Reprinted with permission of Kluwer Academic Publishers

J.S. Byrnes et al., eds., *Wavelets and Their Applications*, 19–38.

Some Stability Properties of Wavelets and scaling functions†

Christopher Heil
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332 USA
heil@math.gatech.edu

Abstract
The property of continuity of an arbitrary scaling function is known to be unstable with respect to the coefficients in the associated dilation equation. That is, if f is a continuous function which is a solution of the dilation equation $f(x) = \sum_{k=0}^{N} c_k f(2x-k)$ then a dilation equation with slightly perturbed coefficients $\{\tilde{c}_0, \dots, \tilde{c}_N\}$ need not have a continuous solution. The convergence of the Cascade Algorithm, an iterative method for solving a dilation equation, is likewise unstable in general. This paper establishes a condition under which stability does occur: both continuity and uniform convergence of the Cascade Algorithm are shown to be stable for those initial choices of coefficients $\{c_0, \dots, c_N\}$ such that the integer translates of the scaling function f are ℓ^∞ linearly independent. In particular, this applies to those scaling functions which can be used to construct orthogonal or biorthogonal wavelets. We show by example that this ℓ^∞ linear independence condition is not necessary for stability to occur.

1. Introduction

Orthogonal and biorthogonal wavelets have become important tools in harmonic analysis, signal processing, and other areas. In this article we bring together some recent results on the characterization of continuous compactly supported scaling functions and on the convergence of algorithms for constructing such scaling functions to conclude that scaling functions which determine continuous compactly supported wavelets satisfy certain stability properties with respect to the coefficients in the associated dilation equation. More generally, these results apply to any scaling function whose integer translates satisfy a certain type of linear independence.

†Supported by NSF Postdoctoral Research Fellowship DMS-9007212.

To explain what we mean by stability, let us first recall some definitions and basic results. An *orthogonal wavelet* is a square-integrable function g such that the collection $\{g_{nk}\}_{n,k\in\mathbf{Z}}$ forms an orthonormal basis for $L^2(\mathbf{R})$, where

$$g_{nk}(x) = 2^{n/2} g(2^n x - k) \tag{1.1}$$

and $L^2(\mathbf{R})$ denotes the space of all square-integrable functions on $\mathbf{R}$. A *biorthogonal wavelet* is a square-integrable function g for which there exists a *dual wavelet* $G \in L^2(\mathbf{R})$ so that $\{g_{nk}\}_{n,k\in\mathbf{Z}}$ and $\{G_{nk}\}_{n,k\in\mathbf{Z}}$ are Riesz bases for $L^2(\mathbf{R})$ which are dual in the sense that $\langle g_{nk}, G_{n'k'}\rangle = \delta_{nn'}\delta_{kk'}$, where δ_{ij} denotes the Kronecker delta. Riesz bases are generalizations of orthonormal bases: every function in the space can still be written as a unique combination of the basis elements and this summation converges unconditionally, but the basis elements need not be orthogonal.

A compactly supported orthogonal wavelet is typically constructed in the following manner. First, select real or complex coefficients $\{c_0, \dots, c_N\}$ satisfying

$$\sum_k c_{2k} = \sum_k c_{2k+1} = 1 \tag{1.2}$$

and

$$\sum_k c_k \bar{c}_{k+2j} = 2\delta_{0j}, \quad j \in \mathbf{Z}, \tag{1.3}$$

where we assume $c_k = 0$ if $k < 0$ or $k > N$. Next, solve the *dilation equation*

$$f(x) = \sum_{k=0}^{N} c_k f(2x - k). \tag{1.4}$$

Assumptions (1.1) and (1.3) guarantee that there will be a nontrivial square-integrable compactly supported solution f to the dilation equation, although this f need not be continuous. Moreover, f is unique up to multiplication by a constant and is supported in the interval $[0, N]$. If f is orthogonal to its integer translates, i.e., if $\int f(x)\overline{f(x+k)}\,dx = \delta_{0k}$, then $\{V_n\}_{n\in\mathbf{Z}}$ will form a *multiresolution analysis* for $L^2(\mathbf{R})$, where

$$V_n = \operatorname{span}\{f(2^n x - k)\}_{k\in\mathbf{Z}}. \tag{1.5}$$

In this case we refer to f as an *orthogonal scaling function.* We then obtain the orthogonal wavelet g by the recipe

$$g(x) = \sum_k (-1)^k c_{N-k} f(2x - k). \tag{1.6}$$

Details of this construction are in the expository survey [12].

The construction of biorthogonal wavelets is similar, except that we begin with two different sets of coefficients $\{c_0, \dots, c_N\}$ and $\{c'_0, \dots, c'_{N'}\}$. These coefficients must each satisfy (1.2) and also

$$\sum_k c'_k \bar{c}_{k+2j} = 2\delta_{0j}, \quad j \in \mathbf{Z}. \tag{1.7}$$

Unlike the orthogonal case, these assumptions do not guarantee the existence of compactly supported square-integrable solutions f, F to the dilation equations determined by $\{c_0, \dots, c_N\}$ and $\{c'_0, \dots, c'_{N'}\}$, respectively. If such solutions do exist, and if

$$\int F(x)\overline{f(x+k)}\, dx = \delta_{0k}, \quad k \in \mathbf{Z}, \tag{1.8}$$

then g defined by (1.6) will be a biorthogonal wavelet, with dual wavelet G derived analogously from F. In this case we refer to f as an *biorthogonal scaling function* and to F as its *dual*. The details of this construction are considerably more involved than in the orthogonal case [6, 8].

Dilation equations with coefficients satisfying (1.2), but not necessarily (1.3) or (1.7), have applications to several areas outside of wavelet theory. For example, they are used in subdivision schemes in computer aided graphics to design curves passing through or near a given set of points [4]. We will refer to any nontrivial compactly supported solution of a dilation equation simply as a *scaling function*. We will always assume in this article that the coefficients $\{c_0, \dots, c_N\}$ of a dilation equation satisfy (1.2) and, to avoid trivialities, that c_0, $c_N \neq 0$. Even for such general choices of coefficients it is still true that if an integrable scaling function exists then it is unique up to multiplication by a constant and is supported in $[0, N]$. However, existence is not guaranteed; some methods for determining which dilation equations have solutions can be found in [15]. If an integrable scaling function exists then it can be shown that (1.2) implies $\sum f(x+k) = C$ a.e. with C a nonzero constant. We therefore can always assume that f is normalized so that

$$\sum_k f(x+k) = 1 \text{ a.e.} \tag{1.9}$$

Since a compactly supported wavelet g is obtained from a compactly supported scaling function f by the finite summation (1.6), it inherits properties such as smoothness from f. The scaling function is itself determined by the coefficients $\{c_0, \dots, c_N\}$, and it is therefore natural to investigate the dependence of the properties of f (and hence of g) on the choice of coefficients. In particular, we wish to know whether the property of continuity is stable under small pertubations of the coefficients. Closely related is the question of whether numerical algorithms for constructing f are stable. For arbitrary choices of coefficients satisfying (1.2), it has recently been shown that continuity is unstable [8]. However, in this article we prove that both continuity and uniform convergence of the Cascade Algorithm (an iterative method for solving dilation equations defined below) are stable when the initial choice of coefficients is one that results in an orthogonal or biorthogonal wavelet, and that in this case the corresponding scaling functions deform uniformly as the coefficients vary. In fact, we obtain this result with the weaker assumption that the integer translates of f satisfy the following kind of linear independence.

