R}^n)$ then we may define the vector $\vec{a}_{f,g}(\vec{\gamma}) = < f, g(\vec{x} - \vec{\gamma}) >$. It follows that

$$\begin{aligned}
\vec{a}_{f_{n+1},g_{n+1}}(\vec{\gamma}) &= < f_{n+1}, g_{n+1}(\vec{x} - \vec{\gamma}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} \sum_{\vec{\beta}\in\Gamma} c_{\vec{\alpha}} \overline{c_{\vec{\beta}}} < f_n(M\vec{x} - \vec{\alpha}), g_n(M\vec{x} - M\vec{\gamma} - \vec{\beta}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} \sum_{\vec{\beta}\in\Gamma} c_{\vec{\alpha}} \overline{c_{\vec{\beta}}} \int f_n(M\vec{x} - \vec{\alpha}) \overline{g_n(M\vec{x} - M\vec{\gamma} - \vec{\beta})}\, d\vec{x} \\
&= \sum_{\vec{\alpha}\in\Gamma} \sum_{\vec{\beta}\in\Gamma} c_{\vec{\alpha}} \overline{c_{\vec{\beta}}} \frac{1}{|\det M|} < f_n(\vec{u} - \vec{\alpha}), g_n(\vec{u} - M\vec{\gamma} - \vec{\beta}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} \sum_{\vec{\beta}\in\Gamma} c_{\vec{\alpha}} \overline{c_{\vec{\beta}}} \frac{1}{|\det M|} < f_n(\vec{u}'), g_n(\vec{u}' + \vec{\alpha} - M\vec{\gamma} - \vec{\beta}) > \\
&= \sum_{\vec{\delta}\in\Gamma} \sum_{\vec{\beta}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}} \frac{1}{|\det M|} < f_n, g_n(\vec{x} - \vec{\delta}) > \\
&= \sum_{\vec{\delta}\in\Gamma} \frac{1}{|\det M|} \sum_{\vec{\beta}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}} \vec{a}_{f_n,g_n}(\vec{\delta}).
\end{aligned}$$

Thus, if we define the matrix T to have entry

$$T_{\vec{\gamma},\vec{\delta}} = \frac{1}{|\det M|} \sum_{\vec{\beta}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}},$$

then $\vec{a}_{f_{n+1},g_{n+1}} = T\vec{a}_{f_n,g_n} = T^{n+1}\vec{a}_{f,g}$.

Notice that even if only a finite number of the $c_{\vec{\gamma}}$'s are nonzero, T is still technically an infinite dimensional matrix. The following lemma justifies using truncated, finite dimensional versions of T in this situation.

Lemma 2.1.1 *Let $K = \{\vec{\gamma} : c_{\vec{\gamma}} \neq 0\}$ be finite, and assume that $s = \max_{\vec{x}\cdot\vec{x}=1} \frac{||M^{-1}\vec{x}||}{||\vec{x}||}$ is less than one. [This will be true for some M^{-j}.] Let $B(\vec{p}, r) = \{\vec{x} : |\vec{x} - \vec{p}| \leq r\}$, the n-dimensional ball of radius r centered at $\vec{p}$, contain all the points $\{(I - M)^{-1}\vec{\gamma}, \vec{\gamma} \in K\}$, and let $r' = \frac{r+sr}{1-s}$. Then*

- *If $\lim f_n$ converges to a compactly supported function f, then the support of f is contained in the ball $B(\vec{p}, r')$,*

- *If the support of f_n is contained in $B(\vec{p}, r')$, then so is the support of f_{n+1}, and*
- *If the support of f_n is contained in $B(\vec{p}, \tilde{r})$, $\tilde{r} > r'$, then the support of f_{n+1} is contained in $B(\vec{p}, r' + s(\tilde{r} - r'))$.*

PROOF: The support of f_{n+1} is $W(\text{support} f_n)$, where

$$W(X) = \bigcup_{\vec{\gamma} \in K} M^{-1}(X + \vec{\gamma})$$

is an operator on the Hausdorff metric space $H(\mathbf{R}^n)$ of all compact subspaces of $\mathbf{R}^n$. W is a contractive operator, and its contractivity factor is $s < 1$. Because $H(\mathbf{R}^n)$ is complete ([Barnsley 88]), W has an unique fixed point A_W, and this compact set will be the support of f if f exists.

To show that $A_W \subset B(\vec{p}, r')$, it suffices to show that $W(B(\vec{p}, r')) \subset B(\vec{p}, r')$. In particular, it suffices to show that $M^{-1}(B(\vec{p}, r') + \vec{\gamma}) \subset B(\vec{p}, r')$ for each $\vec{\gamma} \in K$. If we let $\vec{y}_{\vec{\gamma}}$ be the fixed point of the function $\vec{x} \to M^{-1}(\vec{x} + \vec{\gamma})$ and $\vec{x}$ be an arbitary point in $M^{-1}(B(\vec{p}, r') + \vec{\gamma})$, then

$$\begin{aligned}
d(\vec{p}, \vec{x}) &\leq d(\vec{p}, \vec{y}_{\vec{\gamma}}) + d(\vec{y}_{\vec{\gamma}}, \vec{x}) \\
&\leq r + s(r + r') \\
&= r + sr + \frac{s(r + sr)}{1 - s} \\
&= \frac{r + sr}{1 - s}
\end{aligned}$$

which proves that $\vec{x} \in B(\vec{p}, r')$. [$d(\vec{y}_{\vec{\gamma}}, \vec{x}) < s(r + r')$ because $d(\vec{y}_{\vec{\gamma}}, \vec{z}) < r + r'$ for $\vec{z} \in B(\vec{p}, r')$.]

Finally, the statements on the support of f_{n+1} follow from the contractivity of W.

□

For the rest of this section, we will assume that f_0 is compactly supported and that only a finite number of the $c_{\vec{\gamma}}$'s are non-zero.

Theorem 2.1.2 (The Cascade Algorithm Theorem) *The Cascade algorithm con-*

verges to a compactly supported nonzero function in $L^2(\mathbf{R}^n)$ if and only if the following conditions hold.

- *T has 1 as an eigenvalue.*
- *For every eigenvalue $\lambda \neq 1$ of T with $|\lambda| \geq 1$,*

$$\sum_{i=1}^{t} \sum_{j=1}^{s_i-k} \vec{v}_{i,j}(\vec{0})\vec{w}_{i,k+j} \cdot \vec{a}_{f_0,f_r} = 0$$

  for all $r \geq 0$ and $k = 0 \ldots \max(s_i) - 1$, where

  1. *there are t Jordan blocks with eigenvalue λ in the Jordan form factorization of $T = SJS^{-1}$,*
  2. *the i-th Jordan block is s_i by s_i,*
  3. *$\vec{v}_{i,j}$ is the column of S corresponding to j-th column of the i-th Jordan block with eigenvalue λ ($\vec{v}_{i,j} = 0$ for $j > s_i$), and*
  4. *$\vec{w}_{i,j}$ is the row of S^{-1} corresponding to j-th row of the i-th Jordan block with eigenvalue λ ($\vec{w}_{i,j} = 0$ for $j > s_i$).*

- *For every eigenvalue $\lambda = 1$ of T,*

$$\sum_{i=1}^{t} \sum_{j=1}^{s_i-k} \vec{v}_{i,j}(\vec{0})\vec{w}_{i,k+j} \cdot \vec{a}_{f_0,f_r} = 0$$

  for all $r \geq 0$ and $k = 1 \ldots \max(s_i) - 1$ and

$$\mathrm{Re}(\sum_{i=1}^{t} \sum_{j=1}^{s_i} \vec{v}_{i,j}(\vec{0})\vec{w}_{i,j} \cdot \vec{a}_{f_0,f_r}) \to \sum_{i=1}^{t} \sum_{j=1}^{s_i} \vec{v}_{i,j}(\vec{0})\vec{w}_{i,j} \cdot \vec{a}_{f_0,f_0} \neq 0$$

  as $r \to \infty$, with the same conventions as above.

PROOF: As previously noted, $L^2(\mathbf{R}^n)$ is complete, so the Cascade algorithm converges if and only if the sequence f_n is Cauchy. We thus want to prove that the uniform convergence of

$$|f_n - f_m| = \sqrt{< f_n, f_n > - < f_n, f_m > - < f_m, f_n > + < f_m, f_m >} \to 0$$

as $n, m \to \infty$ is equivalent to our stated conditions.

Let's consider $< f_n, f_n >$ first. We know that it is equal to $\vec{a}_{f_n,f_n}(\vec{0}) = (T^n \vec{a}_{f_0,f_0})(\vec{0})$. If we decompose T into its Jordan factorization $T = SJS^{-1}$, then $< f_n, f_n > = (SJ^nS^{-1}\vec{a}_{f_0,f_0})(\vec{0})$. If we think of this as a scalar sum, then the contribution of each Jordan block to this is

$$([\vec{v}_1 \ldots \vec{v}_s] \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ 0 & & & \lambda \end{bmatrix}^n \begin{bmatrix} \vec{w}_1 \\ \vdots \\ \vec{w}_s \end{bmatrix} \vec{a}_{f_0,f_0})(\vec{0})$$

$$= \left[\vec{v}_1(\vec{0}) \ldots \vec{v}_s(\vec{0})\right] \begin{bmatrix} \lambda^n & n\lambda^{n-1} & \ldots & \binom{n}{s-1}\lambda^{n-s+1} \\ & \lambda^n & \ldots & \vdots \\ & & \ddots & \vdots \\ 0 & & & \lambda^n \end{bmatrix} \begin{bmatrix} \vec{w}_1 \cdot \vec{a}_{f_0,f_0} \\ \vdots \\ \vec{w}_s \cdot \vec{a}_{f_0,f_0} \end{bmatrix}.$$

If $|\lambda| < 1$, then the middle matrix goes to the zero matrix and the contribution is 0. Otherwise, the contribution is

$$\sum_{i=1}^{s} \vec{v}_i(\vec{0}) \sum_{j=i}^{s} \binom{n}{j-i} \lambda^{n-j+i} \vec{w}_j \cdot \vec{a}_{f_0,f_0}.$$

Summing over all Jordan blocks with the same eigenvalue λ gives

$$\sum_{k=1}^{t} \sum_{i=1}^{s_k} \vec{v}_{k,i}(\vec{0}) \sum_{j=i}^{s_k} \binom{n}{j-i} \lambda^{n-j+i} \vec{w}_{k,j} \cdot \vec{a}_{f_0,f_0}$$

$$= \sum_{k=0}^{\max(s_i)-1} \binom{n}{k} \lambda^{n-k} \sum_{i=1}^{t} \sum_{j=1}^{s_i-k} \vec{v}_{i,j}(\vec{0}) \vec{w}_{i,k+j} \cdot \vec{a}_{f_0,f_0}.$$

As $n \to \infty$, the above sum is dominated by the $k = \max(s_i) - 1$ term. For it to converge, either $\binom{n}{k}\lambda^{n-k}$ must be 1 for all k (which is the $\lambda = 1$ and $k = 0$ case), or

$$\sum_{i=1}^{t} \sum_{j=1}^{s_i-k} \vec{v}_{i,j}(\vec{0}) \vec{w}_{i,k+j} \cdot \vec{a}_{f_0,f_0} = 0.$$

But then the sum is dominated by the next lower k term, and we can repeat the argument all the way down to $k = 0$.

Thus, if it exists,

$$\lim_{n\to\infty} < f_n, f_n > = \sum_{i=1}^{t} \sum_{j=1}^{s_i} \vec{v}_{i,j}(\vec{0}) \vec{w}_{i,j} \cdot \vec{a}_{f_0,f_0}$$

where $\vec{v}_{i,j}$ and $\vec{w}_{i,j}$ are associated with Jordan blocks having eigenvalue $\lambda = 1$.

Now let's look at $< f_n, f_m >$. Without loss of generality, we may assume $m > n$ and $m = n + r$. As before, $< f_n, f_m > = (T^n \vec{a}_{f_0,f_{m-n}})(\vec{0}) = (SJ^nS^{-1}\vec{a}_{f_0,f_r})(\vec{0})$. All of the above analysis of the Jordan block contribution to this goes through with $\vec{a}_{f_0,f_0}$ replaced with $\vec{a}_{f_0,f_r}$.

Thus, if it exists,

$$\lim_{n\to\infty} < f_n, f_{n+r} > = \sum_{i=1}^{t} \sum_{j=1}^{s_i} \vec{v}_{i,j}(\vec{0}) \vec{w}_{i,j} \cdot \vec{a}_{f_0,f_r}$$

where $\vec{v}_{i,j}$ and $\vec{w}_{i,j}$ are associated with Jordan blocks having eigenvalue $\lambda = 1$.

Convergence is uniform because each component of the vector $\vec{a}_{f_0,f_r}$ is bounded:

$$\begin{aligned} |\vec{a}_{f_0,f_r}(\vec{\gamma})| &= |< f_0, f_r(\vec{x} - \vec{\gamma}) >| \\ &\le |f_0||f_r(\vec{x} - \vec{\gamma})| \\ &= |f_0||f_r|. \end{aligned}$$

We already proved that $|f_r| = \sqrt{< f_r, f_r >}$ converges as $r \to \infty$; this implies that $|f_r|$ and thus $|\vec{a}_{f_0,f_{m-n}}|$ and $|< f_n, f_m >|$ are bounded.

Finally, for

$$\sqrt{< f_n, f_n > - < f_n, f_m > - < f_m, f_n > + < f_m, f_m >} \to 0$$

as $m, n \to \infty$, we need

$$\mathrm{Re}(< f_n, f_m >) \to < f_n, f_n >.$$

But the final condition in our theorem states just that.

So, just to wrap it all up, we've shown that $\mathrm{Re}(< f_n, f_m >) \rightarrow < f_n, f_n >$, which means that

$$< f_n, f_n > - < f_n, f_m > - < f_m, f_n > + < f_m, f_m > \rightarrow 0$$

as $m, n \rightarrow 0$. The convergence is uniform, so $|f_n - f_m| \rightarrow 0$ and $\lim_{n\rightarrow\infty} f_n$ is in $L^2(\mathbf{R}^n)$ as desired. $\square$

The big problem with this theorem is that while we can easily compute T's Jordan factorization and eigenvalues from the $c_{\vec{\gamma}}$'s, we have no a priori knowledge about $\vec{a}_{f_0,f_r}$. There are two general approaches to getting around this limitation: either assume that $\vec{w}_{i,k+j} \cdot f_0(\vec{x} - \cdot)$ satisfies a rank M dilation equation, or assume that f_0 all by itself does. The following corollaries illustrate both approaches.

Corollary 2.1.3 *Assume that* $\sum_{\vec{\gamma}\in\Gamma} f_0(\vec{x} - \vec{\gamma})$ *is a constant. [The fundamental domain of* Γ *is a popular choice.] Then the Cascade algorithm converges to a compactly supported non-zero function* f *in* $L^2(\mathbf{R}^n)$ *if*

- T *has* 1 *as a simple eigenvalue,*
- *All other eigenvalues of* T *are less than one in absolute value,*
- *The eigenvector of* T *associated with the eigenvalue* 1 *does not have* 0 *as its* $\vec{0}$ *entry, and*
- $\sum_{\vec{k'}\equiv\vec{k}(M)} c_{\vec{k'}} = 1$ *(Condition* $\vec{1}$*) where* $\vec{k'} \equiv \vec{k}(M)$ *if* $\vec{k'} = \vec{k} + M\vec{\gamma}$ *for some* $\gamma \in \Gamma$.

Proof: The idea is that Condition $\vec{1}$ guarantees that $\vec{1}T = \vec{1}$, and the assumption on f_0 guarantees that $\vec{1} \cdot f_0(\vec{x} - \cdot)$ is a constant C. But the constant function C satisfies a very simple rank M dilation equation: $C(\vec{x}) = C(M\vec{x})$. So, $\vec{1} \cdot \vec{a}_{f_0,f_r} = \vec{1} \cdot \vec{a}_{f_0,f_0}$ for all r.

So first let's prove that Condition $\vec{1}$ guarantees that $\vec{1}T = \vec{1}$:

$$\begin{aligned}
(\vec{1}T)(\vec{\delta}) &= \sum_{\vec{\gamma}\in\Gamma} T_{\vec{\gamma},\vec{\delta}} \\
&= \sum_{\vec{\gamma}\in\Gamma} \frac{1}{|\det M|} \sum_{\vec{\beta}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}} \\
&= \frac{1}{|\det M|} \sum_{\vec{\beta}\in\Gamma} \overline{c_{\vec{\beta}}} \sum_{\vec{\gamma}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \\
&= \frac{1}{|\det M|} \sum_{\vec{\beta}\in\Gamma} \overline{c_{\vec{\beta}}} \sum_{\vec{\gamma}'\equiv\vec{\beta}-\vec{\delta}(M)} c_{\vec{\gamma}'} \\
&= 1.
\end{aligned}$$

Next, let's prove that $\vec{1}\cdot\vec{a}_{f_0,f_r} = \vec{1}\cdot\vec{a}_{f_0,f_0} = <1, f_0>$:

$$\begin{aligned}
\vec{1}\cdot\vec{a}_{f_0,f_r} &= \sum_{\vec{k}\in\Gamma} < f_0, f_r(\vec{x}-\vec{k}) > \\
&= < \sum_{\vec{k}\in\Gamma} f_0(\vec{x}+\vec{k}), f_r > \\
&= < C, f_r > \\
&= \sum_{\vec{k}\in\Gamma} c_{\vec{k}} < C, f_{r-1}(M\vec{x}-\vec{k}) > \\
&= \sum_{\vec{k}\in\Gamma} c_{\vec{k}} \frac{1}{|\det M|} < C, f_{r-1} > \\
&= < C, f_{r-1} >
\end{aligned}$$

and the result follows by induction.

Finally, notice that when T has 1 as a simple eigenvalue and all other eigenvalues are smaller in absolute value, the conditions on our Cascade Algorithm theorem for convergence simplify to just $\vec{v}(\vec{0})\vec{w}\cdot\vec{a}_{f_0,f_r} \to \vec{v}(\vec{0})\vec{w}\cdot\vec{a}_{f_0,f_0} \neq 0$ where $\vec{v}$ is the eigenvector and $\vec{w}$ the right eigenvector associated with the eigenvalue 1. But we have shown that $\vec{w} = \vec{1}$, $\vec{1}\cdot\vec{a}_{f_0,f_r} = \vec{1}\cdot\vec{a}_{f_0,f_0}$, and we assumed that $\vec{v}(\vec{0}) \neq 0$. We are done. $\square$

Lemma 2.1.4 *If $\sum_{\vec{k}'\equiv\vec{k}(M)} c_{\vec{k}'} = 1$ (Condition $\vec{1}$), then the Cascade algorithm converges only if $\sum_{\vec{\gamma}\in\Gamma} f_0(\vec{x}-\vec{\gamma})$ is a constant.*

PROOF: Let $P(f) = \sum_{\vec{\gamma}\in\Gamma} f(\vec{x} - \vec{\gamma})$ be the periodization of f. Condition $\vec{1}$ implies that

$$\begin{aligned}
P(f_{n+1})(\vec{x}) &= \sum_{\vec{\gamma}\in\Gamma} f_{n+1}(\vec{x} - \vec{\gamma}) \\
&= \sum_{\vec{\gamma}\in\Gamma} \sum_{\vec{\alpha}\in\Gamma} c_{\vec{\alpha}} f_n(M\vec{x} - M\vec{\gamma} - \vec{\alpha}) \\
&= \sum_{\vec{\gamma}\in\Gamma} \sum_{\vec{\beta}\in\Gamma} \sum_{\vec{\delta}\in\Gamma/M\Gamma} c_{M\vec{\beta}+\vec{\delta}} f_n(M\vec{x} - M(\vec{\gamma} + \vec{\beta}) - \vec{\delta}) \\
&= \sum_{\vec{\gamma}'\in\Gamma} \sum_{\vec{\delta}\in\Gamma/M\Gamma} f_n(M\vec{x} - M\vec{\gamma}' - \vec{\delta}) \\
&= \sum_{\vec{\gamma}''\in\Gamma} f_n(M\vec{x} - \vec{\gamma}'') \\
&= P(f_n)(M\vec{x}).
\end{aligned}$$

Now, if the Cascade algorithm converges, $\lim_{n\to\infty} P(f_n)$ will also converge. But given that $P(f_{n+1}) = P(f_n)(M\vec{x})$, this is impossible unless $P(f_n) = P(f_0)$ is constant. □

The following proposition proves that Condition $\vec{1}$ can be expected to hold under a large variety of situations.

