Reprinted with permission of Kluwer Academic Publishers

J.S. Byrnes et al. (eds.), Probabilistic and Stochastic Methods in Analysis, with Applications, 15–45.

Methods of solving dilation equations†

Christopher Heil
Mathematics Department
Massachusetts Institute of Technology
Cambridge, MA 02139
USA
heil@math.mit.edu

Abstract: A *wavelet basis* is an orthonormal basis for $L^2(\mathbb{R})$, the space of square-integrable functions on the real line, of the form $\{g_{nk}\}_{n,k\in\mathbb{Z}}$, where $g_{nk}(t) = 2^{n/2}\, g(2^n t - k)$ and g is a single fixed function, the *wavelet*. Each *multiresolution analysis* for $L^2(\mathbb{R})$ determines such a basis. To find a multiresolution analysis, one can begin with a *dilation equation* $f(t) = \sum c_k\, f(2t-k)$. If the solution f (the *scaling function*) satisfies certain requirements, then a multiresolution analysis and hence a wavelet basis will follow. This paper surveys methods of achieving this goal. Two separate problems are involved: first, solving a general dilation equation to find a scaling function, and second, determining when such a scaling function will generate a multiresolution analysis. We present two methods for solving dilation equations, one based on the use of the Fourier transform and one operating the time domain utilizing linear algebra. The second method characterizes all continuous, integrable scaling functions. We also present methods of determining when a multiresolution analysis will follow from the scaling function. We discuss simple conditions on the coefficients $\{c_k\}$ which are "almost" sufficient to ensure the existence of a wavelet basis, in particular, they do ensure that $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is a *tight frame*, and we present more complicated necessary and sufficient conditions for the generation of a multiresolution analysis. The results presented are due mainly to Cohen, Colella, Daubechies, Heil, Lagarias, Lawton, Mallat, and Meyer, although several of the results have been independently investigated by other groups, including Berger, Cavaretta, Dahmen, Deslauriers, Dubuc, Dyn, Eirola, Gregory, Levin, Micchelli, Prautzsch, and Wang.

1. Introduction

The *Haar system* is the classical example of an *affine*, or *wavelet*, *orthonormal basis* for the space $L^2(\mathbb{R})$ of square-integrable functions on the real line. It

† We thank David Colella of The MITRE Corporation, McLean, Virginia, for his collaboration on some of the results reported in this paper, and for his review of this document.

consists of a set of translations and dilations of a single function, the *Haar wavelet* $\psi(t) = \chi_{[0,1/2)} - \chi_{[1/2,1)}$, where χ_E is the characteristic function of the set E. Precisely, the Haar system has the form $\{\psi_{nk}\}_{n,k\in\mathbb{Z}}$, where $\psi_{nk}(t) = 2^{n/2}\,\psi(2^n t - k)$. Such a simply-generated orthonormal basis is very appealing; however, the fact that the Haar wavelet is discontinuous severely limits the usefulness of the Haar system in applications. Recently, examples of other, smooth, *wavelets* which generate affine orthonormal bases have been given, the first by Meyer [Mey86]. Meyer's example is an infinitely differentiable function which has a compactly supported Fourier transform. Additional examples have been given by Lemarié [Lem88] and Battle [Bat87] (k-times differentiable with exponential decay), Daubechies [Dau88] (k-times differentiable with compact support), and others. Such smooth wavelets are better suited to applications than the Haar wavelet; for example, they have been used in speech compression [CW], [CMW].

Soon after Meyer's initial example, Mallat and Meyer proved that each *multiresolution analysis* for $L^2(\mathbb{R})$ determines a wavelet basis [Mal89]. Each of the wavelet basis mentioned above are determined by an appropriate multiresolution analysis (although not all wavelet bases are associated with multiresolution analyses). A multiresolution analysis $(\{V_n\}, f)$ is defined as a sequence of subspaces $\{V_n\}_{n\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ such that

1) $V_n \subset V_{n+1}$ for all n,
2) $\cap V_n = \{0\}$,
3) $\cup V_n$ is dense in $L^2(\mathbb{R})$, and
4) $h(t) \in V_n \iff h(2t) \in V_{n+1}$,

together with a function $f \in V_0$ such that the collection of integer translates of f, $\{f(t-k)\}_{k\in\mathbb{Z}}$, forms an orthonormal basis for V_0. Given such a multiresolution analysis we have $f \in V_0 \subset V_1$. As $\{2^{1/2} f(2t-k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V_1, there must therefore exist scalars $\{c_k\}$ such that

$$f(t) = \sum_{k\in\mathbb{Z}} c_k\, f(2t-k). \tag{1.1}$$

This is referred to as the (induced) *dilation equation*, and its solution, f, is the *scaling function*. It can be proved that if we define the *wavelet* g by

$$g(t) = \sum_{k\in\mathbb{Z}} (-1)^k\, c_{N-k}\, f(2t-k) \tag{1.2}$$

(where N is as defined below), then g will generate an orthonormal basis for $L^2(\mathbb{R})$ of the form $\{g_{nk}\}_{n,k\in\mathbb{Z}}$, cf., Section 4. From (1.2), it follows that properties of the wavelet g such as continuity, differentiability, etc., are determined by the corresponding properties of the scaling function f.

Remark. For the Haar system,

$$V_0 = \{h : h \text{ is constant on each interval } [k, k+1)\},$$

the induced dilation equation is $f(t) = f(2t) + f(2t-1)$ (i.e., $c_0 = c_1 = 1$ and all other $c_k = 0$), the scaling function is $f = \chi_{[0,1)}$, and the wavelet is the Haar wavelet $g(t) = \psi(t) = f(2t) - f(2t-1)$.

To find wavelet bases for $L^2(\mathbb{R})$, it suffices to construct multiresolution analyses. One method of achieving this is the following. Choose a set of coefficients $\{c_k\}$ and solve the corresponding dilation equation (1.1) for the scaling function f. If f is orthogonal to each of its integer translates then define V_0 to be the span of the integer translates of f and define V_n for $n \in \mathbb{Z}$ as the appropriate dilation of V_0 (i.e., $V_n = \text{span}\{f_{nk}\}_{k\in\mathbb{Z}}$). If $\cap V_n = \{0\}$ and if $\cup V_n$ is dense in $L^2(\mathbb{R})$ then $(\{V_n\}, f)$ is a multiresolution analysis, and therefore the wavelet g defined by (1.2) will generate an affine orthonormal basis for $L^2(\mathbb{R})$. If this is the case then we say that the coefficients $\{c_k\}$ have *determined* the multiresolution analysis $(\{V_n\}, f)$.

There are obviously two separate difficulties in this approach, namely,

1) solving a given dilation equation to find a scaling function, and
2) determining conditions under which a multiresolution analysis will follow from such a scaling function, i.e., conditions under which f will be orthogonal to its integer translates, etc.

We survey results on these two problems in this paper. A shorter survey, which also includes a discussion of the application of wavelets to fast signal processing algorithms, is [Str89].

The first problem, that of solving a general dilation equation, is not restricted in application to wavelet theory. In particular, dilation equations play a role in spline theory, interpolation and subdivision methods, and smooth curve generation [Ber], [BWb], [CDM], [DD89], [DLd], [DGL91], [MP87], [MP89]. Although we focus in this paper on results by groups involved in wavelet research (including Cohen, Colella, Daubechies, Heil, Lagarias, Lawton, Mallat, and Meyer), many of the same or related results have been independently obtained by groups involved in these other areas (including Berger, Cavaretta, Dahmen, Deslauriers, Dubuc, Dyn, Eirola, Gregory, Levin, Micchelli, Prautzsch, and Wang). In some cases, results by these other groups were obtained earlier or are more complete than the ones we discuss.

In Sections 2 and 3 we consider two methods of solving general dilation equations. The methods in Section 2 are based on the use of the Fourier transform. We prove results due to Daubechies and Lagarias showing that every dilation equation has a solution in the sense of distributions, and that integrable solutions, *if they exist*, are unique up to multiplication by a constant. We then present results of Daubechies and Mallat which show when integrable solutions to dilation equations will exist, and results of Colella and

Heil showing when they will not (these results do not completely characterize those dilation equations which have integrable solutions).

In Section 3 we present a time-domain based method for solving certain dilation equations, due to Daubechies and Lagarias, which utilizes linear algebra. This method produces continuous, integrable scaling functions if appropriate conditions hold. Colella and Heil have proved that this method characterizes those dilation equations which have continuous, integrable solutions.

In Section 4 we consider the second problem. We show that if the coefficients $\{c_k\}$ determine a multiresolution analysis then necessarily

$$\sum_k c_{2k} = \sum_k c_{2k+1} = 1 \tag{1.3}$$

and

$$\sum_k c_k\, c_{k+2l} = 2\,\delta_{0l} \quad \text{for every } l \in \mathbb{Z}, \tag{1.4}$$

where δ_{ij} is the Kronecker delta, i.e., $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. We then prove a result due to Lawton which shows that (1.3) and (1.4) are "almost" sufficient to generate a wavelet orthonormal basis. In particular, Lawton has proved that if (1.3) and (1.4) are satisfied then $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ will be a *tight frame*, i.e., the reconstruction property

$$h = \sum_{n,k} \langle h, g_{nk}\rangle\, g_{nk} \quad \text{for all } h \in L^2(\mathbb{R}) \tag{1.5}$$

will be satisfied, although $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ need not be an orthogonal set. (The general theory of frames was developed by Duffin and Schaeffer in [DS52] in connection with nonharmonic Fourier series. The connection between frames and wavelet theory is surveyed in [HW89], and researched in depth in [Dau90].) We also discuss more complicated conditions, independently derived by Lawton and Cohen, which are both necessary and sufficient to ensure that a multiresolution analysis, and therefore a wavelet orthonormal basis, is generated. Lawton has proved that almost all choices of coefficients $\{c_k\}$ which satisfy (1.3) and (1.4) also satisfy these conditions for orthogonality.

For simplicity of presentation, we assume throughout this paper that coefficients $\{c_k\}$ are given which are real with only $c_0, \dots, c_N$ nonzero, i.e., we consider only Daubechies-type wavelets). In Sections 2 and 3, we assume in addition that (1.3) is satisfied. These conditions are not necessary for many of the proofs, and many of the results in which they are necessary can be modified for more general situations. The fact that the coefficients $\{c_k\}$ are real implies that the scaling function f will be real-valued.