Definition 1.1. Let ℓ^∞ denote the space of all bounded sequences $\{a_k\}_{k\in\mathbf{Z}}$. Then the integer translates $\{h(x+k)\}_{k\in\mathbf{Z}}$ of a function h are said to be *ℓ^∞ linearly independent* if the only sequence $\{a_k\}_{k\in\mathbf{Z}} \in \ell^\infty$ such that $\sum a_k h(x+k) = 0$ is $\{a_k\}_{k\in\mathbf{Z}} = 0$ (i.e., the zero sequence $a_k = 0$ for every k).

A similar definition can be made for ℓ^p linear independence. If $\sum a_k h(x+k) = 0$ implies $\{a_k\}_{k\in\mathbf{Z}} = 0$ for arbitrary sequences $\{a_k\}_{k\in\mathbf{Z}}$ of complex numbers then we say simply that $\{h(x+k)\}_{k\in\mathbf{Z}}$ is *linearly independent*. The implications of these and other types of linear independence of translates have been investigated by several authors, e.g., [4, 5, 16, 17]. We note that there is a closely related concept often referred to as "ℓ^p stability;" however, the use of the term "stability" in that context is unrelated to our use denoting the preservation of some property of a scaling function when the coefficients in the dilation equation are perturbed.

Note that $\{h(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^2 linearly independent for every $h \in L^2(\mathbf{R})$ which is compactly supported; in Example 2.7 we give examples of compactly supported $h \in L^2(\mathbf{R})$ for which $\{h(x+k)\}_{k\in\mathbf{Z}}$ is linearly independent or ℓ^∞ linearly independent.

Our main result can now be stated as follows, where $L^\infty(\mathbf{R})$ denotes the space of all essentially bounded functions with norm $\|h\|_\infty = \text{ess sup}\,|h(x)|$.

Theorem 1.2. *Assume f is a continuous scaling function for a dilation equation determined by coefficients $\{c_0, \dots, c_N\}$ satisfying (1.2). If $\{f(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^∞ linearly independent then the Cascade Algorithm for f converges uniformly, and both the property of continuity and the convergence of the Cascade Algorithm are stable. That is, for every $\varepsilon > 0$ there is a $\delta > 0$ such that if $\{\tilde{c}_0, \dots, \tilde{c}_N\}$ satisfies (1.2) and*

$$\max_{i=0,\dots,N} |c_i - \tilde{c}_i| < \delta \tag{1.10}$$

then:

1) *there exists a continuous scaling function $\tilde{f}$ for the dilation equation determined by $\{\tilde{c}_0, \dots, \tilde{c}_N\}$,*

2) the Cascade Algorithm for $\tilde{f}$ converges uniformly, and
3) $\|f - \tilde{f}\|_\infty < \varepsilon$.

Although ℓ^∞ linear independence is sufficient to ensure stability, we show in Example 2.7 that it is not necessary. That Theorem 1.2 applies to scaling functions used to construct wavelets is established as follows.

Lemma 1.3. If f is an orthogonal or biorthogonal scaling function then $\{f(x+k)\}_{k\in\mathbf{Z}}$ is linearly independent.

Proof. Let F be the dual of f ($F = f$ if f is orthogonal). Assume $\{a_k\}_{k\in\mathbf{Z}}$ is any sequence such that $\sum a_k f(x+k) = 0$. Since f and F have compact support, we can interchange the summation and integration in the following calculation:

$$\bar{a}_j = \sum_k \bar{a}_k \int_{-\infty}^{\infty} F(x+j)\overline{f(x+k)}\,dx = \int_{-\infty}^{\infty} F(x+j)\overline{\sum_k a_k f(x+k)}\,dx = 0, \tag{1.11}$$

the first equality following by (1.8). □

We emphasize that the pertubations allowed in Theorem 1.2 are not completely arbitrary; in particular, they must preserve (1.2). However, we do not assume that either (1.3) or (1.7) is preserved. Therefore, it will not in general be the case that a pertubation $\tilde{f}$ of an orthogonal scaling function f will itself be orthogonal. On the other hand, it is not immediately evident whether the property of biorthogonality is preserved by pertubations; Cohen and Daubechies have recently proved that it is [5].

The proof of Theorem 1.2 occupies all of Section 2, and we obtain it by bringing together results on two methods of studying dilation equations. The first is a matrix-oriented approach which was independently introduced by Micchelli and Prautzsch [18] and Daubechies and Lagarias [11], and used recently by Colella and the author to characterize all dilation equations for which continuous scaling functions exist and to prove that continuity is in general an unstable property [8]. This characterization also precisely bounds the range of Hölder exponents of a continuous scaling function, where we say that a function h is Hölder continuous with exponent α if there exists a constant C_α such that $|h(x) - h(y)| \le C_\alpha |x-y|^\alpha$ for all x, y. We refer to

$$\alpha_{\max} = \sup\{\alpha : h \text{ is Hölder continuous with exponent } \alpha\} \tag{1.12}$$

as the *maximum Hölder exponent* of h, although we note that this supremum need not be attained.

The second method is an iterative algorithm for constructing a scaling function, based on the fact that f is a fixed point of the linear operator

$$Sh(x) = \sum_{k=0}^{N} c_k h(2x-k). \tag{1.13}$$

Therefore, beginning with some "reasonable" function f_0, we define

$$f_{j+1}(x) = Sf_j(x) = \sum_{k=0}^{N} c_k f_j(2x-k). \tag{1.14}$$

When

$$f_0(x) = \chi_{[-1/2,1/2)}(x) = \begin{cases} 1, & -1/2 \le x < 1/2, \\ 0, & \text{otherwise}, \end{cases} \tag{1.15}$$

we call this iteration the *Cascade Algorithm* for computing f. We say that the Cascade Algorithm *converges uniformly* if $\|f - f_j\|_\infty \to 0$ as $j \to \infty$. Berger and Wang have proved results linking the Cascade Algorithm to the matrix methods mentioned above; in particular, they independently obtained a characterization of all continuous compactly supported scaling functions [1, 20].

The Cascade Algorithm plays an important role in applications of dilation equations to subdivision schemes [4], although it is usually appears there in a discrete form. That is, the iteration begins with a function f_0 which can be considered to be defined only at integer points; the first iteration then constructs an f_1 which updates these values and "interpolates" values at the half-integer points, and so forth. These two formulations of the algorithm are equivalent, and where we have used results from the literature of subdivision schemes we have recast them into the form of the Cascade Algorithm defined above.