Proposition 2.1.5 *If $\sum_{\vec{\gamma}\in\Gamma} \phi(\vec{x} - \vec{\gamma})$ is not identically zero, and if there exists a self-M-affine Γ tile $\mathcal{T}$, then condition $\vec{1}$ holds.*

PROOF: Let D be $\mathcal{T}$'s digit set, and consider the function $\tilde{\phi} = \chi_{\mathcal{T}} \sum_{\vec{\gamma}\in\Gamma} \phi(\vec{x} - \vec{\gamma})$. It satisfies the dilation equation

$$\tilde{\phi} = \sum_{\vec{d}\in D} \Big(\sum_{\vec{\gamma}\equiv\vec{d}(M)} c_{\vec{\gamma}} \Big) \tilde{\phi}(M\vec{x} - \vec{d})$$

because

$$\begin{aligned}
\tilde{\phi} &= \chi_{\mathcal{T}} \sum_{\vec{\gamma}\in\Gamma} \phi(\vec{x}-\vec{\gamma}) \\
&= \chi_{\mathcal{T}} \sum_{\vec{\gamma}\in\Gamma} \sum_{\vec{\gamma}'\in\Gamma} c_{\vec{\gamma}'} \phi(M\vec{x}-M\vec{\gamma}-\vec{\gamma}') \\
&= \chi_{\mathcal{T}} \sum_{\vec{\gamma}\in\Gamma} \sum_{\vec{\alpha}\in\Gamma} \sum_{\vec{d}\in D} c_{M\vec{\alpha}+\vec{d}} \phi(M\vec{x}-M\vec{\gamma}-M\vec{\alpha}-\vec{d}) \\
&= \sum_{\vec{d}\in D} \sum_{\vec{\alpha}\in\Gamma} c_{M\vec{\alpha}+\vec{d}} \chi_{\mathcal{T}} \phi(M\vec{x}-M\vec{\gamma}-M\vec{\alpha}-\vec{d}) \\
&= \sum_{\vec{d}\in D} \sum_{\vec{\alpha}\in\Gamma} c_{M\vec{\alpha}+\vec{d}} \tilde{\phi}(M\vec{x}-\vec{d}).
\end{aligned}$$

Let $\tilde{c}_{\vec{d}} = (\sum_{\vec{\gamma}\equiv\vec{d}(M)} c_{\vec{\gamma}})$. We claim that

$$\sum_{\vec{d}\in D} |\tilde{c}_{\vec{d}}|^2 = |\det M|$$

and

$$\sum_{\vec{d}\in D} \tilde{c}_{\vec{d}} = |\det M|$$

must hold. The first equation follows from the Cascade Algorithm Theorem's mandate that $\tilde{\phi}$'s T matrix satisfies $T\vec{e}_{\vec{0}} = \vec{e}_{\vec{0}}$. The second equation — Condition 1 — holds if $\int \tilde{\phi}(\vec{x})\, d\vec{x} \neq 0$ because

$$\int \tilde{\phi}(\vec{x})\, d\vec{x} = \int \sum_{\vec{d}\in D} \tilde{c}_{\vec{d}} \tilde{(}M\vec{x}-\vec{\phi})\, d\vec{x} = \frac{1}{|\det M|} (\sum_{\vec{d}\in D} \tilde{c}_{\vec{d}}) \int \tilde{\phi}(\vec{x})\, d\vec{x}.$$

But say that $\int \tilde{\phi}(\vec{x})\, d\vec{x} = 0$. Then $\int_{\mathcal{S}} \tilde{\phi}(\vec{x})\, d\vec{x} = \tilde{c}_{\vec{d}_{i_1}} \dots \tilde{c}_{\vec{d}_{i_1 k}} \int \tilde{\phi}(\vec{x})\, d\vec{x} = 0$ for any subtile $\mathcal{S} = (M^{-1}\vec{x} + M^{-1}\vec{d}_{i_1}) \dots (M^{-1}\vec{x} + M^{-1}\vec{d}_{i_k})\mathcal{T}$. But the collection of all unions of subtiles is dense in $\mathcal{T}$, so this would imply $\tilde{\phi}(\vec{x}) = 0$, which we assumed was not the case.

Finally, the only solution to the two equations $\sum_{\vec{d}\in D} |\tilde{c}_{\vec{d}}|^2 = |\det M|$ and $\sum_{\vec{d}\in D} \tilde{c}_{\vec{d}} = |\det M|$ is when $\tilde{c}_{\vec{d}} = 1$ for all $\vec{d} \in D$, because the minimum value of $\sum_{\vec{d}\in D} |\tilde{c}_{\vec{d}}|^2$ for $\tilde{c}_{\vec{d}}$ lying on the hyperplane $\sum_{\vec{d}\in D} \tilde{c}_{\vec{d}} = |\det M|$ is when $\tilde{c}_{\vec{d}} = 1$. $\square$

Note that $\sum_{\vec{\gamma}\in\Gamma} \phi(\vec{x}-\vec{\gamma})$ is very rarely identically zero; it can't be if $\int \phi(\vec{x})\, d\vec{x} \neq 0$, for example. It is widely believed but unproven that for every expanding matrix M there exists a self-M-affine Γ tile; Lagarias and Wang [Lagarias 93] prove that this is true for $|\det M| \geq n+1$ and for $n < 4$.

The other approach, as previously mentioned, is to assume that f_0 satisfies a rank M dilation equation. For example, let's say that

$$f_0(\vec{x}) = \sum_{\vec{k}\in\Gamma} b_{\vec{k}} f_0(M\vec{x} - \vec{k})$$

for some set of $b_{\vec{k}}$. Then,

$$\begin{aligned}
\vec{a}_{f_0,f_r}(\vec{\gamma}) &= < f_0, f_r(\vec{x}-\vec{\gamma}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} b_{\vec{\alpha}} \sum_{\vec{\beta}\in\Gamma} \overline{c_{\vec{\beta}}} < f_0(M\vec{x}-\vec{\alpha}), f_{r-1}(M\vec{x} - M\vec{\gamma} - \vec{\beta}) > \\
&= \frac{1}{|\det M|} \sum\sum b_{\vec{\alpha}} \overline{c_{\vec{\beta}}} < f_0(\vec{u}), f_{r-1}(\vec{u} + \vec{\alpha} - M\vec{\gamma} - \vec{\beta}) > \\
&= \frac{1}{|\det M|} \sum_{\vec{\delta}} \sum_{\vec{\beta}} b_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}} < f_0, f_{r-1}(\vec{u} - \vec{\delta} \\
&= \frac{1}{|\det M|} \sum_{\vec{\delta}} \sum_{\vec{\beta}} b_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}} \vec{a}_{f_0,f_{r-1}}(\vec{\gamma}).
\end{aligned}$$

If we define the matrix R to have entry

$$R_{\vec{\gamma},\vec{\delta}} = \frac{1}{|\det M|} \sum_{\vec{\beta}} b_{M\vec{\gamma}+\vec{\beta}-\vec{\delta}} \overline{c_{\vec{\beta}}}$$

then $\vec{a}_{f_0,f_r} = R\vec{a}_{f_0,f_{r-1}} = R^r \vec{a}_{f_0,f_0}$.

Corollary 2.1.6 *Say f_0 satisfies the dilation equation*

$$f_0(\vec{x}) = \sum_{\vec{k}\in\Gamma} b_{\vec{k}} f(M\vec{x} - \vec{k})$$

and that there are only a finite number of non-zero $b_{\vec{k}}$'s and $c_{\vec{\gamma}}$'s. Then the Cascade algorithm converges to a compactly supported nonzero function in $L^2(\mathbf{R}^n)$ if and only *if the following conditions hold.*

- *Both T and R must have 1 as an eigenvalue.*
- *For every eigenvalue $\lambda \neq 1$ of T with $|\lambda| \geq 1$ and eigenvalue $\lambda' \neq 1$ of R with $|\lambda'| \geq 1$,*

$$\sum_{i=1}^{t} \sum_{j=1}^{s_i - k} \sum_{i'=1}^{t'} \sum_{j'=1}^{s'_i - k'} \vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'} \cdot \vec{a}_{f_0,f_0})(\vec{w}_{i,k+j} \cdot \vec{t}_{i',j'}) = 0$$

for $k = 0 \ldots \max(s_i) - 1$ and $k' = 0 \ldots \max(s'_i) - 1$, where

1. *there are t Jordan blocks with eigenvalue λ in the Jordan form factorization of $T = SJS^{-1}$ and t' Jordan blocks with eigenvalue λ' in the Jordan form factorization of $R = QKQ^{-1}$,*
2. *the i-th Jordan block of T is s_i by s_i and the i-th Jordan block of R is s'_i by s'_i,*
3. *$\vec{v}_{i,j}$ is the column of S corresponding to j-th column of the i-th Jordan block with eigenvalue λ ($\vec{v}_{i,j} = 0$ for $j > s_i$),*
4. *$\vec{t}_{i,j}$ is the column of Q corresponding to j-th column of the i-th Jordan block with eigenvalue λ' ($\vec{t}_{i,j} = 0$ for $j > s'_i$),*
5. *$\vec{w}_{i,j}$ is the row of S^{-1} corresponding to j-th row of the i-th Jordan block with eigenvalue λ ($\vec{w}_{i,j} = 0$ for $j > s_i$), and*
6. *$\vec{u}_{i,j}$ is the row of Q^{-1} corresponding to j-th row of the i-th Jordan block with eigenvalue λ' ($\vec{u}_{i,j} = 0$ for $j > s'_i$).*

- *For every eigenvalue $\lambda = 1$ of T and eigenvalue $\lambda' \neq 1$ of R,*

$$\sum_{i=1}^{t} \sum_{j=1}^{s_i - k} \sum_{i'=1}^{t'} \sum_{j'=1}^{s'_i - k'} \vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'} \cdot \vec{a}_{f_0,f_0})(\vec{w}_{i,k+j} \cdot \vec{t}_{i',j'}) = 0$$

for $k = 1 \ldots \max(s_i) - 1$ and $k' = 0 \ldots \max(s'_i) - 1$.

- *For every eigenvalue $\lambda \neq 1$ of T with $|\lambda| \geq 1$ and eigenvalue $\lambda' = 1$ of R,*

$$\sum_{i=1}^{t}\sum_{j=1}^{s_i-k}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i-k'} \vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'}\cdot\vec{a}_{f_0,f_0})(\vec{w}_{i,k+j}\cdot\vec{t}_{i',j'}) = 0$$

  for $k = 0\ldots\max(s_i)-1$ and $k' = 1\ldots\max(s'_i)-1$.

- *For every eigenvalue $\lambda = 1$ of T and eigenvalue $\lambda' = 1$ of R,*

$$\mathrm{Re}(\sum_{i=1}^{t}\sum_{j=1}^{s_i}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i} \vec{v}_{i,j}(\vec{0})(\vec{u}_{i',j'}\cdot\vec{a}_{f_0,f_0})(\vec{w}_{i,j}\cdot\vec{t}_{i',j'})) = \sum_{i=1}^{t}\sum_{j=1}^{s_i}\vec{v}_{i,j}(\vec{0})\vec{w}_{i,j}\cdot\vec{a}_{f_0,f_0} \neq 0$$

  and

$$\sum_{i=1}^{t}\sum_{j=1}^{s_i-k}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i-k'} \vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'}\cdot\vec{a}_{f_0,f_0})(\vec{w}_{i,k+j}\cdot\vec{t}_{i',j'}) = 0$$

  for $k = 1\ldots\max(s_i)-1$ and $k' = 1\ldots\max(s'_i)-1$.

Proof: Let $R = QKQ^{-1}$ be the Jordan form factorization of R. We know that $\vec{a}_{f_0,f_r} = R^r\vec{a}_{f_0,f_0} = QK^rQ^{-1}\vec{a}_{f_0,f_0}$. The contribution of each Jordan block of K to this vector is

$$(\begin{bmatrix}\vec{t}_1\ldots\vec{t}_{s'}\end{bmatrix}\begin{bmatrix}\lambda & 1 & & 0\\ & \lambda & 1 & \\ & & \ddots & \ddots\\ 0 & & & \lambda\end{bmatrix}^r\begin{bmatrix}\vec{u}_1\\ \vdots\\ \vec{u}_{s'}\end{bmatrix}\vec{a}_{f_0,f_0})$$

$$= \begin{bmatrix}\vec{t}_1\ldots\vec{t}_{s'}\end{bmatrix}\begin{bmatrix}\lambda^r & r\lambda^{r-1} & \ldots & \binom{r}{s'-1}\lambda^{r-s'+1}\\ & \lambda^r & \ldots & \vdots\\ & & \ddots & \vdots\\ 0 & & & \lambda^r\end{bmatrix}\begin{bmatrix}\vec{u}_1\cdot\vec{a}_{f_0,f_0}\\ \vdots\\ \vec{u}_{s'}\cdot\vec{a}_{f_0,f_0}\end{bmatrix}.$$

If $|\lambda| < 1$, then the middle matrix goes to the zero matrix and the contribution is $\vec{0}$. Otherwise, the contribution is

$$\sum_{i'=1}^{s'}\vec{t}_{i'}\sum_{j'=i'}^{s'}\binom{r}{j'-i'}\lambda^{r-j'+i'}(\vec{u}_{j'}\cdot\vec{a}_{f_0,f_0}).$$

Summing over all Jordan blocks with the same eigenvalue λ gives

$$\sum_{k'=1}^{t'}\sum_{i'=1}^{s'_{k'}}\vec{t}_{k',i'}\sum_{j'=i'}^{s'_{k'}}\binom{r}{j'-i'}\lambda^{r-j'+i'}(\vec{u}_{k',j'}\cdot\vec{a}_{f_0,f_0})$$

$$=\sum_{k'=0}^{\max(s'_i)-1}\binom{r}{k'}\lambda^{r-k'}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i-k'}\vec{t}_{i',j'}(\vec{u}_{i',k'+j'}\cdot\vec{a}_{f_0,f_0}).$$

Of course, we aren't interested in $\vec{a}_{f_0,f_r}$ as much as we are in the expression

$$\sum_{i=1}^{t}\sum_{j=1}^{s_i-k}\vec{v}_{i,j}(\vec{0})\vec{w}_{i,k+j}\cdot\vec{a}_{f_0,f_r}.$$

The contribution of the λ eigenvalue Jordan blocks of R to this expression is

$$\sum_{k'=0}^{\max(s'_i)-1}\binom{r}{k'}\lambda^{r-k'}\sum_{i=1}^{t}\sum_{j=1}^{s_i-k}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i-k'}\vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'}\cdot\vec{a}_{f_0,f_0})\vec{w}_{i,k+j}\cdot\vec{t}_{i',j'}$$

and if this is to be 0 for all r, then we must either have $\lambda = 1$ and $k' = 0$ or

$$\sum_{i=1}^{t}\sum_{j=1}^{s_i-k}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i-k'}\vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'}\cdot\vec{a}_{f_0,f_0})\vec{w}_{i,k+j}\cdot\vec{t}_{i',j'} = 0.$$

Thus, if it exists,

$$\lim_{r\to\infty}\lim_{n\to\infty} < f_n, f_{n+r} >= \sum_{i=1}^{t}\sum_{j=1}^{s_i}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i}\vec{v}_{i,j}(\vec{0})(\vec{u}_{i',j'}\cdot\vec{a}_{f_0,f_0})(\vec{w}_{i,j}\cdot\vec{t}_{i',j'})$$

where $\vec{v}_{i,j}$ and $\vec{w}_{i,j}$ are associated with Jordan blocks having eigenvalue $\lambda = 1$ and $\vec{t}_{i',j'}$ and $\vec{u}_{i',j'}$ are associated with Jordan blocks having eigenvalue $\lambda' = 1$. $\square$

2.2 Cascade Algorithm Examples

1. $M = 2$, $K = \{0, 1, 2, 3\}$, $f_0 = \chi_{[0,1]}$.

The support of any solutions will be in $[0,3]$.

$$T = \frac{1}{2}\begin{bmatrix} s_{-2} & s_{-3} & 0 & 0 & 0 \\ s_0 & s_{-1} & s_{-2} & s_{-3} & 0 \\ s_2 & s_1 & s_0 & s_{-1} & s_{-2} \\ 0 & s_3 & s_2 & s_1 & s_0 \\ 0 & 0 & 0 & s_3 & s_2 \end{bmatrix}$$

where $s_j = \sum_k c_{k+j}\overline{c_k}$. f_0 satisfies the dilation equation $f_0 = f_0(2x) + f_0(2x-1)$, so

$$R = \frac{1}{2}\begin{bmatrix} \overline{c_2}+\overline{c_3} & \overline{c_3} & 0 & 0 & 0 \\ \overline{c_0}+\overline{c_1} & \overline{c_1}+\overline{c_2} & \overline{c_2}+\overline{c_3} & \overline{c_3} & 0 \\ 0 & \overline{c_0} & \overline{c_0}+\overline{c_1} & \overline{c_1}+\overline{c_2} & \overline{c_2}+\overline{c_3} \\ 0 & 0 & 0 & \overline{c_0} & \overline{c_0}+\overline{c_1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since $\vec{a}_{f_0,f_0} = \vec{e}_0$, only the upper right 3 by 3 submatrix of R matters when computing $R^r\vec{a}_{f_0,f_0}$; call this submatrix R'. R' has the following eigenvalues: $\frac{\overline{c_0+c_1+c_2+c_3}}{2}$, and $\frac{\overline{c_1}+\overline{c_2}\pm\sqrt{(\overline{c_1}+\overline{c_2})^2-4(\overline{c_1c_2}-\overline{c_0c_3})}}{4}$. The requirement that R' have 1 as an eigenvalue leads to two possibilities: either $c_0+c_1+c_2+c_3 = 2$, or $c_0c_3 - c_1c_2 + 2c_1 + 2c_2 = 2$. For both these cases, the corresponding eigenvector is $(\overline{c_3}(2-\overline{c_0}-\overline{c_1}), (2-\overline{c_2}-\overline{c_3})(2-\overline{c_0}-\overline{c_1}), \overline{c_0}(2-\overline{c_2}-\overline{c_3}))$ and the corresponding left eigenvector is $((\overline{c_0}+\overline{c_1})(2-\overline{c_2}-\overline{c_3}), (2-\overline{c_2}-\overline{c_3})(2-\overline{c_0}-\overline{c_1}), (\overline{c_2}+\overline{c_3})(2-\overline{c_2}-\overline{c_3}))$. In the first case, the limit ϕ will satisfy $\int\phi\,dx = 1$ if ϕ exists. By Proposition 2.1.5, condition $\vec{1}$ must then hold, so we know that $c_0+c_2 = c_1+c_3 = 1$.

2. $M = 2$, $K = \{0,1,2,3\}$, $c_0+c_2 = c_1+c_3 = 1$.

   As proven in corollary 2.1.3, these conditions on the c_k's guarantees that T has 1 as an eigenvalue and $\vec{1}$ as a left eigenvector. They also in this case imply that T has $1/2$ as an eigenvalue as well, with $(c_1+c_0-\overline{c_1}-\overline{c_0})\vec{1} + (2,1,0,-1,-2)$ as the left eigenvector.

