

## CONTINUITY OF THE JOINT SPECTRAL RADIUS: APPLICATION TO WAVELETS

CHRISTOPHER HEIL\* AND GILBERT STRANG†

**Abstract.** The joint spectral radius is the extension to two or more matrices of the (ordinary) spectral radius  $\rho(A) = \max |\lambda_i(A)| = \lim \|A^m\|^{1/m}$ . The extension allows matrix products  $\Pi_m$  taken in all orders, so that norms and eigenvalues are difficult to estimate. We show that the limiting process does yield a continuous function of the original matrices—this is their joint spectral radius. Then we describe the construction of wavelets from a dilation equation with coefficients  $c_k$ . We connect the continuity of those wavelets to the value of the joint spectral radius of two matrices whose entries are formed from the  $c_k$ .

**1. Introduction.** The (ordinary) *spectral radius* of a matrix  $A$  is the magnitude of its largest eigenvalue:

$$\rho = \rho(A) = \max \{|\lambda| : \lambda \text{ is an eigenvalue of } A\}.$$

This number controls the growth or decay of the powers  $A^m$ . If  $\rho < 1$  then  $A^m \rightarrow 0$  as  $m \rightarrow \infty$ . If  $\rho > 1$  then the matrix powers are unbounded. The marginal case  $\rho = 1$  leaves boundedness undecided. In any norm the  $m^{\text{th}}$  root of  $\|A^m\|$  always converges to  $\rho$  as  $m \rightarrow \infty$ .

We describe below how the degree of continuity of a wavelet (expressed by its Hölder exponent  $\alpha$ ) is controlled by a spectral radius. But there is a crucial difference from  $\rho(A)$ : *two matrices are involved instead of one*. The number  $\rho(A)$  becomes a *joint spectral radius*  $\hat{\rho}(A, B)$ . It is still defined by a limit of  $m^{\text{th}}$  roots, but  $\|A^m\|$  is replaced by the largest norm  $\|\Pi_m\|$  of products of  $A$ 's and  $B$ 's. The product  $\Pi_m = ABAAB \cdots$  has its  $m$  factors in any order:

$$(1.1) \quad \hat{\rho} = \hat{\rho}(A, B) = \lim_{m \rightarrow \infty} (\max \|\Pi_m\|)^{1/m}.$$

For symmetric or normal or commuting or upper-triangular matrices, this joint spectral radius is the larger of  $\rho(A)$  and  $\rho(B)$ . Always  $\hat{\rho}(A, B) \geq \rho(A)$  and  $\hat{\rho}(A, B) \geq \rho(B)$ , since the product  $\Pi_m$  might be  $A^m$  or  $B^m$ . An extreme case of inequality is

$$A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \quad \text{and} \quad AB = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}.$$

The eigenvalues of  $A$  and  $B$  are all zero, so that  $\rho(A) = 0 = \rho(B)$ . In fact  $A^2$  and  $B^2$  are zero matrices, so the product  $\Pi_m$  is nonzero only when

---

\*School of Mathematics, Skiles Building, Room 242, Georgia Institute of Technology, Atlanta, Georgia 30332.

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. Partially supported by National Science Foundation Grant DMS-9006220.

factors  $A$  and  $B$  alternate. The key to  $\hat{\rho}$  is that  $\rho(AB)^{1/2} = 2 = \|AB\|^{1/2}$ . Therefore  $\hat{\rho}(A, B)$ , which is between eigenvalues and norms, also equals 2.

The eigenvalues of  $A$  and  $B$  fail to control the eigenvalues of products—not to mention the norms of those products. We cannot compute  $\hat{\rho}$  from  $\rho(A)$  and  $\rho(B)$ .

