

## TWO-SCALE DIFFERENCE EQUATIONS II. LOCAL REGULARITY, INFINITE PRODUCTS OF MATRICES AND FRACTALS\*

INGRID DAUBECHIES‡§ AND JEFFREY C. LAGARIAS†

**Abstract.** This paper studies solutions of the functional equation

$$f(x) = \sum_{n=0}^N c_n f(kx - n),$$

where  $k \geq 2$  is an integer, and  $\sum_{n=0}^N c_n = k$ . Part I showed that equations of this type have at most one  $L^1$ -solution up to a multiplicative constant, which necessarily has compact support in  $[0, N/k - 1]$ . This paper gives a time-domain representation for such a function  $f(x)$  (if it exists) in terms of infinite products of matrices (that vary as  $x$  varies). Sufficient conditions are given on  $\{c_n\}$  for a continuous nonzero  $L^1$ -solution to exist. Additional conditions sufficient to guarantee  $f \in C^r$  are also given. The infinite matrix product representations is used to bound from below the degree of regularity of such an  $L^1$ -solution and to estimate the Hölder exponent of continuity of the highest-order well-defined derivative of  $f(x)$ . Such solutions  $f(x)$  are often smoother at some points than others. For certain  $f(x)$  a hierarchy of fractal sets in  $\mathbb{R}$  corresponding to different Hölder exponents of continuity for  $f(x)$  is described.

**Key words.** Hölder continuity, subdivision schemes, wavelets, infinite matrix products

**AMS(MOS) subject classifications.** 26A15, 26A18, 39A10, 42A05

**1. Introduction.** This is the second part of a series of two papers concerning functional equations of the type

$$(1.1) \quad f(x) = \sum_{n=0}^N c_n f(kx - n).$$

In Part I (Daubechies and Lagarias (1991)) we discussed existence and uniqueness of  $L^1$ -solutions. We saw that nontrivial  $L^1$ -solutions to (1.1) can only exist if  $|\sum c_n| \geq k$ . If  $\sum c_n = k$ , then the solution, if it exists at all, is unique and furthermore it has compact support contained in  $[0, N/(k-1)]$ . We shall assume that  $\sum c_n = k$  in this paper. Functions satisfying equations of type (1.1) arise in many contexts. Our own motivation lies in the role played by such two-scale difference equations in the construction of orthonormal bases of compactly supported wavelets (see § 6.3 in part I or Daubechies (1988)). Similar equations characterize certain nowhere differentiable functions constructed by De Rham (1956, 1957, 1959). They also play an important role in interpolating subdivision schemes with applications to computer aided design, on which an important body of work exists; papers in this subject treating 2-scale difference equations are, e.g., Dubuc (1986), Micchelli (1986), Dyn, Gregory, and Levin (1987, 1991), Deslauriers and Dubuc (1987, 1989), Micchelli and Prautzsch (1987, 1989), and Dyn and Levin (1989). All these examples satisfy  $\sum c_n = k$ .

For the sake of simplicity we shall usually choose  $k=2$ , even though all our techniques can be applied for general integer values of  $k \geq 2$ . We shall also restrict ourselves to real coefficients  $c_n$  and, correspondingly, real functions  $f$ . Our analysis can, however, also be used for complex  $c_n$ , without any changes. We are thus mainly

---

\* Received by the editors November 30, 1988; accepted for publication (in revised form) September 16, 1991.

† AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974.

‡ Mathematics Department, Rutgers University, New Brunswick, New Jersey 08903.

§ This author is "Bevoegdverklaard Navorser" at the Belgian National Science Foundation (on leave); also on leave from Department of Theoretical Physics, Vrije Universiteit Brussel, Belgium.

concerned with equations of the type

$$(1.2) \quad f(x) = \sum_{n=0}^N c_n f(2x - n)$$

with real  $c_n$  and  $\sum c_n = 2$ .

Not all equations of type (1.2), with  $\sum c_n = 2$  have a nontrivial  $L^1$ -solution. There exist several possible approaches to the study of existence and smoothness of solutions to (1.2). One approach uses the trigonometric polynomial  $p(\xi) = \frac{1}{2} \sum_{n=0}^N c_n e^{in\xi}$ . In particular, if  $p(\xi)$  can be factored as

$$(1.3) \quad p(\xi) = [(1 + e^{i\xi})/2]^{L+1} q(\xi)$$

and

$$(1.4) \quad \sup_{\xi \in \mathbb{R}} |q(\xi)| < 2^{L+1-m},$$

then there exists a nontrivial  $L^1$ -solution  $f$  to (1.2), which is, moreover,  $m$  times continuously differentiable (Daubechies (1988)). Other sufficient conditions on  $p$  guaranteeing existence and smoothness can be found in Deslauriers and Dubuc (1987, 1989). Typically these conditions all achieve regularity of  $f$  by imposing decay on its Fourier transform  $\hat{f}(\xi)$ . These methods work best when  $p(\xi)$  is a nonnegative function, as illustrated by Deslauriers and Dubuc (1987), who obtain very sharp information about the regularity of functions  $f$  constructed via a symmetric Lagrangian interpolation scheme. For more general examples, lacking the positivity of  $p(\xi)$ , this analysis leads to less than optimal results (Daubechies (1988), Deslauriers and Dubuc (1989)).