   That leaves three other eigenvalues to worry about. They are solutions to the following cubic:

$$
\begin{aligned}
&-\frac{c_1\,\overline{c_1}}{8}-\frac{c_0\,\overline{c_0}}{8}+\frac{\overline{c_1}x}{4}-\frac{x^2}{2}+\frac{c_0\,\overline{c_1}}{8}+\frac{c_1\,\overline{c_0}}{8}-\frac{xc_0}{4}+\frac{xc_1}{4}-\frac{\overline{c_0}c_1\,\overline{c_1}}{4}+\frac{{c_1}^2\left(\overline{c_1}\right)^3c_0}{2}+\frac{\overline{c_1}c_0\,\overline{c_0}}{4}\\
&+\frac{3\,c_0\,\left(\overline{c_1}\right)^2c_1}{4}-\frac{c_0\,\overline{c_1}c_1}{4}+{c_0}^2\left(\overline{c_1}\right)^2\overline{c_0}c_1+\frac{\overline{c_1}{c_0}^2\left(\overline{c_0}\right)^2}{4}-\frac{5\,{c_1}^2\left(\overline{c_1}\right)^2c_0\,\overline{c_0}}{4}\\
&+\frac{{c_1}^3\left(\overline{c_0}\right)^2}{4}-\frac{{c_1}^2\left(\overline{c_0}\right)^3c_0}{4}-\frac{{c_1}^3\overline{c_1}\left(\overline{c_0}\right)^2}{4}+\frac{{c_0}^2\left(\overline{c_1}\right)^3}{4}\\
&-\frac{{c_0}^3\left(\overline{c_1}\right)^2\overline{c_0}}{4}-\frac{{c_0}^2\left(\overline{c_1}\right)^3c_1}{4}-\frac{{c_1}^3\left(\overline{c_1}\right)^3}{4}+\frac{{c_0}^2\left(\overline{c_0}\right)^3c_1}{2}\\
&+\frac{{c_0}^2\left(\overline{c_0}\right)^2c_1}{4}+c_0\,\left(\overline{c_0}\right)^2{c_1}^2\overline{c_1}-\frac{{c_1}^2\overline{c_0}\left(\overline{c_1}\right)^2}{2}-\frac{c_0\,\left(\overline{c_1}\right)^2{c_1}^2}{2}+\frac{{c_0}^3\overline{c_1}\left(\overline{c_0}\right)^2}{2}-\frac{\overline{c_0}x}{4}\\
&+\frac{7\,\left(\overline{c_1}\right)^2c_0\,\overline{c_0}c_1}{8}-\frac{3\,{c_0}^2\overline{c_1c_0}c_1}{8}+\frac{3\,{c_1}^3\left(\overline{c_1}\right)^2}{8}+\frac{{c_1}^3\overline{c_0}}{8}-\frac{\left(\overline{c_1}\right)^3c_1}{8}-\frac{{c_1}^2\left(\overline{c_0}\right)^2}{8}\\
&-\frac{{c_1}^3\overline{c_1}}{8}+\frac{{c_0}^2\left(\overline{c_0}\right)^2}{8}-\frac{{c_0}^2\left(\overline{c_0}\right)^3}{8}-\frac{{c_0}^2\left(\overline{c_1}\right)^2}{8}-\frac{5\,{c_1}^2\left(\overline{c_1}\right)^2}{8}+\frac{3\,{c_1}^2\left(\overline{c_1}\right)^3}{8}-\frac{{c_0}^3\left(\overline{c_0}\right)^2}{8}\\
&+\frac{c_0\,\left(\overline{c_1}\right)^3}{8}+\frac{7\,{c_1}^2\overline{c_1}c_0\,\overline{c_0}}{8}-\frac{3\,c_0\,\overline{c_0}c_1\,\overline{c_1}}{4}+\frac{c_0\,\overline{c_1}{c_1}^2}{8}-\frac{{c_1}^2\overline{c_0}c_0}{8}-\frac{3\,c_0\,\left(\overline{c_0}\right)^2c_1\,\overline{c_1}}{8}\\
&-\frac{3\,{c_1}^2\left(\overline{c_0}\right)^2c_0}{8}-\frac{5\,{c_1}^3\overline{c_0c_1}}{8}-\frac{\left(\overline{c_1}\right)^2c_0\,\overline{c_0}}{8}+\frac{c_1\,\left(\overline{c_0}\right)^3c_0}{8}+\frac{{c_1}^2\left(\overline{c_0}\right)^2\overline{c_1}}{8}-\frac{3\,{c_0}^2\left(\overline{c_1}\right)^2\overline{c_0}}{8}\\
&-\frac{5\,c_0\,\left(\overline{c_1}\right)^3c_1}{8}+\frac{\overline{c_0}\left(\overline{c_1}\right)^2c_1}{8}+\frac{{c_0}^3\overline{c_1c_0}}{8}+\frac{{c_0}^2\left(\overline{c_1}\right)^2c_1}{8}+\frac{xc_0\,\overline{c_0}}{4}-\frac{3\,xc_1\,\overline{c_1}}{4}\\
&+\frac{\overline{c_1}c_0\,x}{4}+\frac{c_1\,\overline{c_0}x}{4}+\frac{xc_0\,c_1}{4}+\frac{c_0\,x^2}{2}+x^3-\frac{x{c_1}^2}{4}+\frac{\overline{c_0c_1}x}{4}-\frac{3\,{c_0}^2\left(\overline{c_0}\right)^2x}{4}-\frac{3\,{c_1}^2\left(\overline{c_1}\right)^2x}{4}\\
&+\frac{c_1\,\left(\overline{c_0}\right)^2x}{4}-\frac{\overline{c_1}c_0\,x^2}{2}-\frac{c_1\,\overline{c_0}x^2}{2}+\frac{3\,c_1\,\left(\overline{c_1}\right)^2c_0\,x}{4}-\frac{c_0\,\overline{c_0c_1}x}{4}\\
&+\frac{\overline{c_0}x^2}{2}-\frac{\left(\overline{c_1}\right)^2x}{4}+\frac{3\,xc_1\,\left(\overline{c_0}\right)^2c_0}{4}+\frac{3\,x{c_1}^2\overline{c_0c_1}}{4}-\frac{xc_1\,c_0\,\overline{c_0}}{4}\\
&+\frac{3\,{c_0}^2\overline{c_0c_1}x}{4}-\frac{c_0\,\overline{c_0}xc_1\,\overline{c_1}}{2}+\frac{3\,{c_1}^2\overline{c_0c_1}}{4}+\frac{c_1\,c_0\,\overline{c_0}}{4}-\frac{c_0\,\left(\overline{c_1}\right)^2}{4}+\frac{c_1\,\left(\overline{c_1}\right)^2}{4}\\
&+\frac{{c_1}^2\overline{c_1}}{4}-\frac{{c_1}^2\overline{c_0}}{4}-\frac{3\,x\overline{c_0}c_1\,\overline{c_1}}{4}-\frac{3\,xc_0\,\overline{c_1}c_1}{4}+\frac{3\,x{c_1}^2\overline{c_1}}{4}+\frac{3\,xc_1\,\left(\overline{c_1}\right)^2}{4}+\frac{x\overline{c_1}{c_0}^2}{4}\\
&-\frac{{c_0}^3\left(\overline{c_0}\right)^3}{4}+\frac{{c_1}^3\left(\overline{c_1}\right)^2\overline{c_0}}{2}-\frac{x{c_1}^2\left(\overline{c_0}\right)^2}{2}-\frac{x\left(\overline{c_1}\right)^2{c_0}^2}{2}-\frac{5\,\overline{c_1}{c_0}^2\left(\overline{c_0}\right)^2c_1}{4}.
\end{aligned}
$$

3. $M = 2$, $K = \{0, 1, 2, 3\}$, $c_k \in \mathbf{R}$.

When the coefficients are real, two simplifications occur. First, $s_j = s_{-j}$. Second, $< f, f(x+j) > \; = \; < f, f(x-j) >$, so when examining the behavior of $< f_n, f_n(x-j) >$ it suffices to consider the three by three matrix

$$T' = \frac{1}{2} \begin{bmatrix} s_0 & 2s_1 & 2s_2 \\ s_2 & s_1 + s_3 & s_0 \\ 0 & s_3 & s_2 \end{bmatrix}.$$

The requirement that 1 be an eigenvalue of T' is a quintic in c_0 and c_3, so may be solved for at least one solution in these when the other four coefficients are fixed.

2.3 Orthogonality of the Scaling Function

Proposition 2.3.1 *If ϕ is orthogonal to $\phi(\vec{x}-\vec{\gamma})$ for all $\vec{\gamma} \in \Gamma$, then*

$$\sum_{\vec{\beta}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}}\overline{c_{\vec{\beta}}} = \delta_{\vec{0},\vec{\gamma}}|\det M|$$

for all $\vec{\gamma} \in \Gamma$.

Proof: By hypothesis, we have that $<\phi, \phi(\vec{x}-\vec{\gamma})> = g\delta_{\vec{0},\vec{\gamma}}$ for some g. On the other hand,

$$\begin{aligned}
<\phi, \phi(\vec{x}-\vec{\gamma})> &= <\sum_{\vec{\alpha}\in\Gamma} c_\alpha \phi(M\vec{x}-\vec{\alpha}), \sum_{\vec{\beta}\in\Gamma} c_\beta \phi(M\vec{x}-M\vec{\gamma}-\vec{\beta})> \\
&= \frac{1}{|\det M|}\sum_{\vec{\alpha}\in\Gamma}\sum_{\vec{\beta}\in\Gamma} c_\alpha \overline{c_{\vec{\beta}}} <\phi(\vec{u}), \phi(\vec{u}-M\vec{\gamma}-\vec{\beta}+\vec{\alpha})> \\
&= \frac{1}{|\det M|}\sum_{\vec{\alpha}\in\Gamma}\sum_{\vec{\beta}\in\Gamma} c_\alpha \overline{c_{\vec{\beta}}} g\delta_{\vec{0},M\vec{\gamma}+\vec{\beta}-\vec{\alpha}} \\
&= \frac{g}{|\det M|}\sum_{\vec{\beta}\in\Gamma} c_{M\vec{\gamma}+\vec{\beta}}\overline{c_{\vec{\beta}}}.
\end{aligned}$$

Setting this equal to $g\delta_{\vec{0},\vec{\gamma}}$ gives the result. □

Corollary 2.3.2 *If ϕ is orthogonal to $\phi(\vec{x}-\vec{\gamma})$ for all $\vec{\gamma} \in \Gamma$, then T has 1 as an eigenvalue and $e_{\vec{0}}$ as an associated eigenvector.*

Proof:

$$(Te_{\vec{0}})(\vec{\gamma}) = T_{\vec{0},\vec{\gamma}} = \frac{1}{|\det M|}\sum_{\vec{\beta}} c_{M\vec{\gamma}+\vec{\beta}}\overline{c_{\vec{\beta}}} = \delta_{0,\gamma}$$

□

Proposition 2.3.3 *If T has 1 as a simple eigenvalue and $e_{\vec{0}}$ as the associated eigenvector, then ϕ and $\phi(\vec{x}-\vec{\gamma})$ are orthogonal for all $\vec{\gamma} \in \Gamma$.*

Proof: As noted in the previous section, $T\vec{a}_{\phi,\phi} = \vec{a}_{\phi,\phi}$. On the other hand, if T has 1 as a simple eigenvalue and $e_{\vec{0}}$ as the associated eigenvector, then $\vec{a}_{\phi,\phi} = \lambda e_{\vec{0}}$

for some constant λ. This implies that $\vec{a}_{\phi,\phi}(\vec{\gamma}) = <\phi, \phi(\vec{x} - \vec{\gamma})> = 0$ for $\vec{\gamma} \neq \vec{0}$, as desired. $\square$

2.4 Orthogonality of Wavelets

In this section, we show how to verify that a given set of wavelets are both mutually orthonormal and orthogonal to a given scaling function. Recall that the purpose of wavelets is to give a basis for W_0, the orthogonal complement of V_0 in V_1. ($V_1 = V_0 \oplus W_0$). In particular, $W_0 \subset V_1$, so a wavelet ψ_i satisfies

$$\psi_i = \sum_{\vec{\gamma} \in \Gamma} d_{i,\vec{\gamma}} \phi(M\vec{x} - \vec{\gamma})$$

for some set of $d_{i,\vec{\gamma}}$ because the $\phi(M\vec{x} - \vec{\gamma})$ are an orthogonal basis for V_1.

The $d_{i,\vec{\gamma}}$'s are similar to the $c_{\vec{\gamma}}$'s in that they give a constructive handle on the wavelets. We will thus be concerned about imposing conditions on the $d_{i,\vec{\gamma}}$'s that ensure orthonormality, and later when we are trying to construct wavelets, we will be interested in finding sets of $d_{i,\vec{\gamma}}$'s that satisfy these conditions. Our analysis is simplified if we adopt the convention that $\psi_0 = \phi$, so that for example, $d_{0,\vec{k}} = c_{\vec{k}}$. A scaling function along with a complete set of wavelets is referred to as a *wavelet system.*

Proposition 2.4.1 (Shifted Orthogonality Condition) *If* $<\psi_i, \psi_j(\vec{x} - \vec{k})> = g\delta_{\vec{0},\vec{k}}\delta_{i,j}$ *for all* $i, j \in 0 \ldots |\det M| - 1$ *and* $\vec{k} \in \Gamma$*, then*

$$\sum_{\vec{\beta} \in \Gamma} d_{i,\vec{\beta}+M\vec{\gamma}} \overline{d_{j,\vec{\beta}}} = \delta_{i,j}\delta_{\vec{0},\vec{\gamma}} |\det M|$$

for all $i, j \in 0 \ldots |\det M| - 1$ *and* $\gamma \in \Gamma$.

PROOF:

$$
\begin{aligned}
<\psi_i, \psi_j(\vec{x}-\vec{\gamma})> &= \sum_{\vec{\alpha}\in\Gamma}\sum_{\vec{\beta}\in\Gamma} d_{i,\vec{\alpha}}\overline{d_{j,\vec{\beta}}} < \phi(M\vec{x}-\vec{\alpha}), \phi(M\vec{x}-M\vec{\gamma}-\vec{\beta}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} \frac{1}{|\det M|} \sum_{\vec{\beta}\in\Gamma} d_{i,\vec{\alpha}}\overline{d_{j,\vec{\beta}}} < \phi(\vec{u}), \phi(\vec{u}+\vec{\alpha}-M\vec{\gamma}-\vec{\beta}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} \frac{1}{|\det M|} \sum_{\vec{\beta}\in\Gamma} d_{i,\vec{\alpha}}\overline{d_{j,\vec{\beta}}} g \delta_{\vec{0},M\vec{\gamma}+\vec{\beta}-\vec{\alpha}} \\
&= \frac{g}{|\det M|} \sum_{\vec{\beta}\in\Gamma} d_{i,M\vec{\gamma}+\vec{\beta}}\overline{d_{j,\vec{\beta}}}
\end{aligned}
$$

Setting this equal to $g\delta_{i,j}\delta_{\vec{0},\vec{\gamma}}$ gives the result. $\square$

The above orthogonality condition can be rephrased more elegantly if, following [Kautsky 95], we put it into matrix form. Define the matrix $A_{\vec{x}}$ by saying that its (i,j)-th element is $d_{i,M\vec{x}+\vec{y}_j}$ for some fixed ordering of the elements $\vec{y}_i$ of $\Gamma/M\Gamma$, $i = 0 \ldots |\det M| - 1$.

Proposition 2.4.2 (Shifted Orthogonality Condition: Matrix Version) *The shifted orthogonality condition is equivalent to*

$$\sum_{\vec{l}\in\Gamma} A_{\vec{l}} A^*_{\vec{k}+\vec{l}} = |\det M| \delta_{\vec{0},\vec{k}} I$$

for every $\vec{k} \in \Gamma$.

PROOF:

$$
\begin{aligned}
(\sum_{\vec{l}\in\Gamma} A_{\vec{l}} A^*_{\vec{k}+\vec{l}})_{(i,j)} &= \sum_{\vec{l}\in\Gamma}\sum_{k} (A_{\vec{l}})_{(i,k)} (A^*_{\vec{k}+\vec{l}})_{(k,j)} \\
&= \sum_{\vec{l}\in\Gamma}\sum_{k} d_{i,M\vec{l}+\vec{y}_k} \overline{d_{j,M\vec{k}+M\vec{l}+\vec{y}_k}} \\
&= \sum_{\vec{\gamma}\in\Gamma} d_{i,\vec{\gamma}} \overline{d_{j,\vec{\gamma}+M\vec{k}}}
\end{aligned}
$$

On the other hand, the (i,j)-th entry of $|\det M|\delta_{\vec{0},\vec{k}} I$ is $|\det M|\delta_{\vec{0},\vec{k}}\delta_{i,j}$. $\square$

Definition 2.4.3 *A matrix valued function $A(\vec{x})$ is called* paraunitary by rows *if $A(\vec{x})A(\vec{x})^* = I$ for $\vec{x}$ on the $|\det M|$ dimensional torus $T = \{\vec{x} : x_i\overline{x_i} = 1, i = 1, \ldots, n\}$. If $A(\vec{x})^*A(\vec{x}) = I$ for $\vec{x} \in T$, then $A(\vec{x})$ is called* paraunitary by columns, *and if $A(\vec{x})$ is paraunitary by both rows and columns, then it is simply called* paraunitary.

Lemma 2.4.4 *The* polyphase *power series matrix*

$$A(\vec{x}) = \frac{1}{\sqrt{|\det M|}} \sum_{\vec{k}\in\Gamma} A_{\vec{k}} \vec{x}^{\vec{k}}$$

is paraunitary by rows if and only if the matrices $A_{\vec{k}}$ satisfy the shifted orthogonality condition.

Proof:

$$\begin{aligned}
A(\vec{x})A(\vec{x})^* &= \frac{1}{|\det M|}(\sum_{\vec{k}\in\Gamma} A_{\vec{k}}\vec{x}^{\vec{k}})(\sum_{\vec{j}\in\Gamma} A_{\vec{j}}^*\overline{\vec{x}^{\vec{j}}}) \\
&= \frac{1}{|\det M|}(\sum_{\vec{k}\in\Gamma} A_{\vec{k}}\vec{x}^{\vec{k}})(\sum_{\vec{j}\in\Gamma} A_{\vec{j}}^*\vec{x}^{-\vec{j}}) \\
&= \frac{1}{|\det M|}\sum_{\vec{i}\in\Gamma}\sum_{\vec{k}-\vec{j}=\vec{i}} A_{\vec{k}}A_{\vec{j}}^*\vec{x}^{\vec{i}} \\
&= \frac{1}{|\det M|}\sum_{\vec{i}\in\Gamma}\vec{x}^{\vec{i}}\sum_{\vec{k}\in\Gamma} A_{\vec{k}}A_{\vec{k}-\vec{i}}^* \\
&= \frac{1}{|\det M|}\sum_{\vec{i}\in\Gamma}\vec{x}^{\vec{i}}|\det M|\delta_{\vec{0},\vec{i}}I \\
&= I
\end{aligned}$$

□

Theorem 2.4.5 *The set of wavelet systems satisfying the shifted orthogonality condition and forming a basis of V_1 is in one to one correspondence with the set of paraunitary power series matrices.*

Proof: [Adapted from [Strichartz 93].] By lemma 2.4.4, the set of wavelet systems satisfying the shifted orthogonality condition is in one to one correspondence with power series matrices that are paraunitary by rows. We will now show that wavelet systems that form a complete basis of V_1 are in one to one correspondence with the set of power series matrices which are paraunitary by columns.