The spectral radius is bounded above by norms and below by eigenvalues. For a single matrix we have (for each  $m$ ) an equality and an inequality:

$$\rho(A^m)^{1/m} = \rho(A) \leq \|A^m\|^{1/m}.$$

For two matrices we maximize over products in all orders and we expect two inequalities:

$$(1.2) \quad (\max \rho(\Pi_m))^{1/m} \leq \hat{\rho}(A, B) \leq (\max \|\Pi_m\|)^{1/m}.$$

The proof of the first inequality comes from the corresponding result for a single matrix, by considering repetitions  $\Pi_m \Pi_m \cdots \Pi_m$  of any fixed product. For the second inequality, break any product  $\Pi_n$  into pieces of fixed length  $m$  with a remainder: if  $n = mq + r$  then  $\Pi_n = \Pi_m^{(1)} \Pi_m^{(2)} \cdots \Pi_m^{(q)} \Pi_r$ . For  $n = mq$  and  $r = 0$  we have  $\|\Pi_n\|^{1/n} \leq (\max \|\Pi_m\|)^{1/m}$ . The extra factor  $\Pi_r$  has no effect in the limit as  $n \rightarrow \infty$ .

The right side of (1.2) approaches equality as  $m \rightarrow \infty$  by the definition of the joint spectral radius. Whether the left side also approaches equality is much less clear. It is a beautiful theorem of Berger and Wang [BW1] that this does occur:

**THEOREM 1.1.** [BW1].  $\limsup_{m \rightarrow \infty} (\max \rho(\Pi_m))^{1/m} = \hat{\rho}(A, B)$ .

Thus  $\hat{\rho}$  can be approximated from above and from below, by computing the norms and eigenvalues of finitely many matrix products. The convergence as  $m \rightarrow \infty$  may be quite slow. Examples are given by Colella and Heil [CH1], [HC], with a recursive algorithm that significantly reduces the calculation on the norm side.

**EXAMPLE 1.1.** Set

$$A = \frac{1}{5} \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad B = \frac{1}{5} \begin{pmatrix} 3 & -3 \\ 0 & -1 \end{pmatrix}.$$

Then by actual computations up to  $m = 30$  factors,

$$\hat{\rho} \geq \max_{1 \leq m \leq 30} (\max \rho(\Pi_m))^{1/m} = \rho(A^{12}B)^{1/13} \approx 0.659679$$

and

$$\hat{\rho} \leq \min_{1 \leq m \leq 30} (\max \|\Pi_m\|)^{1/m} = \|A^5 B A^{11} B A^{12}\|^{1/30} \approx 0.671271.$$

Less than two decimal places of accuracy are achieved after computing  $2^{31} - 2$  matrix products. With 612944 products in the recursive algorithm the estimate is  $\hat{\rho} \leq 0.660025$ .

It is frustrating that this fundamental number  $\hat{\rho}$ , a function only of the entries in  $A$  and  $B$ , should be so difficult to compute. The same is true in the more highly developed theory of products of *random matrices* [CKN]. The corresponding number  $\delta$  is an expected value instead of a maximum, and the only weakness is its resistance to calculation. In our present problem, Lagarias and Wang [LW] conjecture that equality holds on the left side of (1.2) for some finite  $m$ . Thus an eigenvalue of a finite product  $\Pi_m$  (for unknown and possibly large  $m$ ) may reveal the exact value of the joint spectral radius.

Little is known about  $\hat{\rho}$ . From its definition (1.1), we were not even certain that  $\hat{\rho}$  was a continuous function. Certainly each norm  $\|\Pi_m\|^{1/m}$  depends continuously on the entries of  $A$  and  $B$ . This assures us that the limit  $\hat{\rho}$  is at least upper-semicontinuous. (It is an infimum of continuous functions, the norms of products.) The Berger–Wang theorem yields the opposite result, that  $\hat{\rho}$  is at the same time lower-semicontinuous. (It is a supremum of continuous functions, the eigenvalues of products.) Eventually we realized, with this essential help from Berger and Wang, that continuity does hold. This is the unique novel result of the present note:

**COROLLARY 1.1.**  *$\hat{\rho}(A, B)$  is a continuous function of the matrices  $A$  and  $B$ .*

Allow us to write out the proof in what may be unnecessary detail. We are given matrices  $A$  and  $B$  and a fixed  $\varepsilon > 0$ . For sufficiently large  $m$ , both inequalities in (1.2) are within  $\varepsilon/2$  of equality. Then if  $\|A - C\|$  and  $\|B - D\|$  are small enough, each side of (1.2) is within  $\varepsilon/2$  of the corresponding inequality for  $C$  and  $D$ . Therefore  $\hat{\rho}(C, D)$ , which is caught in between, is within  $\varepsilon$  of  $\hat{\rho}(A, B)$ .