In this paper we use a different technique to study the regularity of solutions of (1.2). It is essentially a “time domain” method, in the sense that it does not involve Fourier transforms at any stage. This time domain approach hinges on the fact that if  $f$  satisfies (1.2), then the values  $f(x)$  can be easily calculated, recursively, for all dyadic  $x$ , i.e., all  $x$  of the type  $m2^{-j}$  ( $m \in \mathbb{Z}, j \in \mathbb{N}$ ), if the values  $f(m)$  at the integers are known. Provided suitable conditions are satisfied, we can then show that there exists a continuous extension of these  $f(m2^{-j})$  to all of  $\mathbb{R}$ , thus defining  $f(x)$ . In fact, an explicit formula for  $f(x)$  can be expressed using an infinite product of matrices (depending on  $x$ ). It is then possible to discuss the Hölder continuity, differentiability, etc., of this extension, which is the desired solution of (1.2).

We explain the matrix technique in detail in § 2. We show how to choose a “correct” initialization  $\{f(m); m \in \mathbb{Z}\}$  for the iterative spline approximation to  $f$ , and we formulate sufficient conditions on the  $c_n$  guaranteeing that the process converges to a continuous  $L^1$ -solution of (1.2). We also compute lower bounds for the Hölder exponent of the resulting function  $f(x)$ . In § 3 we show how similar ideas, using the sum rules (1.5) together with some technical conditions, can be used to prove that  $f \in C^L$ .

The time domain method proposed here leads in many instances to sharper results than the Fourier transform methods mentioned above. One such example is the function  $\phi$  plotted in Fig. 1(a) (or in Fig. 3(a) in Part I). Even when handled with care, the Fourier transform method only establishes that the Hölder exponent of  $\phi$  is at least  $.5 - \varepsilon$  (see Daubechies (1988)). In fact  $\phi$  is Hölder continuous with exponent  $2 - \ln(1 + \sqrt{3})/\ln 2 = .5500\dots$ ; the method presented in this paper achieves this (optimal) result. Moreover, the Fourier transform method typically only controls global regularity properties: the detailed, local analysis of the regularity of solutions to (1.2), accessible

via our matrix method as explained in § 4, is wholly outside the reach of Fourier transform based techniques. Using the time domain approach we can show, e.g., that the function  $\phi$  in Fig. 1a is almost everywhere differentiable, and that its Hölder exponent .5500... is determined by a dense set of “bad” points which has zero measure. In fact, there exists a whole hierarchy of fractal sets, all with zero measure, corresponding to the Hölder exponent between .5500... and 1. Another interesting example is the basic function associated to the dyadic interpolation scheme first studied in Dubuc (1986) and Dyn, Gregory and Levin (1987). In this case the function  $p(\xi) = \sum c_n e^{in\xi}$  is positive, allowing the Fourier transform method to achieve the already optimal result  $f \in C^{2-\epsilon}$  for the global regularity of  $f$ . Our time domain method recovers this result (although by a more involved analysis than via the Fourier transform method), but it also establishes that  $f$  is almost everywhere twice differentiable, which is a result outside the reach of the Fourier transform method. A detailed discussion of these and other examples can be found in § 5.

When we developed this technique, we were unaware of related and at that time unpublished work of Micchelli and Prautzsch (1989) and Dyn and Levin (1989), to which referees drew our attention. Let us give a short overview of the situation, pointing out the overlap and the differences between our work and theirs. The two papers by Micchelli and Prautzsch (1989) and the two by Dyn and Levin (1989) (hereafter called [MP] and [DL]) are both motivated by the applications of interpolation subdivision schemes to computer aided design (see Micchelli (1986), Micchelli and Prautzsch (1987) and Dyn, Gregory, and Levin (1987, 1989)). The coefficients  $c_n$  in an interpolation subdivision scheme typically satisfy  $c_{2n+n_0} = \delta_{b_0}$  for some  $n_0$ , but both [MP] and [DL] are applicable to more general coefficient choices. [MP] focuses on the existence and smoothness of solutions to (1.2), whereas [DL] is more concerned with the convergence of the corresponding interpolation scheme, that is, in the language of § 4 in part I, with existence of smooth solutions *and* convergence of the cascade algorithm to these solutions. It is proved in [DL] that the cascade algorithm converges to a  $C^L$  solution only if the coefficients  $c_n$  satisfy the  $L+1$  “sum rules”

$$(1.5) \quad \sum_n c_n n^l (-1)^n = 0, \quad l = 0, \dots, L.$$

Moment conditions of the type (1.5) also turn up in a different context. If the solution  $f$  to (1.2) has the property that the integer translates  $f(\cdot - n)$ ,  $n \in \mathbb{Z}$  are all independent, then the moment condition (1.5) is equivalent to saying that all the polynomials of degree at most  $L$  can be written as combinations of the  $f(\cdot - n)$ , and can therefore be obtained exactly by the associated subdivision scheme. See Cavaretta, Dahmen, and Micchelli (1989) for a general multivariate discussion of this aspect. Sum rules of type (1.5) are in fact satisfied by all interesting examples (wavelets as well as interpolating schemes); it is easy to check that they are equivalent to the requirement that  $p(\xi)$  can be factored as in (1.3). If we are concerned with only the existence of a reasonably smooth solution to (1.2), without convergence of the cascade algorithm to this solution, then conditions (1.5) are not necessary, and [MP] do not, in fact, use them. The discussion of possible solutions  $f(x)$  to (1.2) in [MP] makes use of the same matrices as our approach. In particular, for the choice  $k = 2$ , we deal with two matrices  $T_0$  and  $T_1$ , both determined by the  $c_n$ , and  $f(x)$  is then given by the limit