Let f be an arbitrary function in V_1. Then

$$f = \sum_{\vec{\alpha}\in\Gamma} f_{\vec{\alpha}}\phi(M\vec{x} - \vec{\alpha})$$

for some set of $f_{\vec{\alpha}}$'s because $\{\phi(M\vec{x} - \vec{\alpha})\}_{\vec{\alpha}\in\Gamma}$ is a basis for V_1. We want to see under what conditions

$$f = \sum_{i=0}^{r}\sum_{\vec{\gamma}\in\Gamma} g_{\vec{\gamma},i}\psi_i(\vec{x} - \vec{\gamma})$$

for some set of $g_{\vec{\gamma},i}$'s. But, because of orthogonality,

$$\begin{aligned}
g_{\vec{\gamma},i} &= < f, \psi_i(\vec{x} - \vec{\gamma}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} f_{\vec{\alpha}} < \phi(M\vec{x} - \vec{\alpha}), \psi_i(\vec{x} - \vec{\gamma}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} f_{\vec{\alpha}} \sum_{\vec{\beta}\in\Gamma} \overline{d_{i,\vec{\beta}}} < \phi(M\vec{x} - \vec{\alpha}), \phi(M\vec{x} - M\vec{\gamma} - \vec{\beta}) > \\
&= \sum_{\vec{\alpha}\in\Gamma} f_{\vec{\alpha}} \sum_{\vec{\beta}\in\Gamma} \overline{d_{i,\vec{\beta}}} \frac{1}{|\det M|}\delta_{\vec{\alpha},M\vec{\gamma}+\vec{\beta}} \\
&= \frac{1}{|\det M|}\sum_{\vec{\beta}\in\Gamma} f_{M\vec{\gamma}+\vec{\beta}}\overline{d_{i,\vec{\beta}}}.
\end{aligned}$$

So, we want to see when

$$\begin{aligned}
\sum_{\vec{\alpha}\in\Gamma} f_{\vec{\alpha}}\phi(M\vec{x} - \vec{\alpha}) &= \frac{1}{|\det M|}\sum_{i=0}^{r}\sum_{\vec{\gamma}\in\Gamma}\sum_{\vec{\beta}\in\Gamma} f_{M\vec{\gamma}+\vec{\beta}}\overline{d_{i,\vec{\beta}}}\psi_i(\vec{x} - \vec{\gamma}) \\
&= \frac{1}{|\det M|}\sum_{i=0}^{r}\sum_{\vec{\gamma}\in\Gamma}\sum_{\vec{\beta}\in\Gamma} f_{M\vec{\gamma}+\vec{\beta}}\overline{d_{i,\vec{\beta}}}\sum_{\vec{\gamma}'\in\Gamma} d_{i,\vec{\gamma}'}\phi(M\vec{x} - M\vec{\gamma} - \vec{\gamma}')
\end{aligned}$$

which is equivalent to

$$f_{M\vec{\gamma}+\vec{\alpha}} = \frac{1}{|\det M|}\sum_{i=0}^{r}\sum_{\vec{\gamma}\in\Gamma}\sum_{\vec{\beta}\in\Gamma} f_{M\vec{\gamma}+\vec{\beta}}\overline{d_{i,\vec{\beta}}}d_{i,\vec{\alpha}-M\vec{\gamma}}$$

and also to

$$\sum_{i=0}^{|\det M|-1} \sum_{\vec{\gamma}\in\Gamma} \overline{d_{i,\vec{\beta}}} d_{i,\vec{\alpha}-M\vec{\gamma}} = |\det M| \delta_{\alpha,\beta}.$$

On the other hand, the polyphase matrix $A(\vec{x})$ is paraunitary by columns when

$$\begin{aligned}
(A(\vec{x})^* A(\vec{x}))_{(\alpha,\beta)} &= \frac{1}{|\det M|}((\sum_{\vec{k}\in\Gamma} A^*_{\vec{k}} \vec{x}^{-\vec{k}})(\sum_{\vec{l}\in\Gamma} A_{\vec{l}} \vec{x}^{\vec{l}}))_{(\alpha,\beta)} \\
&= \frac{1}{|\det M|} \sum_{\vec{i}\in\Gamma} \vec{x}^{\vec{i}} \sum_{\vec{k}} (A^*_{\vec{k}} A_{\vec{i}+\vec{k}})_{(\alpha,\beta)} \\
&= \frac{1}{|\det M|} \sum_{\vec{i}\in\Gamma} \vec{x}^{\vec{i}} \sum_{\vec{k}} \sum_{j=0}^{|\det M|-1} (A^*_{\vec{k}})_{(\alpha,j)} (A_{\vec{i}+\vec{k}})_{(j,\beta)} \\
&= \frac{1}{|\det M|} \sum_{\vec{i}\in\Gamma} \vec{x}^{\vec{i}} \sum_{\vec{k}} \sum_{j} \overline{d_{j,M\vec{k}+\vec{y}_\alpha}} d_{j,M\vec{i}+M\vec{k}+\vec{y}_\beta} \\
&= \delta_{\alpha,\beta}
\end{aligned}$$

which is equivalent to the above expression. □

Corollary 2.4.6 *If $\phi, \psi_1, \ldots, \psi_r$ and their translates form an orthonormal basis for V_1, then $r = |\det M| - 1$.*

PROOF: By the preceeding theorem, the $r+1$ by $|\det M|$ polyphase matrix associated with $\phi, \psi_1, \ldots, \psi_r$ is paraunitary. But only a square matrix can be paraunitary, so $r = |\det M| - 1$. □

2.5 Putting it All Together

We are now ready to return full circle and show how good scaling functions and wavelets give rise to multiresolution approximations. We use the abbreviation $m = |\det M|$ throughout.

Lemma 2.5.1 *If a scaling function ϕ and associated wavelets ψ_i are orthonormal to each other and their Γ translates, then they along with the rescaled wavelets $m^{j/2}\psi_i(M^j\vec{x})$ and their Γ translates are all orthonormal.*

PROOF: First we show that the translate of any rescaled wavelet is orthogonal to the scaling function; that is

$$< \phi, m^{j/2}\psi_i(M^j\vec{x} - \vec{k}) > = 0.$$

By hypothesis this is true for $j = 0$; we prove the general case by induction:

$$\begin{aligned} < \phi, m^{j/2}\psi_i(M^j\vec{x} - \vec{k}) > &= m^{j/2}\sum_{\vec{\gamma}} c_{\vec{\gamma}} < \phi(M\vec{x} - \vec{\gamma}, \psi_i(M^j\vec{x} - \vec{k}) > \\ &= m^{j/2+1}\sum_{\vec{\gamma}} c_{\vec{\gamma}} < \phi(\vec{u}), \psi_i(M^{j-1}\vec{u} + M^{j-1}\vec{\gamma} - \vec{k}) > . \end{aligned}$$

But since $\vec{k} - M^{j-1}\vec{\gamma} \in \Gamma$, each inner product in the above sum is zero by induction.

Second we show that the rescaled wavelets and their translates are orthonormal. By rescaling and translation, all we need to prove is that

$$< \psi_i, m^{k/2}\psi_j(M^k\vec{x} - \vec{\gamma}) > = \delta_{i,j}\delta_{0,k}\delta_{\vec{0},\vec{\gamma}}.$$

We are given that this is true for $k = 0$, so we may assume $k > 0$. But then

$$\begin{aligned} < \psi_i, m^{k/2}\psi_j(M^k\vec{x} - \vec{\gamma}) > &= m^{k/2}\sum_{\vec{\alpha}} d_{\vec{\alpha},i} < \phi(M\vec{x} - \vec{\alpha}), \psi_j(M^k\vec{x} - \vec{\gamma}) > \\ &= m^{k/2+1}\sum_{\vec{\alpha}} d_{\vec{\alpha},i} < \phi(\vec{u}), \psi_j(M^{k-1}\vec{u} + M^{k-1}\vec{\alpha} - \vec{\gamma}) > \end{aligned}$$

and every inner product in the above sum is zero by our first result. $\square$

Lemma 2.5.2 *A compactly supported scaling function ϕ orthonormal to its Γ translates generates a multiresolution analysis of $L^2(\mathbf{R}^n)$ if and only if*

$$\int \phi(\vec{x})\, d\vec{x} = \sqrt{\mathrm{vol}(\mathbf{R}^n/\Gamma)}.$$

PROOF: [INSPIRED BY A PROOF SKETCH IN [STRICHARTZ 94].] We want to show that $\cup_{i\in\mathbf{Z}} V_i = L^2(\mathbf{R}^n)$. If $P_i f$ is the projection of a function f into V_i, then what we want is equivalent to $\lim_{i\to\infty} P_i f = f$ for all $f \in L^2(\mathbf{R}^n)$. Since we are in L^2, this in turn is the same as saying $\lim_{i\to\infty} ||P_i f - f||_2^2 = 0$ or even $\lim_{i\to\infty} ||P_i f||_2^2 = ||f||_2^2$ since $||P_i f||_2^2 + ||P_i f - f||_2^2 = ||f||_2^2$ by the Pythagorean theorem.

Now,

$$P_i f = \sum_{\vec{\gamma}\in\Gamma} < f, |\det M|^{i/2} \phi(M^i\vec{x} - \vec{\gamma}) > |\det M|^{i/2} \phi(M^i \vec{x} - \vec{\gamma})$$

which implies

$$||P_i f||_2^2 = < P_i f, P_i f > = \sum_{\vec{\gamma}\in\Gamma} |< f, \phi(M^i\vec{x} - \vec{\gamma}) >|^2 |\det M|^i.$$

Now let χ_B be the characteristic function of a n dimensional ball B with unspecified radius and center. Linear combinations of such functions are dense in $L^2(\mathbf{R}^n)$, so it suffices to prove that $\lim_{i\to\infty} ||P_i\chi_B||_2^2 = ||\chi_B||_2^2$. But

$$\begin{aligned} ||P_i\chi_B||_2^2 &= |\det M|^i \sum_{\vec{\gamma}\in\Gamma} |< \chi_B, \phi(M^i\vec{x} - \vec{\gamma}) >|^2 \\ &= |\det M|^i \sum_{\vec{\gamma}\in\Gamma} |\int_B \phi(M^i\vec{x} - \vec{\gamma}) > \, d\vec{x}|^2 \\ &= |\det M|^{-i} \sum_{\vec{\gamma}\in\Gamma} |\int_{M^i B} \phi(\vec{u} - \vec{\gamma})\, d\vec{u}|^2. \end{aligned}$$

Now, for large i, $M^i B$ will be a very large hyper-ellipsoid. The support of $\phi(\vec{u} - \vec{\gamma})$ will either be entirely outside of this hyper-ellipsoid (in which case it contributes 0 to the sum), partly inside the hyper-ellipsoid, or entirely inside the hyper-ellipsoid, in which case its contribution will be $|\det M|^{-i} |\int \phi(\vec{x})\, d\vec{x}|^2$.

Specifically,

$$\left| ||P_i\chi_B||^2 - \frac{|\det M|^{-i}\operatorname{vol}(M^iB)|\int\phi(\vec{x})\,d\vec{x}|^2}{\operatorname{vol}(\mathbf{R}^n/\Gamma)}\right|$$

$$\leq \frac{|\det M|^{-i}}{\operatorname{vol}(\mathbf{R}^n/\Gamma)}\left[(R_1+R_2)\operatorname{area}(\partial M^iB))\left|\int\phi(\vec{x})\,d\vec{x}\right|^2+(2R_1+2R_2)K\right]$$

where

- The support of $\phi(\vec{x})$ is contained by a ball of radius R_1,
- $\mathbf{R}^n/\Gamma$, the fundamental region of the lattice Γ, is contained by a ball of radius R_2,
- K is the maximum of $|\int_X \phi(\vec{x})\,d\vec{x}|^2$ over all compact subsets X of $\mathbf{R}^n$,
- ∂M^iB is the boundary of M^iB, and
- $\operatorname{vol}(M^iB)$ is the n-dimensional measure of M^iB, while $\operatorname{area}(\partial M^iB)$ is the $n-1$-dimensional measure of ∂M^iB.

The proof of this bound on $||P_i\chi_B||^2$ is simple. Let M^iB-R_1 be the hyper-ellipsoid obtained by removing a ball of radius R_1 from all the boundary points of M^i. The number of $\phi(\vec{u}-\vec{\gamma})$ whose support is entirely inside M^iB will be less than or equal to the number of lattice points $\vec{\gamma}$ inside M^iB-R_1. But this will be approximately equal to $\frac{\operatorname{vol}(M^iB-R_1)}{\operatorname{vol}(\mathbf{R}^n/\Gamma)}$, with the absolute value of the error bounded by $\frac{R_2\operatorname{area}(M^iB-R_1)}{\operatorname{vol}(\mathbf{R}^n/\Gamma)}$.

Next, we need to consider the $\phi(\vec{u}-\vec{\gamma})$ which are only partially inside M^iB. The contribution of these to the sum will be at most $|\det M|^{-i}K$, by the definition of K. To find the number of these functions, we need to count lattice points inside the hyper-annulus consisting of those points within distance R_1 of ∂M^iB. But there will certainly be less than $(2R_1+2R_2)\frac{\operatorname{area}(\partial M^iB)}{\operatorname{vol}(\mathbf{R}^n/\Gamma)}$ of these. Combining this with the previous bound (and a liberal application of the triangle inequality), gives the desired bound.

Finally, note that $\operatorname{vol}(M^iB)=|\det M|^i\operatorname{vol}(B)=|\det M|^i||\chi_B||_2^2$ and that

$$\lim_{i\to\infty}\frac{\operatorname{area}(\partial M^iB)}{|\det M|^{-i}}\to 0.$$

Thus,

$$\lim_{i\to\infty} ||P_i\chi_B||_2^2 = ||\chi_B||_2^2 \frac{|\int \phi(\vec{x})\, d\vec{x}|^2}{\mathrm{vol}(\mathbf{R}^n/\Gamma)}$$

and the theorem is proven. $\Box$

Theorem 2.5.3 *Given any compactly supported function ϕ which is orthonormal to its Γ translates and satisfies $\int \phi(\vec{x})\, d\vec{x} = \sqrt{\mathrm{vol}(\mathbf{R}^n/\Gamma)}$, then*

$$V_j = \{\sum_{\vec{\gamma}\in\Gamma} a_{\vec{\gamma}}\phi(M^j\vec{x}-\vec{\gamma}) \mid a_{\vec{\gamma}} \in \mathbf{C}, \sum_{\vec{\gamma}\in\Gamma} |a_{\vec{\gamma}}|^2 < \infty\}$$

is a rank M multiresolution analysis of $L^2(\mathbf{R}^n)$.

PROOF: [ADOPTED FROM [STRICHARTZ 94].] The nesting, scaling, and orthonormality conditions follow immediately from the definition of V_j, and completeness by the preceding lemma. The only thing left to prove is separation. This follows from the bound

$$f \in V_j \implies \max |f(\vec{x})|^2 \leq C|\det M|^j < f, f >$$

because $f \in V_j$ for all j would imply (as $j \to -\infty$) $\max |f(\vec{x})|^2 = 0$ and thus $f(\vec{x}) = 0$. To prove the bound, write

$$\begin{aligned}
|f(\vec{x})|^2 &= \left|\sum_{\vec{\gamma}\in\Gamma} < f, |\det M|^{j/2}\phi(M^j\vec{x}-\vec{\gamma}) > |\det M|^{j/2}\phi(M^j\vec{x}-\vec{\gamma})\right|^2 \\
&= |\det M|^{2j}\left|\sum_{\vec{\gamma}\in\Gamma} \phi(M^j\vec{x}-\vec{\gamma}) \int f(\vec{y})\overline{\phi(M^j\vec{y}-\vec{\gamma})}\, d\vec{y}\right|^2 \\
&= |\det M|^{2j}\left|\int f(\vec{y}) \left(\sum_{\vec{\gamma}\in\Gamma} \phi(M^j\vec{x}-\vec{\gamma})\overline{\phi(M^j\vec{y}-\vec{\gamma})}\right) d\vec{y}\right|^2 \\
&\leq |\det M|^{2j}\left(\int |f(\vec{y})|^2\, d\vec{y}\right)\left(\int \left|\sum_{\vec{\gamma}\in\Gamma} \phi(M^j\vec{x}-\vec{\gamma})\overline{\phi(M^j\vec{y}-\vec{\gamma})}\right|^2 d\vec{y}\right) \\
&= |\det M|^{2j} < f, f > \\
&\quad \int \left(\sum_{\vec{\gamma}\in\Gamma} \phi(M^j\vec{x}-\vec{\gamma})\overline{\phi(M^j\vec{y}-\vec{\gamma})}\right)\left(\sum_{\vec{\gamma}'\in\Gamma} \overline{\phi(M^j\vec{x}-\vec{\gamma}')}\phi(M^j\vec{y}-\vec{\gamma}')\right) d\vec{y}
\end{aligned}$$

$$
\begin{aligned}
&= |\det M|^{2j} < f,f > \\
&\quad \sum_{\vec{\gamma}\in\Gamma}\sum_{\vec{\gamma}'\in\Gamma} \phi(M^j\vec{x}-\vec{\gamma})\overline{\phi(M^j\vec{x}-\vec{\gamma}')}\int \overline{\phi(M^j\vec{y}-\vec{\gamma})}\phi(M^j\vec{y}-\vec{\gamma}')\,d\vec{y} \\
&= |\det M|^{j} < f,f > \sum_{\vec{\gamma}\in\Gamma}\sum_{\vec{\gamma}'\in\Gamma} \phi(M^j\vec{x}-\vec{\gamma})\overline{\phi(M^j\vec{x}-\vec{\gamma}')}\delta_{\vec{\gamma},\vec{\gamma}'} \\
&= |\det M|^{j} < f,f > \sum_{\vec{\gamma}\in\Gamma} |\phi(M^j\vec{x}-\vec{\gamma})|^2 \\
&\leq |\det M|^{2j} < f,f > C
\end{aligned}
$$

where the first inequality comes from the CBS inequality and the second from the fact that ϕ is compactly supported. $\square$

Theorem 2.5.4 *The Cascade algorithm applied to a starting function ϕ_0 converges to a compactly supported function ϕ which generates a rank M multiresolution analysis of $L^2(\mathbf{R}^n)$ if*

- $\sum_{\vec{\gamma}\in\Gamma}\phi_0(\vec{x}-\vec{\gamma}) = 1$,
- $\int \phi_0\,d\vec{x} = \sqrt{\mathrm{vol}(\mathbf{R}^n/\Gamma)}$,
- *Only a finite number of the $c_{\vec{\gamma}}$'s are non-zero,*
- *T has 1 as a simple eigenvalue and $e_{\vec{0}}$ as the associated eigenvector,*
- *All other eigenvalues of T are less than one in absolute value, and*
- $\sum_{\vec{k}'\equiv\vec{k}(M)} c_{\vec{k}'} = 1$ *(Condition $\vec{1}$).*

Proof: By corollary 2.1.3, the Cascade algorithm converges under these conditions to a non-zero compact function ϕ with $\int \phi\,d\vec{x} = \int \phi_0\,d\vec{x} = \sqrt{\mathrm{vol}(\mathbf{R}^n/\Gamma)}$. By proposition 2.3.3, T having 1 as a simple eigenvalue and $e_{\vec{0}}$ as the associated eigenvector guarantees that ϕ with be orthonormal to its Γ translates. But theorems 2.5.2 and 2.5.3 imply that a compactly supported function, orthonormal to its Γ translates, and with $\sqrt{\mathrm{vol}(\mathbf{R}^n/\Gamma)}$ as its integral, will generate a rank M multiresolution analysis of $L^2(\mathbf{R}^n)$. $\square$

Theorem 2.5.5 *Assume that*

$$\phi_0(\vec{x}) = \sum_{\vec{k}\in\Gamma} b_{\vec{k}}\phi(M\vec{x} - \vec{k})$$

for some finite set of $b_{\vec{k}}$. *Then the Cascade algorithm applied to a starting function* ϕ_0 *converges to a compactly supported function* ϕ *which generates a rank* M *multiresolution analysis of* $L^2(\mathbf{R}^n)$ *if and only if*

- $\int \phi_0 \, d\vec{x} = \sqrt{\mathrm{vol}(\mathbf{R}^n/\Gamma)}$,
- *Both* T *and* R *must have* 1 *as an eigenvalue.*
- *Only finite number of the* $c_{\vec{\gamma}}$*'s are non-zero,*
- T *has* 1 *as an eigenvalue and* $e_{\vec{0}}$ *as the associated eigenvector,*
- $\sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}} = |\det M|$ *(Condition 1),*
- *For every eigenvalue* $\lambda \neq 1$ *of* T *with* $|\lambda| \geq 1$ *and eigenvalue* $\lambda' \neq 1$ *of* R *with* $|\lambda'| \geq 1$,

  $$\sum_{i=1}^{t}\sum_{j=1}^{s_i-k}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i-k'} \vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'} \cdot \vec{a}_{\phi_0,\phi_0})(\vec{w}_{i,k+j} \cdot \vec{t}_{i',j'}) = 0$$

  for $k = 0 \ldots \max(s_i) - 1$ *and* $k' = 0 \ldots \max(s'_i) - 1$,
- *For every eigenvalue* $\lambda \neq 1$ *of* T *with* $|\lambda| \geq 1$ *and eigenvalue* $\lambda' = 1$ *of* R,

  $$\sum_{i=1}^{t}\sum_{j=1}^{s_i-k}\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i-k'} \vec{v}_{i,j}(\vec{0})(\vec{u}_{i',k'+j'} \cdot \vec{a}_{\phi_0,\phi_0})(\vec{w}_{i,k+j} \cdot \vec{t}_{i',j'}) = 0$$

  for $k = 0 \ldots \max(s_i) - 1$ *and* $k' = 1 \ldots \max(s'_i) - 1$, *and*
- *For every eigenvalue* $\lambda' = 1$ *of* R,

  $$\sum_{i'=1}^{t'}\sum_{j'=1}^{s'_i} (\vec{u}_{i',j'} \cdot \vec{a}_{\phi_0,\phi_0})\vec{t}_{i',j'}(\vec{0}) = \vec{a}_{\phi_0,\phi_0}(\vec{0})$$

and

$$\sum_{i'=1}^{t'} \sum_{j'=1}^{s'_i-k'} (\vec{u}_{i',k'+j'} \cdot \vec{a}_{\phi_0,\phi_0})\vec{t}_{i',j'}(\vec{0}) = 0$$

for $k' = 1 \ldots \max(s'_i) - 1$.