Given these restrictions, the Haar system is of course the only example with $N = 1$. It can be shown that multiresolution analyses can be produced only when N is odd. We will use the case $N = 3$ to illustrate many of the results in this paper. For this case, assumption (1.3) reduces to the statement $c_0 + c_2 = c_1 + c_3 = 1$, i.e., the collection of four-coefficient dilation equations with the given restrictions is a two-parameter family. We select the independent parameters to be c_0 and c_3, and represent this collection of four-coefficient dilation equations as the (c_0, c_3)-plane. Figure 1.1 shows several geometrical objects in the (c_0, c_3)-plane. The following results regarding these geometrical objects are discussed in this paper.

1) There are no integrable solutions to dilation equations corresponding to points on or outside the ellipse, with the single exception of the point $(1, 1)$.
2) There do exist integrable solutions to dilation equations corresponding to points on and inside the circle, and inside the shaded region.
3) There are continuous, integrable solutions to dilation equations in a large portion of the triangle, and no continuous, integrable solutions outside the triangle.
4) There are differentiable, integrable solutions to dilation equations on the solid portion of the dashed line.
5) Each point on the circle, with the single exception of the point $(1, 1)$, determines a multiresolution analysis and therefore a wavelet basis for $L^2(\mathbb{R})$. We refer to this circle as the *circle of orthogonality*.

Throughout this paper, $L^p(\mathbb{R})$ will denote the *Lebesgue space* of p-integrable functions on the real line, with norm $\|f\|_p = \left(\int |f(t)|^p\, dt\right)^{1/p}$ for $1 \le p < \infty$ and $\|f\|_\infty = \operatorname{esssup} |f(t)|$. The *inner product* of functions f, g is $\langle f, g\rangle = \int f(t)\, \overline{g(t)}\, dt$. The Fourier transform of an integrable function f is $\hat{f}(\gamma) = \int f(t)\, e^{i\gamma t}\, dt$. Integrals with unspecified limits are over the entire real line.

2. Fourier methods

By considering the Fourier transform of the dilation equation, we can prove that every dilation equation has a solution in the sense of distributions. Consideration of the smoothness and decay of the Fourier transforms of these distributions can indicate whether or not these distributions are given by functions on the real line. We assume throughout this section that (1.3) is satisfied.

Some notation is required to adequately describe distributions. We let $S(\mathbb{R})$ denote the Schwartz space of infinitely differentiable, rapidly decreasing functions on the real line, and let $S'(\mathbb{R})$ denote its topological dual, the space of tempered distributions. For functions φ we define the *translation* operator

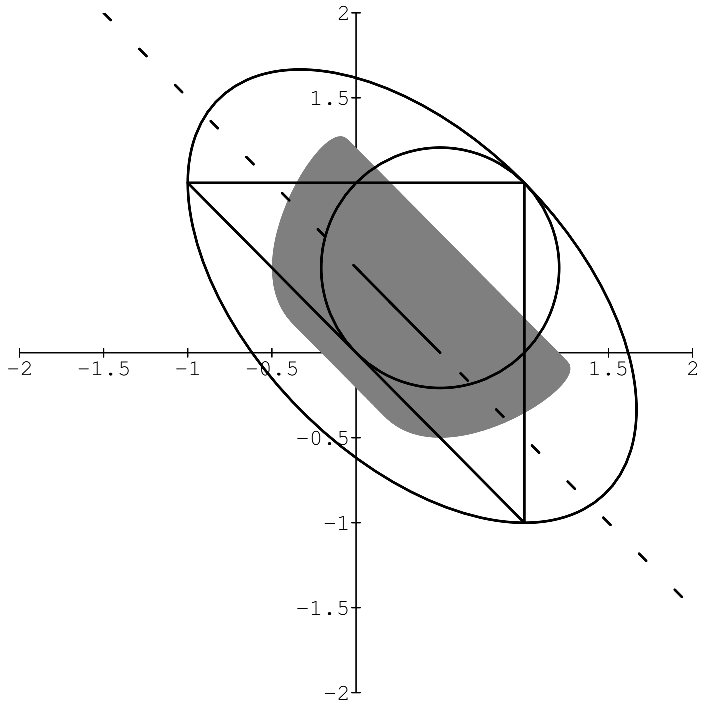

Figure 1.1: Circle of orthogonality, ellipse, line, triangle, and shaded region.

$T_a\varphi(t) = \varphi(t-a)$ and the *dilation* operator $D_a\varphi(t) = \varphi(at)$. Translation and dilation of a distribution $\nu \in S'(\mathbb{R})$ is defined by duality, i.e., $\langle T_a\nu, \varphi\rangle = \langle \nu, T_{-a}\varphi\rangle$ and $\langle D_a\nu, \varphi\rangle = a^{-1}\langle \nu, D_{a^{-1}}\varphi\rangle$. With this notation, the dilation equation (1.1) has the form $f = \sum c_k\, D_2T_kf$. Therefore, we say that $\nu \in S'(\mathbb{R})$ is a *scaling distribution* if

$$\nu = \sum_k c_k\, D_2T_k\nu,$$

i.e., if $\langle \nu, \varphi\rangle = \sum c_k\, \langle D_2T_k\nu, \varphi\rangle$ for all $\varphi \in S(\mathbb{R})$. By taking Fourier transforms, we therefore have that ν is a scaling distribution if and only if

$$D_2\hat{\nu} = m_0\, \hat{\nu},$$

where $m_0(\gamma) = (1/2) \sum c_k \, e^{ik\gamma}$. If it is the case that $\hat{\nu}$ is a function on $\mathbb{R}$ then this is equivalent to

$$\hat{\nu}(2\gamma) = m_0(\gamma)\, \hat{\nu}(\gamma) \quad \text{for a.e. } \gamma \in \mathbb{R}. \tag{2.1}$$

Assume now that $\hat{\nu}$ is a continuous function on $\mathbb{R}$. Then we can iterate (2.1), obtaining (formally) $\hat{\nu}(\gamma) = \hat{\nu}(\gamma/2^n) \prod_1^n m_0(\gamma/2^j) \to \hat{\nu}(0) \prod_1^\infty m_0(\gamma/2^j)$. Daubechies established that this infinite product converges, and proved with Lagarias the following result, cf., [Dau88], [DLb].

Theorem 2.1.

1) $P(\gamma) = \prod_1^\infty m_0(\gamma/2^j)$ converges uniformly on compact sets to a continuous function which has polynomial growth at infinity.
2) Define f to be the tempered distribution such that $\hat{f} = P$. Then $\hat{f}(2\gamma) = m_0(\gamma)\, \hat{f}(\gamma)$ for all γ, so f is a scaling distribution. The support of f is contained in $[0, N]$.
3) If ν is another scaling distribution such that $\hat{\nu}$ is a function on $\mathbb{R}$ which is continuous at zero then $\nu = \hat{\nu}(0)\, f$.
4) If a nonzero integrable solution to (1.1) exists then it is f, up to multiplication by a constant, and $\int f(t)\, dt = 1$.

We call the distribution f defined in Theorem 2.1 the *canonical scaling distribution*. Other solutions to the dilation equation are given in [DLb], and certain classes of solutions are characterized in [CHc].

The proof of Theorem 2.1 requires only that $\sum c_k = 2$; if this is not the case then a canonical solution of the dilation equation can still be defined, but the uniqueness results of Thoerem 2.1 will not hold. Even with the assumption $\sum c_k = 2$, uniqueness in function spaces other than $L^1(\mathbb{R})$ may not hold. For example, the Hilbert transform $H\nu$ of any solution ν of a dilation equation is also a solution of the same dilation equation. Since H maps $L^p(\mathbb{R})$ into $L^p(\mathbb{R})$ for $1 < p < \infty$, uniqueness cannot hold in any of these spaces. Additional uniqueness criteria and methods of generating new solutions to dilation equations from known solutions are given in [CHc].

Existence of an integrable solution of a dilation equation is not guaranteed. The following, from [CHc], is an easily checkable necessary condition for the existence of such solutions, based on the fact that the Fourier transform of an integrable solution must decay at infinity.

Theorem 2.2. Given $x \in [0, 2\pi)$. Assume that the set

$$\{x \bmod 2\pi, 2x \bmod 2\pi, \dots, 2^{n-1}x \bmod 2\pi\}$$

is invariant mod 2π under multiplication by 2. If

$$\prod_{j=1}^{n-1} |m_0(2^j x)| \geq 1 \qquad \text{and} \qquad m_0(2^{-j}x) \neq 0 \quad \text{for all } j \geq 1$$

then the canonical scaling distribution is not an integrable function, and therefore there do not exist any integrable solutions to (1.1).

Remark. Consider the case $N = 3$. The set $\{2\pi/3, 4\pi/3\}$ is invariant mod 2π under multiplication by 2, and $|m_0(2\pi/3)\, m_0(4\pi/3)| \geq 1$ for all (c_0, c_3) on and outside the ellipse shown in Figure 1.1. The additional hypotheses of Theorem 2.2 are also satisfied for all but countably many of these points, and therefore for almost no point on or outside the ellipse can an integrable solution to the corresponding dilation equation exist. All but one of the countably many remaining points are also eliminated when the 3-cycle $\{2\pi/7, 4\pi/7, 8\pi/7\}$ is checked in addition [CHc]. The remaining single point is $(1, 1)$; the integrable solution to the dilation equation corresponding to this point is $f = (1/3)\chi_{[0,3)}$.

Theorem 2.2 deals with non-existence of integrable solutions by establishing conditions under which the Fourier transform $P = \hat{f}$ of the canonical scaling distribution f will not decay at infinity. Alternatively, by imposing sufficient decay on $\hat{f}$ we can obtain $f \in L^2(\mathbb{R})$, and therefore $f \in L^1(\mathbb{R})$ since f has compact support. This is made precise in the next theorem, due to Daubechies [Dau88]. The notation used in the theorem is as follows. Since $2\, m_0(\pi) = \sum (-1)^k c_k = \sum c_{2k} - \sum c_{2k+1} = 0$, we can factor a term of the form $1 + e^{i\gamma}$ from $m_0(\gamma)$. If the zero at π has multiplicity at least L then $m_0(\gamma) = ((1 + e^{i\gamma})/2)^L\, Q(\gamma)$, and therefore

$$\hat{f}(\gamma) = \prod_{j=1}^{\infty} m_0(\gamma/2^j) = \left(\frac{\sin \gamma/2}{\gamma/2}\right)^L \prod_{j=1}^{\infty} Q(\gamma/2^j).$$

Theorem 2.3.

1) If $\|Q\|_\infty < 2^{L-1/2}$ then the canonical scaling distribution f is an integrable function.
2) If $\|Q\|_\infty < 2^{L-1}$ then the canonical scaling distribution f is a continuous, integrable function.