$$\lim_{n \rightarrow \infty} [T_{d_1(x)} T_{d_2(x)} \cdots T_{d_n(x)} \cdots] \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(N-1) \end{bmatrix},$$

where the index  $d_j(x)$  is the  $j$ th digit in the binary expansion for  $x$  (see § 2). A very special structure of the matrices  $T_0, T_1$  is then required for  $f$  to be a  $C^L$ -function: both matrices have to have the eigenvalues  $1, \frac{1}{2}, \dots, 2^{-L}$ , and the corresponding left eigenvectors  $e_l^0, e_l^1$  ( $l = 0, \dots, L$ ) have to define flags of nested subspaces with very particular properties (see § 5 in [MP] or § 2, 3 below). In our paper, we require that the sum rules (1.5) hold; as shown by examples and by [MP], they are not necessary, but they make it possible to write very explicit expressions for the eigenvectors  $e_l^0, e_l^1$ . It turns out that (1.5) forces the  $e_l^0, e_l^1$  to satisfy many of the conditions in [MP], which explains why we have only one technical condition (as compared to the four conditions in Theorem 5.2 in [MP]). If this condition is satisfied, then a simple spline approximation argument proves that the function  $f$  is indeed  $L$  times continuously differentiable. This simple argument was partly suggested to us by one of the anonymous referees, who we would like to thank here. In an appendix we shortly sketch our original, longer approach, which does not use splines but which can be generalized to certain higher-dimensional situations where splines would be hard to use.

Another difference between our paper and [MP] is that we also exploit the structure of the matrices  $T_0, T_1$  to study local regularity properties of  $f$ : the importance of the binary expansion of  $x$  in determining  $f(x)$  means that points with different frequencies of the digits 1 or 0 in their binary expansion correspond to different local Hölder exponents for  $f$ . In fact, this feature, first observed experimentally in graphs of orthonormal wavelet bases with compact support, was one of our main motivations for undertaking this study.

**2. Continuity and Hölder continuity.** If  $f$  satisfies (1.2), then the values of  $f(x)$  can be computed explicitly, recursively, for all dyadic rationals  $x$ , i.e., for all  $x$  of the type  $m2^{-j}$  ( $m \in \mathbb{Z}, j \in \mathbb{N}$ ), if the values  $f(m)$  at the integers  $m$  are known. If  $f$  is continuous, then the  $f(m2^{-j})$  suffice to determine  $f$  everywhere; a sequence of convergent spline approximations was constructed in Theorem 4.1 in part I. However, for general  $c_n$  and for arbitrarily chosen  $f(m)$  this procedure will typically diverge and not lead to a continuous function at all. The following proposition gives a necessary condition that the  $c_n$  have to satisfy and specifies restrictions on the choice of the initial values  $f(m)$  for the iteration scheme.

PROPOSITION 2.1. *Assume that  $\sum_{n=0}^N c_n = 2$ . If  $f$  is a nontrivial continuous  $L^1$ -solution to (1.2), then:*

(1<sup>0</sup>) *The  $(N - 1) \times (N - 1)$  dimensional matrix  $\mathbf{M}$  defined by*

$$\mathbf{M}_{ij} = c_{2i-j}, \quad 1 \leq i, j \leq N - 1$$

*has 1 as an eigenvalue.*

(2<sup>0</sup>) *The  $(N - 1)$  dimensional vector  $(f(1), \dots, f(N - 1))$  is a right eigenvector of  $\mathbf{M}$  with eigenvalue 1, and  $f(m) = 0$  for all  $m \leq 0$  and  $m \geq N$ .*

*Proof.* By Corollary 2.2 in part I,  $\text{support}(f) \subset [0, N]$ . This implies that  $f(m) = 0$  for  $m \leq 0, m \geq N$ . The other conclusions follow from Theorem 5.1 in part I.  $\square$

In what follows we shall impose that the  $c_n$  satisfy the first sum rule in (1.5); this condition is sufficient (but not necessary) for (1<sup>0</sup>) to hold. Explicitly,

$$(2.1) \quad \sum_{n=0}^N (-1)^n c_n = 0$$

or, equivalently, since  $\sum_n c_n = 2$ ,

$$(2.2) \quad \sum_n c_{2n} = \sum_n c_{2n+1} = 1.$$

This condition is satisfied in all the practical examples discussed in part I. It also has a very simple interpretation in the cascade algorithm (see § I.4). In this algorithm, the  $j$ th level consists of  $(2^j - 1)N + 1$  different coefficients  $a_m^j$ , which are computed from the  $(j - 1)$ th level via

$$a_{2m}^j = \sum_k c_{2(m-k)} a_k^j, \quad a_{2m+1}^j = \sum_k c_{2(m-k)+1} a_k^j$$

(see, e.g., Micchelli and Prautzsch (1987), Dyn, Gregory, and Levin (1989) or Daubechies (1988)). The condition (2.2) states, therefore, that the total weight of the  $a_l^{j-1}$  in the computation of  $a_m^j$  is independent of  $m$ . In the case where the cascade algorithm simplifies to an interpolation subdivision scheme, this “equal opportunity” condition is automatically satisfied, since then

$$c_{l+2n} = \delta_{n0} \quad \text{for some } l \quad (\text{see } \S \text{ I.4}),$$

hence

$$\sum_m c_{2m+1+l} = \sum_n c_{n+l} - \sum_m c_{2m+l} = \sum_{n=0}^N c_n - 1 = 1 = \sum c_{2m+l}.$$

If the scaling factor  $k$  is an integer larger than 2, then (2.2) has to be rewritten as

$$(2.2') \quad \sum_n c_{kn} = \sum_n c_{kn+1} = \dots = \sum c_{kn+(k-1)} = 1.$$

This is again an “equal opportunity” condition when looked at from the point of view of the cascade algorithm. Note that (2.2') is equivalent to requiring that  $p(\xi) = \sum_{n=0}^N c_n e^{in\xi}$  is divisible by  $(1 + e^{i\xi} + \dots + e^{i(k-1)\xi})$ .