PROOF: Same as the previous theorem, but using the conditions associated with corollary 2.1.6 instead of corollary 2.1.3. $\square$

2.6 Regularity

There are many different yet related ways to require that a multiresolution analysis be "regular" to degree $\vec{p} \in \mathbf{Z}^n$. Below are listed some of the more important or popular. We use the notation that $\vec{a} \leq \vec{b}$ if $a_i \leq b_i$ for all i.

- **[Derivatives]** The $\vec{p}$-th derivative $\partial^{\vec{p}}\phi$ of the scaling function exists.

- **[Polynomial Representation]** For every compact set X and polynomial $q(\vec{x})$ of degree less than or equal to $\vec{p}$, there is a function $f \in V_0$ which agrees with $q(\vec{x})$ on X.

- **[Vanishing Moments]** $< \vec{x}^{\vec{s}}, \psi_i >= 0$ for all $\vec{0} \leq \vec{s} \leq \vec{p}$ and $i = 1 \ldots |\det M| - 1$.

- **[Strang-Fix]** $\hat{\phi}(\vec{v})$, the Fourier transform of ϕ, has a 0 of order $\vec{p}$ for all $\vec{v} = 2\pi\Gamma - \vec{0}$: $(\frac{\partial}{\partial \vec{x}})^{\vec{s}}\hat{\phi}(\vec{v}) = 0$ for $\vec{0} \leq \vec{s} \leq \vec{p}$.

- **[Sum Rule I]**
$$\sum_{\vec{\gamma} \in \Gamma} c_{\vec{\gamma}} \vec{\gamma}^{\vec{s}} e^{2\pi i \vec{v} \cdot \vec{\gamma}} = 0$$
for every $\vec{v} \in \Gamma / M^*\Gamma - 0$ and $\vec{0} \leq \vec{s} \leq \vec{p}$.

- **[Sum Rule II]**
$$\sum_{\vec{k} \equiv \vec{k}'(M)} c_{\vec{k}} \vec{k}^{\vec{s}} = C_{\vec{s}}$$
for every $\vec{k}' \in \frac{\Gamma}{M\Gamma}$ and $\vec{0} \leq \vec{s} \leq \vec{p}$, where $C_{\vec{s}}$ is a constant not dependent on $\vec{k}'$.

(These are by no means the only measures of regularity possible. For examinations of other measures of regularity [such as Sobolev smoothness or Holder continuity], see [Cohen 96], [Karoui 94], and [Jia 97].)

Some of these criteria are obviously related. For example, Polynomial Representation implies Vanishing Moments, and they are equivalent if ϕ generates a complete multiresolution analysis of $L^2(\mathbf{R}^n)$. The following sections examine, explain, and motivate each of these notions more closely.

2.6.1 Derivatives

In this section we investigate the $L^2(\mathbf{R}^n)$ existence of derivatives of the scaling function. It is clear that if a particular derivative exists for the scaling function ϕ then it exists for all the wavelets, since these are just linear combinations of translates of $\phi(M\vec{x})$.

By the repeated application of the chain rule to the dilation equation, we get

$$\partial^{\vec{i}}\phi = \sum_{\vec{k}\in\Gamma}\sum_{|\vec{j}|=|\vec{i}|} \mathcal{M}_{[r]_{\vec{j}\vec{i}}}(\partial^{\vec{j}}\phi)(M\vec{x}-\vec{k})$$

where $r = |\vec{i}|$,

$$\mathcal{M}_{[r]_{\vec{j}\vec{i}}} = \binom{r}{\vec{i}}^{-1} \sum_{\vec{e}_{k_1}+\ldots+\vec{e}_{k_r}=\vec{i}} \; \sum_{\vec{e}_{k'_1}+\ldots+\vec{e}_{k'_r}=\vec{j}} M_{k'_1k_1}\ldots M_{k'_rk_r},$$

and M_{ij} is the (i,j)-th entry of M. The matrix $\mathcal{M}_{[r]}$ was studied by Cabrelli, Heil, and Molter in connection with vanishing moments; they prove the following result.

Lemma 2.6.1 *[Cabrelli 96] Let $\vec{\lambda}$ be the vector of eigenvalues of an arbitrary matrix A. Then the eigenvalues of $\mathcal{A}_{[s]}$ are $\vec{\lambda}^{\vec{q}}$ for every vector $\vec{q}$ such that $|\vec{q}| = s$.*

Proof: [Sketch] Choose a basis such that A is upper-triangular; this does not change $\mathcal{A}_{[s]}$'s eigenvalues. In fact, there is an ordering of monomials of total degree s such that $\mathcal{A}_{[s]}$ will also be upper-triangular when A is. But then the values $\vec{\lambda}^{\vec{q}}$ appear on $\mathcal{A}_{[s]}$'s diagonal, and we are done. $\square$

Armed with this result, we return to the study of $\partial^{\vec{i}}\phi$. To investigate the $L^2(\mathbf{R}^n)$ existence of these derivatives, it is natural to consider the vector

$$\vec{a}_{[r]}(\vec{i},\vec{j},\vec{\gamma}) = <\partial^{\vec{i}}\phi, \partial^{\vec{j}}\phi(\vec{x}-\vec{\gamma})>$$

where $|\vec{i}| = |\vec{j}| = r$ and $\vec{\gamma} \in \Gamma$. We have the recurrence relation

$$\vec{a}_{[r]}(\vec{i}, \vec{j}, \vec{\gamma}) = \sum_{\vec{l} \in \Gamma} \sum_{\vec{l}' \in \Gamma} \sum_{|\vec{i}'|=r} \sum_{|\vec{j}'|=r} \frac{c_{\vec{l}} \overline{c_{\vec{l}'}}}{|\det M|} \mathcal{M}^{[r]}_{\vec{i}'\vec{i}} \mathcal{M}^{[r]}_{\vec{j}'\vec{j}} \vec{a}_{[r]}(\vec{i}', \vec{j}', \vec{l} - \vec{l}' - M\vec{\gamma})$$

and if we define the matrix $T_{[r]}$ to have entry

$$T_{[r]_{(\vec{i},\vec{j},\vec{\gamma}),(\vec{i}',\vec{j}',\vec{\gamma}')}} = \sum_{\vec{l} \in \Gamma} \frac{c_{M\vec{\gamma}+\vec{l}-\vec{\gamma}'} \overline{c_{\vec{l}}}}{|\det M|} \mathcal{M}_{[r]_{\vec{i}'\vec{i}}} \mathcal{M}_{[r]_{\vec{j}'\vec{j}}}$$

we get the equation $\vec{a}_{[r]} = T_{[r]} \vec{a}_{[r]}$, which may be analyzed using the techniques of section 2.1. Notice in particular that

$$T_{[r]_{(\vec{i},\vec{j},\vec{\gamma}),(\vec{i}',\vec{j}',\vec{\gamma}')}} = \mathcal{M}_{[r]_{\vec{i}'\vec{i}}} \mathcal{M}_{[r]_{\vec{j}'\vec{j}}} T_{\vec{\gamma},\vec{\gamma}'}$$

which means that $T_{[r]}$ is a Kronecker (tensor) product: $T_{[r]} = \mathcal{M}_{[r]} \otimes \mathcal{M}_{[r]} \otimes T$.

Proposition 2.6.2 *If $\partial^{\vec{i}}\phi \in L^2(\mathbf{R}^n)$ for all $|\vec{i}| = r$ and if M has $\vec{\lambda}$ as its vector of eigenvalues, then T has $\vec{\lambda}^{-\vec{j}-\vec{k}}$ as an eigenvalue, for some $\vec{j}$ and $\vec{k}$ such that $|\vec{j}| = |\vec{k}| = r$.*

Proof: If $\partial^{\vec{i}}\phi \in L^2(\mathbf{R}^n)$ for all $|\vec{i}| = r$, then by the Cascade Algorithm theorem, $T_{[r]}$ must have 1 as an eigenvector. Let $\vec{v}$ be the corresponding eigenvector. Without loss of generality, we may assume $\vec{v} = \vec{a} \otimes \vec{b} \otimes \vec{w}$, where $\vec{a}$ and $\vec{b}$ are eigenvectors of $\mathcal{M}_{[r]}$ and $\vec{w}$ is an eigenvector of T. Specifically, let $\mathcal{M}_{[r]}\vec{a} = \alpha\vec{a}$, $\mathcal{M}_{[r]}\vec{b} = \beta\vec{b}$ and $T\vec{w} = \lambda\vec{w}$. Then $T_{[r]}\vec{v} = (\mathcal{M}_{[r]} \otimes \mathcal{M}_{[r]} \otimes T)(\vec{a} \otimes \vec{b} \otimes \vec{w}) = (\mathcal{M}^{[r]}\vec{a}) \otimes (\mathcal{M}^{[r]}\vec{b}) \otimes (T\vec{w}) = (\alpha\vec{a}) \otimes (\beta\vec{b}) \otimes (\lambda\vec{w}) = \alpha\beta\lambda(\vec{a} \otimes \vec{b} \otimes \vec{w}) = \alpha\beta\lambda\vec{v}$. Thus, $\lambda = \frac{1}{\alpha\beta}$. But by lemma 2.6.1, $\alpha = \vec{\lambda}^{\vec{j}}$ and $\beta = \vec{\lambda}^{\vec{k}}$ for some $\vec{j}$ and $\vec{k}$ such that $|\vec{j}| = |\vec{k}| = r$. $\square$

In the one dimensional case, $\mathcal{M}_{[r]} = M^r$, and T must have M^{-2r} as an eigenvalue in order for ϕ to have r derivatives. In higher dimensions, the situation becomes more complicated. For example, $M = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ has $1 \pm i$ as eigenvalues, so T must have $i/2$, $1/2$, or $-i/2$ as an eigenvalue for $\frac{\partial\phi}{\partial x}$ and $\frac{\partial\phi}{\partial y}$ to both exist. For $\partial^{\vec{i}}\phi$ to exist for all $|\vec{i}| = r$, T must have one of $\pm 2^{-r}$ or $\pm 2^{-r}i$ as an eigenvalue.

2.6.2 Vanishing Moments

It is possible to calculate explicitly the moments of a given wavelet or scaling function. This is because we can assume that $\int \phi \, d\vec{x}$ is known and because

$$\begin{aligned}
\int \vec{x}^{\vec{p}} \phi \, d\vec{x} &= \sum_{\vec{\gamma} \in \Gamma} \int \vec{x}^{\vec{p}} c_{\vec{\gamma}} \phi(M\vec{x} - \vec{\gamma}) \, d\vec{x} \\
&= \frac{1}{|\det M|} \sum_{\vec{\gamma} \in \Gamma} c_{\vec{\gamma}} \int (M^{-1}\vec{u} - \vec{\gamma})^{\vec{p}} \phi(\vec{u}) \, d\vec{u} \\
&= \frac{1}{|\det M|} \sum_{\vec{\gamma} \in \Gamma} c_{\vec{\gamma}} \int ((M^{-1}\vec{u})^{\vec{p}} + P_{\vec{p},\vec{\gamma}}(\vec{x}))\phi(\vec{u}) \, d\vec{u}
\end{aligned}$$

where $P_{\vec{p},\vec{\gamma}}(\vec{x})$ is a polynomial of degree strictly less than $\vec{p}$. If we assume that we have already calculated $\int \vec{x}^{\vec{q}} \phi(\vec{x}) \, d\vec{x}$ for $\vec{q} < \vec{p}$, then we can calculate $\int P_{\vec{p},\vec{\gamma}}(\vec{x})\phi(\vec{u}) \, d\vec{u}$ and call it $C_{\vec{p},\vec{\gamma}}$. Thus,

$$\int \vec{x}^{\vec{p}} \phi \, d\vec{x} = \frac{1}{|\det M|} \sum_{\vec{\gamma} \in \Gamma} c_{\vec{\gamma}} (\int (M^{-1}\vec{x})^{\vec{p}} \phi(\vec{x}) \, d\vec{x} + C_{\vec{p},\vec{\gamma}}).$$

Now consider the matrix $\mathcal{M}_{[s]}^{-1}$ where $s = |\vec{p}|$. If we let $X_{[s]}(\vec{x})$ be the vector, indexed by vectors $\vec{q}$ such that $|\vec{q}| = s$, then [Cabrelli 96] shows that $\mathcal{M}_{[s]}^{-1}$ satisfies the equation

$$X_{[s]}(M^{-1}\vec{x}) = \mathcal{M}_{[s]}^{-1} X_{[s]}(\vec{x}).$$

Thus, we can write

$$\int X_{[s]}(\vec{x}) \, d\vec{x} = C_{[s]} + \frac{1}{|\det M|} \sum_{\vec{\gamma}} c_{\vec{\gamma}} \int \mathcal{M}_{[s]}^{-1} X_{[s]}(\vec{x}) \, d\vec{x}$$

where $C_{[s]}$ is the vector with $\vec{p}$ entry

$$\frac{1}{|\det M|} \sum_{\vec{\gamma}} c_{\vec{\gamma}} C_{\vec{p},\vec{\gamma}}.$$

But then, assuming that $\sum_{\vec{\gamma}} c_{\vec{\gamma}} = |\det M|$ (Condition 1), we can solve this equation and get

$$\int X_{[s]}(\vec{x})\, d\vec{x} = (I - \mathcal{M}_{[s]}^{-1})^{-1} C_{[s]}.$$

$I - \mathcal{M}_{[s]}^{-1}$ is always invertible because all of M^{-1}'s eigenvalues, and hence all of $M_{[s]}^{-1}$'s eigenvalues as well, are less than one.

2.6.3 Fourier Analysis of Wavelets

The Fourier transform of a function f in $L^2(\mathbf{R}^n)$ is defined as

$$\hat{f}(\vec{v}) = \int_{\mathbf{R}^n} f(\vec{x})e^{-i\vec{v}\cdot\vec{x}}\, d\vec{x}.$$

The dilation equation allows us to deduce a beautiful form for $\hat{\phi}$:

$$\begin{aligned}
\hat{\phi}(\vec{v}) &= \sum_{\vec{\gamma}\in\Gamma} c_\gamma \widehat{\phi(M\vec{x} - \vec{\gamma})} \\
&= \sum_{\vec{\gamma}} c_\gamma \int_{\mathbf{R}^n} f(M\vec{x} - \vec{\gamma})e^{-i\vec{v}\cdot\vec{x}}\, d\vec{x} \\
&= \sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} \int f(\vec{u} - \vec{\gamma})e^{-i\vec{v}\cdot M^{-1}\vec{u}}\, d\vec{u} \\
&= \sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} \int f(\vec{u'})e^{-i\vec{v}\cdot M^{-1}(\vec{u'}+\vec{\gamma})}\, d\vec{u'} \\
&= \sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} e^{-i\vec{v}\cdot M^{-1}\vec{\gamma}} \int f(\vec{u'})e^{-i\vec{v}\cdot M^{-1}\vec{u'}}\, d\vec{u'} \\
&= (\sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} e^{-i\vec{v}\cdot M^{-1}\vec{\gamma}})\hat{\phi}(M^{-*}\vec{v})
\end{aligned}$$

where M^{-*} is the adjoint (i.e., the conjugate transpose) of M^{-1}. If $N = M^{-*}$ and $m(\vec{v}) = \sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} e^{-i\vec{v}\cdot\vec{\gamma}}$, then we can write

$$\hat{\phi}(\vec{v}) = m(N\vec{v})\hat{\phi}(N\vec{v})$$

In particular,

$$\begin{aligned}
\hat{\phi}(\vec{v}) &= \hat{\phi}(N^r\vec{v}) \prod_{j=1}^{r} m(N^j\vec{v}) \\
&= \hat{\phi}(\vec{0}) \prod_{j=1}^{\infty} m(N^j\vec{v})
\end{aligned}$$

since N is eventually contractive. Note that $\hat{\phi}(\vec{0}) = \int_{\mathbf{R}^n} \phi(\vec{x})\, d\vec{x} = \sqrt{\mathrm{vol}(\mathbf{R}^n/\Gamma)}$ since ϕ is assumed to be the scaling function of a complete multiresolution analysis of $L^2(\mathbf{R}^n)$.

The zeros of $\hat{\phi}$ will play an important role in determining the regularity of ϕ. The following proposition tells us where they are.

Proposition 2.6.3 *If ϕ satisfies Condition 1, then $\hat{\phi}(\vec{v}) = 0$ for $\vec{v} \in 2\pi\Gamma - \vec{0}$*

PROOF: We saw above that $\hat{\phi}(\vec{v}) = m(N\vec{v})\hat{\phi}(N\vec{v})$, so the zeros of $\hat{\phi}(\vec{v})$ are either zeros of $m(N\vec{v})$ or N^{-j} times such a zero. Now,

$$
\begin{aligned}
(2.1) \qquad m(N\vec{v}) &= \sum_{\vec{\gamma}\in\Gamma} \frac{c_{\vec{\gamma}}}{|\det M|} e^{-i\vec{v}\cdot M^{-1}\vec{\gamma}} \\
(2.2) \qquad &= \sum_{\vec{\gamma}'\in\frac{\Gamma}{M\Gamma}} \sum_{\vec{\alpha}\in M\Gamma} \frac{c_{\vec{\gamma}'+\vec{\alpha}}}{|\det M|} e^{-i\vec{v}\cdot M^{-1}\vec{\gamma}'} e^{-i\vec{v}\cdot M^{-1}\vec{\alpha}} \\
(2.3) \qquad &= \sum_{\vec{\gamma}'\in\frac{\Gamma}{M\Gamma}} e^{-i\vec{v}\cdot M^{-1}\vec{\gamma}'}
\end{aligned}
$$

because $e^{-i\vec{v}\cdot M^{-1}\vec{\alpha}} = 1$ for $\vec{v} \in 2\pi\Gamma$ and $\vec{\alpha} \in M\Gamma$, and $\sum_{\vec{\alpha}\in M\Gamma} \frac{c_{\vec{\gamma}'+\vec{\alpha}}}{|\det M|} = 1$ by Condition 1.