2. A proof of the stability result

We begin by showing how a matrix-oriented technique for studying continuous scaling functions, independently introduced by Micchelli and Prautzsch [18] and Daubechies and Lagarias [11] (and extended by other groups, e.g., [8, 1]), can be used to characterize continuous solutions of dilation equations.

Assume that coefficients $\{c_0, \dots, c_N\}$ are given, and that an associated continuous scaling function f exists. Since $\mathrm{supp}(f) \subset [0, N]$, define $v(x)$ for each $x \in [0,1]$ as the vector

$$v(x) = \begin{pmatrix} f(x) \\ f(x+1) \\ \vdots \\ f(x+N-1) \end{pmatrix}. \tag{2.1}$$

Clearly v and f contain "equivalent" information. By considering the dilation equation at the points $f(x), \ldots, f(x+N-1)$ in turn and using the fact that $f(x) = 0$ if $x \leq 0$ or $x \geq N$, we find that if $x \in [0, 1/2]$ then the value of $v(x)$ is obtained from the value of $v(2x)$ in a fixed linear manner. Hence there is an $N \times N$ matrix T_0 such that

$$v(x) = T_0 v(2x), \quad x \in [0, 1/2]. \tag{2.2}$$

Similarly, there is a matrix T_1 such that

$$v(x) = T_1 v(2x-1), \quad x \in [1/2, 1]. \tag{2.3}$$

In fact, $(T_0)_{ij} = c_{2i-j-1}$ and $(T_1)_{ij} = c_{2i-j}$. Note that there is consistency at $x = 1/2$, i.e.,

$$v(1/2) = T_0 v(1) = T_1 v(0). \tag{2.4}$$

Letting τ denote the operator

$$\tau x = 2x \bmod 1 = \begin{cases} 2x, & 0 \leq x < 1/2, \\ 2x-1, & 1/2 < x \leq 1, \end{cases} \tag{2.5}$$

we can summarize (2.2)–(2.4) as

$$v(x) = T_{d_1} v(\tau x), \quad x = .d_1 d_2 \cdots \in [0, 1], \tag{2.6}$$

where $x = .d_1 d_2 \cdots$ denotes any binary expansion of x (note that $\tau x = .d_2 d_3 \cdots$). When $x = 1/2$ there is ambiguity in (2.6), but we resolve this by allowing either of the two expansions $1/2 = .100\cdots$ or $1/2 = .011\cdots$ to be used as long as T_{d_1} and $\tau(1/2)$ are then interpreted consistently. Points x which have nonunique binary expansions will be called *dyadic*; they have the form $x = k/2^j$ where $k, j \in \mathbf{Z}$. Each dyadic point has one binary expansion ending in infinitely many zeros, and one expansion ending in infinitely many ones. For simplicity, we write the former expansion for dyadic $x \in (0, 1)$ as $x = .d_1 \cdots d_m$.

Assume now that $x = .d_1 d_2 \cdots \in [0, 1]$ is arbitrary, and that y is "close" to x. If x and y are close enough (and their binary expansions are chosen appropriately if one or both is dyadic) then they will share the same first few digits in their binary expansions. So, we can write $y = .d_1 \cdots d_m d'_{m+1} d'_{m+2} \cdots$ for some $m > 0$. Applying (2.4) recursively, we find

$$v(y) - v(x) = T_{d_1}(v(\tau y) - v(\tau x)) = \cdots = T_{d_1} \cdots T_{d_m}(v(\tau^m y) - v(\tau^m x)). \tag{2.7}$$

Since $y \to x$ as $m \to \infty$, this implies that all products $T_{d_1} \cdots T_{d_m}$ must converge to zero as $m \to \infty$, at least when restricted to differences $v(w) - v(z)$.

Since v is continuous it suffices to consider dyadic w, z, i.e., we can restrict $T_{d_1} \cdots T_{d_m}$ to the subspace

$$W = \operatorname{span}\{v(w) - v(z) : \text{dyadic } w, z \in [0,1]\}. \tag{2.8}$$

Note that (2.6) implies that W is right-invariant under both T_0 and T_1. Some explicit methods for determining W are given in [8].

Now, v and W are determined by the values that v takes on the set of dyadic points in $[0,1]$. For a typical dyadic $x = .d_1 \cdots d_m \in (0,1)$, we have by (2.6) that

$$v(x) = T_{d_1} \cdots T_{d_m} v(\tau^m x) = T_{d_1} \cdots T_{d_m} v(0). \tag{2.9}$$

Moreover, by the continuity of f we have $v(0) = (0, f(1), \dots, f(N-1))^{\mathrm{t}}$ and $v(1) = (f(1), \dots, f(N-1), 0)^{\mathrm{t}}$. Thus v and W are completely determined by the vector $a = (f(1), \dots, f(N-1))^{\mathrm{t}}$. Note that (2.6) implies that $v(0)$ is a right 1-eigenvector for T_0. Writing T_0 in block form as

$$T_0 = \begin{pmatrix} c_0 & 0 \\ * & M \end{pmatrix} \tag{2.10}$$

with M the $(N-1) \times (N-1)$ submatrix $M_{ij} = c_{2i-j}$, it follows that a is a right 1-eigenvector for M.

The above analysis shows that if f is a continuous scaling function then all restricted products $(T_{d_1} \cdots T_{d_m})|_W$ converge to zero as $m \to \infty$, and that W is determined by an appropriate right 1-eigenvector a for the matrix M. In the converse direction, it can be shown that if there exists some right 1-eigenvector a for M such that every $(T_{d_1} \cdots T_{d_m})|_W \to 0$, then there exists a continuous scaling function f. Specifically, the construction is as follows. Let $a = (a_1, \dots, a_{N-1})^{\mathrm{t}}$ be any right 1-eigenvector for M and define $v(0) = (0, a_1, \dots, a_{N-1})^{\mathrm{t}}$ and $v(1) = (a_1, \dots, a_{N-1}, 0)^{\mathrm{t}}$. Note that such an a always exists since (1.2) implies that $(1, \dots, 1)$ is a left 1-eigenvector for M. For dyadic $x = .d_1 \cdots d_m \in (0,1)$ define $v(x)$ by (2.9), and then define W by (2.8). Note that W is right-invariant under both T_0 and T_1 by construction, and that it depends on the choice of eigenvector a. We have now defined v on the dyadics, a dense subset of $[0,1]$. If it is the case that every $(T_{d_1} \cdots T_{d_m})|_W \to 0$ as $m \to \infty$ then it follows that v is uniformly continuous on this dense set. Therefore v can be extended continuously to all of $[0,1]$. An obvious "unfolding" of v then gives a continuous scaling function f. With a little attention to detail the maximum Hölder exponent of continuity of f can also be determined by this method.