We return to the dyadic case ( $k=2$ ), and we assume that (2.2) is satisfied. It immediately follows that for all  $j, 1 \leq j \leq N - 1$ ,

$$\sum_{i=1}^{N-1} \mathbf{M}_{ij} = \sum_{i=1}^{N-1} c_{2i-j} = 1,$$

where  $\mathbf{M}_{ij} = c_{2i-j}$  (see above, or § I.5) and where we use the convention  $c_l = 0$  if  $l < 0$  or  $l > N$ . Consequently, 1 is an eigenvalue of  $\mathbf{M}$ , with left eigenvector  $(1, 1, \dots, 1)$ . The necessary condition in Proposition 2.1 ( $1^0$ ) is therefore satisfied.

Let us assume that the eigenvalue 1 of  $\mathbf{M}$  is nondegenerate. (This will be guaranteed by a technical condition below.) Then there is a unique (up to normalization) right eigenvector  $\mathbf{a}$  with eigenvalue 1 for  $\mathbf{M}$ . This eigenvector cannot be orthogonal to the left eigenvector  $(1, 1, \dots, 1)$ , i.e.,  $\sum_{i=1}^{N-1} \mathbf{a}_i \neq 0$ , so that we can normalize  $\mathbf{a}$  to have  $\sum_{i=1}^{N-1} \mathbf{a}_i = 1$ . We then pick the  $f(n)$  to be

$$\begin{aligned} f(n) &= \mathbf{a}_n & n = 1, \dots, N - 1 \\ &= 0 & n \leq 0 \quad \text{or} \quad n \geq N. \end{aligned}$$

We define successive spline approximations  $f_j$  to  $f$  as in § I.4:

- (1)  $f_0(x)$  is linear on every  $[n, n + 1]$ ,  
 $f_0(n) = f(n)$ ,
- (2)  $f_j = V^j f_0$ ,

where  $V$  is the linear operator

$$(2.3) \quad (Vg)(x) = \sum_{n=0}^N c_n g(2x - n).$$

If  $f$  is continuous, then the  $f_j$  converge to  $f$  (see Theorem 4.1 in part I). A priori we have, however, no reason to expect a continuous solution to (1.2); in fact, for many choices of the  $c_n$ , even if they satisfy (2.2),  $f$  will not be continuous, and the  $f_j$  will converge to  $f$  only in some distributional sense. We shall impose further conditions on the  $c_n$  which will enable us to prove that the  $f_j$  constitute a Cauchy sequence in  $L^\infty$ , which then automatically leads to continuity for  $f$ . In order to do this, we introduce a “vector” notation.

Finding solutions to (1.2) is the same as finding fixed points for the linear operator  $V$  defined in (2.3). We shall only be concerned here with functions  $g$  supported on  $[0, N] = \text{support}(f)$ . For such functions (2.3) can be rewritten in “vector” form. Define the vector valued function  $\mathbf{w}: [0, 1] \rightarrow \mathbb{R}^N$  by

$$(2.4) \quad \mathbf{w}(x)_n = g(x + n - 1), \quad n = 1, \dots, N.$$

Knowing the function  $g$  is equivalent to knowing  $\mathbf{w}: [0, 1] \rightarrow \mathbb{R}^N$ . Note that necessarily

$$(2.5) \quad \mathbf{w}(0)_j = \mathbf{w}(1)_{j-1} \quad \text{for } 2 \leq j \leq N.$$

Moreover,  $g$  is continuous if and only if  $\mathbf{w}$  is continuous on  $[0, 1]$  and if

$$(2.6) \quad \mathbf{w}(0)_1 = 0 = \mathbf{w}(1)_N.$$

Let us now define the linear operator  $\mathbf{V}$  on vector valued functions  $\mathbf{w}$  satisfying (2.5), (2.6) by

$$(\mathbf{V}\mathbf{w})(x)_n = (\mathbf{V}g)(x + n - 1),$$

where we assume  $\mathbf{w}$  and  $g$  are linked as in (2.4). Then  $\mathbf{V}$  is given explicitly by

$$(2.7) \quad \mathbf{V}\mathbf{w}(x) = \begin{cases} \mathbf{T}_0\mathbf{w}(2x) & \text{if } x \leq \frac{1}{2}, \\ \mathbf{T}_1\mathbf{w}(2x - 1) & \text{if } x \geq \frac{1}{2}, \end{cases}$$

where  $\mathbf{T}_0, \mathbf{T}_1$  are the  $N \times N$ -matrices defined by

$$(2.8) \quad \begin{aligned} (\mathbf{T}_0)_{ij} &= c_{2i-j-1}, & 1 \leq i, j \leq N, \\ (\mathbf{T}_1)_{ij} &= c_{2i-j}, & 1 \leq i, j \leq N, \end{aligned}$$

or

$$\mathbf{T}_0 = \begin{bmatrix} c_0 & 0 & 0 & 0 & \cdots & & 0 & 0 \\ c_2 & c_1 & c_0 & 0 & \cdots & & 0 & 0 \\ c_4 & c_3 & c_2 & c_1 & c_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & c_N & c_{N-1} \end{bmatrix},$$

$$\mathbf{T}_1 = \begin{bmatrix} c_1 & c_0 & 0 & 0 & \cdots & & 0 & 0 \\ c_3 & c_2 & c_1 & c_0 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & c_N \end{bmatrix}.$$