Now, $\frac{\Gamma}{M\Gamma}$ is a finite abelian group under addition and $\vec{\gamma}' \to e^{-i\vec{v}\cdot M^{-1}\vec{\gamma}'}$ is a homomorphism from $\frac{\Gamma}{M\Gamma}$ to a multiplicative subgroup of the roots of unity, just as long as $\vec{v} \in 2\pi\Gamma$. In other words, $e^{-i\vec{v}\cdot M^{-1}\vec{\gamma}'}$ is a character, and it is well known that the sum of a nontrivial character over the elements of its group is zero. This proves that $m(N\vec{v})$ is zero whenever $\vec{v} \in 2\pi\Gamma$, $\vec{v} \notin 2\pi M^*\Gamma$.

Finally, every $\vec{v} \in 2\pi\Gamma - \vec{0}$ is either not in $2\pi M\Gamma$ or is N^{-j} times such a $\vec{v}$, so $\hat{\phi}$ has zeros at all those places. $\square$

Lemma 2.6.4 *The scaling function coefficients satisfy the sum rule*

$$\sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}} \vec{\gamma}^{\vec{p}} e^{2\pi i\vec{v}\cdot\vec{\gamma}} = 0$$

for all $\vec{v} \in \Gamma/M^*\Gamma - 0$ if and only if $\partial^{\vec{p}} m(w) = 0$ for $\vec{w} \in \frac{2\pi\Gamma}{M^*\Gamma} - 2\pi\Gamma$.

Proof: $m(\vec{w}) = \sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} e^{-i\vec{w}\cdot\vec{\gamma}}$, so

$$\begin{aligned}
\partial^{\vec{p}} m(w) &= \sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} \partial^{\vec{p}} e^{-i\vec{w}\cdot\vec{\gamma}} \\
&= \sum_{\vec{\gamma}} \frac{c_\gamma}{|\det M|} (-i)^{|\vec{p}|} \vec{\gamma}^{\vec{p}} e^{-i\vec{w}\cdot\vec{\gamma}} \\
&= 0.
\end{aligned}$$

$\square$

Theorem 2.6.5 *The sum rule*

$$\sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}} \vec{\gamma}^{\vec{s}} e^{2\pi i\vec{v}\cdot\vec{\gamma}} = 0$$

for all $\vec{0} \leq \vec{s} \leq \vec{p}$ *and* $\vec{v} \in \Gamma/M^*\Gamma - 0$ *implies the Strang-Fix condition*

$$(\frac{\partial}{\partial\vec{x}})^{\vec{s}} \hat{\phi}(\vec{v}) = 0$$

for all $\vec{0} \leq \vec{s} \leq \vec{p}$ *and* $\vec{v} \in 2\pi\Gamma - \vec{0}$.

Proof: Since $\hat{\phi}(\vec{v}) = m(N\vec{v})\hat{\phi}(N\vec{v})$,

$$\partial^{\vec{s}} \hat{\phi}(\vec{v}) = \sum_{\vec{i}+\vec{j}=\vec{s}} \partial^{\vec{i}} m(N\vec{v}) \partial^{\vec{j}} \hat{\phi}(N\vec{v}).$$

Assume that $\vec{v} \notin 2\pi M^*\Gamma$. Then by lemma 2.6.4, $\partial^{\vec{i}} m(N\vec{v}) = 0$ for $\vec{0} \leq \vec{i} \leq \vec{p}$, so $\partial^{\vec{s}} \hat{\phi}(\vec{v}) = 0$. If on the other hand, $\vec{v} \in 2\pi M^*\Gamma$, then $N^k\vec{v} \notin 2\pi M^*\Gamma$ for some k. Also, $\hat{\phi}(N\vec{v}) = 0$ for $\vec{v} \in 2\pi M^*\Gamma - \vec{0}$, and by induction we may assume that $\partial^{\vec{s}} \hat{\phi}(N\vec{v}) = 0$ for $\vec{0} \leq \vec{s} < \vec{p}$ and $\vec{v} \in 2\pi M^*\Gamma - \vec{0}$. Thus,

$$\partial^{\vec{p}} \hat{\phi}(\vec{v}) = m(N\vec{v}) \partial^{\vec{p}} \hat{\phi}(N\vec{v}) = \partial^{\vec{p}} \hat{\phi}(N^k\vec{v}) m(N^k\vec{v}) \prod_{l=1}^{k-1} m(N^l\vec{v}).$$

But then $N^K\vec{v} \notin 2\pi M^*\Gamma$ and $m(N^k\vec{v}) = 0$ so $\partial^{\vec{p}} \hat{\phi}(\vec{v}) = 0$. $\square$

2.6.4 Sum Rules

Proposition 2.6.6 *Sum Rule I*

$$\sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}}\vec{\gamma}^{\vec{s}} e^{2\pi i\vec{v}\cdot\vec{\gamma}} = 0$$

for $\vec{v} \in \Gamma/M^*\Gamma - \vec{0}$ *and the Sum Rule II*

$$\sum_{\vec{\gamma}\equiv\vec{\gamma}'(M)} c_{\vec{\gamma}}\vec{\gamma}^{\vec{s}} = C_{\vec{s}}$$

are equivalent.

Proof:

$$\begin{aligned}
\sum_{\vec{\gamma}\in\Gamma} c_{\vec{\gamma}}\vec{\gamma}^{\vec{s}} e^{2\pi i\vec{v}\cdot\vec{\gamma}} &= \sum_{M\vec{\alpha}+\vec{\beta}=\vec{\gamma}} c_{M\vec{\alpha}+\vec{\beta}}(M\vec{\alpha}+\vec{\beta})^{\vec{s}} e^{2\pi i\vec{v}\cdot(M\vec{\alpha}+\vec{\beta})} \\
&= \sum_{\vec{\alpha}\in\Gamma}\sum_{\vec{\beta}\in\Gamma} c_{M\vec{\alpha}+\vec{\beta}}(M\vec{\alpha}+\vec{\beta})^{\vec{s}} e^{2\pi i\vec{v}\cdot\vec{\beta}} \\
&= \sum_{\vec{\beta}\in\Gamma} e^{2\pi i\vec{v}\cdot\vec{\beta}} \sum_{\vec{\alpha}\in\Gamma} c_{M\vec{\alpha}+\vec{\beta}}(M\vec{\alpha}+\vec{\beta})^{\vec{s}}
\end{aligned}$$

Now, if Sum Rule II holds, then $\sum_{\vec{\alpha}\in\Gamma} c_{M\vec{\alpha}+\vec{\beta}}(M\vec{\alpha}+\vec{\beta})^{\vec{s}} = C_{\vec{s}}$ and then $\sum_{\vec{\beta}\in\Gamma} e^{2\pi i\vec{v}\cdot\vec{\beta}} C_{\vec{s}} = 0$. On the other hand, say $\sum_{\vec{\beta}\in\Gamma} c_{M\vec{\alpha}+\vec{\beta}} e^{2\pi i\vec{v}\cdot\vec{\beta}} \sum_{\vec{\alpha}\in\Gamma}(M\vec{\alpha}+\vec{\beta})^{\vec{s}} = 0$ for all $\vec{v} \in \Gamma/M^*\Gamma - \vec{0}$. We can write this in matrix form by letting F be the Fourier matrix with $(\vec{v},\vec{\beta})$ entry $e^{2\pi i\vec{v}\cdot\vec{\beta}}$ $[\vec{v} \in \Gamma/M^*\Gamma, \vec{\beta} \in \Gamma/M\Gamma]$ and $\vec{z}$ be the vector with $\vec{\beta}$ entry $\sum_{\vec{\alpha}\in\Gamma}(M\vec{\alpha}+\vec{\beta})^{\vec{s}}$. Then $F\vec{z} = C'\vec{e}_1$ for some constant C'. Since F is a Fourier matrix, we can invert it to get that $\vec{z} = C'F^{-1}\vec{e}_1 = \frac{1}{|\det M|}C'\vec{1}$. But this is the same as saying $\sum_{\vec{\alpha}\in\Gamma}(M\vec{\alpha}+\vec{\beta})^{\vec{s}} = C'$ for all $\vec{\beta} \in \Gamma/M\Gamma$, which is precisely Sum Rule II.
□

Notice that this proposition gives us an explicit formula for $C_{\vec{s}}$:

$$C_{\vec{s}} = \frac{1}{|\det M|}\sum_{\vec{\beta}\in\Gamma} e^{2\pi i\vec{0}\cdot\vec{\beta}} \sum_{\vec{\alpha}\in\Gamma} c_{M\vec{\alpha}+\vec{\beta}}(M\vec{\alpha}+\vec{\beta})^{\vec{s}} = \frac{1}{|\det M|}\sum_{\vec{\gamma}} c_{\vec{\gamma}}\vec{\gamma}^{\vec{s}}.$$

Theorem 2.6.7 *If the scaling function coefficients satisfy the sum rule*

$$\sum_{\vec{k}\equiv\vec{k}'(M)} c_{\vec{k}}\vec{k}^{\vec{s}} = C_{\vec{s}}$$

for $\vec{0} \leq \vec{s} \leq \vec{p}$ *then* ϕ *has* $\vec{p}$ *vanishing moments.*

PROOF: By proposition 2.6.6 and theorem 2.6.5, this sum rule implies the Strang-Fix condition. But by [Strang 73], the Strang-Fix condition implies the Polynomial Representation property, which coupled with the fact that the wavelets are orthogonal to the scaling function, implies the vanishing moments condition. $\square$

Chapter 3

Constructing a Multiresolution Analysis

3.1 Constructing Wavelets

The analysis in section 2.4 tells us that the problem of finding wavelets given a scaling function is the same as the problem of given a paraunitary (by rows) power series vector $\vec{p}$, finding a paraunitary power series matrix $A(\vec{x})$ with $\vec{p}$ as its first row.

For certain values of $|\det M|$, this problem is not hard. For example, if $|\det M| = 2$, then $A(\vec{x}) = \begin{bmatrix} a(\vec{x}) & b(\vec{x}) \\ -\overline{b(\vec{x})} & \overline{a(\vec{x})} \end{bmatrix}$ is paraunitary whenever $\vec{p} = (a(\vec{x}), b(\vec{x}))$ is. There are similar formulas for $|\det M| = 4$ and $|\det M| = 8$ [Meyer 92], but not for any other value of $|\det M|$. The obstruction is topological: if $\vec{p}(\vec{x})$ covers all of the $2|\det M| - 1$ dimensional sphere S (which it certainly can do if n is large enough), and if the rows of A are given by a formula which only depends on $\vec{p}$'s values (i.e., if A is an "orthogonal design" [Geramita 79]), then rows of A determine $|\det M| - 1$ linearly independent vector fields on the S. But by [Adams 62], this can only occur when $2|\det M| - 1$ equals 3, 7, or 15.

On the other hand, there is a solution to the problem if it is not required that the rows of $A(\vec{x})$ follow a simple pattern.

Theorem 3.1.1 *For every smooth paraunitary power series vector $\vec{p}$, there is parau-*

nitary power series matrix $A(\vec{x})$ with $\vec{p}$ as its first row.

PROOF: First, find an invertible power series matrix $B(\vec{x})$ with $\vec{p}$ as its first row. [Madych 94] shows how to do this when $\vec{p}(\vec{x})$ is smooth; in section 3.1.3, we will show how to do this when ϕ is compact. Next, apply Gram-Schmidt to B. This preserves the first row, and turns B into a paraunitary matrix. If we interpret all of the divisions and square roots involved in the Gram-Schmidt process as their formal power series analogues, the answer $A(\vec{x})$ will be a paraunitary power series matrix. $\square$

The main problem with this construction is that $A(\vec{x})$ will not be polynomial even when $\vec{p}$ is. To put it another way, the wavelets it constructs will not be compactly supported even when ϕ is. The following three subsections solve this still open problem in the following special cases: when $\vec{p}$ is one dimensional, when the number of monomials in $\vec{p}$ is less than $|\det M|$, and when a certain quadratic form can be diagonalized.

3.1.1 The Method of Kautsky and Turcajová

In [Kautsky 95], Kautsky and Turcajová give a method for obtaining compactly supported wavelets from a compactly supported scaling function in the one dimensional case.[1] This section describes their approach.

Lemma 3.1.2 *Every paraunitary matrix $A(\vec{x})$ of the form $B + Cx$ is equal to $(I - P + Px)H$ with H unitary and P a symmetric projection matrix satisfying $P^* = P$ and $P^2 = P$.*

Proof: Let $H = A(1) = B + C$. Because $A(\vec{x})$ is paraunitary, H must be unitary. The shifted orthogonality conditions imply that $BB^* + CC^* = (H - C)(H^* - C^*) + CC^* = I - CH^* - HC^* + 2CC^* = I$ and that $BC^* = (H - C)C^* = HC^* - CC^* = 0$. Together, this means that $HC^* = CC^* = CH^*$, which means that CH^* must be a symmetric projection matrix, and $B + Cx = ((I - P) + Px)H$ for $P = CH^*$.

Lemma 3.1.3 *Every paraunitary polynomial matrix $A(x)$ of one variable x in degree d may be factored as*

$$A(\vec{x}) = (I - P_1 + P_1 x)(I - P_2 + P_2 x) \dots (I - P_d + P_d x)H$$

with H unitary and P_i a symmetric projection matrix.

Proof: Let $A(\vec{x}) = A_0 + A_1 x + \dots A_n x^d$. By the shifted orthogonality conditions we know that $A_0 A_d^* = 0$. Let P be a symmetric projection matrix such that $PA_0 = 0$ and $(I - P)A_d = 0$. For example, let P project onto A_d's column space. Then

$$\begin{aligned} (Px^{-1} + (I - P))A(\vec{x}) &= PA_0 x^{-1} + (PA_1 + (I - P)A_1) + \dots \\ &\quad + (PA_d + (I - P))x^{d-1} + (I - P)A_n \\ &= (PA_1 + (I - P)A_1) + \dots + (PA_d + (I - P))x^{d-1}. \end{aligned}$$

[1] [Lawton 96] gives a different algorithm, also based on a factorization of $A(\vec{x})$, for solving this problem.

But now we have a paraunitary polynomial matrix of degree $d-1$; it is paraunitary because the product of two paraunitary matrices is paraunitary. Applying the procedure iteratively until $d=1$ we are left with degree 0 paraunitary matrix H which must be unitary. Since then

$$H = (P_d x^{-1} + (I - P_d)) \ldots (P_1 x^{-1} + (I - P_1)) A(\vec{x})$$

and since $(I - P_d + P_d x)(P_d x^{-1} + (I - P_d)) = I$, we get

$$(I - P_1 + P_1 x)(I - P_2 + P_2 x) \ldots (I - P_d + P_d x) H = A(\vec{x})$$

as desired. $\square$

Corollary 3.1.4 *For every paraunitary vector polynomial $\vec{p}(x)$ of one variable, there is a paraunitary matrix polynomial $A(x)$ with $\vec{p}(x)$ as its first row.*

Let $\vec{p}(x) = \vec{p}_0 + \vec{p}_1 x + \ldots + \vec{p}_d x^d$. Because $\vec{p}$ is paraunitary, $\vec{p}_0 \vec{p}_d^* = 0$. Thus we can find a symmetric projection matrix P such that $P\vec{p}_0 = 0$ and $(I-P)\vec{p}_d = 0$. For example, let P project onto $\vec{p}_d$. Then $(Px^{-1} + (I-P))\vec{p}(x)$ will be a paraunitary vector polynomial of degree $d-1$, and we may iterate. When we are done, we will have a vector $\vec{\tilde{p}}$ such that $|\vec{\tilde{p}}| = 1$. Choose a unitary matrix H with $\vec{\tilde{p}}$ as its first row; this can always be done. But then

$$(I - P_1 + P_1 x)(I - P_2 + P_2 x) \ldots (I - P_d + P_d x) H$$

is a paraunitary matrix polynomial with $\vec{p}$ as its first row. $\square$

Unfortunately, this construction does not extend to higher dimensions because not every $A(\vec{x})$ is factorizable in this way. For example, $A + Bx + Cy + Dxy$ where

$$A = \frac{1}{6}\begin{bmatrix} 2+\sqrt{2} & -1-\sqrt{2} \\ -2-\sqrt{2} & 1+\sqrt{2} \end{bmatrix}, \; B = \frac{1}{6}\begin{bmatrix} 2-\sqrt{2}-a & \frac{-1-\sqrt{2}-a}{3+2\sqrt{2}} \\ \frac{2+\sqrt{2}+a}{3+2\sqrt{2}} & 1-\sqrt{2}+a \end{bmatrix}$$

$$C = \frac{1}{6}\left[\begin{array}{cc} 1-\sqrt{2}+a & \frac{-2-\sqrt{2}+a}{3+2\sqrt{2}} \\ \frac{1+\sqrt{2}+a}{3+2\sqrt{2}} & 2-\sqrt{2}-a \end{array}\right],\ D = \frac{1}{6}\left[\begin{array}{cc} 1+\sqrt{2} & 2+\sqrt{2} \\ 1+\sqrt{2} & 2+\sqrt{2} \end{array}\right]$$

is paraunitary for all a but not a product of degree 0 and 1 paraunitary matrix polynomials unless $a = 0$ or $a = 1$. This can be verified using the following lemma.

Lemma 3.1.5 (The Projection Lemma) *If a wavelet system $A_{\vec{x}}$ satisfies the shifted orthogonality condition, then so does $B_{\vec{x}}$, where*

$$B_{\vec{x}} = \sum_{\vec{x}=R\vec{y}} A_{\vec{y}}$$

PROOF:

$$\begin{aligned}
\sum_{\vec{l}\in\Gamma} B_{\vec{l}} B^*_{\vec{k}+\vec{l}} &= \sum_{\vec{l}\in\Gamma} \sum_{\vec{l}=R\vec{x}} \sum_{\vec{k}+\vec{l}=R\vec{y}} A_{\vec{x}} A^*_{\vec{y}} \\
&= \sum_{\vec{l}\in\Gamma} \sum_{\vec{l}=R\vec{x}} \sum_{\vec{k}=R\vec{w}=R(\vec{y}-\vec{x})} A_{\vec{x}} A^*_{\vec{x}+\vec{w}} \\
&= \sum_{\vec{k}=R\vec{w}} \sum_{\vec{l}\in\Gamma} \sum_{\vec{l}=R\vec{x}} A_{\vec{x}} A^*_{\vec{x}+\vec{w}} \\
&= \sum_{\vec{k}=R\vec{w}} \sum_{\vec{x}} A_{\vec{x}} A^*_{\vec{x}+\vec{w}} \\
&= \sum_{\vec{k}=R\vec{w}} \delta_{\vec{0},\vec{w}} I \\
&= \delta_{\vec{0},\vec{k}} I
\end{aligned}$$

□

3.1.2 The Rotation Method

Consider the matrix which rotates $\vec{e}_1$ onto $\vec{p}(\vec{x})$ and leaves everything orthogonal to the plane containing these vectors unchanged. The transpose of this matrix will have first row $\vec{p}$ and will be paraunitary; it is thus a good candidate for $A(\vec{x})$.