Proof. We prove only the first statement.

Set $M(\gamma) = \prod_1^\infty Q(\gamma/2^j)$; this is a continuous function. Define $R = \|M \cdot \chi_{[-1,1]}\|_\infty$; then since $M(2\gamma) = Q(\gamma)\, M(\gamma)$ we have $\|M \cdot \chi_{[-2^n, 2^n]}\|_\infty \leq \|Q\|_\infty^n\, R$, whence $|M(\gamma)| \leq C\, |\gamma|^{\log_2 \|Q\|_\infty}$ for some constant C. Therefore,

$$|\hat{f}(\gamma)|^2 \leq C' \left(\frac{\sin \gamma/2}{\gamma/2}\right)^{1+p},$$

where $p = 2L - 1 - 2\log_2 \|Q\|_\infty$ and C' is another constant. Since $p > 0$, $((\sin\gamma/2)/(\gamma/2))^{1+p}$ is integrable, and therefore $\hat{f} \in L^2(\mathbb{R})$. Hence $f \in L^2(\mathbb{R})$, and therefore f is integrable since it has compact support. ■

Remark. For $N = 3$, the multiplicity L is one except for those points on the dashed line shown in Figure 1.1; for those points, $L = 2$.

The region of points (c_0, c_3) for which the hypotheses of the first part of Theorem 2.3 is satisfied with $L = 1$ is the shaded region shown in Figure 1.1, i.e., integrable scaling functions exist for all points in this region [CHc] (see also the remark following Theorem 2.4 for an additional region).

No points in the (c_0, c_3)-plane satisfy the second part of Theorem 2.3 with $L = 1$. For $L = 2$, i.e., $c_3 = 1/2 - c_0$, Theorem 2.3 implies that continuous solutions exist for $-1/4 < c_0 < 3/4$. This result is inferior to the one obtained in Section 3, where it is shown that continuous scaling functions occur on this line precisely when $-1/2 < c_0 < 1$, and in fact are differentiable if $0 < c_0 < 1/2$ (i.e., on the solid portion of the line shown in Figure 1.1). Moreover, it is shown in Section 3 that continuous scaling functions occur over a large region of the (c_0, c_3)-plane, including the regions shown in Figures 3.1–3.6.

Eirola has taken a different (but still Fourier-based) approach in [Eir]. He obtains conditions under which scaling functions will be continuous and estimates for the *Sobolev* exponent of continuity for these scaling functions. In Section 3 we discuss a time-domain method for obtaining estimates for the Hölder exponent of continuity of scaling functions.

We end this section with an adaptation of an existence result due to Mallat [Mal89]; part of the proof we give is due to Lawton [Law90].

Theorem 2.4. If

$$|m_0(\gamma)|^2 + |m_0(\gamma + \pi)|^2 \le 1 \quad \text{for all } \gamma,$$

then the canonical scaling distribution is an integrable function.

Proof. Set

$$u_n(\gamma) = \chi_{[-2^n\pi, 2^n\pi]}(\gamma) \cdot \prod_{j=1}^{n} m_0(\gamma). \tag{2.2}$$

By Theorem 2.1, u_n converges uniformly on compact sets to the Fourier transform of the canonical scaling distribution f. Now,

$$\begin{aligned}
\|u_n\|_2^2 &= \int_{-2^n\pi}^{2^n\pi} \left| \prod_{j=1}^{n} m_0\left(\frac{\gamma}{2^j}\right) \right|^2 d\gamma \\
&= \int_0^{2^n\pi} \left| \prod_{j=1}^{n} m_0\left(\frac{\gamma}{2^j}\right) \right|^2 d\gamma \quad + \quad \int_0^{2^n\pi} \left| \prod_{j=1}^{n} m_0\left(\frac{\gamma + 2^n\pi}{2^j}\right) \right|^2 d\gamma \\
&= \int_0^{2^n\pi} \left(\left| m_0\left(\frac{\gamma}{2^n}\right) \right|^2 + \left| m_0\left(\frac{\gamma}{2^n} + \pi\right) \right|^2 \right) \left| \prod_{j=1}^{n-1} m_0\left(\frac{\gamma}{2^j}\right) \right|^2 d\gamma \\
&\leq \int_{-2^{n-1}\pi}^{2^{n-1}\pi} \left| \prod_{j=1}^{n-1} m_0\left(\frac{\gamma}{2^j}\right) \right|^2 d\gamma \\
&= \|u_{n-1}\|_2^2,
\end{aligned}$$

and, by a similar argument, $\|u_1\|_2^2 = 2\pi$. Therefore $\{u_n\}$ is contained in the ball in $L^2(\mathbb{R})$ of radius $\sqrt{2\pi}$ and therefore has a weak* accumulation point. Since $u_n(\gamma) \to \hat{f}(\gamma)$ pointwise, this accumulation point must be $\hat{f}$, whence $f \in L^2(\mathbb{R})$. Since f has compact support, it is therefore integrable as well. ■

Remark. For $N = 3$, equation (2.2) is satisfied for all points (c_0, c_3) on and inside the circle shown in Figure 1.1. Therefore, there exist integrable solutions for all dilation equations corresponding to such points. By the remark following Theorem 2.3, integrable solutions also exist for points in the shaded region in Figure 1.1. The union of these two regions does not exhaust the set of four-coefficient dilation equations which have integrable solutions, cf., [CHc].

3. Matrix methods

In [DLc], Daubechies and Lagarias proved sufficient conditions under which a dilation equation has a continuous, integrable solution (or, more generally, an integrable and n-times differentiable solution). In particular, they proved that if the *joint spectral radius* $\rho(T_0|_V, T_1|_V)$ of two $N \times N$ matrices T_0, T_1 (whose entries contain only the coefficients $\{c_k\}$) restricted to a certain $N-1$ dimensional subspace V is less than one then the canonical scaling distribution f is a continuous and integrable function, and, moreover, is Hölder continuous with Hölder exponent $\alpha \geq -\log_2 \rho(T_0|_V, T_1|_V)$. We outline this result in this section. This result is extended to a necessary and sufficient condition in [CHb], i.e., it is shown there that the canonical scaling distribution f is a continuous and integrable function if and only if $\rho(T_0|_W, T_1|_W) < 1$, where W is an appropriate subspace of V, and that in this case $\alpha = -\log_2 \rho(T_0|_W, T_1|_W)$. It is conjectured in [CHb] that $W = V$ in general except for a set of coefficients of measure zero, and it is proved in [CHa] that $\rho(T_0|_W, T_1|_W) = \rho(T_0|_V, T_1|_V)$ for all choices of coefficients with $N \le 3$. We therefore consider in this survey only the subspace V. We assume throughout this section that (1.3) is satisfied.

Given the coefficients $\{c_k\}$, define the $N \times N$ matrices T_0 and T_1 by $(T_0)_{ij} = c_{2i-j-1}$ and $(T_1)_{ij} = c_{2i-j}$. For example, for $N = 3$ we have

$$T_0 = \begin{pmatrix} c_0 & 0 & 0 \\ c_2 & c_1 & c_0 \\ 0 & c_3 & c_2 \end{pmatrix} \qquad \text{and} \qquad T_1 = \begin{pmatrix} c_1 & c_0 & 0 \\ c_3 & c_2 & c_1 \\ 0 & 0 & c_3 \end{pmatrix}.$$

For $x \in [0,1]$, $x \ne 1/2$, define

$$\tau x = \begin{cases} 2x, & 0 \le x < 1/2, \\ 2x - 1, & 1/2 < x \le 1, \end{cases}$$

i.e., if $x = .d_1d_2d_3\dots$ is the binary decimal expansion of x then $\tau x = .d_2d_3\dots$. Although $\tau(1/2)$ is not uniquely defined, this ambiguity will not pose any problems in the analysis.

We say that a function f is *Hölder continuous* if there exist constants K, α such that $|f(x) - f(y)| \le K\,|x - y|^\alpha$ for all x, $y \in \mathbb{R}$. The largest such exponent α is the *Hölder exponent* and the corresponding smallest constant K is the *Hölder constant*.

The relationship between the dilation equation (1.1) and the matrices T_0, T_1 is given in the following result from [DLc].

Proposition 3.1.

1) Assume f is a continuous and integrable scaling function. Define the vector-valued function $v(x)$ for $x \in [0,1]$ by

$$v(x) = \begin{pmatrix} f(x) \\ f(x+1) \\ \vdots \\ f(x+N-1) \end{pmatrix}. \tag{3.1}$$

Then v is continuous on $[0,1]$ and satisfies

$$\begin{aligned} &v_1(0) = v_N(1) = 0, \\ &v_{i+1}(0) = v_i(1) \quad \text{for } i = 1,\dots,N-1, \\ &v(x) = T_{d_1} v(\tau x) \quad \text{for } x = .d_1d_2\dots \in [0,1],\ x \ne 1/2, \\ &v(1/2) = T_0 v(1) = T_1 v(0), \end{aligned} \tag{3.2}$$

where $v_i(x)$ is the i^{th} component of $v(x)$. Moreover, if f is Hölder continuous with Hölder exponent α then the same is true of v.

2) Assume v is a continuous vector-valued function on $[0, 1]$ satisfying (3.2). Define the function f by

$$f(x) = \{ \begin{matrix} 0, & x \le 0 \text{ or } x \ge N, \\ v_i(x), & i-1 \le x \le i,\ i = 1, \ldots, N. \end{matrix} \tag{3.3}$$

Then f is a continuous and integrable scaling function. Moreover, if v is Hölder continuous with Hölder exponent α then the same is true of f.

The fundamental theorem on the existence of continuous, integrable scaling functions is the following result from [DLc]. The notation used in the theorem is as follows. Let V denote the subspace

$$V = \{u \in \mathbb{R}^N : u_1 + \cdots + u_N = 0\},$$

and let M be the $(N-1) \times (N-1)$ matrix $M_{ij} = c_{2i-j}$. A point $x = .d_1 \ldots d_m \in [0, 1]$ with a finite binary decimal expansion is called a *dyadic point*.

Theorem 3.2. Fix any norm $\|\cdot\|$ on $\mathbb{R}^N$, and assume there exist constants $C > 0$ and $0 < \lambda < 1$ such that

$$\|(T_{d_1} \cdots T_{d_m})|_V\| \le C\, \lambda^m \tag{3.4}$$

for every choice of $d_1, \ldots, d_m \in \{0, 1\}$ and every $m > 0$. Then the following statements are true.