**2. Wavelets and linear algebra.** A wavelet is a function  $\psi$  whose dilates and translates  $\psi(2^j x - k)$  form an orthogonal basis for  $L^2(\mathbf{R})$ . (More general definitions are possible. There are biorthogonal wavelets and even nonorthogonal wavelets.) We briefly describe the classical construction. The discrete version yields a particularly attractive orthonormal basis for  $\mathbf{R}^n$ . The expansion of a vector in terms of this discrete basis is achieved by a very fast algorithm—the discrete wavelet transform operates in  $\mathcal{O}(n)$  steps while the discrete Fourier transform requires  $\mathcal{O}(n \log n)$ . The expository paper [S] describes analogies and competitions between these transforms. In most of signal processing the standard methods are Fourier-based. In the compression of fingerprint images (the FBI has 25 million to digitize and compare) wavelet bases now seem to be superior.

The joint spectral radius enters in determining the smoothness (the order of Hölder continuity) of wavelets. The matrices  $A$  and  $B$  contain coefficients from the *dilation equation*. This has become the starting point for the construction of  $\psi$ , and we briefly outline the steps. The next section explains the connection to the joint spectral radius.

First, choose coefficients  $(c_0, \dots, c_N)$ . Second, solve the dilation equation for the *scaling function*  $\varphi$ :

$$(2.1) \quad \varphi(x) = \sum_{k=0}^N c_k \varphi(2x - k).$$

Third, construct  $\psi$  directly from  $\varphi$  and the  $c_k$  (in reverse order and with alternating signs):

$$(2.2) \quad \psi(x) = \sum_{k=0}^N (-1)^k c_{N-k} \varphi(2x - k).$$

(Other choices of indexing are possible; this gives  $\varphi$  and  $\psi$  supported in  $[0, N]$ ). The properties of  $\psi$  clearly depend on the choice of the  $c_k$ . The condition for “minimal accuracy” is

$$(2.3) \quad \sum_k c_{2k} = \sum_k c_{2k+1} = 1.$$

The condition for orthogonality of the family  $\psi(2^j x - k)$  is

$$(2.4) \quad \sum_k c_k \bar{c}_{k+2j} = 2 \delta_{0j}.$$

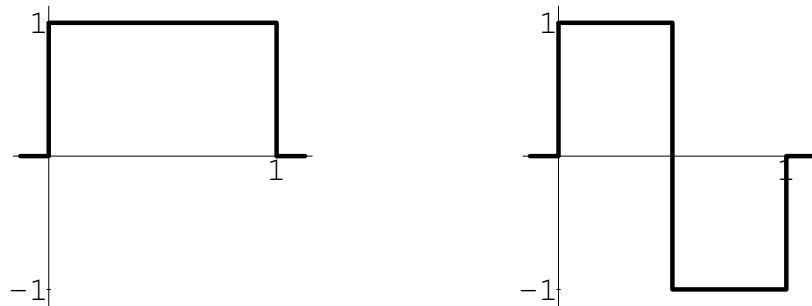


FIG. 1. *Box function*  $\varphi$  (left) and *Haar wavelet*  $\psi$  (right).

EXAMPLE 2.1. Make the choice  $c_0 = c_1 = 1$ . Then the dilation equation (2.1) requires  $\varphi(x)$  to be the sum of  $\varphi(2x)$  and  $\varphi(2x - 1)$ . The solution is the box function  $\varphi = \chi_{[0,1]}$ . Both the minimal accuracy and orthogonality conditions hold, and  $\psi(x) = \varphi(2x) - \varphi(2x - 1)$  generates a wavelet

orthonormal basis. This property of  $\psi$  was first observed by Haar [H].