The matrix which performs a rotation of θ from $\vec{c}$ to $\vec{d}$ while leaving everything orthogonal to that plane unchanged is given by

$$I - (1-\cos\theta)(\vec{c}\vec{c}^T + \vec{d}\vec{d}^T) + \sin\theta(\vec{d}\vec{c}^T - \vec{c}\vec{d}^T)$$

under the assumption that $\vec{c}$ and $\vec{d}$ are orthonormal. To rotate $\vec{e}_0$ onto $\vec{p}(\vec{x})$ (both of which have unit length), set $\vec{c} = \vec{e}_1$, $\vec{d} = \frac{\vec{p}(\vec{x}) - <\vec{p}(\vec{x}),\vec{e}_1>\vec{e}_1}{||\vec{p}(\vec{x}) - <\vec{p}(\vec{x}),\vec{e}_1>\vec{e}_1||}$, $\cos\theta = <\vec{p}(\vec{x}), \vec{e}_1>$, and $\sin\theta = ||\vec{p}(\vec{x}) - <\vec{p}(\vec{x}), \vec{e}_1> \vec{e}_1||$. The matrix becomes

$$\begin{aligned} I \quad &- \quad (1- <\vec{p}(\vec{x}), \vec{e}_1>)(\vec{e}_1\vec{e}_1^T + \frac{(\vec{p}(\vec{x}) - <\vec{p}(\vec{x}), \vec{e}_1> \vec{e}_1)(\vec{p}(\vec{x})^T - <\vec{e}_1, \vec{p}(\vec{x})> \vec{e}_1^T)}{||\vec{p}(\vec{x}) - <\vec{p}(\vec{x}), \vec{e}_1> \vec{e}_1||^2}) \\ &+ \quad (\vec{p}(\vec{x}) - <\vec{p}(\vec{x}), \vec{e}_1> \vec{e}_1)\vec{e}_1^T - \vec{e}_1(\vec{p}(\vec{x})^T - <\vec{e}_1, \vec{p}(\vec{x})> \vec{e}_1^T). \end{aligned}$$

Notice that this expression contains no square roots, and the only division is by $||\vec{p}(\vec{x}) - <\vec{p}(\vec{x}), \vec{e}_1> \vec{e}_1||^2 = 1 - |<\vec{p}(\vec{x}), \vec{e}_1>|^2$.

Theorem 3.1.6 *If $\vec{p}(\vec{x})$ contains fewer than $|\det M|$ monomials, then there exists a paraunitary polynomial matrix with $\vec{p}(\vec{x})$ as its first row.*

Proof: Since $\vec{p}(\vec{x})$ contains fewer than $|\det M|$ monomials, there is some unit vector $\vec{v}$ such that $<\vec{p}(\vec{x}), \vec{v}> = 0$. With out loss of generality, let $\vec{v} = \vec{e}_1$ and let R be the matrix constructed above which rotates $\vec{e}_1$ onto $\vec{p}$. Then R^t is a paraunitary polynomial matrix with $\vec{p}$ as its first row. $\square$

Theorem 3.1.7 *If $2|\det M| > n+1$, then the rotation method produces $|\det M| - 1$ wavelets from every compactly supported scaling function.*

Proof: Consider $\vec{p}(\vec{x})$ as a mapping from the torus $T = \{\vec{x} \in \mathbf{C}^n : |x_i| = 1, i = 1, \ldots, n\}$ to the sphere $S = \{\vec{y} \in \mathbf{C}^m : \vec{y} \cdot \overline{\vec{y}} = 1\}$. Since ϕ is compactly supported, $\vec{p}$

is polynomial and hence Lipschitz. T is n (real) dimensional and S is $2|\det M| - 1$ dimensional, so by Sard's theorem, the set $\vec{p}(T)$ is not dense in S. Thus we can find a point $\vec{y} \in S$ such that the distance between $\vec{y}$ and $\vec{p}(T)$ is greater than some $\epsilon > 0$. Without loss of generality, let $\vec{y} = \vec{e}_1$. Then $| < \vec{p}(\vec{x}), \vec{e}_1 > | \leq \sqrt{1-\epsilon^2} < 1$, and the matrix which rotates $\vec{e}_1$ onto $\vec{p}(\vec{x})$ has rows which are bounded by $\frac{C}{1-|<\vec{p}(\vec{x}),\vec{e}_1>|^2} \leq \frac{C}{\epsilon}$.

$\square$

3.1.3 The ABC Method

In this section, we show how to find compactly supported wavelets from a compactly support scaling function, assuming a certain quadratic form can be diagonalized. The proof relies on the following version of the famous Quillen-Suslin theorem.

Theorem 3.1.8 *If $A(\vec{x})$ is an r by m ($r < m$) Laurent polynomial matrix which is paraunitary by rows, then there exists an invertible Laurent polynomial matrix $B(\vec{x})$ with $A(\vec{x})$ as its first r rows.*

Proof: [Rotman 79] Let P be the ring of Laurent polynomials $\mathbf{C}[x_1, x_1^{-1}, \ldots x_n, x_n^{-1}]$. Define $f : P^m \to P^r$ by $f(q_1, \ldots q_m) = [q_1 \ \ldots \ q_m]A^*$. Since $f(A) = I$, we have the exact sequence

$$0 \to \ker f \to P^m \xrightarrow{f} P^r \to 0$$

P^r is projective, so the sequence splits and we have $P^m = \ker f \oplus < A_1, \ldots, A_r >$ where $A_1, \ldots, A_r$ are the rows of A. Also, $\ker f$ is a finitely generated projective P-module, so by the Quillen-Suslin theorem, it is free and has a basis $\vec{p}_{r+1}, \ldots, \vec{p}_m$. Let $B(\vec{x})$ be the matrix with $A(\vec{x})$ as its first r rows and $\vec{p}_i$ as its i-th row, $i > r$. Then B is a basis for P^m, and thus invertible, and we are done. □

Fitchas and Galligo [Fitchas 89] and Logar and Sturmfels [Logar 92] give constructive proofs of the above theorem using Grobner bases. In particular, Fitchas and Galligo's algorithm gives a bound of $O((md)^{n^2})$ on the degrees of the polynomials in B when the polynomials in A have degree less than d.

We should note that finding an invertible Laurent polynomial matrix with a given top row is equivalent to the problem of finding compactly supported *biorthogonal* wavelets given a compactly supported scaling function; see [Karoui 94] or [Strang 95] for more details about biorthogonal wavelets. Once biorthogonal wavelets are constructed, it is easy to make "pre-wavelets" [Riemenschneider 92], functions which span W_0 with their Γ translates, but aren't mutually orthogonal. The proof of our next theorem starts with this construction, in fact.

Theorem 3.1.9 *Each of the following statements are implied by the ones under it.*

1. *For every paraunitary Laurent polynomial vector $\vec{p}(\vec{x})$, there is a paraunitary Laurent polynomial matrix $A(\vec{x})$ with $\vec{p}$ as its first row.*

2. *For every Laurent polynomial matrix $B(\vec{x})$ such that $B(\vec{x})B(\vec{x})^*$ has determinant 1, there is a Laurent polynomial row vector $\vec{v}(\vec{x})$ such that $\vec{v}B$ is paraunitary (i.e., $\vec{v}BB^*\vec{v}^* = 1$).*

3. *For every Laurent polynomial matrix $B(\vec{x})$ such that $B(\vec{x})B(\vec{x})^*$ has determinant 1, there is an invertible Laurent polynomial C such that CBB^*C^* is diagonal.*

Proof: (2) $\Longrightarrow$ (1): Our strategy is to build $A(\vec{x})$ one row at a time. We start with $s = 1$ rows of a paraunitary by rows matrix $A(\vec{x})$ already built.

By theorem 3.1.8, there is an invertible Laurent polynomial matrix $B(\vec{x})$ with $A(\vec{x})$ as its first s rows. Let B_i be the i-th row of B, considered as a row vector. We can assume that $B_iA^* = \vec{0}^t$ for $i > r$, otherwise just set $B_i' = B_i - B_iA^*A$. [Remember that $AA^* = I$.]

Since $B(\vec{x})$ is invertible, BB^* has determinant 1. If we let $C(\vec{x})$ be the $s+1$ through m rows of $B(\vec{x})$, then $\det(BB^*) = \det(AA^*)\det(CC^*) = \det(CC^*) = 1$. Now apply part (2) to C to get a paraunitary Laurent polynomial vector $\vec{v}C$. This vector is paraunitary, Laurent polynomial, and orthogonal to all the rows of $A(\vec{x})$, so we may add as the $s+1$-th row of $A(\vec{x})$, and begin again. Stop when $s = m-1$, because given a $m-1$ by m paraunitary matrix $A(\vec{x})$, it may be easily extended to a m by m matrix by letting the last row be the exterior product of the first $m-1$.

(3) $\Longrightarrow$ (2): C is invertible, so its determinant must be a monomial. In particular, the determinant of CBB^*C^* must be a positive real constant. But CBB^*C^* is a diagonal matrix, so its determinant is the product of its diagonal entries. On the other hand, the only way a product of Laurent polynomials can be a constant is if each one is a monomial and hence invertible. Let α be the first entry of CBB^*C^*; then $\frac{\vec{e}_1}{\sqrt{\alpha}}CBB^*C^*\frac{\vec{e}_1}{\sqrt{\alpha}}^* = 1$ and $\vec{v} = \frac{\vec{e}_1}{\sqrt{\alpha}}CB$ is paraunitary. $\square$

Now, consider the following algorithm

The ABC Algorithm

Input: A k by m Laurent polynomial matrix $B(\vec{x})$ such that $\det B(\vec{x})B(\vec{x})^* = 1$.
Output: A Laurent polynomial matrix $B(\vec{x})'$.

1. Set r to 1.

2. If $r = k$, then terminate with output $B(\vec{x})' = B(\vec{x})$.

3. If head $|B_r|^2 >$ head $|B_{r+1}|^2$, then switch B_r and B_{r+1}, set r equal to $\max(1, r - 1)$, and goto Step 3.

4. For s from 1 to k, $s \neq r$, and while head $< B_s, B_r >$ is divisible by head $|B_r|^2$, replace B_s with $B_s - \alpha B_r$ where $\alpha = \frac{\text{head} < B_s, B_r >}{\text{head} \, |B_r|^2}$.

5. If head $|B_r|^2 \leq$ head $|B_{r+1}|^2$, then set r equal to $r + 1$.

6. Go to Step 2.

Theorem 3.1.10 *The ABC algorithm always terminates in a finite number of steps, and its output has the following properties:*

1. *There exists an invertible Laurent polynomial matrix $C(\vec{x})$ such that $B' = CB$.*

2. $\det B(\vec{x})'(B(\vec{x})')^* = 1$, *and*

3. head $|B_i'|^2$ *does not divide* head $< B_j', B_i' >$ *for all $i, j = 1 \ldots k$, where B_i' is the i-th row of B'.*

Proof: First, note that the matrix B is only changed in steps 3 and 4, and that both of these changes (switching two rows, or subtracting a multiple of a row from another row) are easily inverted. Second, the existence of an invertible C such that $B' = BC$ implies $\det B(\vec{x})'(B(\vec{x})')^* = 1$. Third, the only time the algorithm repeats a step with the same r value is when one of the head $|B_i|$'s is reduced, and since this can only happen a finite number of times the algorithm must always terminate.

To prove the last claim, notice that at all steps of the algorithm, the following two properties hold:

- head $|B_1|^2 \leq$ head $|B_2|^2 \leq \ldots \leq$ head $|B_r|^2$, and

- head $|B_i'|^2$ does not divide head $< B_j', B_i' >$, for all $i < r$.

These properties are true initially, and the only possible place they could be broken is at steps 3 or 4. Step 3 is safe because it only switches rows of B and decrements r when it does. Step 4 doesn't break the first property because

$$\text{head}\,|B_s - \frac{\text{head} < B_s, B_r >}{\text{head}\,|B_r|^2} B_r|^2 = \text{head}\,|B_s|^2 - \frac{\text{head} < B_s, B_r >}{\text{head}\,|B_r|^2}\,\text{head} < B_r, B_s >$$

$$-\frac{\text{head}\,|B_r|^2}{\text{head} < B_s, B_r >}\,\text{head} < B_s, B_r > + \frac{\text{head} < B_s, B_r > \text{head}\,|B_r|^2}{\text{head}\,|B_r|^2\,\text{head} < B_s, B_r >}\,\text{head}\,|B_r|^2$$

$$= \text{head}\,|B_s|^2 - \frac{\text{head} < B_s, B_r >}{\text{head}\,|B_r|^2}\,\text{head} < B_r, B_s >$$

which is less than or equal to head $|B_s|^2$ by Cauchy-Schwartz. It doesn't break the second property because

$$\text{head} < B_k, B_s - \frac{\text{head} < B_s, B_r >}{\text{head}\,|B_r|^2} B_r >$$

$$= \text{head} < B_k, B_s > - \frac{\text{head}\,|B_r|^2}{\text{head} < B_s, B_r >}\,\text{head} < B_k, B_r >$$

and for $k, s < r$, this always is equal to head $< B_k, B_s >$ by a similar argument. $\square$

Corollary 3.1.11 *For every paraunitary Laurent polynomial vector $\vec{p}(x)$ of one variable, there is a paraunitary Laurent polynomial matrix $A(x)$ with $\vec{p}$ as its first row.*

Proof: Take any Laurent polynomial matrix $B(x)$ such that $\det B(x)B(x)^* = 1$, and run the ABC algorithm on it. Since the entries of the output B' will all be Laurent polynomials of one variable, head $|B_i'|^2$ does not divide head $< B_j', B_i' >$ implies that (head $< B_j', B_i' >$) $<$ head $|B_i'|^2$ in the monomial ordering. Now consider the equation $\det B(x)'(B(x)')^* = 1$. Because $\det B(x)'(B(x)')^* = |B_1'|^2 \dots |B_k'|^2 + E(x)$, where E is the sum of terms, each of which has head monomial less than head $|B_1'|^2 \dots |B_k'|^2$, this implies that head $|B_1'|^2 \dots |B_k'|^2 = 1$. But this implies that $|B_1'|^2 \dots |B_k'|^2 = 1$, which in turn implies that each $|B_i'|^2$ is a constant. This further implies that $B'(B')^*$ is a diagonal matrix, which by Theorem 3.1.9, implies the existence of $A(x)$. $\square$

3.1.4 Unitary Coordinates

In this section, we discuss a method of parameterizing wavelet systems satisfying the shifted orthogonality condition.

Recall that wavelet systems satisfying the shifted orthogonality condition are in an one-to-one correspondence with paraunitary matrices $A(\vec{x})$. Our parameters, the unitary coordinates, will essentially just be the discrete fourier transform of $A(\vec{x})$.

Specifically, let

$$A(\vec{x}) = \sum_{i_1=0}^{s_1-1} \sum_{i_2=0}^{s_2-1} \cdots \sum_{i_n=0}^{s_n-1} A_{i_1,i_2,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$$

and let

$$\vec{\zeta} = (e^{\frac{2\pi i}{s_1}}, e^{\frac{2\pi i}{s_2}}, \ldots e^{\frac{2\pi i}{s_n}}).$$

Then we define the *unitary coordinate* $U_{\vec{k}}$ to be $A(\vec{\zeta}^{\vec{k}})$. $U_{k_1,k_2,\ldots k_n}$ is unitary because $A(\vec{x})$ is paraunitary and $|\vec{\zeta}^{\vec{k}}| = 1$.

Conversely, given unitary coordinates, we can reconstruct the wavelet system with the formula

$$A_{\vec{j}} = \frac{1}{s_1 s_2 \ldots s_n} \sum_{\vec{k}\in S} U_{\vec{k}}$$

where $S = [0, s_1) \times [0, s_2) \times \cdots \times [0, s_n)$.

Unfortunately, the $U_{k_1,k_2,\ldots,k_n}$ are not independent. The shifted orthogonality condition $\sum A_{\vec{j}} A^*_{\vec{j}+\vec{k}} = \delta_{\vec{0},\vec{k}} I$ implies that

$$\begin{aligned}
\sum_{\vec{j}\in S, \vec{j}+\vec{l}\in S} \left(\sum_{\vec{k}\in S} \vec{\zeta}^{-\vec{j}\vec{k}} U_{\vec{k}} \right) \left(\sum_{\vec{k'}\in S} \vec{\zeta}^{(\vec{j}+\vec{l})\vec{k'}} U^*_{\vec{k}} \right) &= |S|^2 \delta_{\vec{0},\vec{l}} \\
\sum_{\vec{k}\in S} \sum_{\vec{k'}\in S} U_{\vec{k}} U^*_{\vec{k'}} \sum_{\vec{j}\in S, \vec{j}+\vec{l}\in S} \vec{\zeta}^{\vec{j}\vec{k'}+\vec{l}\vec{k'}-\vec{j}\vec{k}} &= |S|^2 \delta_{\vec{0},\vec{l}}.
\end{aligned}$$

This equation can be simplified in two ways. First, if $\vec{l} = \vec{0}$, the inner sum becomes $\sum_{\vec{j}\in S} (\vec{\zeta}^{\vec{k'}-\vec{k}})^{\vec{j}}$ which equals $|S| \delta_{\vec{k'}-\vec{k},\vec{0}}$. Also, if $\vec{k} = \vec{k'}$, $U_{\vec{k}} U^*_{\vec{k'}} = I$ and the inner sum becomes $\sum_{\vec{j}\in S, \vec{j}+\vec{l}\in S} \vec{\zeta}^{\vec{l}\vec{k'}} = 0$.

We are left with

$$\sum_{\vec{k}\in S}\sum_{\vec{k}'\neq\vec{k}\in S} U_{\vec{k}} U^*_{\vec{k}'} \sum_{\vec{j}\in S, \vec{j}+\vec{l}\in S} \vec{\zeta}^{(\vec{j}+\vec{l})\vec{k}'-\vec{j}\vec{k}} = 0$$

The set of unitary matrices satisfying this condition for all $\vec{l}$ is on a one-to-one correspondence with the set of orthonormal wavelet systems.

3.2 Constructing Haar Tiles

This section investigates the simplest class of multidimensional wavelets, those generated by self-affine tilings. These are the natural generalizations of the Haar scaling function $\chi_{[0,1]}$.

A Γ-*tiling of* $\mathbf{R}^n$ is a closed subset $\mathcal{T}$ (called the *tile*) of $\mathbf{R}^n$ such that

- $\cup_{\vec{\gamma}\in\Gamma}\mathcal{T}+\vec{j}=\mathbf{R}^n$, and
- $(\text{int } T+\vec{\alpha})\cap(\text{int } \mathcal{T}+\vec{\beta})=\emptyset$ for $\vec{\alpha},\vec{\beta}\in\Gamma$, $\vec{\alpha}\neq\vec{\beta}$

where $\text{int}\,\mathcal{T}$ is the interior of $\mathcal{T}$.

A tile is said to be *self-affine* if there is an expanding matrix M such that $M\mathcal{T}=\cup_{\vec{k}\in K}\mathcal{T}+\vec{k}$. K is said to be $\mathcal{T}$'s *digit set*. The fact that $\text{vol}(M\mathcal{T})=\text{vol}(\cup_{\vec{k}\in K}\mathcal{T}+\vec{k})=\sum_{\vec{k}\in K}\text{vol}(\mathcal{T})$ implies that $|K|=|\det M|$.

Given a self-affine Γ-tiling of $\mathbf{R}^n$, it is easy to see how it generates a scaling function. Just set $\phi=\chi_{\mathcal{T}}$, the characteristic function of the tile $\mathcal{T}$, then the self-affineness of $\mathcal{T}$ implies that ϕ satisfies the dilation equation $\phi=\sum_{\vec{a}\in A}1\phi(M\vec{x}-\vec{a})$. The fact that $\mathcal{T}$ creates a Γ-tiling guarantees that ϕ is orthogonal to its Γ translates and that the multiresolution analysis so generated is complete.