1) 1 is a simple eigenvalue of T_0, T_1, and M.
2) M has a right eigenvector $(a_1, \ldots, a_{N-1})^{\text{t}}$ for the eigenvalue 1 such that $a_1 + \cdots + a_{N-1} = 1$.
3) Set $v(0) = (0, a_1, \ldots, a_{N-1})^{\text{t}}$ and define $v(x)$ for $x = .d_1 \ldots d_m \in [0, 1]$ by

$$v(x) = T_{d_1} \cdots T_{d_m} v(0). \tag{3.5}$$

Then $v_1(x) + \cdots + v_N(x) = 1$ for every such x.
4) v is bounded on the set of dyadic points in $[0, 1]$.
5) v is Hölder continuous on the set of dyadic points in $[0, 1]$ with Hölder exponent $\alpha \ge -\log_2 \lambda$, and has a unique continuous extension to $[0, 1]$ which is Hölder continuous with the same exponent α.
6) v satisfies (3.2), and therefore the function f defined by (3.3) is a continuous, integrable scaling function and is Hölder continuous with exponent˜α.

Proof. Full details of the proof can be found in [DLc]; we sketch selected points below.

1) Follows from the fact that V has dimension $N-1$ in $\mathbb{R}^N$, and that M is a submatrix of both T_0 and T_1.

2) Follows from 1) and the fact that $(1,\dots,1)$ is a left eigenvector for M for the eigenvalue 1.

3) Since $(1,\dots,1)$ is a common left eigenvector for T_0 and T_1 for the eigenvalue 1,

$$\begin{aligned} v_1(x)+\cdots+v_N(x) &= (1,\dots,1)\,v(x) \\ &= (1,\dots,1)\,T_{d_1}\cdots T_{d_m}v(0) \\ &= (1,\dots,1)\,v(0) \\ &= 1. \end{aligned}$$

5) Choose any dyadic $x = d_1\dots d_k \in [0,1]$ and assume $y > x$ is also dyadic. If $2^{-m-1} \le y - x < 2^{-m}$ with $m > k$ then $x = .d_1\dots d_m$ and $y = .d_1\dots d_m d_{m+1}\dots d_{m+j}$ for some j. From 3), $v(\tau^m y) - v(0) \in V$, so

$$\begin{aligned} \|v(y)-v(x)\| &= \|T_{d_1}\cdots T_{d_m}\big(v(\tau^m y)-v(0)\big)\| \\ &\le \|(T_{d_1}\cdots T_{d_m})|_V\|\,\|v(\tau^m y)-v(0)\| \\ &\le 2\,L\,C\,\lambda^m \\ &= 2\,L\,C\,\lambda^{-1}\,(2^{-m-1})^{-\log_2\lambda} \\ &\le 2\,L\,C\,\lambda^{-1}\,|y-x|^{-\log_2\lambda}, \end{aligned}$$

where $L = \sup\{\|v(t)\| : \text{dyadic } t \in [0,1]\} < \infty$ by 4). Thus v is Hölder continuous from the right on the set of dyadic points in $[0,1]$ with Hölder exponent $\alpha \ge -\log_2\lambda$. A similar proof establishes Hölder continuity from the left.

6) Given $x = .d_1\dots d_m$ dyadic, we have $v(x) = T_{d_1}(T_{d_2}\cdots T_{d_m}v(0)) = T_{d_1}v(\tau x)$. By continuity, this holds for all $x \in [0,1]$. ■

Examples of norms on $\mathbb{R}^N$ are $\|u\|_p = (|u_1|^p + \cdots + |u_N|^p)^{1/p}$ for $1 \le p < \infty$ and $\|u\|_\infty = \max\{|u_1|,\dots,|u_N|\}$.

Condition (3.4) is most easily analyzed in a spectral form, as follows.

The *joint spectral radius* of a set of $N\times N$ matrices $\{A_0,\dots,A_n\}$ is the straightforward generalization of the usual spectral radius of a single matrix, namely,

$$\rho(A_0,\dots,A_n) = \limsup_{m\to\infty} \lambda_m,$$

where

$$\lambda_m = \max_{d_j\in\{0,\dots,n\}} \|A_{d_1}\cdots A_{d_m}\|^{1/m}.$$

The joint spectral radius was first introduced by Rota and Strang [RS60]. Recent articles are [BWa], [DLa].

Lemma 3.3.

1) For every $\lambda > \rho(A_0, \ldots, A_n)$ there exists a constant $C > 0$ such that $\lambda_m^m \le C\,\lambda^m$ for every m.
2) If there exist $C, \lambda > 0$ such that $\lambda_m^m \le C\,\lambda^m$ for every m then $\rho(A_0, \ldots, A_n) \le \lambda$.

It follows from Lemma 3.3 that (3.4) is equivalent to $\rho(T_0|_V, T_1|_V) < 1$ (however, $\rho(T_0|_V, T_1|_V) = 1$ is not equivalent to $\lambda_m^m \le C$ for every m).

The joint spectral radius can be difficult to compute, except in special cases. For a single matrix A, $\rho(A)$ is simply the usual spectral radius of A and is therefore the largest of the absolute values of the eigenvalues of A. This is not true in general, i.e., if we define

$$\sigma_m = \max_{d_j \in \{0,\ldots,n\}} \rho(A_{d_1} \cdots A_{d_m})^{1/m},$$

then $\rho(A_0, \ldots, A_n) \ne \sigma_1 = \max\{\rho(A_0), \ldots, \rho(A_n)\}$. However, we do have the following, cf., [DLc].

Lemma 3.4.

1) $\sigma_m \le \rho(A_0, \ldots, A_n) \le \lambda_m$ for every m.
2) $\rho(A_0, \ldots, A_n)$ is independent of the choice of basis, i.e., if B is any invertible matrix then $\rho(BA_0B^{-1}, \ldots, BA_nB^{-1}) = \rho(A_0, \ldots, A_n)$.
3) If there exists an invertible matrix B such that $BA_0B^{-1}, \ldots, BA_nB^{-1}$ are all simultaneously symmetric, then $\rho(A_0, \ldots, A_n) = \sigma_1$.

Berger and Wang have proved that $\rho(A_0, \ldots, A_n) = \limsup \sigma_m$, and therefore $\rho(A_0, \ldots, A_n) = \sup \sigma_m$ [BWa].

We return now to consideration of the matrices T_0, T_1. Since V has dimension $N-1$, an appropriate change of basis gives $\rho(T_0|_V, T_1|_V) = \rho(S_0, S_1)$, where S_0, S_1 are $(N-1) \times (N-1)$ matrices (not necessarily unique).

Remark. For $N = 3$ we can set

$$S_0 = \begin{pmatrix} c_0 & 0 \\ -c_3 & 1 - c_0 - c_3 \end{pmatrix} \qquad \text{and} \qquad S_1 = \begin{pmatrix} 1 - c_0 - c_3 & -c_0 \\ 0 & c_3 \end{pmatrix},$$

cf., [CHa]. The shaded area in Figure 3.1 shows the set SS of points (c_0, c_3) for which S_0 and S_1 can be simultaneously symmetrized with $\rho(S_0, S_1) < 1$. Continuous, integrable scaling functions therefore exist for all points in this region.

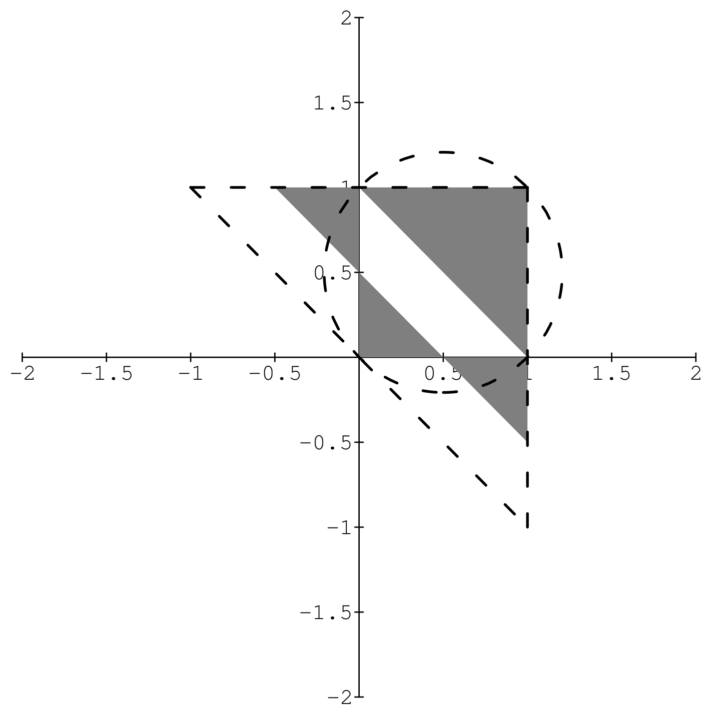

Figure 3.1: Region SS where simultaneous symmetrization is possible and leads to continuous scaling functions (shaded area).

In the regions where simultaneous symmetrization is not possible, Lemma 5.4 can be used to estimate the joint spectral radius.

Remark. Set $N = 3$, and let $C_{p,m}$ be the set of points (c_0, c_3) such that $\rho(S_0, S_1) \le \lambda_m < 1$ with the choice of norm $\|\cdot\|_p$. By Theorem 3.2, continuous, integrable scaling functions exist for all points in any $C_{p,m}$. Figures 3.2–3.4 show $C_{p,1}$ for several choices of p, i.e., the sets obtained by considering the matrices S_0, S_1 directly (since $\lambda_1 = \max\{\|S_0\|_p, \|S_1\|_p\}$).

Figure 3.5 shows the region $C_{2,16}$ obtained by considering, for each point (c_0, c_3), the Euclidean space norm $\|\cdot\|_2$ of all 65536 possible products $S_{d_1} \cdots S_{d_{16}}$ of S_0 and S_1 of length 16.

The union of the regions shown in Figures 3.2–3.5, plus the region SS shown in Figure 3.1, is shown in Figure 3.6. Continuous, integrable scaling functions therefore exist for all points in the shaded area in Figure 3.6. By the remark following Theorem 3.6, there are no continuous, integrable scaling functions on or outside the solid boundary shown in Figure 3.6.