In summary, a continuous scaling function f exists if and only if there exists a right 1-eigenvector a for M such that every $(T_{d_1} \cdots T_{d_m})|_W \to 0$ as $m \to \infty$, where W is determined by the choice of a. This convergence property can be recast as a "joint" spectral property of $T_0|_W$ and $T_1|_W$, as follows.

First, recall that for a single matrix A, powers A^m of A converge to zero as $m \to \infty$ if and only if $\rho(A) < 1$, where $\rho(A)$ is the spectral radius of A:

$$\rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\} = \limsup_{m\to\infty} \|A^m\|^{1/m}. \qquad (2.11)$$

This value is independent of the choice of norm $\|\cdot\|$.

For two matrices A_0, A_1, we follow Rota and Strang [19] and define their *joint spectral radius* $\hat\rho(A_0, A_1)$ as

$$\hat\rho(A_0, A_1) = \limsup_{m\to\infty} \max_{d_j=0,1} \|A_{d_1} \cdots A_{d_m}\|^{1/m}. \qquad (2.12)$$

It can be shown that all products $A_{d_1} \cdots A_{d_m}$ converge to zero as $m \to \infty$ if and only if $\hat\rho(A_0, A_1) < 1$ [2]. Unlike the spectral radius of a single matrix, the joint spectral radius can be difficult to compute. For example,

$$\hat\rho(A_0, A_1) \geq \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A_0 \text{ or } A_1\}, \qquad (2.13)$$

but strict inequality may occur. Equality occurs, for example, if A_0, A_1 commute, are both upper-triangular, or are both Hermitian. Some methods of evaluating or approximating a joint spectral radius can be found in [13].

The joint spectral radius of matrices A_0, A_1 restricted to a subspace U which is invariant under both A_0 and A_1 is defined analogously. By an appropriate change of basis we can always write $\hat\rho(A_0|_U, A_1|_U) = \hat\rho(B_0, B_1)$, where B_0, B_1 are $P \times P$ matrices with $P = \dim(U)$.

Thus we can formalize a characterization of all continuous scaling functions as follows [82]. Recall that each continuous scaling function f determines a subspace W and a right 1-eigenvector a for the matrix M; conversely, each such 1-eigenvector a for M determines a subspace W.

Theorem 2.1. Given coefficients $\{c_0, \ldots, c_N\}$ satisfying (1.2). Then there exists a continuous scaling function f for the dilation equation determined by $\{c_0, \ldots, c_N\}$ if and only if

$$\hat\rho(T_0|_W, T_1|_W) < 1 \qquad (2.14)$$

for some right 1-eigenvector a for M. Moreover, in this case f is Hölder continuous with maximum Hölder exponent $\alpha_{\max} = -\log_2 \hat\rho(T_0|_W, T_1|_W)$.

In general, the subspace W depends implicitly on the choice of eigenvector a. However, for most coefficients the eigenvalue 1 for M will be simple, in which case W is independent of a. By a slight abuse of terminology we therefore usually refer to W as if it was unique. In any case, since integrable scaling functions are unique up to multiplication by a constant there can be at most one a such that $\hat\rho(T_0|_W, T_1|_W) < 1$.

Theorem 2.1 extends to coefficients not satisfying (1.2) [8] and generalizes to a characterization of n-times differentiable scaling functions [13]. Compactly supported infinitely differentiable scaling functions are impossible [10]. Although Theorem 2.1 considers continuity and Hölder exponents only in a global sense, one of the main advantages of this matrix approach to the study of dilation equations over other approaches is its ability to characterize the local behavior of scaling functions. This was investigated in detail in [11].

To apply Theorem 2.1 to the proof of Theorem 1.2, we must examine the stability of the property $\hat{\rho}(T_0|_W, T_1|_W) < 1$ with respect to changes in the coefficients. Consider therefore the following result, proved in [14].

Proposition 2.2. Given $N \times N$ matrices A_0, A_1, the joint spectral radius $\hat{\rho}(A_0, A_1)$ is a continuous function of the entries of A_0 and A_1.

Since the entries of T_0 and T_1 consist only of the coefficients $\{c_0, \ldots, c_N\}$, it is tempting to conclude that $\hat{\rho}(T_0|_W, T_1|_W)$ is a continuous function of $\{c_0, \ldots, c_N\}$. In fact this is not the case, and counterexamples were constructed in [8], as follows.

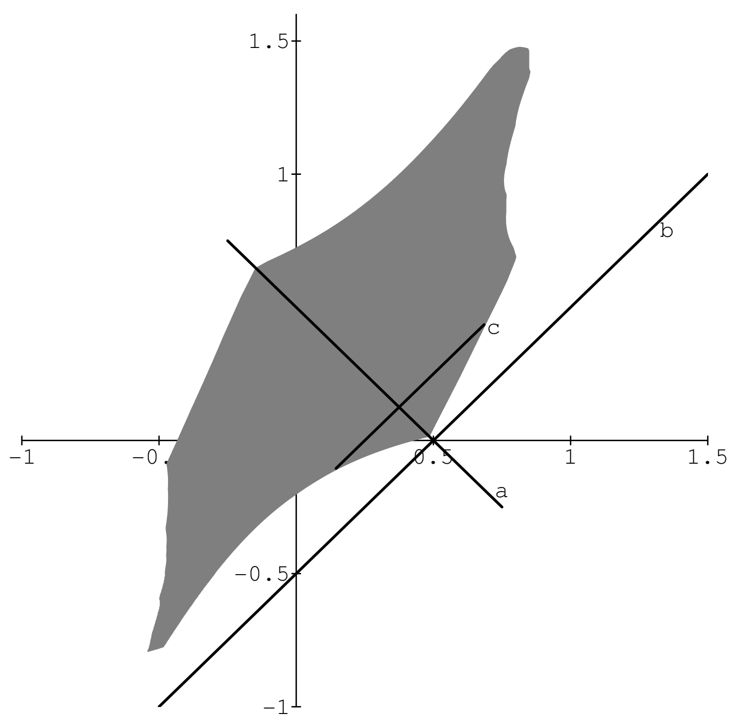

Figure 2.1: The (c_0, c_1)-plane, identified with symmetric real-valued seven-coefficient dilation equations.

Example 2.3. Set $N = 6$, and consider the class of real-valued coefficients $\{c_0, \ldots, c_6\}$ which satisfy (1.2) and are symmetric, i.e., $c_0 = c_6$, $c_1 = c_5$, and $c_2 = c_4$. This leaves only two free parameters, which we select to be c_0 and c_1. We identify this class of coefficients with the (c_0, c_1)-plane;

this plane is depicted in Figure 2.1. Continuous scaling functions exist and $\hat{\rho}(T_0|_W, T_1|_W) < 1$ for dilation equations along line segment a, with $\hat{\rho}(T_0|_W, T_1|_W) = 1/2$ at the point $(1/2, 0)$. Continuous scaling functions do not exist and $\hat{\rho}(T_0|_W, T_1|_W) > 1$ for dilation equations on line segment b, except at the intersection point $(1/2, 0)$ with line a. As a consequence, $\hat{\rho}(T_0|_W, T_1|_W)$ is discontinuous at $(1/2, 0)$. Continuous scaling functions also exist along line segment c; however, both $\hat{\rho}(T_0|_W, T_1|_W)$ and the maximum Hölder exponent $\alpha_{\max}$ are discontinuous at the intersection point $(3/8, 1/8)$ with line a.