The accuracy and orthogonality conditions (2.3) and (2.4) ensure that equation (2.1) yields a unique, compactly supported, integrable scaling function  $\varphi$ . Without these conditions, existence is not guaranteed. (If a compactly supported scaling function does exist then it is unique.) We only consider compact support in this note.

**3. The matrix form of the construction.** To make the connection with the joint spectral radius, we convert the dilation equation (2.1) into a matrix form. The key fact is that if a compactly supported scaling function  $\varphi$  exists then it is supported in the interval  $[0, N]$ . Therefore, the vector-valued function  $v$  defined by

$$v(x) = \begin{pmatrix} \varphi(x) \\ \varphi(x+1) \\ \vdots \\ \varphi(x+N-1) \end{pmatrix}, \quad \text{for } x \in [0, 1],$$

“captures” all the essential information about  $\varphi$ . Assume  $\varphi$  is continuous, so that  $\varphi(0) = 0 = \varphi(N)$ . Choose  $x$  in the interval  $[0, 1/2]$ , and apply the dilation equation to each of  $\varphi(x), \varphi(x+1), \dots, \varphi(x+N-1)$ . Working through some algebra, we find that these values of  $\varphi$  are determined by the values of  $\varphi(2x), \varphi(2x+1), \dots, \varphi(2x+N-1)$  in a fixed linear manner. In other words, there is a linear transformation  $T_0$  which takes  $v(2x)$  to  $v(x)$  for each  $0 \leq x \leq 1/2$ :

$$(3.1) \quad v(x) = T_0 v(2x), \quad \text{for } 0 \leq x \leq 1/2.$$

Similarly, there is a  $T_1$  such that

$$(3.2) \quad v(x) = T_1 v(2x-1), \quad \text{for } 1/2 \leq x \leq 1.$$

In fact,  $T_0$  and  $T_1$  are the  $N \times N$  matrices whose entries are  $(T_0)_{ij} = c_{2i-j-1}$  and  $(T_1)_{ij} = c_{2i-j}$ . There is *consistency* at  $x = 1/2$ :

$$(3.3) \quad v(1/2) = T_0 v(1) = T_1 v(0).$$

To simplify the notation, let  $\tau x$  be the fractional part of  $2x$ :

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

We purposely leave  $\tau(1/2)$  undefined. If  $x = k/2^j$  then it has two possible binary expansions: a “terminating” expansion ending in infinitely many zeros, and another expansion ending in infinitely many ones. We call such points *dyadic*. Otherwise,  $x$  has a unique binary expansion. In any case,

the first digit  $d_1$  in the binary expansion  $x = .d_1d_2\cdots$  is unique as long as  $x \neq 1/2$ :

$$d_1 = \begin{cases} 0, & 0 \leq x < 1/2, \\ 1, & 1/2 < x \leq 1, \end{cases} \quad \text{and} \quad \tau x = .d_2d_3\cdots.$$

Therefore, except for  $x = 1/2$ , we can summarize (3.1) and (3.2) as a single equation:

$$(3.4) \quad v(x) = T_{d_1}v(\tau x), \quad \text{for } x = .d_1d_2\cdots.$$

Because of the consistency (3.3), this formula also applies to  $x = 1/2$ : either of the two binary expansions  $1/2 = .100\cdots$  or  $1/2 = .011\cdots$  may be used as long as  $T_{d_1}$  and  $\tau(1/2)$  are interpreted consistently.