Note that if we strip either  $\mathbf{T}_0$  of its first row and column, or  $\mathbf{T}_1$  of its last row and column, then the resulting  $(N - 1) \times (N - 1)$  matrix is exactly  $\mathbf{M}$ . Moreover, (2.2) implies that the  $N$ -dimensional row vector  $\mathbf{e}_1 = (1, \dots, 1)$  is a left eigenvector of both  $\mathbf{T}_0$  and  $\mathbf{T}_1$ , with eigenvalue 1,

$$\mathbf{e}_1 \cdot \mathbf{T}_0 = \mathbf{e}_1 = \mathbf{e}_1 \cdot \mathbf{T}_1.$$

Formula (2.7) can be written more succinctly if we use binary expansions. For any  $x \in [0, 1]$  we can write

$$x = \sum_{j=1}^{\infty} d_j 2^{-j}, \quad \text{where } d_j = 0 \text{ or } 1 \text{ for all } j.$$

We then define the action of the *shift operator*  $\tau$  on  $x$  by

$$(2.9) \quad \tau x = \sum_{j=2}^{\infty} d_j 2^{-j+1}.$$

This corresponds to shifting the decimal point in the binary expansion of  $x$  one step to the right and dropping the old first decimal. In fact,

$$\tau x = \begin{cases} 2x & \text{if } 0 \leq x < \frac{1}{2}, \\ 2x - 1 & \text{if } \frac{1}{2} < x \leq 1. \end{cases}$$

Note that this does not yet define  $\tau x$  for  $x = \frac{1}{2}$ ; we shall come back below to what happens at  $x = \frac{1}{2}$ . For  $x \neq \frac{1}{2}$ , (2.3) can therefore be rewritten as

$$(2.10) \quad \mathbf{w}(x) = \mathbf{T}_{d_1(x)} \mathbf{w}(\tau x),$$

where we have introduced the notation  $d_j(x)$  for the  $j$ th digit in the binary expansion of  $x$ . For  $x = \frac{1}{2}$  the first digit in the binary expansion (2.9) is not well defined, since we can take either  $d_1 = 0, d_j = 1$  for  $j \geq 2$  or  $d_1 = 1, d_j = 0$  for  $j \geq 2$ . Extending (2.10) to  $x = \frac{1}{2}$  leads, therefore, to two different equations, depending on which choice was made for the binary expansion of  $\frac{1}{2}$ ,

$$\mathbf{w}(\frac{1}{2}) = \mathbf{T}_0 \mathbf{w}(1) \quad \text{or} \quad \mathbf{w}(\frac{1}{2}) = \mathbf{T}_1 \mathbf{w}(0).$$

For  $\mathbf{w}$  satisfying the restrictions (2.5), (2.6), there is, however, no contradiction between these two equations, since  $(\mathbf{T}_1)_{i,j} = (\mathbf{T}_0)_{i+1,j+1}$  for  $1 \leq i, j \leq N - 1$ . Equation (2.10) therefore holds true for any  $x \in [0, 1]$ .

Let us apply all this to the spline functions  $f_j$  and define

$$[\mathbf{v}_j(x)]_n = f_j(x + n - 1), \quad j \in \mathbb{N}, \quad n = 1, \dots, N.$$

It follows that

$$\begin{aligned} \mathbf{v}_j(x) &= (\mathbf{V}^j \mathbf{v}_0)(x) \\ &= \mathbf{T}_{d_1(x)} \mathbf{T}_{d_2(x)} \cdots \mathbf{T}_{d_j(x)} \mathbf{v}_0(\tau^j x). \end{aligned}$$

Now that we have introduced all the necessary notation, we are in a position to formulate the main result of this section.

We will use the operator norm for matrices defined by  $\|\mathbf{T}\| = \sup_{\mathbf{v} \neq 0} \|\mathbf{T}\mathbf{v}\| / \|\mathbf{v}\|$ , where  $\|\mathbf{v}\|$  is the Euclidean norm of  $\mathbf{v}$ ,  $\|\mathbf{v}\|^2 = \sum_{j=1}^N v_j^2$ .

**THEOREM 2.2.** *Assume that the  $c_n, n = 0, \dots, N$  satisfy*

$$\sum_n c_{2n} = \sum_n c_{2n+1} = 1.$$

*Define the  $N \times N$ -matrices  $\mathbf{T}_0$  and  $\mathbf{T}_1$  by*

$$(\mathbf{T}_0)_{ij} = c_{2i-j-1}, \quad (\mathbf{T}_1)_{ij} = c_{2i-j}, \quad 1 \leq i, j \leq N.$$

*Define  $E_1$  to be the  $(N - 1)$ -dimensional subspace orthogonal to  $\mathbf{e}_1 = (1, \dots, 1)$ , the common left eigenvector of  $\mathbf{T}_0, \mathbf{T}_1$  for the eigenvalue 1. Assume that there exist  $\lambda < 1$  and  $C > 0$  such that, for all  $m \in \mathbb{N}$ ,*

$$(2.11) \quad \max_{d_j=0 \text{ or } 1, j=1, \dots, m} \|\mathbf{T}_{d_1} \mathbf{T}_{d_2} \cdots \mathbf{T}_{d_m}|_{E_1}\| \leq C \lambda^m.$$

Then the following hold:

(1) The eigenvalue 1 is of the  $(N - 1) \times (N - 1)$ -dimensional matrix  $\mathbf{M}$  defined by  $\mathbf{M}_{ij} = c_{2i-j}$ ,  $1 \leq i, j \leq N - 1$ , is simple and there is an associated right eigenvector  $\mathbf{a}$  with  $\sum_{i=1}^{N-1} \mathbf{a}_i = 1$ .