The failure of $\hat{\rho}(T_0|_W, T_1|_W)$ to be continuous as a function of $\{c_0, \dots, c_N\}$ is due to the fact that the subspace W depends explicitly on the coefficients $\{c_0, \dots, c_N\}$ and therefore can change dimension abruptly as the coefficients vary. Thus, although we can write $\hat{\rho}(T_0|_W, T_1|_W) = \hat{\rho}(B_0, B_1)$, where B_0, B_1 are $P \times P$ matrices with $P = \dim(W)$, the value of P depends on $\{c_0, \dots, c_N\}$, so Proposition 2.2 does not apply. For example, in Example 2.3 we have $\dim(W) = 1$ at the point $(1/2, 0)$, $\dim(W) = 3$ on line segment a except at $(1/2, 0)$, and $\dim(W) = 5$ on line segments b and c except at the two intersection points with line a.

Although W depends on $\{c_0, \dots, c_N\}$, we can embed it in a larger subspace which is independent of the coefficients. In particular, note that (1.2) implies that $(1, \dots, 1)$ is a common left 1-eigenvector of T_0 and T_1. Therefore, if $x = .d_1 \cdots d_m \in (0, 1)$ is dyadic then

$$(1, \dots, 1)v(x) = (1, \dots, 1)T_{d_1} \cdots T_{d_m} v(0) = (1, \dots, 1)v(0). \tag{2.15}$$

Hence

$$W \subset V = \{u \in \mathbf{C}^N : u_1 + \cdots + u_N = 0\}. \tag{2.16}$$

Note that $\dim(V) = N - 1$, and that V is right-invariant under both T_0 and T_1 since it is the orthogonal complement of $(1, \dots, 1)$ in $\mathbf{C}^N$. Clearly $\hat{\rho}(T_0|_W, T_1|_W) \le \hat{\rho}(T_0|_V, T_1|_V)$, so the condition

$$\hat{\rho}(T_0|_V, T_1|_V) < 1 \tag{2.17}$$

is sufficient, although not necessary, to imply the existence of a continuous scaling function f. Moreover, this condition is stable since the independence of V from the coefficients implies by Proposition 2.2 that $\hat{\rho}(T_0|_V, T_1|_V)$ is a continuous function of $\{c_0, \dots, c_N\}$. For example, we can write $\hat{\rho}(T_0|_V, T_1|_V) = \hat{\rho}(S_0, S_1)$, where S_0, S_1 are the $(N-1) \times (N-1)$ matrices formed by computing

$$BT_iB^{-1} = \begin{pmatrix} 1 & 0 \\ * & S_i \end{pmatrix}, \quad i = 0, 1, \tag{2.18}$$

with

$$B = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \tag{2.19}$$

$$B^{-1} = \begin{pmatrix} 1 & -1 & -1 & \cdots & -1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{2.20}$$

Note that the entries of S_0, S_1 are simple linear combinations of the coefficients $\{c_0, \dots, c_N\}$. Returning to the class of dilation equations discussed in Example 2.3, the shaded region in Figure 2.1 is a numerical appoximation of $\{(c_0, c_1) : \hat{\rho}(T_0|_V, T_1|_V) < 1\}$. Note that the instability of the condition $\hat{\rho}(T_0|_W, T_1|_W) < 1$ occurs on the portion of line segment a which lies outside this region.

Thus, if we could establish that for a continuous scaling function f,

$$\{f(x+k)\}_{k \in \mathbf{Z}} \text{ is } \ell^\infty \text{ linearly independent} \implies \hat{\rho}(T_0|_V, T_1|_V) < 1, \tag{2.21}$$

then we would have proved part of Theorem 1.2. Another part of the theorem would follow from the next result, which was proved in [8].

Proposition 2.4. Given coefficients $\{c_0, \dots, c_N\}$ satisfying (1.2), such that $\hat{\rho}(T_0|_V, T_1|_V) < 1$. Let f be the associated continuous scaling function. Then for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\{\tilde{c}_0, \dots, \tilde{c}_N\}$ satisfies (1.2) and (1.10) then there exists a continuous scaling function $\tilde{f}$ for the dilation equation determined by $\{\tilde{c}_0, \dots, \tilde{c}_N\}$, such that $\|f - \tilde{f}\|_\infty < \varepsilon$.

Moreover, if (2.21) is true for continuous f then the final conclusion of Theorem 1.2 would follow from the next result linking the joint spectral radius and the Cascade Algorithm, proved by Berger and Wang [1].

Theorem 2.5. If the coefficients $\{c_0, \dots, c_N\}$ satisfy (1.2), then the Cascade Algorithm converges uniformly to a continuous scaling function f if and only if $\hat{\rho}(T_0|_V, T_1|_V) < 1$.

Our goal therefore is to establish (2.21) for continuous scaling functions f. Now, by Theorem 2.1, if f is continuous then $\hat{\rho}(T_0|_W, T_1|_W) < 1$, so if it was true that ℓ^∞ linear independence implied $W = V$ then (2.21) would be proved. However, this is false; the next proposition shows that $W = V$ is equivalent to linear independence, not ℓ^∞ linear independence. In Example 2.7 we give specific examples of scaling functions f such that $\{f(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^∞ linearly independent but $W \neq V$.

Proposition 2.6. Assume f is a continuous scaling function for a dilation equation determined by coefficients $\{c_0, \dots, c_N\}$ satisfying (1.2). Define

$$\mathcal{W} = \operatorname{span}\{v(x) : x \in [0,1]\}. \tag{2.22}$$

Then the following statements are equivalent.

a: $W = V$.
b: $\mathcal{W} = \mathbf{C}^N$.
c: $\{f(x+k)\}_{k\in\mathbf{Z}}$ is linearly independent.

Proof. Item a $\iff$ Item b. Note first that W can be written

$$W = \operatorname{span}\{v(x) - v(0) : x \in [0,1]\}. \tag{2.23}$$

In particular, the restriction to dyadic x can be dropped since v is continuous. Therefore,

$$\mathcal{W} = W + \operatorname{span}\{v(0)\}. \tag{2.24}$$

Since $W \subset V$ and $\dim(V) = N-1$, the equivalence between Item a and Item b will follow if we show $v(0) \notin V$. However, by (1.9) and the compact support of f we have $v_1(x) + \cdots + v_N(x) = 1$ for every $x \in [0,1]$. Taking $x = 0$, we conclude $v(0) \notin V$, as desired.