Now we use the assumption that  $\varphi$  is continuous. Let  $x < y$  be dyadic points in  $[0, 1]$ . If  $y$  is close enough to  $x$  then the first few digits in its terminating binary expansion will coincide with the first few digits in the corresponding expansion of  $x$ . That is,  $x = .d_1\cdots d_m d_{m+1} d_{m+2}\cdots$  and  $y = .d_1\cdots d_m d'_{m+1} d'_{m+2}\cdots$  for some  $m$ . The closer  $y$  is to  $x$ , the larger  $m$  will be. Applying (3.4) repeatedly, we obtain

$$(3.7) \quad \begin{aligned} v(y) - v(x) &= T_{d_1}(v(\tau y) - v(\tau x)) \\ &= T_{d_1}T_{d_2}(v(\tau^2 y) - v(\tau^2 x)) \\ &\vdots \\ &= T_{d_1}\cdots T_{d_m}(v(\tau^m y) - v(\tau^m x)). \end{aligned}$$

Since  $m \rightarrow \infty$  and  $v(y) \rightarrow v(x)$  as  $y \rightarrow x$ , the products  $\Pi_m = T_{d_1}\cdots T_{d_m}$  must converge to zero in the limit, at least when applied to vectors of the form  $v(w) - v(z)$ . To ensure that only vectors of this form are considered, we restrict our attention to the subspace

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

which is invariant under both  $T_0$  and  $T_1$ . Then all the restricted products  $(\Pi_m)|_W$  must converge to zero as  $m \rightarrow \infty$ . As Berger and Wang [BW1] observed, this is equivalent to  $\hat{\rho} < 1$  (on  $W$ ). We therefore have the following necessary condition for  $\varphi$  to be continuous:

**THEOREM 3.1.** [CH2]. *If  $\varphi$  is a continuous scaling function then*

$$\hat{\rho}(T_0|_W, T_1|_W) < 1.$$

The subspace  $W$  is not as difficult to determine explicitly as it may appear. It is the smallest subspace which contains the vector  $v(1) - v(0)$  and is invariant under both  $T_0$  and  $T_1$ . By an appropriate change of basis we can always realize the action of  $T_0, T_1$  on  $W$  as the action of two smaller matrices  $A, B$  on  $\mathbf{R}^n$  with  $n = \dim(W)$ . In this case,  $\hat{\rho}(T_0|_W, T_1|_W) = \hat{\rho}(A, B)$ .

The arguments leading to Theorem 3.1 did not make use of the accuracy or orthogonality conditions (2.3) or (2.4). Theorem 3.1 therefore applies to completely arbitrary dilation equations. As an implicit corollary, we obtain a necessary condition for  $\varphi$  to be continuously differentiable. This follows since  $\varphi'$  is itself a scaling function for the dilation equation with coefficients doubled to  $2c_k$ :

$$\varphi'(x) = \sum_{k=0}^N 2c_k \varphi'(2x - k).$$

A necessary condition for  $\varphi \in C^n$  follows by repetition. These corollaries are given explicitly in [HC].

When the minimal accuracy condition (2.3) does hold, the matrices  $T_0$  and  $T_1$  are column stochastic in the sense that each column sums to one—although the entries need not be nonnegative. Therefore the row vector  $(1, \dots, 1)$  is a common left eigenvector for  $T_0$  and  $T_1$  for the eigenvalue 1. The subspace

$$(3.6) \quad V = \{u \in \mathbf{C}^N : u_1 + \dots + u_N = 0\},$$

is therefore invariant under both  $T_0$  and  $T_1$ , and  $V$  contains  $W$ .

We can also work out a converse to Theorem 3.1 in terms of a joint spectral radius. This time we must first construct the vector-valued function  $v$  and from it obtain a scaling function  $\varphi$ . Our inspiration again comes from applying (3.4) recursively: if  $x$  is a dyadic point with terminating binary expansion  $x = .d_1 \dots d_m$  then  $\tau^m x = 0$ , so

$$(3.7) \quad \begin{aligned} v(x) &= T_{d_1} v(\tau x) \\ &\vdots \\ &= T_{d_1} \dots T_{d_m} v(\tau^m x) \\ &= T_{d_1} \dots T_{d_m} v(0) \\ &= \Pi_m v(0). \end{aligned}$$

Thus  $v(0)$  determines  $v(x)$  for each dyadic  $x$ . To find  $v(0)$ , solve the eigenvector problem  $v(0) = T_0 v(0)$ . Our recursion (3.7) amounts to the observation that once the values of  $\varphi$  are known at the integers, the dilation equation gives the values at the half-integers, then the quarter-integers, and so forth to all dyadic points. This is a fast recursive algorithm for graphing any scaling function  $\varphi$ .