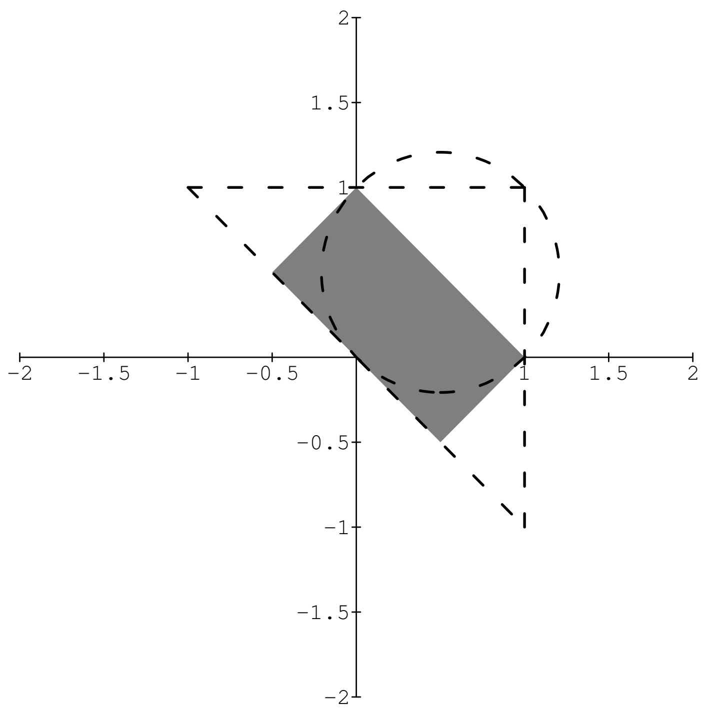

Figure 3.2: The set $C_{1,1}$ (shaded area).

Note that half of the circle of orthogonality lies inside the shaded area in Figure 3.6, and half lies outside the solid line. Therefore there exist many wavelet bases with $N = 3$ for which the wavelet is continuous, cf., Figures 3.7 and 3.8.

For large m, direct computation of λ_m is impractical. The following algorithm can be used to select a subset of matrices which can be used to estimate $\rho(A_0, \ldots, A_n)$ [CHb], cf., [DLc].

Proposition 3.5. Given $\rho > \rho(A_0, \ldots, A_n)$. For each of the matrices $A_0, \ldots, A_n$ in turn, implement the following recursion.

- Given a product $P = A_{d_1} \cdots A_{d_m}$. If $\|P\|^{1/m} < \rho$ then keep P as a *building block*. Otherwise, repeat this step with each of the products $PA_0, \ldots, PA_n$ in turn.

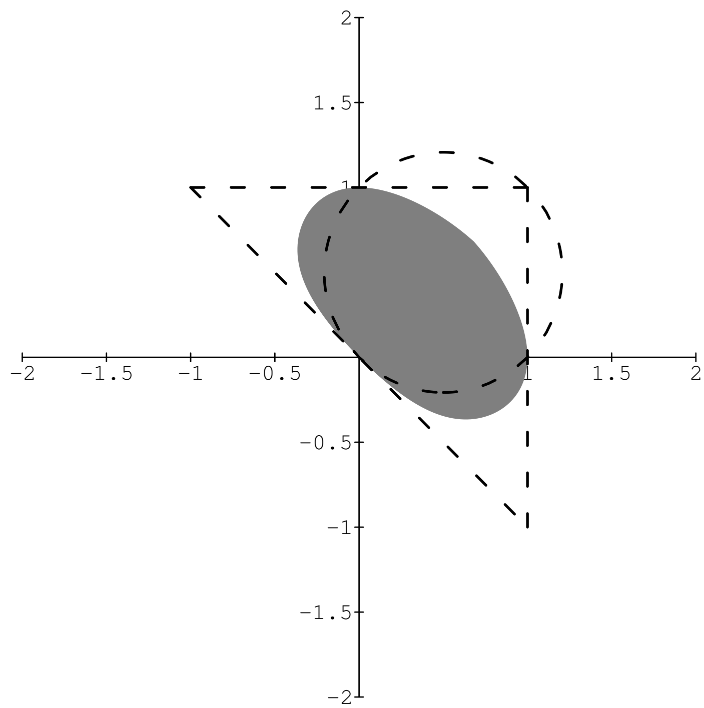

Figure 3.3: The set $C_{2,1}$ (shaded area).

Label the resulting set of building blocks $P_1, \ldots, P_l$, and let m_j be the length of the product P_j. Then the following statements hold.

1) There is an $r \geq 0$ such that if $P = A_{d_1} \cdots A_{d_m}$ is any product of the matrices $A_0, \ldots, A_n$, then $P = P_{j_1} \cdots P_{j_k} R$ where R is some product of at most r of the matrices $A_0, \ldots, A_n$.
2) $\rho(A_0, \ldots, A_n) \leq \max\{\|P_1\|^{1/m_1}, \ldots, \|P_l\|^{1/m_l}\}$.

This algorithm can be used to significantly shorten the time required to estimate a joint spectral radius.

Remark. For $N = 3$ and $(c_0, c_3) = (.6, -.2)$, for which simultaneous symmetrization is not possible, we compute (using the norm $\|\cdot\|_1$) $\lambda_1 = .737$ and $\lambda_{13} = .682$. The computation of λ_{13} required the calculation of of 8192 matrix products; however, the algorithm given in Proposition 3.5 equals this estimate after only 94 matrix product computations. A deeper search, with a maximum matrix product length of 73, required only 14156 matrix product computations and resulted in the estimate $\rho(S_0, S_1) \leq .661$.

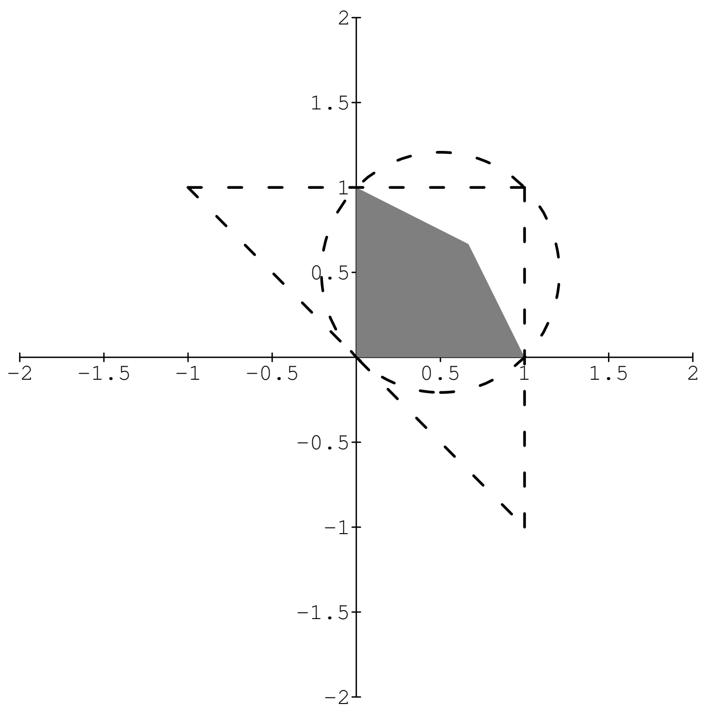

Figure 3.4: The set $C_{\infty,1}$ (shaded area).

Even if λ_{73} could be computed it would not improve this estimate, e.g., $\|S_0^2 S_1 S_0^{14} S_1 S_0^{14} S_1 S_0^{13} S_1 S_0^{12} S_1 S_0^{12} S_1\|_1^{1/73} = .663$. These computations, and the significance of the point $(c_0, c_3) = (.6, -.2)$, are explained in detail in [CHa]; note, however, that the Hölder exponent of continuity for the scaling function determined by the coefficients $(c_0, c_3) = (.6, -.2)$ is at least $-\log_2 .661 \approx .598$, and therefore this scaling function is smoother than the standard four-coefficent example, the Daubechies scaling function D_4, which is determined by the coefficients $(c_0, c_3) = ((1+\sqrt{3})/4, (1-\sqrt{3})/4)$, and whose Hölder exponent of continuity is approximately .550. These two scaling functions are shown in Figures 3.7 and 3.8. Each of these two choices of coefficients lies on the circle of orthogonality and determines a multiresolution analysis for $L^2(\mathbb{R})$.

Theorem 3.2 is extended to a necessary and sufficient condition in [CHb]. We indicate now the method used to obtain the converse. Given an $N \times N$ matrix A and an eigenvalue λ of A, set $U_\lambda = \{u \in \mathbb{C}^N : (A - \lambda)^k u = 0 \text{ for some } k > 0\}$. By standard Jordan decomposition techniques we can

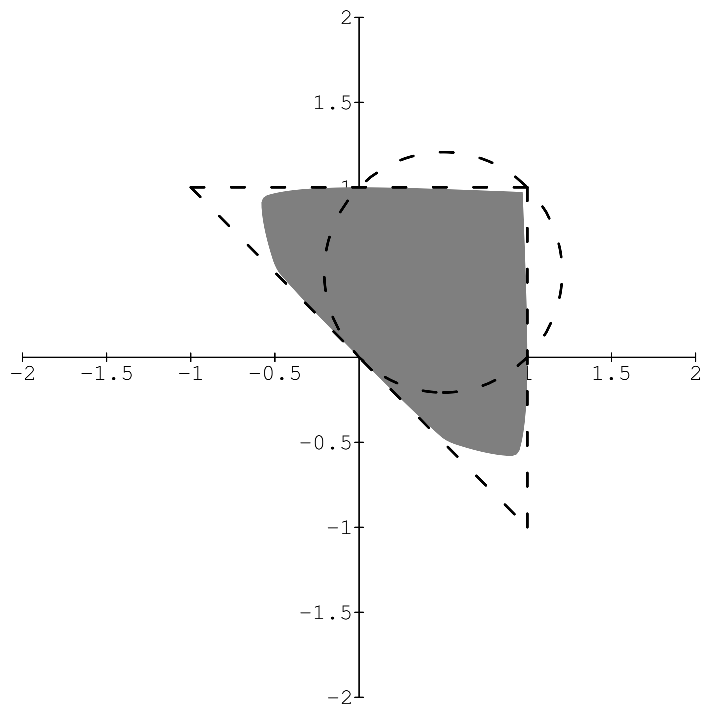

Figure 3.5: The set $C_{2,16}$ (shaded area).

write $\mathbb{C}^N = U_\lambda \oplus W$, where W is a unique A-invariant subspace of $\mathbb{C}^N$. Given $v \in \mathbb{C}^N$ we say that v *has a component in* U_λ if $v = u + w$ where $u \in U_\lambda$, $w \in W$, and $u \neq 0$. The following result is from [CHb].