Item b $\implies$ Item c. Assume that $\mathcal{W} = \mathbf{C}^N$, and that $\sum a_k f(x+k) = 0$ for some sequence $\{a_k\}_{k\in\mathbf{Z}}$. Then the fact that $\operatorname{supp}(f) \subset [0,N]$ implies

$$a_k f(y+k) + \cdots + a_{k+N-1} f(y+k+N-1) = 0, \quad y \in [-k, -k+1], \tag{2.25}$$

for every $k \in \mathbf{Z}$. Hence

$$(a_k, \dots, a_{k+N-1}) \cdot v(x) = 0, \quad x \in [0,1]. \tag{2.26}$$

Since $\mathcal{W} = \mathbf{C}^N$ we must therefore have $a_k = \cdots = a_{k+N-1} = 0$ for every k, so $\{f(x+k)\}_{k\in\mathbf{Z}}$ is linearly independent.

Item c $\implies$ Item b. Assume $\mathcal{W} \neq \mathbf{C}^N$; then the subspace

$$\mathcal{W}^\perp = \{u \in \mathbf{C}^N : u \cdot w = 0 \text{ for all } w \in \mathcal{W}\} \tag{2.27}$$

is nontrivial. Let $(a_1, \dots, a_N) \in \mathcal{W}^\perp$ be nonzero, and consider the line

$$\ell = \{(a_2, \dots, a_N, t) : t \in \mathbf{C}\}. \tag{2.28}$$

If ℓ does not intersect $\mathcal{W}^\perp$ then $(0,\ldots,0,t) \in \mathcal{W}^\perp$ for every t. From this it follows that $v_N(x) = 0$ for every $x \in [0,1]$, i.e., $f(x) = 0$ for $x \in [N-1,N]$. However, this is impossible since $c_N \neq 0$. Therefore, there must be some $a_{N+1} \in \mathbf{C}$ such that $(a_2,\ldots,a_N,a_{N+1}) \in \mathcal{W}^\perp$. Similarly, since $c_0 \neq 0$ there exists an $a_0 \in \mathbf{C}$ such that $(a_0,a_1,\ldots,a_{N-1}) \in \mathcal{W}^\perp$. By induction, we obtain a sequence $\{a_k\}_{k\in\mathbf{Z}}$ such that (2.26) holds. Therefore $\sum a_k f(x+k) = 0$, so $\{f(x+k)\}_{k\in\mathbf{Z}}$ is linearly dependent. □

In particular, Proposition 2.6 is sufficient to prove Theorem 1.2 if we require only linear independence of $\{f(x+k)\}_{k\in\mathbf{Z}}$ rather than ℓ^∞ linear independence. We will complete the full proof of Theorem 1.2 after the following example, which shows that ℓ^∞ linear independence does not imply $W = V$, even if $\hat\rho(T_0|_V, T_1|_V) < 1$.

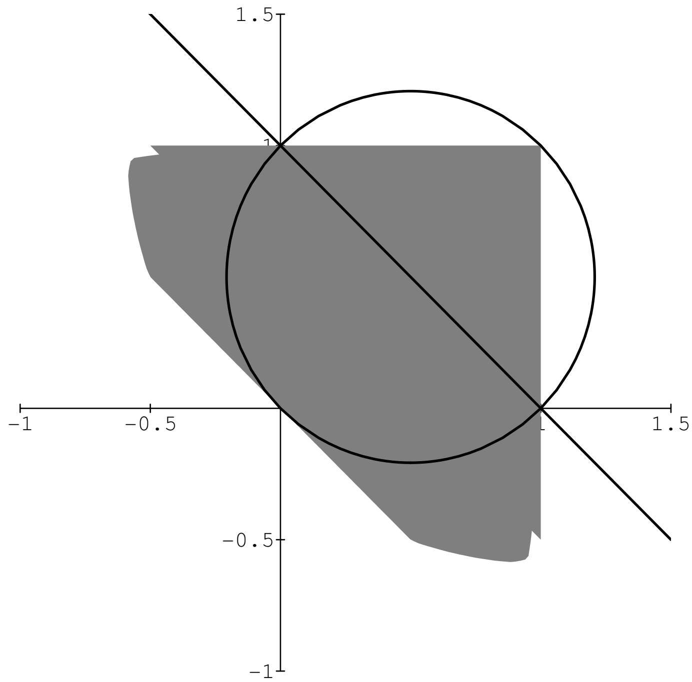

Figure 2.2. The (c_0,c_3)-plane, identified with real-valued four-coefficient dilation equations.

Example 2.7. Set $N = 3$, and consider the class of real-valued coefficients $\{c_0,c_1,c_2,c_3\}$ which satisfy (1.2). This leaves only two free parameters, which we select to be c_0 and c_3. We identify this class of coefficients with the (c_0,c_3)-plane; this plane is depicted in Figure 2.2. It is shown in [7] that $W = V$ for all dilation equations except those satisfying $1 - c_0 - c_3 = 0$, and that $\hat\rho(T_0|_W, T_1|_W) = \hat\rho(T_0|_V, T_1|_V)$ even when $W \neq V$. This latter fact results from the small number of coefficients involved; compare Example 2.3 where $N = 6$ and $\hat\rho(T_0|_W, T_1|_W) \neq \hat\rho(T_0|_V, T_1|_V)$ occurs. The shaded region in Figure 2.2 is a numerical approximation of the region where $\hat{\rho}(T_0|_V, T_1|_V) < 1$; continuous scaling functions therefore occur within this area. For comparison, the line is the set of coefficients such that $1 - c_0 - c_3 = 0$, and the circle consists of those coefficients for which (1.3) is satisfied. Recall that (1.3) is a necessary condition for a scaling function to be orthogonal; in fact, all of the points on the circle do produce orthogonal scaling functions except for the single point $(1, 1)$. The scaling function for this exceptional point is $f = (1/3)\chi_{[0,3]}$; note that the integer translates of this f are ℓ^∞ linearly dependent.

Consider now those coefficients satisfying $1 - c_0 - c_3 = 0$. In this case $\dim(W) = 1$ while $\dim(V) = 2$. With $\mathcal{W}$ as in Proposition 2.6 we have $\dim(\mathcal{W}) = 2$. One basis for $\mathcal{W}$ is $\{v(0), v(1)\}$. Since $v(0) = (0, c_0, c_3)$ and $v(1) = (c_0, c_3, 0)$, the space $\mathcal{W}^\perp$ is one-dimensional and is spanned by $u = (-c_3/c_0, 1, -c_0/c_3)$. Fix now any c_0 in the range $0 < c_0 < 1$; then $\hat{\rho}(T_0|_V, T_1|_V) < 1$ and a continuous scaling function f exists. Assume that $\sum a_k f(x+k) = 0$ for some sequence $\{a_k\}_{k\in\mathbf{Z}}$. Then the fact that $\mathrm{supp}(f) \subset [0, 3]$ implies

$$a_k f(y+k) + a_{k+1} f(y+k+1) + a_{k+2} f(y+k+2) = 0, \quad y \in [-k, -k+1]. \tag{2.29}$$