(2) The vector-valued functions  $\mathbf{v}_j(x)$  defined above satisfy  $\mathbf{e}_1 \cdot \mathbf{v}_j(x) = 1$  for all  $j \in \mathbb{N}$ , all  $x \in [0, 1]$ .

(3) The corresponding functions  $f_j$  converge uniformly to a continuous function  $f$ ,

$$\|f_j - f\|_{L^\infty} \leq C 2^{-j|\ln \lambda|/\ln 2}.$$

(4) The limit function  $f$  is an  $L^1$ -solution to (1.2); it is normalized so that  $\int_0^N dx f(x) = 1$ , and it is Hölder continuous,

$$(2.12) \quad |f(x) - f(y)| \leq C|x - y|^\alpha,$$

with  $\alpha = |\ln \lambda|/\ln 2$ .

*Proof.* (1) The constraint (2.16) automatically implies that 1 is a simple eigenvalue of  $\mathbf{T}_0$  and  $\mathbf{T}_1$ . Indeed, if 1 were not a simple eigenvalue of, e.g.,  $\mathbf{T}_0$ , then there would exist a right eigenvector  $\mathbf{e}'_1$  for  $\mathbf{T}_0$  in  $E_1$ , with eigenvalue 1 (regardless of whether the eigenvalue 1 is degenerate or not, in which last case the matrix  $\mathbf{T}_0$ , restricted to the invariant subspace for the eigenvalue 1, can be brought in Jordan normal form but not diagonalized). It would then immediately follow that

$$\|\mathbf{T}_0^m|_{E_1}\| \geq \|\mathbf{T}_0^m \mathbf{e}'_1\| / \|\mathbf{e}'_1\| = 1,$$

contradicting (2.11).

(2) We already know that 1 is an eigenvalue of  $\mathbf{M}$  and that the  $(N - 1)$ -dimensional vector  $(1, \dots, 1)$  is a left eigenvector for this eigenvalue. Since  $(\mathbf{T}_0)_{i+1, j+1} = \mathbf{M}_{ij}$ ,  $1 \leq i, j \leq N - 1$  and  $(\mathbf{T}_0)_{1j} = 0$  for  $j \geq 2$ , we find that any eigenvalue of  $\mathbf{M}$  is an eigenvalue of  $\mathbf{T}_0$ , with at least the same multiplicity. Since 1 is a simple eigenvalue of  $\mathbf{T}_0$ , it is, therefore, also a simple eigenvalue of  $\mathbf{M}$ . It then follows from arguments presented above that the right eigenvector  $\mathbf{a}$  for the eigenvalue 1 of  $\mathbf{M}$  can be normalized so that  $\sum_{n=1}^{N-1} \mathbf{a}_n = 1$ . As previously agreed, we then define

$$[\mathbf{v}_0(x)]_n = \mathbf{a}_{n-1}(1 - x) + \mathbf{a}_n x, \quad n = 1, \dots, N,$$

with the convention  $\mathbf{a}_0 = 0 = \mathbf{a}_N$ .

(3) From the normalization of  $\mathbf{a}$ , it now follows that, for all  $x \in [0, 1]$ ,

$$\mathbf{e}_1 \cdot \mathbf{v}_0(x) = \sum_{n=1}^N [\mathbf{v}_0(x)]_n = 1.$$

We now prove, by induction, that the same is true for all  $\mathbf{v}_j$ . Suppose  $\mathbf{e}_1 \cdot \mathbf{v}_j(x) = 1$ . Then, for all  $x \in [0, 1]$ ,

$$\mathbf{e}_1 \cdot \mathbf{v}_{j+1}(x) = \mathbf{e}_1 \cdot \mathbf{T}_{d_1(x)} \mathbf{v}_j(\tau x) = \mathbf{e}_1 \cdot \mathbf{v}_j(\tau x) = 1,$$

where we have used  $\mathbf{e}_1 \cdot \mathbf{T}_0 = \mathbf{e}_1 = \mathbf{e}_1 \cdot \mathbf{T}_1$ .