Theorem 3.6. Assume v is a continuous vector-valued function on $[0,1]$ such that (3.2) holds, and let $T = T_{d_1} \cdots T_{d_m}$ be any fixed product of the matrices T_0, T_1. Let $x \in [0,1]$ be that point whose binary decimal expansion is $x = .d_1 \dots d_m d_1 \dots d_m \dots$. If

1) λ is an eigenvalue of $T|_V$, and
2) there is some $z \in [0,1]$ such that $v(x) - v(z)$ has a component in U_λ,

then $|\lambda| < 1$ and the Hölder exponent of continuity of v is at most $-\log_2 |\lambda|^{1/m}$.

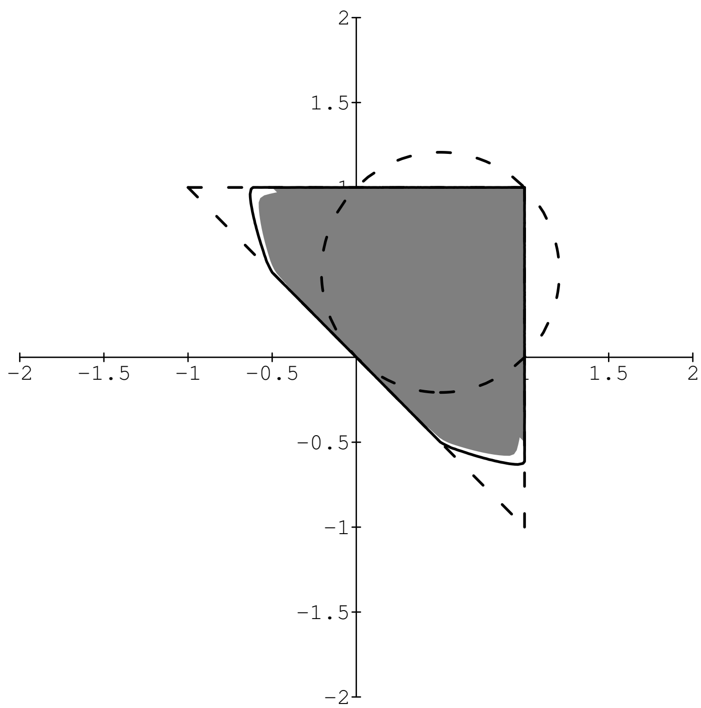

Figure 3.6: Union of the sets SS, $C_{1,1}$, $C_{2,1}$, $C_{\infty,1}$, and $C_{2,16}$ (shaded area); boundary of the set E_{16} (solid line).

Since $\rho(T_0|_V, T_1|_V) = \sup \sigma_m$ is the supremum of the absolute values of the eigenvalues of every $(T_{d_1} \cdots T_{d_m})|_V$, it follows that if the hypotheses of Theorem 3.6 are satisfied for each product $T = T_{d_1} \cdots T_{d_m}$ then $\rho(T_0|_V, T_1|_V) \leq 1$ with $\sigma_m < 1$ for all m, and the Hölder exponent of v satisfies $\alpha \leq -\log_2 \rho(T_0|_V, T_1|_V)$. Therefore, if the hypotheses of Theorem 3.6 are satisfied for each product $T = T_{d_1} \cdots T_{d_m}$ then Theorem 3.6 is the converse to Theorem 3.2, except for the possibility of one special case, namely,

$$\sup_m \sigma_m = 1 \qquad \text{and} \qquad \sigma_m < 1 \text{ for all } m.$$

It is unknown whether this special case can actually occur. It is proven in [CHa] that the hypotheses of Theorem 3.6 are always satisfied if $N \leq 3$ and it is conjectured in [CHb] that they are always satisfied in general except for a set of coefficients of measure zero. Methods for determining the validity of the hypotheses of Theorem 3.6 for any specific choice of coefficients are given in [CHb].

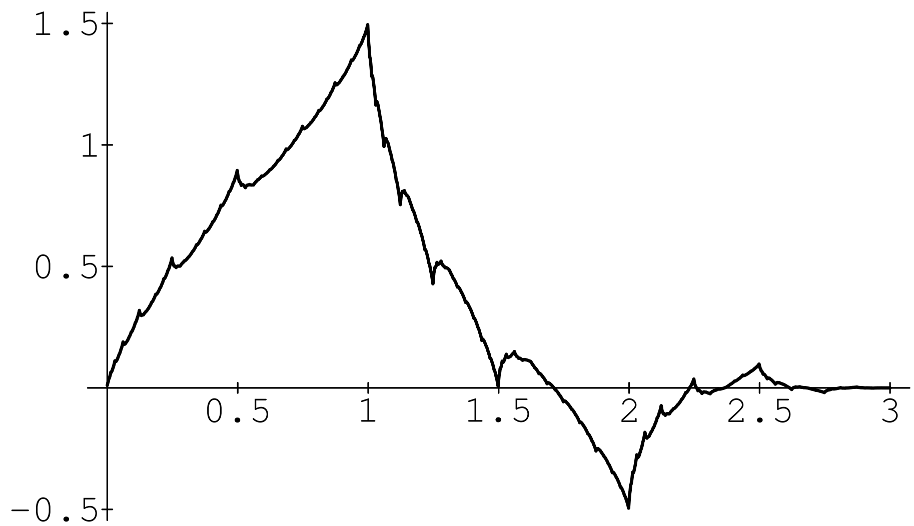

Figure 3.7: Scaling function corresponding to $(c_0, c_3) = (.6, -.2)$.

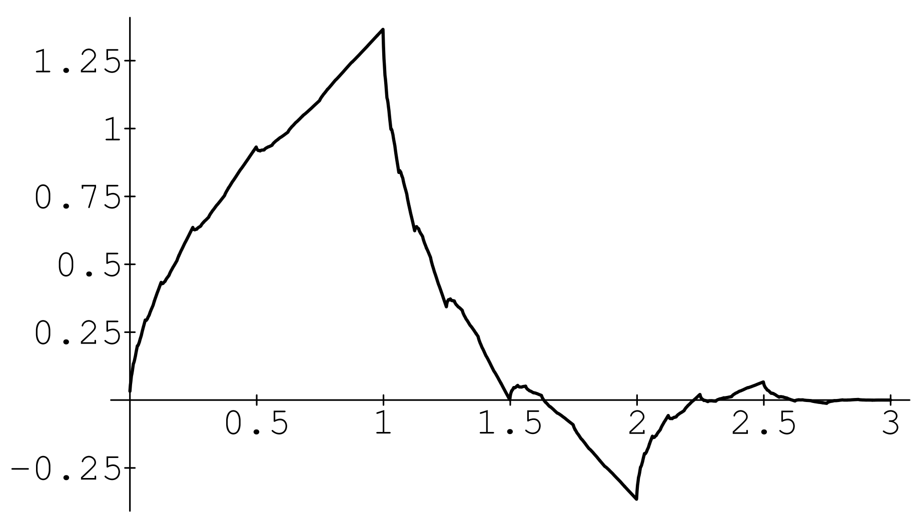

Figure 3.8: Daubechies scaling function D_4.

Remark. Set $N = 3$ and define $E_m = \{(c_0, c_3) : \sigma_m \geq 1\}$. By Theorem 3.6, no dilation equation determined by a point in E_m can have a continuous, integrable solution. The set E_1 is precisely the boundary and exterior of the triangle shown in Figure 1.1. The solid line in Figure 3.6 shows a numerical approximation of the boundary of E_{16} [CHa]. By previous remarks, continuous, integrable scaling functions do exist in the shaded region in Figure 3.6.

The results of this section can be extended from consideration of continuous solutions to n-times differentiable solutions. If f is such a solution then its derivatives $f^{(j)}$ satisfy the dilation equations

$$f^{(j)}(t) = \sum_k 2^j \, c_k \, f^{(j)}(2t - k).$$

Therefore the vector $(f^{(j)}(1), \ldots, f^{(j)}(N-1))^{\mathrm{t}}$ is a right eigenvector for the matrix M for the eigenvalue 2^{-j}. As M is an $(N-1) \times (N-1)$ matrix, f can therefore possess at most $N - 2$ derivatives. This can always be achieved for an appropriate choice of coefficients [DLc].

The following modification of Theorem 3.2 for the case of higher derivatives is from [DLc].

Theorem 3.7. Assume that the coefficients $\{c_k\}$ satisfy the *sum rules* $\sum (-1)^k \, k^j \, c_k = 0$ for $j = 0, \ldots, n$. Define $V_n = \{u \in \mathbb{R}^N : e_j \, u = 0, \, j = 0, \ldots, n\}$, where $e_j = (1^j, 2^j, \ldots, N^j)$. If $\rho(T_0|_{V_n}, T_1|_{V_n}) < 2^{-n}$ then there exists an n-times differentiable solution f to (1.1), and the n^{th} derivative $f^{(n)}$ of f is Hölder continuous with exponent $\alpha \geq -\log_2 2^n \, \rho(T_0|_{V_n}, T_1|_{V_n})$.

Remark. For the case $N = 3$, differentiable solutions can exist only on the solid portion of the line shown in Figure 1.1. None of these solutions can be twice differentiable. In particular, for $N = 3$, no wavelet which generates an affine orthonormal basis can be differentiable since wavelets must be derived from points lying of the circle of orthogonality.

4. Orthogonality

In this section we consider the relationship between the choice of coefficents $\{c_k\}$ and frame or basis properties of the associated wavelet. We assume N is odd in this section.

We require the following lemmas. $C_c(\mathbb{R})$ denotes the space of all continuous functions on $\mathbb{R}$ which have compact support. The proof of the first lemma can be found in [Mal89].

Lemma 4.1. If $(\{V_n\}, f)$ is a multiresolution analysis then $\int f(t)\,dt = 1$.

Lemma 4.2. If (1.3) holds then the canonical scaling function f satisfies $\sum f(t-k) = 1$ a.e.

Proof. Set $\theta_0 = \chi_{[0,1)}$ and $\theta_j(t) = \sum c_k\, \theta_{j-1}(2t-k)$. Since $\hat{\theta}$ is continuous, $\hat{\theta}(0) = 1$, and $\hat{\theta}_j(2\gamma) = m_0(\gamma)\,\hat{\theta}_{j-1}(\gamma)$, it follows that $\theta_j \to f$ weakly in $L^2(\mathbb{R})$, i.e., $\langle \theta_j, h\rangle \to \langle f, h\rangle$ for all $h \in L^2(\mathbb{R})$. Note that $\sum \theta_0(t-k) = 1$ a.e.; by induction, the same is true of θ_j, and hence of f. ■

Next, we establish necessary conditions on the coefficients $\{c_k\}$ in order that a multiresolution analysis exist.