Now that the values of  $v(x)$  have been constructed for dyadic  $x$ , we can fill in the values at intermediate points by taking limits—if this  $v$  is continuous on the set of dyadic points. In this case we “unfold”  $v$  to get a continuous scaling function  $\varphi$  by defining

$$\varphi(x) = \begin{cases} 0, & x \leq 0 \text{ or } x \geq N, \\ v_i(x), & i-1 \leq x \leq i, \quad i = 1, \dots, N. \end{cases}$$

Here  $v_i(x)$  is the  $i^{\text{th}}$  component of  $v(x)$ . So, the problem is to ensure continuity of  $v$  on the dyadics. The argument is similar to the one used for Theorem 3.1. If  $x = .d_1 \cdots d_m d_{m+1} d_{m+2} \cdots$  and  $y = .d_1 \cdots d_m d'_{m+1} d'_{m+2} \cdots$  with  $x, y$  both dyadic then (3.5) holds. If we assume that all the restricted products  $(\Pi_m)|_W$  converge to zero as  $m \rightarrow \infty$  then we will have  $v(y) \rightarrow v(x)$  as  $y \rightarrow x$ . But the convergence of all products to zero is equivalent to the assumption  $\hat{\rho} < 1$ . So, we have the following sufficient condition for the existence of a continuous scaling function:

**THEOREM 3.2.** [CH2]. *If  $\hat{\rho}(T_0|_W, T_1|_W) < 1$  then the dilation equation yields a continuous scaling function  $\varphi$ .*

The actual details involved in the proofs of Theorems 3.1 and 3.2 reveal precise bounds for the possible Hölder exponents of  $\varphi$  and its derivatives. (The number  $\alpha \leq 1$  is a Hölder exponent if there exists  $C_\alpha$  so that  $|\varphi(x) - \varphi(y)| \leq C_\alpha |x - y|^\alpha$  for all  $x, y$ .) In fact,  $\varphi$  is Hölder continuous for each exponent  $\alpha$  in the range  $0 \leq \alpha < -\log_2 \hat{\rho}$  and not for any  $\alpha > -\log_2 \hat{\rho}$ . (For our matrices  $T_0$  and  $T_1$ , the joint spectral radius  $\hat{\rho}$  is never below  $1/2$  so that  $-\log_2 \hat{\rho} \leq 1$ .) A condition can also be given for the marginal case  $\alpha = -\log_2 \hat{\rho}$ . As with Theorem 3.1, Theorem 3.2 implicitly leads to a sufficient condition for  $\varphi \in C^n$ .

**EXAMPLE 3.1.** Make the choice  $c_0 = 3/5$ ,  $c_1 = 6/5$ ,  $c_2 = 2/5$ ,  $c_3 = -1/5$ . The accuracy and orthogonality conditions (2.3) and (2.4) are met. The scaling function  $\varphi$  and wavelet  $\psi$  are shown in Figure 2. The subspace  $W$  equals the subspace  $V$ , and  $\hat{\rho}(T_0|_W, T_1|_W) = \hat{\rho}(A, B)$  with  $A, B$  as in Example 1.1. Thus  $\hat{\rho}(T_0|_W, T_1|_W) \approx 0.66 < 1$ , so  $\varphi$  and  $\psi$  are continuous. The maximum Hölder exponent is  $\alpha \approx -\log_2 0.66 \approx 0.60$ . Since  $\alpha < 1$ ,  $\varphi$  and  $\psi$  are not differentiable.