Hence $(a_k, a_{k+1}, a_{k+2}) \cdot v(x) = 0$ for all $x \in [0, 1]$, so $(a_k, a_{k+1}, a_{k+2}) \in \mathcal{W}^\perp$. Therefore $a_k = -(c_3/c_0) a_{k+1}$ and $a_{k+2} = -(c_0/c_3) a_{k+1}$, so by recursion we obtain

$$a_k = \left(-\frac{c_0}{c_3}\right)^k a_0, \quad k \in \mathbf{Z}. \tag{2.30}$$

Conversely, if $\{a_k\}_{k\in\mathbf{Z}}$ satisfies (2.30) then $\sum a_k f(x+k) = 0$. Thus the integer translates of each of these scaling functions are linearly dependent. However, if $c_0 \neq c_3$ (which occurs if $c_0 \neq 1/2$) then $\{f(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^∞ linearly independent, even though $W \neq V$ and $\hat{\rho}(T_0|_V, T_1|_V) < 1$. On the other hand, $\{f(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^∞ linearly dependent for the case $c_0 = c_3$; in particular, the scaling function corresponding to that point is

$$f(x) = \begin{cases} x/2, & x \in [0,1], \\ 1/2, & x \in [1,2], , \\ (3-x)/2, & x \in [2,3], \end{cases} \tag{2.31}$$

and $\sum (-1)^k f(x+k) = 0$. Note that $W \neq V$ and $\hat{\rho}(T_0|_V, T_1|_V) < 1$ in this case as well.

We return now to the proof of Theorem 1.2. By our previous discussion, it remains only to show that ℓ^∞ linear independence of the integer translates of a continuous scaling function implies $\hat{\rho}(T_0|_V, T_1|_V) < 1$. By Theorem 2.5, to prove $\hat{\rho}(T_0|_V, T_1|_V) < 1$ it suffices to prove uniform convergence of the Cascade Algorithm, and this is the approach we take. In particular, we adapt several results from the analysis of subdivision schemes in [4].

Recall first that the Cascade Algorithm is defined by

$$f_{j+1}(x) = \sum_{k=0}^{N} c_k f_j(2x-k), \tag{2.32}$$

where $f_0 = \chi_{[-1/2,1/2)}$. The following lemma implies that each f_j is piecewise constant on each interval $I_{jk} = [2^{-j}(k-1/2), 2^{-j}(k+1/2))$ for $k \in \mathbf{Z}$.

Lemma 2.8. $f_j(x) = \sum_k f_j(2^{-j}k) f_0(2^j x - k)$.

Proof. The base step $j = 0$ is clear, so assume that the assertion holds for some $j \geq 0$. Then we compute

$$\begin{aligned}
\sum_k f_{j+1}(2^{-j-1}k) f_0(2^{j+1}x-k) &= \sum_k \sum_l c_l f_j(2^{-j}k - l) f_0(2^{j+1}x-k) \\
&= \sum_l c_l \sum_k f_j(2^{-j}k) f_0(2^{j+1}x - k - 2^l) \\
&= \sum_l c_l \sum_k f_j(2^{-j}k) f_0(2^j(2x-l)-k) \\
&= \sum_l c_l f_j(2x-l) \\
&= f_{j+1}(x),
\end{aligned} \tag{2.33}$$

so the result follows by induction. $\square$

Assume now that f is a continuous scaling function. Then the fact that the f_j are piecewise constant on the intervals I_{jk} implies that the f_j will converge uniformly to f if

$$\lim_{j\to\infty} \sup_k |f(2^{-j}k) - f_j(2^{-j}k)| = 0. \tag{2.34}$$

The following lemma is a first step towards comparing the values of $f_j(2^{-j}k)$ and $f(2^{-j}k)$.

Lemma 2.9. If f is a continuous scaling function for a dilation equation determined by coefficients $\{c_0, \dots, c_N\}$ satisfying (1.2), then

a: $f(x) = \sum_k f_j(2^{-j}k) f(2^j x - k)$ for every j, and
b: $\lim_{j\to\infty} \|f(x) - \sum_k f(2^{-j}k) f(2^j x - k)\|_\infty = 0$.

Proof. Item a: This follows from an induction similar to the one used to prove Lemma 2.8.

Item b: Choose $\varepsilon > 0$, and note that f is uniformly continuous since $\mathrm{supp}(f) \subset [0, N]$. Hence, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$

whenever $|x-y|<\delta$. Fix j_0 so that $2^{-j_0}N<\delta$, and choose $j>j_0$. Given $x\in\mathbf{R}$, define

$$K_x = \{k\in\mathbf{Z} : f(2^j x-k)\neq 0\}, \tag{2.35}$$

and note that if $k\in K_x$ then $2^j x-k\in[0,N]$, whence $|x-2^{-j}k|<\delta$. Using (1.9), we therefore compute

$$\begin{aligned}
&\bigl|f(x)-\sum_k f(2^{-j}k)f(2^j x-k)\bigr| \\
&= \bigl|f(x)\sum_k f(2^j x-k)-\sum_k f(2^{-j}k)f(2^j x-k)\bigr| \\
&\le \sum_{k\in K_x} |f(x)-f(2^{-j}k)|\,|f(2^j x-k)| \\
&\le \varepsilon N\|f\|_\infty,
\end{aligned} \tag{2.36}$$

from which the result follows. □

Next, the following fact from [4] about ℓ^∞ linear independence will allow us to compare $f(2^{-j}k)$ and $f_j(2^{-j}k)$ directly.

Lemma 2.10. If h is continuous and compactly supported and $\{h(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^∞ linearly independent then there exists a constant $C>0$ such that

$$\sup_k |a_k| \le C\Bigl\|\sum_k a_k h(x+k)\Bigr\|_\infty \tag{2.37}$$

for every sequence $\{a_k\}_{k\in\mathbf{Z}}\in\ell^\infty$.

Proof. Suppose no such positive C existed; then we could find sequences $\{a_k^j\}_{k\in\mathbf{Z}}\in\ell^\infty$, each with $\sup_k |a_k^j|=1$, such that

$$\lim_{j\to\infty}\Bigl\|\sum_k a_k^j h(x+k)\Bigr\|_\infty = 0. \tag{2.38}$$

Now, for each j there must be some k_j such that $|a_{k_j}^j|\ge 1/2$. Set $b_k^j = a_{k+k_j}^j$, and define

$$h_j(x) = \sum_k b_k^j h(x+k) = \sum_k a_k^j h(x-k_j+k). \tag{2.39}$$

Then $\lim_{j\to\infty}\|h_j\|_\infty = 0$, and $|b_0^j|\ge 1/2$ for every j. Since each sequence $\{b_k^j\}_{k\in\mathbf{Z}}$ has ℓ^∞ norm 1, there exist j_i such that the sequences $\{b_k^{j_i}\}_{k\in\mathbf{Z}}$ converge as $i\to\infty$, i.e., for each k there exists b_k such that $\lim_{i\to\infty} b_k^{j_i} = b_k$.