(4) Next we show that the  $\mathbf{v}_k(x)$  are uniformly bounded. Since  $\mathbf{e}_1 \cdot \mathbf{v}_k(x) = 1$  for all  $k, x$ , it follows that  $\mathbf{v}_k(x) - \mathbf{v}_l(x) \in E_1$  for all  $k, l, x$ . Hence

$$\begin{aligned} \|\mathbf{v}_{k+1}(x) - \mathbf{v}_{k(x)}\| &= \|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_k(x)} [\mathbf{v}_1(\tau^k x) - \mathbf{v}_0(\tau^k x)]\| \\ &\leq C \lambda^k \sup_{y \in [0, 1]} \|\mathbf{v}_1(y) - \mathbf{v}_0(y)\|. \end{aligned}$$

Consequently,

$$\begin{aligned} \|\mathbf{v}_k(x)\| &\leq \|\mathbf{v}_0(x)\| + \sum_{j=1}^k \|\mathbf{v}_j(x) - \mathbf{v}_{j-1}(x)\| \\ &\leq \sup_{y \in [0,1]} \|\mathbf{v}_0(y)\| + C(1-\lambda)^{-1} \sup_{y \in [0,1]} \|\mathbf{v}_1(y) - \mathbf{v}_0(y)\|, \end{aligned}$$

so that  $\|\mathbf{v}_k(x)\|$  is bounded uniformly in  $k$  and in  $x$ .

(5) Together with  $\mathbf{e}_1 \cdot \mathbf{v}_k(x) = 1$ , the uniform boundedness of the  $\mathbf{v}_k(x)$  implies that the  $\mathbf{v}_k$  constitute a Cauchy sequence in  $L^\infty$ -norm. Indeed, since

$$\begin{aligned} \|\mathbf{v}_{j+k}(x) - \mathbf{v}_j(x)\| &= \|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_j(x)}[\mathbf{v}_k(x) - \mathbf{v}_0(x)]\| \\ &\leq 2C\lambda^j \sup_{l,y} \|\mathbf{v}_l(y)\|, \end{aligned}$$

we find that  $\sup_{x \in [0,1]} \|\mathbf{v}_{j+k}(x) - \mathbf{v}_j(x)\|$  can be made arbitrarily small by choosing  $j$  large enough, independently of  $k$ . It follows that there exists a limit,

$$\mathbf{v}(x) = \lim_{j \rightarrow \infty} \mathbf{v}_j(x),$$

which is continuous since all the  $\mathbf{v}_j$  are continuous and the convergence is uniform in  $x$ ,

$$\sup_{x \in [0,1]} \|\mathbf{v}(x) - \mathbf{v}_j(x)\| \leq C\lambda^j.$$

Since every  $\mathbf{v}_j$  satisfies (2.5), (2.6), so does  $\mathbf{v}$ . It follows that the function  $f$  defined on  $[0, N]$  by

$$f(x) = [\mathbf{v}(x - [x])]_{[x]+1}$$

(where  $[x]$  denotes the largest integer not exceeding  $x$ ) is continuous, and that

$$(2.13) \quad \|f - f_j\|_{L^\infty} \leq C\lambda^j = C2^{-j\alpha}$$

with  $\alpha = |\ln \lambda| / \ln 2$ .

(6) Since  $f_j = Vf_{j-1}$ , the limit function  $f$  satisfies  $f = Vf$ , i.e.,  $f$  is a solution to (1.2), which is necessarily  $L^1$  since  $f$  is bounded and compactly supported. We have, moreover,  $\mathbf{e}_1 \cdot \mathbf{v}(x) = \lim_{j \rightarrow \infty} \mathbf{e}_1 \cdot \mathbf{v}_j(x) = 1$ , so that

$$\int_0^N dx f(x) = \int_0^1 dx \sum_{n=1}^N [\mathbf{v}(x)]_n = 1.$$

(7) The Hölder continuity follows from (2.13) and standard spline results translating estimates on how well a function can be approximated by piecewise polynomials into Hölder estimates on the function itself. (See, e.g., Theorem 6.10 in Schumaker (1981), which uses approximations by piecewise constant functions rather than piecewise linear  $f_j$ ; an estimate similar to (2.13) can also be proved for piecewise constant approximations to  $f$ .) For the sake of convenience, we also give a direct proof by the following short argument. Suppose that  $2^{-(j+1)} \leq y - x \leq 2^{-j}$ . Then there exists  $l \in \mathbb{N}$  so that one of the two following alternatives holds:  $(l-1)2^{-j} \leq x \leq y \leq l2^{-j}$  or  $(l-1)2^{-j} \leq x \leq l2^{-j} \leq y \leq (l+1)2^{-j}$ . We shall only discuss the second case; the first one is similar. We then have

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_j(x)| + |f_j(x) - f_j(l2^{-j})| \\ &\quad + |f_j(l2^{-j}) - f_j(y)| + |f_j(y) - f(y)| \\ &\leq 2C2^{-\alpha j} + |f_j(x) - f_j(l2^{-j})| + |f_j(y) - f_j(l2^{-j})|, \end{aligned}$$

by (2.13). Because of the choice of  $l$ , there exists  $k \in \mathbb{N}$  so that  $x' = x - k$  and  $l'2^{-j} = l2^{-j} - k$  are both in  $[0, 1]$ . We can, moreover, choose binary expansions for  $x'$  and  $l'2^{-j}$  with coinciding first  $j$  digits (choose the expansion ending in ones for  $l'2^{-j}$ , and if  $x'$  is dyadic, the expansion ending in zeros for  $x'$ ). It follows that

$$\begin{aligned} |f_j(x) - f_j(l'2^{-j})| &\leq \|v_j(x') - v_j(l'2^{-j})\| \\ &= \|\mathbf{T}_{d_1(x')} \cdots \mathbf{T}_{d_j(x')} [v_0(\tau^j x') - v_0(\tau^j(l'2^{-j}))]\| \\ &\leq C\lambda^j = C2^{-aj}, \end{aligned}$$

where we have used (2.11), the boundedness of  $v_0$ , and  $v_0(u) - v_0(u') \in E_1$  for all  $u, u'$ . Similarly we can bound  $|f_j(y) - f_j(l'2^{-j})|$ ; putting it all together leads to