We close by mentioning some connections, observed in [CH2], between Corollary 1.1 and the continuity of the joint spectral radius. By Corollary 1.1,  $\hat{\rho}(A, B)$  is a continuous function of the entries of  $A$  and  $B$ . The entries of  $T_0$  and  $T_1$  consist of the coefficients  $(c_0, \dots, c_N)$ . Despite this,  $\hat{\rho}(T_0|_W, T_1|_W)$  is not in general a continuous function of  $(c_0, \dots, c_N)$  because the dimension of  $W$  can change abruptly as the coefficients vary. However, if the minimal accuracy condition (2.3) holds then  $W \subset V$ , and  $V$  is *independent* of the coefficients. Therefore  $\hat{\rho}(T_0|_V, T_1|_V)$  is continuous, and the condition  $\hat{\rho}(T_0|_V, T_1|_V) < 1$  is stable under small perturbations of the coefficients. Then  $\hat{\rho}(T_0|_W, T_1|_W) \leq \hat{\rho}(T_0|_V, T_1|_V) < 1$  ensures the existence of a continuous scaling function—which deforms uniformly as the coefficients vary. But the maximum Hölder exponents need not vary continuously since they depend critically on the value of  $\hat{\rho}(T_0|_W, T_1|_W)$ .



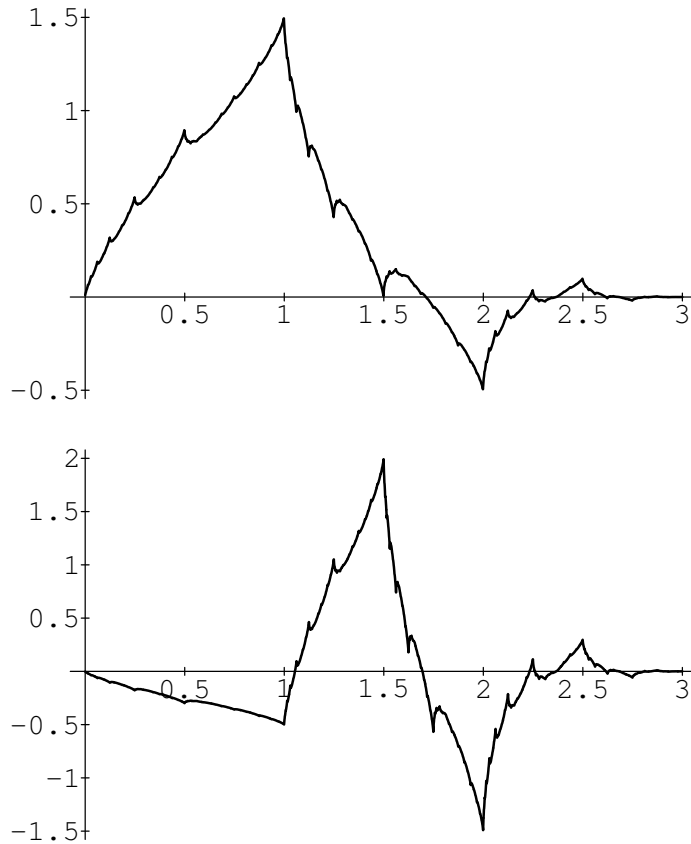


FIG. 2. *Scaling function  $\varphi$  (top) and wavelet  $\psi$  (bottom) from Example 3.1.*

**4. Historical notes.** The joint spectral radius was first described in [RS], for subsets of normed algebras. The first examples of continuous wavelets were found by Strömberg [St] and Meyer [Me]. The Meyer wavelet is  $C^\infty$ , but not compactly supported. Mallat and Meyer later developed multiresolution analysis [Ma], which results in the wavelet construction outlined in Section 2. Daubechies [D] discovered the first examples of continuous, compactly supported wavelets, including  $C^n$  wavelets for arbitrarily large  $n$ . Later, with Lagarias, she proved that compactly supported wavelets cannot be infinitely differentiable [DL1].