Since h has compact support the summations in (2.39) are all finite, and therefore for each x,

$$\sum_k b_k h(x+k) = \lim_{i\to\infty} \sum_k b_k^{j_i} h(x+k) = \lim_{i\to\infty} h_{j_i}(x) = 0. \tag{2.40}$$

However, $\{h(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^∞ linearly independent, so this implies $b_k = 0$ for every k, which is a contradiction since $|b_0| = \lim_{i\to\infty} |b_0^{j_i}| \geq 1/2$. $\square$

Putting this all together, we obtain the following.

Theorem 2.11. Given a continuous scaling function f. If $\{f(x+k)\}_{k\in\mathbf{Z}}$ is ℓ^∞ linearly independent then $\hat\rho(T_0|_V, T_1|_V) < 1$, but not conversely.

Proof. Since we showed in Example 2.7 that $\hat\rho(T_0|_V, T_1|_V) < 1$ does not imply ℓ^∞ linear independence, we need only show that ℓ^∞ linear independence implies $\hat\rho(T_0|_V, T_1|_V) < 1$. For this, it suffices by our previous discussions to establish (2.34). So, choose $\varepsilon > 0$; then by Lemma 2.9 there exists a j_0 such that

$$\Big\| \sum_k \big(f(2^{-j}k) - f_j(2^{-j}k)\big) f(2^j x - k) \Big\|_\infty < \varepsilon \tag{2.41}$$

for all $j > j_0$. By Lemma 2.10 there exists a constant $C > 0$ such that

$$\sup_k |a_k| \leq C \Big\| \sum_k a_k f(y-k) \Big\|_\infty \tag{2.42}$$

for every $\{a_k\}_{k\in\mathbf{Z}} \in \ell^\infty$. Taking $a_k = f(2^{-j}k) - f_j(2^{-j}k)$ we conclude that

$$\sup_k |f(2^{-j}k) - f_j(2^{-j}k)| < \frac{\varepsilon}{C} \tag{2.43}$$

for all $j > j_0$, and therefore (2.34) follows. $\square$

The proof of Theorem 1.2 is now complete. In particular, ℓ^∞ linear independence of the integer translates of a continuous scaling function f implies $\hat\rho(T_0|_V, T_1|_V) < 1$ by Theorem 2.11, which is equivalent to uniform convergence of the Cascade Algorithm by Theorem 2.5. By Proposition 2.2, the condition $\hat\rho(T_0|_V, T_1|_V) < 1$ is stable under small pertubations of the coefficients and by Theorem 2.1 it is sufficient to imply the existence of continuous scaling functions. Finally, Proposition 2.4 implies that scaling functions deform uniformly when $\hat\rho(T_0|_V, T_1|_V) < 1$.

However, ℓ^∞ linear independence of $\{f(x+k)\}_{k\in\mathbf{Z}}$ is not a necessary condition for the conclusions of Theorem 1.2 to hold. In particular, this requires only $\hat\rho(T_0|_V, T_1|_V) < 1$, and we showed in Example 2.7 that there exist scaling functions with $\hat\rho(T_0|_V, T_1|_V) < 1$ whose integer translates are ℓ^∞ linearly dependent.

3. Bibliography

[1] M.A. Berger and Y. Wang, *Two-scale dilation equations and the cascade algorithm*, Random Comp. Dynamics (to appear).

[2] M.A. Berger and Y. Wang, *Bounded semi-groups of matrices*, Lin. Alg. Appl. **166** (1992), 21–27.

[3] J.S. Byrnes, Jennifer L. Byrnes, Karl Berry, and Kathyrn A. Hargreaves, editors, *Probabilistic and Stochastic Methods in Analysis, with Applications: Proceedings of the NATO Advanced Study Institute,* volume 372 of *NATO ASI series C: Matheatical and physical sciences,* Il Ciocco Resort, Tuscany, Italy, 1992, Kluwer Academic Publishers Group, Dordrecht, The Netherlands.

[4] A. Cavaretta, W. Dahmen, and C.A. Micchelli, *Stationary Subdivision*, Mem. Amer. Math. Soc. **93** (1991), 1–186.

[5] A. Cohen and I. Daubechies, *A stability criterion for biorthogonal wavelet bases and their related subband coding schemes*, Duke Math. J. **68** (1992), 313–335.

[6] A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **65** (1992), 485–560.

[7] 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 **38** (1992), 876–881.

[8] D. Colella and C. Heil, *Characterizations of scaling functions: Continuous solutions*, SIAM J. Matrix Anal. Appl. **15** (1994) (to appear).

[9] I. Daubechies, *Ten Lectures on Wavelets.* Number 61 in CBMS/NSF Series in Applied Math., SIAM Publ., 1992.

[10] I. Daubechies and J. Lagarias, *Two-scale difference equations: I. Existence and global regularity of solutions*, SIAM J. Math. Anal. **22** (1991), 1388–1410.

[11] I. Daubechies and J. Lagarias, *Two-scale difference equations: II. Local regularity, infinite products of matrices and fractals*, SIAM J. Math. Anal. **23** (1992), 1031–1079.

[12] C. Heil, *Methods of solving dilation equations*, In Byrnes et al. [3], pp. 15–46.

[13] C. Heil and D. Colella, *Dilation equations and the smoothness of compactly supported wavelets*, Wavelets: Mathematics and Applications (J.

Benedetto and M. Frazier, eds.), CRC Press, Boca Raton, FL, 1993, pp. 161–200.

[14] C. Heil and G. Strang, *Continuity of the joint spectral radius: Application to wavelets*, Linear Algebra for Signal Processing (A. Bojanczyk and G. Cybenko, eds.), IMA Vol. Math. Appl., Springer–Verlag, New York (to appear).

[15] Christopher Heil, *Methods of solving dilation equations*, In Byrnes et al. [3], pp. 15–46.

[16] R.-Q. Jia and J. Wang, *Stability and linear independence associated with wavelet decompositions*, Proc. Amer. Math. Soc. **117** (1993), 1115–1124.

[17] P.G. Lemarié, *Fonctions a support compact dans les analyses multirésolutions*, Rev. Mat. Iberoamericana **7** (1991), 157–182.

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

[19] G.C. Rota and G. Strang, *A note on the joint spectral radius*, Kon. Nederl. Akad. Wet. Proc. A **63** (1960), 379–381.

[20] Y. Wang, *On two-scale dilation equations*, preprint.

ERRATA

Note: This errata listing is not included in the published version of this paper.

Page 31. The statement of Proposition 2.6 is correct, but the proof of the implication c $\Rightarrow$ b is incorrect. An independent proof was given by Sun [Sun] (available at `http://gauss.math.ucf.edu/~qsun/`).

ERRATA REFERENCES

[Sun] Q. Sun, *Two-scale difference equation: Local and global linear independence*, unpublished manuscript (1991).