Proposition 4.3. If the coefficients $\{c_k\}$ determine a multiresolution analysis then (1.3) and (1.4) hold. The converse is not true.

Proof. Integrating both sides of the dilation equation implies that $\sum c_k = 2$, since $\int f(t)\,dt$ is nonzero by Lemma 4.1. Since f is orthogonal to its integer translates,

$$\begin{aligned} 2\,\delta_{0l} &= 2 \int f(t)\, f(t+l)\, dt \\ &= 2 \sum_{j,k} c_j\, c_k \int f(2t-j)\, f(2t+2l-k)\, dt \\ &= \sum_k c_k\, c_{k+2l}, \end{aligned}$$

so (1.4) holds. This, combined with the fact $\sum c_k = 2$, implies (1.3).

To see that (1.3) and (1.4) are not sufficient, consider the coefficient choice $c_0 = 1$, $c_1 = \cdots = c_{N-1} = 0$, $c_N = 1$. These coefficients satisfy (1.3) and (1.4), yet the canonical scaling function $f = (1/N)\,\chi_{[0,N)}$ is not orthogonal to its integer translates if $N > 1$. ■

Remark. For $N = 3$, the set of points in the (c_0, c_3)-plane which satisfy both (1.3) and (1.4) is precisely the circle of orthogonality shown in Figure 1.1.

Equations (1.3) and (1.4) are equivalent to

$$m_0(0) = 1 \qquad \text{and} \qquad m_0(\pi) = 0 \tag{4.1}$$

and

$$|m_0(\gamma)|^2 + |m_0(\gamma+\pi)|^2 = 1 \quad \text{for all } \gamma. \tag{4.2}$$

Equation (4.2) implies that, in signal processing terms, $m_0(\gamma)$ and $m_0(\gamma+\pi)$ form a *quadrature mirror filter pair*. Such filter pairs induce fast digital signal processing algorithms, e.g., *subband coding*. Daubechies has characterized those trigonometric polynomials m_0 which satisfy (4.1) and (4.2) in [Dau88].

Although (1.3) and (1.4) are not sufficient to ensure that $\{c_k\}$ will generate a multiresolution analysis (and therefore that $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ will be an orthonormal basis), Lawton has proven that (1.3) and (1.4) are sufficient to ensure that the sequence $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ will satisfy the reconstruction property (1.5) of an orthonormal basis. Such a sequence is called a *tight frame*. (1.5) alone does not imply that $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is an orthogonal sequence or a basis, i.e., in general the summation in (1.5) is not unique. See [DS52] or [HW89] for exposition on frames and their properties.

The following theorem and proof are from [Law90].

Theorem 4.4. If the coefficients $\{c_k\}$ satisfy (1.3) and (1.4) then $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is a tight frame for $L^2(\mathbb{R})$.

Proof. We proceed in four steps.

1) From (4.2) and Theorems 2.4 and 2.1, the canonical scaling distribution f is an integrable function with support contained in $[0, N]$ and satisfies $\int f(t)\,dt = 1$.

2) Define the operator $P_n : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by

$$P_n h = \sum_k \langle h, f_{nk}\rangle\, f_{nk}. \tag{4.3}$$

We claim that $P_n \to I$ as $n \to +\infty$ and $P_n \to 0$ as $n \to -\infty$, where I is the identity operator on $L^2(\mathbb{R})$.

First, however, we show that the operators $\{P_n\}$ are uniformly bounded in norm. Since $\mathrm{supp}(f_{nk}) \subset I_{nk} = [k/2^n, (k+N)/2^n]$, with n fixed each set $\mathrm{supp}(f_{nk})$ can intersect at most N other $\mathrm{supp}(f_{nj})$. Therefore, for any scalars $\{a_k\}$,

$$\begin{aligned} \Big\| \sum_k a_k\, f_{nk} \Big\|_2 &\le N^{1/2} \Big(\sum_k |a_k|^2\, \|f_{nk}\|_2^2 \Big)^{1/2} \\ &= N^{1/2}\, \|f\|_2 \Big(\sum_k |a_k|^2 \Big)^{1/2}, \end{aligned} \tag{4.4}$$

e.g., [Hei90, Prop. 2.4.10]. Therefore,

$$\|P_n h\|_2 \le N^{1/2}\, \|f\|_2 \Big(\sum_k |\langle h, f_{nk}\rangle|^2 \Big)^{1/2}. \tag{4.5}$$

Now, for each $r = 0, \dots, N-1$, the sequence $\{f_{n(lN+r)}\}_{l\in\mathbb{Z}}$ is an orthogonal collection of functions since their supports are disjoint. Therefore, by Bessel's inequality, $\sum |\langle h, f_{n(lN+r)}\rangle|^2 \le \|h\|_2^2\, \|f\|_2^2$. Combining this with (4.5) we obtain $\|P_n h\|_2 \le N\, \|f\|_2^2\, \|h\|_2$, and therefore $\sup \|P_n\|_2 \le N\, \|f\|_2^2 < \infty$.

Because the operators $\{P_n\}$ are uniformly bounded in norm, to prove $P_n h \to h$ as $n \to +\infty$ for all $h \in L^2(\mathbb{R})$ it suffices to consider h in a dense subset of $L^2(\mathbb{R})$, say $h \in C_c(\mathbb{R})$. For such an h, since $\sum_k f_{nk}(t) = 2^{n/2}$ a.e. (Lemma 4.2), we can write

$$\begin{aligned} \|h - P_n h\|_2 &= \left(\int \left| \sum_k (2^{-n/2}\, h(t) - \langle h, f_{nk}\rangle)\, f_{nk}(t) \right|^2 dt \right)^{1/2} \\ &\le N^{1/2} \left(\sum_k \|(2^{-n/2}\, h(t) - \langle h, f_{nk}\rangle)\, f_{nk}(t)\|_2^2 \right)^{1/2} \\ &\le N^{1/2}\, \|f\|_2 \left(\sum_k \alpha_{nk}^2 \right)^{1/2}, \end{aligned} \tag{4.6}$$

where we have used (4.4) again and where

$$\alpha_{nk} = \sup_{t \in I_{nk}} |2^{-n/2}\, h(t) - \langle h, f_{nk}\rangle|.$$

To see that $\sum \alpha_{nk}^2 \to 0$ as $n \to +\infty$, define

$$\beta_{nk} = \sup_{s,t \in I_{nk}} |h(s) - h(t)| \qquad \text{and} \qquad \tilde{h}_n = \sum_k \beta_{nk}^2\, \chi_{I_{nk}}.$$

Note that $\tilde{h}_n(t) \to 0$ pointwise as $n \to +\infty$ since $h \in C_c(\mathbb{R})$. Further, $\beta_{nk} \le \beta_{n0}$ and $I_{nk} \subset I_{n0}$ for all k, so $\tilde{h}_n \le \tilde{h}_0$ for $n \ge 0$. As $\tilde{h}_0$ is clearly integrable, it follows from the Lebesgue Dominated Convergence Theorem that $\int \tilde{h}_n(t)\, dt \to 0$ as $n \to +\infty$. Now, since $\int f_{nk}(t)\, dt = 2^{-n/2}$ (Lemma 4.1), we have for $t \in I_{nk}$ that

$$\begin{aligned} |2^{-n/2}\, h(t) - \langle h, f_{nk}\rangle| &= \left| \int_{I_{nk}} (h(t) - h(s))\, f_{nk}(s)\, ds \right|^2 \\ &\le \left(\int_{I_{nk}} |h(t) - h(s)|^2\, ds \right) \left(\int_{I_{nk}} |f_{nk}(s)|^2\, ds \right) \\ &\le \|f\|_2^2\, \beta_{nk}^2 \int \chi_{I_{nk}}(s)\, ds. \end{aligned}$$

Therefore, $\sum_k \alpha_{nk}^2 \le \|f\|_2^2 \int \tilde{h}_n(s)\, ds \to 0$ as $n \to +\infty$, which, combined with (4.6), implies that $P_n h \to h$ in $L^2(\mathbb{R})$ as $n \to +\infty$.

A similar proof shows that $P_n h \to 0$ as $n \to -\infty$.

3) Define the operator $F_n : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by

$$F_n h = \sum_k \langle h, g_{nk}\rangle\, g_{nk}.$$

We claim then that $F_n = P_{n+1} - P_n$ for each $n \in \mathbb{Z}$. Using the dilation equation (1.1) and the definition (1.2) we compute

$$P_n h + F_n h = \sum_k \langle h(t), 2^{n/2} f(2^n t - k)\rangle\, 2^{n/2} f(2^n t - k)$$
$$+ \sum_k \langle h(t), 2^{n/2} g(2^n t - k)\rangle\, 2^{n/2} g(2^n t - k)$$

$$= 2^n \sum_{k,p,q} \langle h(t), c_p\, f(2^{n+1}t - 2k - p)\rangle\, c_q\, f(2^{n+1}t - 2k - q)$$
$$+ 2^n \sum_{k,p,q} \langle h(t), (-1)^p\, c_{N-p}\, f(2^{n+1}t - 2k - p)\rangle\, (-1)^q\, c_{N-q}\, f(2^{n+1}t - 2k - q)$$

$$= 2^n \sum_{p,q,k} \langle h(t), (c_p\, c_q + (-1)^{p+q}\, c_{N-p}\, c_{N-q})\, f(2^{n+1}t - 2k - p)\rangle\, f(2^{n+1}t - 2k - q)$$

$$= \sum_{j,l} \frac{1}{2} \sum_k (c_{j-2k}\, c_{l-2k} + (-1)^{j+l}\, c_{N-j+2k}\, c_{N-l+2k})$$
$$\times \langle h(t), 2^{(n+1)/2}\, f(2^{n+1}t - j)\rangle\, 2^{(n+1)/2}\, f(2^{n+1}t - l)$$

$$= \sum_{j,l} C(j,l)\, \langle h, f_{(n+1)j}\rangle\, f_{(n+1)l},$$

where

$$C(j,l) = \frac{1}{2} \sum_k (c_{j-2k}\, c_{l-2k} + (-1)^{j+l}\, c_{N-j+2k}\, c_{N-l+2k}).$$

It suffices, therefore, to show that $C(j,l) = \delta_{jl}$. Note that by making the change of index $m = -k + j + l - (N-1)/2$ (recall N is odd) in the second summation, we obtain

$$C(2j,2l) = \frac{1}{2} \sum_k c_{2j-2k}\, c_{2l-2k} + \frac{1}{2} \sum_m c_{2j-2m+1}\, c_{2l-2m+1} = \frac{1}{2} \sum_k c_{2j-k}\, c_{2l-k} = \delta_{jl},$$

because of hypothesis (1.4). Similar calculations show that $C(2j, 2l+1) = C(2j+1, 2l) = 0$ and $C(2j+1, 2l+1) = \delta_{jl}$, as desired.