Conversely, given a multiresolution analysis where $\phi=c\chi_Q$ for some set Q, then $\sum_{\vec{j}\in\Gamma}|c|^2|Q|\chi_Q(\vec{x}-\vec{j})=1$ and $Q\cup(Q+\vec{\gamma})$ has measure 0. Also, $\chi_Q=\sum_{\vec{k}\in K}\chi_Q(M\vec{x}-\vec{k})$, so $MQ=\cup_{\vec{k}\in K}Q+\vec{k}$. Thus, Q is a self-affine Γ-tiling. Our previous results on multiresolution approximations imply that $\text{vol}(Q)=\text{vol}(\mathbf{R}^n/\Gamma)$ and $c=(\text{vol}(\mathbf{R}^n/\Gamma))^{-1/2}$.

Given this equivalence between self-affine tiles and multiresolution approximations where the scaling function is the characteristic function of a set, we define a *Haar* multiresolution approximation to be one arising from a self-affine tile. Every Haar multiresolution approximation has degree $\vec{0}$ vanishing moments, and so satisfies the degree $\vec{0}$ sum rule

$$\sum_{\vec{k}\equiv\vec{k}'(M)} c_{\vec{k}}=C$$

Since $c_{\vec{k}}$ is a constant not depending on $\vec{k}$, this implies that each of the $\vec{k} \in K$ comes from a different coset of $\Gamma/M\Gamma$. Since $|K| = |\det M|$, K in fact forms a full set of representatives of $\Gamma/M\Gamma$.

Not every full set of cosets of $\Gamma/M\Gamma$ generates a self-affine tile, however. In fact, it is unknown whether for every expanding matrix M there is a set K such that K is the digit set for a self-affine tile [Lagarias 93]. One way to check whether a particular set K generates a self-affine tile is to use the Cascade algorithm.

3.3 Constructing Smooth Scaling Functions

In this section, we show how to construct compactly supported wavelet systems satisfying the sum rule

$$\sum_{\vec{k} \equiv \vec{k}'(M)} c_{\vec{k}} \vec{k}^{\vec{p}} = C_{\vec{p}}$$

where $C_{\vec{p}}$ is a constant not dependent on $\vec{k}'$. As shown in section 2.6.4, this sum rule implies degree $\vec{p}$ vanishing moments.

Given a wavelet system's polyphase matrix $A(\vec{x})$, we can write the sum rule as

$$\vec{e}_0[A(\vec{x})D_{\vec{p}}(\vec{x})](\vec{0}) = C_{\vec{p}}\vec{1}$$

where $D_{\vec{p}}(\vec{x}) = \sum_{\vec{i}} D_{\vec{p},-\vec{i}} \vec{x}^{\vec{i}}$ and $D_{\vec{p},\vec{i}}$ is the $|\det M|$ by $|\det M|$ diagonal matrix with (j,j) entry $(M\vec{i}+\vec{y}_j)^{\vec{p}}$ for the same fixed ordering of coset representatives $\vec{y}_0, \dots, \vec{y}_{|\det M|-1}$ of $\Gamma/M\Gamma$ used to define $A(\vec{x})$. For example, $D_{\vec{0}}(\vec{x}) = \sum_{\vec{i}} I\vec{x}^{\vec{i}}$. One advantage of trying to solve the sum rules in this form is that it explicitly brings the wavelets into the problem; as we saw in the last section, it is not currently known how to get compactly supported wavelets from a compactly supported scaling function.

First we consider the case where $A(\vec{x}) = B + C\vec{x}^{\vec{i}}$. By the results in section 3.1.1, we know that we may write $A(\vec{x}) = H(I - P + P\vec{x}^{\vec{i}})$ with H unitary and P a symmetric projection matrix. The degree $\vec{0}$ sum rules imply that

$$\vec{e}_0[H(I - P + P\vec{x}^{\vec{x}})(I + I\vec{x}^{-\vec{x}})](\vec{0}) = C_{\vec{0}}\vec{1}$$

or

$$\vec{e}_0 H = C_{\vec{0}}\vec{1}.$$

Because H is unitary, we see that $C_{\vec{0}} = \frac{1}{\sqrt{|\det M|}}$ and that H should be chosen to be the wavelet system associated with a Haar tile.

The degree $\vec{e}_j$ sum rules imply

$$\begin{aligned}
\vec{e}_0[H(I-P+P\vec{x}^{\vec{i}})(D_{\vec{e}_j,\vec{0}}+D_{\vec{e}_j,\vec{i}}\vec{x}^{\vec{i}})](\vec{0}) &= C_{\vec{e}_j}\vec{1} \\
\vec{1}[(I-P)D_{\vec{e}_j,\vec{0}}+PD_{\vec{e}_j,\vec{i}}] &= C'_{\vec{e}_j}\vec{1} \\
\vec{1}P(D_{\vec{e}_j,\vec{i}}-D_{\vec{e}_j,\vec{0}}) &= \vec{1}(C'_{\vec{e}_j}I-D_{\vec{e}_j,\vec{0}}).
\end{aligned}$$

The matrix $D_{\vec{e}_j,\vec{i}}-D_{\vec{e}_j,\vec{0}}$ is diagonal with (k,k) entry $(M\vec{i}+\vec{y}_k)^{\vec{e}_j}-\vec{y}_k^{\vec{e}_j}=M_j\cdot\vec{i}$ where M_j is the $\vec{j}$-th row of M. Thus,

$$(M_j\cdot\vec{i})\vec{1}P=(C'_{\vec{e}_j}\vec{1}-\vec{1}D_{\vec{e}_j,\vec{0}}).$$

P is a projection matrix, so this equation is solvable if and only if $(C'_{\vec{e}_j}\vec{1}-\vec{1}D_{\vec{e}_j,\vec{0}})$ and $(C'_{\vec{e}_j}\vec{1}-\vec{1}D_{\vec{e}_j,\vec{0}})-(M_j\cdot\vec{i})\vec{1}$ are orthogonal. This implies

$$\begin{aligned}
C'_{\vec{e}_j}\vec{1}\cdot C'_{\vec{e}_j}\vec{1}-2C'_{\vec{e}_j}\vec{1}\cdot\vec{1}D_{\vec{e}_j,\vec{0}}+\vec{1}D_{\vec{e}_j,\vec{0}}\cdot\vec{1}D_{\vec{e}_j,\vec{0}}-(M_j\cdot\vec{i})(C'_{\vec{e}_j}\vec{1}\cdot\vec{1}+\vec{1}D_{\vec{e}_j,\vec{0}}\cdot\vec{1}) &= 0 \\
|\det M|(C'_{\vec{e}_j})^2-2\operatorname{tr}D_{\vec{e}_j,\vec{0}}C'_{\vec{e}_j}-|\det M|(M_j\cdot\vec{i})C'_{\vec{e}_j}+\operatorname{tr}(D_{\vec{e}_j,\vec{0}})^2+(M_j\cdot\vec{i})\operatorname{tr}D_{\vec{e}_j,\vec{0}} &= 0
\end{aligned}$$

which is a quadratic in $C'_{\vec{e}_j}$. It has a real solution when

$$4(\operatorname{tr}D_{\vec{e}_j,\vec{0}})^2+|\det M|^2(M_j\cdot\vec{i})^2-4|\det M|\operatorname{tr}(D_{\vec{e}_j,\vec{0}})^2\geq 0$$

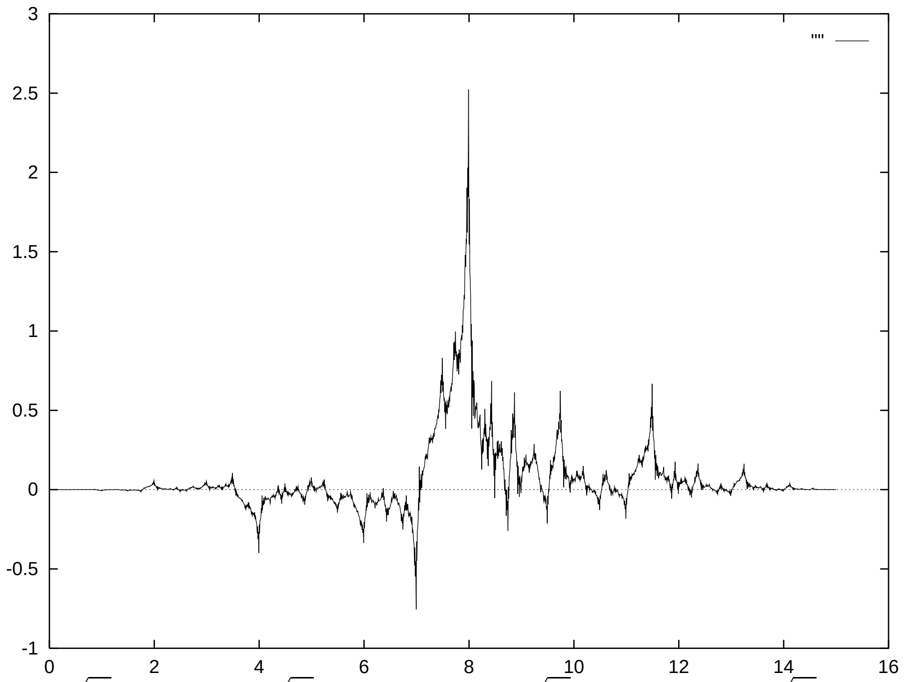
which, assuming that $M_j\cdot\vec{i}\neq 0$, can easily be satisfied by choosing $\vec{i}$ to be large. We have thus proven

Theorem 3.3.1 *The wavelet system $A(\vec{x})=H(I-P+P\vec{x}^{\vec{i}})$ can be made to have degree $\vec{0}$ and $\vec{e}_j$ vanishing moments whenever*

$$4(\sum_{\vec{y}_k\in\Gamma/M\Gamma}\vec{y}_k^{\vec{e}_j})^2+|\det M|^2(M_j\cdot\vec{i})^2\geq 4|\det M|\sum_{\vec{y}_k\in\Gamma/M\Gamma}\vec{y}_k^{2\vec{e}_j}.$$

3.3.1 Examples

1. $n=1$, $y_k=k$, $M>1$. Then $M_j=M=|\det M|$, and to have degree 1



$\phi(x) = \frac{1-\sqrt{15}}{16}\phi(2x) + \frac{15-\sqrt{15}}{16}\phi(2x-7) + \frac{15+\sqrt{15}}{16}\phi(2x-8) + \frac{1+\sqrt{15}}{16}\phi(2x-15)$

Figure 3-1: A Rank 2 Scaling Function with One Vanishing Moment

vanishing moments i must satisfy

$$4(\frac{(M-1)M}{2})^2 + M^4 i^2 \geq 4M\frac{M(M-1)(2M-1)}{6}.$$

But this is the same as $i^2 \geq \frac{1}{3}$, which always holds for non-zero, integer i.

2. $n = 1$, $M = 2$, $y_0 = 0$, $y_1 = 7$. i must satisfy $4(49) + 4(2i)^2 \geq 8(49)$, which means that $|i| \geq 4$. Working out the case $i = 4$ gives $C_1' = \frac{15 \pm \sqrt{15}}{2}$ and the matrix $P = \frac{1}{16}\begin{bmatrix} 15 & \sqrt{15} \\ \sqrt{15} & 1 \end{bmatrix}$. Figure 3-1 shows the resulting scaling function.

3. $n = 2$, $M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, $y_0 = (0,0)$, $y_1 = (1,0)$. If we want degree $(1,0)$ vanishing moments, then $\vec{i} = (i_1, i_2)$ must satisfy $4 + 4(i_1+i_2)^2 \geq 8$ which is true whenever $i_1 + i_2 \neq 0$. For degree $(0,1)$ vanishing moments, then $\vec{i}$ must satisfy $4(-i_1+i_2)^2 \geq 0$ which is always true.

Note that under the these conditions, a wavelet system of the form $B + C\vec{x}^{\vec{i}}$

can't have both degree $(1,0)$ and $(0,1)$ vanishing moments at the same time, because that would imply $\vec{1}P = \frac{1}{i_1+i_2}(C_1, C_1 - 1) = \frac{1}{-i_1+i_2}(C_2, C_2)$ which is impossible.

4. $n = 2$, $M = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $y_0 = (0,0)$, $y_1 = (1,0)$, $y_2 = (0,1)$, and $y_3 = (1,1)$. If we want degree $(1,0)$ vanishing moments, then $\vec{i}$ must satisfy $16 + 16(2i_1)^2 \geq 32$ or $i_1 \neq 0$. Similarly for degree $(0,1)$ vanishing moments, we must have $i_2 \neq 0$. Again, it is impossible for both conditions to hold at once.

Bibliography

[Adams 62] Adams, J. F. "Vector Fields on Spheres." Ann. Math. 75 (1962), 603-622.

[Barnsley 88] Barnsley, Michael. *Fractals Everywhere.* San Diego: Academic Press, 1988.

[Berger 92] Berger, M. and Y. Wang. "Multidimensional Two-Scale Dilation Equations." *Wavelets — A Tutorial In Theory and Applications*, C.K. Chui, editor. Boston: Academic Press, 1992.

[Cabrelli 96] Cabrelli, Carlos, Christopher Heil and Ursula Molter. "Accuracy of Lattice Translates of Several Multidimensional Refinable Functions." ftp://ftp.math.gatech.edu/pub/users/heil/accmulti.ps.Z. Preprint, 8/6/96.

[Cohen 96] Cohen, Albert, Karlheinz Grochenig and Lars Villemoes. "Regularity of Multivariate Refinable Functions." http://lynx.ann.jussieu.fr/~cohen/multivariate.tex. Preprint, 1996.

[Daubechies 92a] Daubechies, Ingrid. *Ten Lectures on Wavelets.* CBMS-NSF Regional Conference Series in Applied Mathematics, # 61. Philadelphia: SIAM, 1992.

[Daubechies 92b] Daubechies, Ingrid and J. C. Lagarias. "Two-scale difference equations II: Local regularity, infinite products of matrices and fractals." SIAM J. Math. Analysis 23 (1992), 1031-1079.

[Fitchas 89] Fitchas, N. and A. Galligo. "Le Nullstellensatz effectif et la conjecture de Serre en calcul formel." *Seminaire Structures Algebriques Ordonnees*, UER de Math. Universite Paris VII, 1989.

[Geramita 79] Geramita, Anthony and Jennifer Seberry. *Orthogonal Designs: Quadratic Forms and Hadamard Matrices.* New York: M. Dekker, 1979.

[Grochenig 87] Grochenig, Karlheinz. "Analyse multiechelle et bases d'ondelettes." C.R. Acad. Sci. Paris Series I # 305 (1987), 13-17.

[Grochenig 92] Grochenig, Karlheinz and W. R. Madych. "Multiresolution Analysis, Haar Bases, and Self-Similar Tilings." IEEE Trans. Info. Th IT 38 # 2, Part II (1992), 556-568.

[Han 97] Han, Bin and Rong-Qing Jia. "Multivariate Refinement Equations and Subdivision Schemes." Preprint, 1997.

[Heil 92] Heil, Christopher. "Methods of solving dilation equations." *Probabilistic and Stochastic Methods in Analysis, with Applications*, J. S. Byrnes et al. editors. Kluwer Academic, 1992.

[Heller 94] Heller, Peter Niels. "Rank M Wavelets with N Vanishing Moments." SIAM J. Matrix Analysis and Applications 16 #2 (1995), 502 - 519.

[Jia 91] Jia, Rong-Qing and C. A. Micchelli. "Using the Refinement Equation for the Construction of Pre-Wavelets II: Powers of Two." *Curves and Surfaces*, P.J. Laurent, A. Le Mehaute and L. L. Schumaker, editors. Boston: Academic Press, 1991.

[Jia 94] Jia, Rong-Qing and Z. Shen. "Multiresolution and Wavelets." Proc. Edinburgh Math. Soc. 37 (1994), 271-300.

[Jia 97] Jia, Rong-Qing. "Characterization of Smoothness of Multivariate Refinable Functions in Sobolev Spaces." Preprint, 1997.

[Karoui 94] Karoui, Abderrazek and Rémi Vaillancourt. "Nonseparable Biorthogonal Wavelet Bases of $L^2(\mathbf{R}^n)$." Preprint, 1994.

[Kautsky 95] Kautsky, Jaroslav and Radka Turcajová. "A Matrix Approach to Discrete Wavelets." *Wavelets: Theory, Algorithms, and Applications*, C.K. Chui, editor. San Diego: Academic Press, 1994.

[Lagarias 93] Lagarias, J.C. and Y. Wang. "Haar Bases for $L^2(\mathbf{R}^n)$ and Algebraic Number Theory." Journal of Number Theory 57 (1996), 181-197.

[Lawton 96] Lawton, W., S. L. Lee and Zuowei Shen. "An Algorithm for Matrix Extension and Wavelet Construction." Mathematics of Computation 65, # 214 (1996), 723-737.

[Logar 92] Logar, Alessandro and Bernd Sturmfels. "Algorithms for the Quillen-Suslin Theorem." Journal of Algebra 145 (1992), 231-239.

[Madych 92] Madych, W.R. "Some Elementary Properties of Multiresolution Analyses of $L^2(\mathbf{R}^n)$." *Wavelets—A Tutorial in Theory and Applications*, C. K. Chui, editor. Boston: Academic Press, 1992.

[Madych 94] Madych, W.R. "Orthogonal Wavelet Bases for $L^2(\mathbf{R}^n)$." *Fourier Analysis: Analytic and Geometric Aspects*, William O. Bray, P.S. Milojevic, and Caslav V. Stanojevic, editors. Lecture Notes in Pure & Appl. Math, 157. New York: M. Dekker, 1994.

[Massopust 94] Massopust, Peter R. *Fractal Functions, Fractal Surfaces, and Wavelets.* San Diego: Academic Press, 1994.

[Meyer 92] Meyer, Yves. *Wavelets and Operators.* Cambridge Studies in Advanced Mathematics, # 37. New York: Cambridge University Press, 1992.

[Polya 71] Polya, George and Gabor Szego. *Aufgaben und Lehrsatze aus def Analysis.* Berlin: Springer, 1970.

[Riemenschneider 92] Riemenschneider, S.D. and Z. Shen. "Wavelets and Prewavelets in Low Dimensions." J. Approx. Theory 71 (1992), 18-38.

[Rotman 79] Rotman, J.J. *An Introduction to Homological Algebra.* New York: Academic Press, 1979.

[Stockler 92] Stockler, Joachim. "Multivariate Wavelets." *Wavelets — A Tutorial in Theory and Applications*, C. K. Chui, editor. Boston: Academic Press, 1992.

[Strang 73] Strang, Gilbert and G. Fix. "A Fourier Analysis of the Finite-Element Variational Method." *Constructive Aspects of Functional Analysis*, G. Geymonat, editor. Roma: C.I.M.E., 1973.

[Strang 94] Strang, Gilbert and Vasily Strela. "Orthogonal Multiwavelets with Vanishing Moments." J. Optical Engineering 33 (1994), 2104 - 2107.

[Strang 95] Strang, Gilbert and Truong Nguyen. *Wavelets and Filter Banks.* Wellesley, MA: Wellesley-Cambridge Press, 1996.

[Strang 97] Strang, Gilbert. "Eigenvalues of $(\downarrow 2)$H and Convergence of the Cascade Algorithm." IEEE Trans. Sig. Proc 44 (1996), 233 - 238.

[Strichartz 90] Strichartz, Robert S. "Self-Similar Measures and Their Fourier Transforms I." Indiana U. Math. J. 39 (1990), 797-817.

[Strichartz 91] Strichartz, Robert S. "Wavelets and Self-Affine Tilings." Constructive Approximation 9 (1993), 327-346.

[Strichartz 93] Strichartz, Robert S. "How to Make Wavelets." American Mathematics Monthly 100 (1993), 539-556.

[Strichartz 94] Strichartz, Robert S. "Construction of Orthonormal Wavelets." *Wavelets: Mathematics and Applications*, John J. Benedetto and Michael W. Frazier, editors. Ann Arbor: CRC Press, 1994.