Daubechies and Lagarias also derived the matrix method for estimating the smoothness of wavelets, on which our discussion in Section 3 is based. In [DL2], they derived the matrix form (3.1)–(3.1) of the dilation equation and proved that if the coefficients satisfy the minimal accuracy condition (2.3) and if  $\hat{\rho}(T_0|_V, T_1|_V) < 1$  then  $\varphi$  is continuous with Hölder exponent at least  $-\log_2 \hat{\rho}(T_0|_V, T_1|_V) - \varepsilon$ . Theorem 3.2 is an extension of this result to arbitrary dilation equations, and Theorem 3.1 is its converse. These

two theorems were proved by Colella and Heil [CH2] together with corresponding sharp bounds for the Hölder exponent. Berger and Wang [BW2] independently obtained the same results. Daubechies and Lagarias also derived sufficient conditions for  $n$ -times differentiability and used the matrix approach to study the *local* smoothness of scaling functions and wavelets [DL2].

Dilation equations play a key role in subdivision or refinement schemes used in computer aided graphics. This is a separate application, using different coefficients  $c_k$  and different functions  $\varphi$ , but there is substantial overlap with wavelet theory. In particular, Micchelli and Prautzsch [MP] published the matrix form of the dilation equation before Daubechies and Lagarias, and they proved a necessary and sufficient condition for continuity of a scaling function [MP]. This condition was not in terms of a joint spectral radius, and did not lead to the Hölder exponent. An excellent survey of subdivision schemes is [CDM].

## REFERENCES

- [BW1] M.A. BERGER AND Y. WANG, *Bounded semi-groups of matrices*, Lin. Alg. Appl. **166** (1992), 21–27.
- [BW2] M.A. BERGER AND Y. WANG, *Two-scale dilation equations and the cascade algorithm*, Random Comp. Dynamics (to appear).
- [CDM] A. CAVARETTA, W. DAHMEN, AND C.A. MICCHELLI, *Stationary Subdivision*, Mem. Amer. Math. Soc. **93** (1991), 1–186.
- [CH1] 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.
- [CH2] D. COLELLA AND C. HEIL, *Characterizations of scaling functions: Continuous solutions*, SIAM J. Matrix Anal. Appl. **15** (1994), 496–518.
- [CKN] J.E. COHEN, H. KESTEN, AND C.M. NEWMAN, EDS., *Random Matrices and Their Applications*, Contemporary Math. 50, Amer. Math. Soc., Providence, 1986.
- [D] I. DAUBECHIES, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **41** (1988), 909–996.
- [DL1] I. DAUBECHIES AND J. LAGARIAS, *Two-scale difference equations: I. Existence and global regularity of solutions*, SIAM J. Math. Anal. **22** (1991), 1388–1410.
- [DL2] 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.
- [H] A. HAAR, *Zur Theorie der orthogonalen Funktionensysteme*, Math. Ann. **69** (1910), 331–371.
- [HC] 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 (1993), 161–200.
- [LW] J.C. LAGARIAS AND Y. WANG, *The finiteness conjecture for the joint spectral radius*, Lin. Alg. Appl. (to appear).
- [Ma] S.G. MALLAT, *Multiresolution approximations and wavelet orthonormal bases for  $L^2(\mathbf{R})$* , Trans. Amer. Math. Soc. **315** (1989), 69–87.

- [Me] Y. MEYER, *Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs*, Séminaire Bourbaki **662** (1985–1986).
- [MP] C.A. MICCHELI AND H. PRAUTZSCH, *Uniform refinement of curves*, Lin. Alg. Appl. **114/115** (1989), 841–870.
- [RS] G.C. ROTA AND G. STRANG, *A note on the joint spectral radius*, Kon. Nederl. Akad. Wet. Proc. A **63** (1960), 379–381.
- [S] G. STRANG, *Wavelet transforms versus Fourier transforms*, Bull. Amer. Math. Soc. **28** (1993), 288–305.
- [St] J.O. STRÖMBERG, *A modified Franklin system and higher-order spline systems on  $\mathbf{R}^n$  as unconditional bases for hardy spaces*, Conf. on Harm. Anal. in Honor of A. Zygmund, Vol. II, W. Beckner et al., eds., Wadsworth (1981), 475–494.