4) From steps 2) and 3), $\sum_n F_n = \lim_{n\to\infty}(P_n - P_{-n}) = I$. That is, for $h \in L^2(\mathbb{R})$, $h = \sum_n F_n h = \sum_{n,k} \langle h, g_{nk}\rangle\, g_{nk}$, whence $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is a tight frame. ■

Corollary. The coefficients $\{c_k\}$ determine a multiresolution analysis if and only if

1) (1.3) and (1.4) are satisfied, and
2) f is orthogonal to each of its integer translates.

In this case, $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is an orthonormal basis for $L^2(\mathbb{R})$.

Proof. Because of Proposition 4.3 we need only prove that if 1) and 2) hold then $\{c_k\}$ determines a multiresolution analysis. From (1.4) and the orthogonality of the integer translates of f,

$$\|f\|_2^2 = \int |f(t)|^2\, dt = \sum_{j,k} c_j\, c_k \int f(2t-j)\, f(2t-k)\, dt = \frac{1}{2}\sum_k c_k^2 = 1.$$

Therefore $\{f(t-k)\}_{k\in\mathbb{Z}}$ is an orthonormal set and hence is an orthonormal basis for $V_0 = \operatorname{span}\{f(t-k)\}_{k\in\mathbb{Z}}$. Defining $V_n = \operatorname{span}\{f_{nk}\}_{k\in\mathbb{Z}}$, we have $V_n \subset V_{n+1}$ because f is a scaling function. The operator P_n defined by (4.3) is then the orthogonal projection of $L^2(\mathbb{R})$ onto V_n. Since $P_n \to 0$ as $n \to -\infty$ we have $\cap V_n = \{0\}$, and similarly $\cup V_n$ is dense in $L^2(\mathbb{R})$ since $P_n \to I$ as $n \to +\infty$. Thus $(\{V_n\}, f)$ is a multiresolution analysis.

To prove that $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is an orthonormal basis, note that from (1.2), (1.4), and the orthogonality of the integer translates of f,

$$\|g\|_2^2 = \sum_{j,k} (-1)^{j+k}\, c_{N-j}\, c_{N-k} \int f(2t-j)\, f(2t-k)\, dt = \frac{1}{2}\sum_k c_k^2 = 1.$$

From the theorem, we know that $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is a tight frame, so for $m, j \in \mathbb{Z}$ fixed,

$$\begin{aligned}
1 &= \|g_{mj}\|_2^2 \\
&= \langle g_{mj}, g_{mj}\rangle \\
&= \Big\langle g_{mj}, \sum_{n,k} \langle g_{mj}, g_{nk}\rangle\, g_{nk} \Big\rangle \\
&= \sum_{n,k} |\langle g_{mj}, g_{nk}\rangle|^2.
\end{aligned}$$

Thus $\langle g_{mj}, g_{nk}\rangle = \delta_{mn}\, \delta_{jk}$, i.e., $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ forms an orthonormal set. This, combined with the tight frame property, implies that $\{g_{nk}\}_{n,k\in\mathbb{Z}}$ is an orthonormal basis. ■

Lawton and Cohen have independently established necessary and sufficient conditions under which f will be orthogonal to its integer translates. Lawton's formulation is the following [Law90], [Law91]. ℓ^2 denotes the space of all square-summable sequences.

Theorem 4.5. Define the operator $G : \ell^2 \to \ell^2$ by

$$(Ga)_l = \frac{1}{2} \sum_{j,k} c_j \, c_k \, a_{2l+j-k} \quad \text{for } a \in \ell^2.$$

Then the coefficients $\{c_k\}$ determine a multiresolution analysis if and only if

1) (1.3) and (1.4) are satisfied, and
2) δ_{0l} is the only eigenvector for G for the eigenvalue 1.

Proof. Note that δ_{0l} is an eigenvector for G for the eigenvalue 1 because of (1.4), and the sequence a defined by $a_l = \int f(t) \, f(t+l) \, dt$ is also an eigenvector for G for the eigenvalue 1 since

$$\begin{aligned} (Ga)_l &= \frac{1}{2} \sum_{j,k} c_j \, c_k \int f(t) \, f(t - 2l + j - k) \, dt \\ &= \frac{1}{2} \int \Bigl(\sum_j c_j \, f(t-j) \Bigr) \Bigl(\sum_k c_k \, f(t - 2l - k) \Bigr) \, dt \\ &= \frac{1}{2} \int f\Bigl(\frac{t}{2}\Bigr) \, f\Bigl(\frac{t}{2} - l\Bigr) \, dt \\ &= a_l. \end{aligned}$$

Therefore, if δ_{0l} is the only eigenvector for G for the eigenvalue 1 then $a_l = c \, \delta_{0l}$ for some constant c, so f is orthogonal to its integer translates. The converse of this statement is proved in [Law91]. The proof is therefore complete by the corollary to Theorem 4.4. ■

Lawton has proved, using a result of Pollen [Pol], that except for a set of measure zero, coefficients which satisfy (1.3) and (1.4) also satisfy the condition that δ_{0l} be the only eigenvector for G for the eigenvalue 1. Therefore almost all choices of coefficients satisfying (1.3) and (1.4) will determine a multiresolution analysis.

Cohen's formulation, which has been shown to be equivalent to Lawton's, is the following [Coh90], cf., [Law91].

Theorem 4.6. The coefficients $\{c_k\}$ determine a multiresolution analysis if and only if

1) (1.3) and (1.4) are satisfied, and
2) there exists a $\gamma \in [-\pi/2, \pi/2]$ such that $\hat{f}(\gamma + 2k\pi) = 0$ for every $k \in \mathbb{Z}$.

Remark. For $N = 3$, the set of points satisfying (1.3) and (1.4) is the circle shown in Figure 1.1. Of these, every point with the single exception of the point $(1,1)$ does determine a multiresolution analysis [CHc].

5. Bibliography

[Bat87] G. Battle. A block spin construction of ondelettes. *Comm. Math. Phys.*, 110:601–615, 1987.

[Ber] M.A. Berger. Random affine iterated function systems: smooth curve generation. Preprint.

[BWa] M.A. Berger and Y. Wang. Bounded semi-groups of matrices. Preprint.

[BWb] M.A. Berger and Y. Wang. Multi-scale dilation equations and iterated function systems. Preprint.

[CDM] A. Cavaretta, W. Dahmen, and C.A. Micchelli. Stationary subdivision. *Memoirs Amer. Math. Soc.* To appear.

[CHa] D. Colella and C. Heil. The characterization of continuous, four-coefficient scaling functions and wavelets. *IEEE Trans. Inf. Th., Special issue on Wavelet Transforms and Multiresolution Signal Analysis.* To appear.

[CHb] D. Colella and C. Heil. Characterizations of scaling functions, I. Continuous solutions. Preprint.

[CHc] D. Colella and C. Heil. Characterizations of scaling functions, II. Distributional and functional solutions. Preprint.

[CMW] R. Coifman, Y. Meyer, and M.V. Wickerhauser. Signal compression with wave packets. Preprint.

[Coh90] A. Cohen. Ondelettes, analyses multirésolutions et filtres miroirs en quadrature. *Ann. Inst. H. Poincare*, 7:439–459, 1990.

[CW] R. Coifman and M.V. Wickerhauser. Best-adapted wave packet bases. Preprint.

[Dau88] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 41:909–996, 1988.

[Dau90] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Th.*, 1990.

[DD89] G. Deslauriers and S. Dubuc. Symmetric iterative interpolation processes. *Constr. Approx.*, 5:49–68, 1989.

[DGL91] N. Dyn, J.A. Gregory, and D. Levin. Analysis of uniform binary subdivision schemes for curve design. *Const. Approx.*, 7:127–147, 1991.

[DLa] I. Daubechies and J. Lagarias. Sets of matrices all infinite products of which converge. *Lin. Alg. Appl.* To appear.

[DLb] I. Daubechies and J. Lagarias. Two-scale difference equations: I. Gobal regularity of solutions. *SIAM J. Math. Anal.* To appear.

[DLc] I. Daubechies and J. Lagarias. Two-scale difference equations: II. Local regularity, infinite products and fractals. *SIAM J. Math. Anal.* To appear.

[DLd] N. Dyn and D. Levin. Interpolating subdivision schemes for the generation of curves and surfaces. Preprint.

[DS52] R.J. Duffin and A.C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952.

[Eir] T. Eirola. Sobolev characterization of solutions of dilation equations. *SIAM J. Math. Anal.* Submitted.

[Hei90] C. Heil. *Wiener amalgam spaces in generalized harmonic analysis and wavelet theory.* Ph.D. thesis, University of Maryland, College Park, MD, 1990.

[HW89] C. Heil and D. Walnut. Continuous and discrete wavelet transforms. *SIAM Rev.*, 31:628–666, 1989.

[Law90] W. Lawton. Tight frames of compactly supported affine wavelets. *J. Math. Phys.*, 31:1898–1901, 1990.

[Law91] W. Lawton. Necessary and sufficient conditions for constructing orthonormal wavelet bases. *J. Math. Phys.*, 32:57–61, 1991.

[Lem88] P.G. Lemarié. Ondelettes à localisation exponentielle. *Journ. de Math. Pures et Appl.*, 67:227–236, 1988.

[Mal89] S.G. Mallat. Multiresolution approximations and wavelet orthonormal bases for $L^2(R)$. *Trans. Amer. Math. Soc.*, 315:69–87, 1989.

[Mey86] Y. Meyer. Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. *Séminaire Bourbaki*, 662, 1985–1986.

[MP87] C.A. Micchelli and H. Prautzsch. Refinement and subdivision for spaces of integer translates of compactly supported functions. In C.F. Griffiths and G. A. Watson, editors, *Numerical Analysis*, pages 192–222, 1987.

[MP89] C.A. Micchelli and H. Prautzsch. Uniform refinement of curves. *Lin. Alg. Appl.*, 114/115:841–870, 1989.

[Pol] D. Pollen. $SU_I(2, F[z, 1/z])$ for F a subfield of C. Preprint.

[RS60] G.C. Rota and G. Strang. A note on the joint spectral radius. *Indag. Math.*, 22:379–381, 1960.

[Str89] G. Strang. Wavelets and dilation equations: a brief introduction. *SIAM Rev.*, 31:614–627, 1989.