$$|f(x) - f(y)| \leq C'2^{-aj} \leq C''|x - y|^\alpha,$$

which is (2.12).  $\square$

*Remarks.* (1) Note that the argument in point (7) would also work, in principle, if  $\lambda < \frac{1}{2}$ . In that case  $|f(x) - f(y)| \leq C|x - y|^{1+\varepsilon}$  would follow, with  $\varepsilon > 0$ , which is only possible if  $f \equiv \text{constant}$ . Since  $\text{support}(f) = [0, N]$  and  $f$  is continuous, this implies  $f \equiv 0$ . It follows that  $\lambda$  in (2.11) necessarily satisfies  $\lambda \geq \frac{1}{2}$ .

(2) Under the conditions of this theorem, for  $0 \leq x \leq 1$  all infinite products  $\mathbf{T}_{d_1(x)}\mathbf{T}_{d_2(x)}\mathbf{T}_{d_3(x)} \cdots$  of the matrices  $\mathbf{T}_0$  and  $\mathbf{T}_1$  converge to the limit matrices

$$\mathbf{T}_{d_1(x)}\mathbf{T}_{d_2(x)}\mathbf{T}_{d_3(x)} \cdots = \begin{bmatrix} f(x) & f(x) & \cdots & f(x) \\ f(x+1) & f(x+1) & \cdots & f(x+1) \\ \vdots & \vdots & \ddots & \vdots \\ f(x+N-1) & f(x+N-1) & \cdots & f(x+N-1) \end{bmatrix}.$$

Theorem 2.2 is similar to Theorem 5.1 in Micchelli and Prautzsch (1989), with the following differences. On one hand, [MP] are only interested in continuity, and technical condition (b) in their Theorem 5.1 is a little less tight than our (2.11), although both are very similar. On the other hand, they have extra conditions (a), (c), (d), which are here automatically satisfied because we have restricted ourselves to the case where the sum rule (2.2) holds.

To apply Theorem 2.2 we have to verify the technical condition (2.11). It might seem impossible to check in practice, since it involves the norms of infinitely many products of matrices. It turns out, however, that (2.11) can be reduced to a criterion that uses only a finite-time computer search. This is the constant of the next lemma.

LEMMA 2.3. *Define*

$$(2.14) \quad \lambda_m = \max_{\substack{d_j=0 \text{ or } 1 \\ j=1, \dots, m}} \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_1}\|^{1/m}.$$

*A necessary and sufficient condition for (2.11) to hold is that*

$$(2.15) \quad \lambda_m < 1 \quad \text{for some } m \in \mathbb{N}.$$

*Proof.* Suppose  $\lambda_m < 1$ . Write  $n = qm + r$ , with  $q, r \in \mathbb{N}$ ,  $0 \leq r < m$ . Then

$$\|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_n}|_{E_1}\| \leq \lambda_m^{mq} \lambda_r^r \leq \max(1, \lambda_1, \dots, \lambda_{m-1}^{m-1}) \lambda_m^{-r} \lambda_m^n.$$

This implies (2.11), with  $\lambda = \lambda_m < 1$ , and  $C = \lambda_m^{-m+1} \max(1, \dots, \lambda_{m-1}^{m-1})$ . Conversely, if (2.11) is satisfied, then  $\lambda_m \leq C^{1/m} \lambda$ , hence  $\lambda_m < 1$  for large enough  $m$ .  $\square$

In some examples, the smallest value  $m$  for which (2.14) holds may still be too large: up to a formidable  $2^m$  matrix norms may have to be checked for every candidate  $m$ . In the next section, we shall see some additional tricks to simplify the search.

The technical condition (2.11) or (2.14) on  $\mathbf{T}_0, \mathbf{T}_1$  can be interpreted as a “spectral” constraint. In fact, given two matrices  $\mathbf{S}_0, \mathbf{S}_1$ , we can define (Rota and Strang (1960), Daubechies and Lagarias (1991)) the joint spectral radius for  $\mathbf{S}_0, \mathbf{S}_1$  by

$$(2.16) \quad \hat{\rho}(\mathbf{S}_0, \mathbf{S}_1) = \limsup_{m \rightarrow \infty} [ \max_{\substack{d_j=0 \text{ or } 1 \\ j=1, \dots, m}} \|\mathbf{S}_{d_1} \cdots \mathbf{S}_{d_m}\|^{1/m} ].$$

In the case where the two matrices are identical, it is well known that (2.16) reduces to the spectral radius. We have the following.

LEMMA 2.4. *A necessary and sufficient condition for (2.11) to hold is that the joint spectral radius  $\hat{\rho}(\mathbf{T}_0|_{E_1}, \mathbf{T}_1|_{E_1}) < 1$ .*

*Proof.* We have  $\hat{\rho}(\mathbf{T}_0|_{E_1}, \mathbf{T}_1|_{E_1}) = \limsup_{m \rightarrow \infty} \lambda_m$ . If  $\hat{\rho}(\mathbf{T}_0|_{E_1}, \mathbf{T}_1|_{E_1}) < 1$ , then  $\lambda_m < 1$  follows for large enough  $m$ . On the other hand, the proof of Lemma 2.3 above shows that  $\lambda_m < 1$  implies  $\lambda_n \leq C^{1/n} \lambda_m$  for all  $n \in \mathbb{N}$ , which leads to  $\hat{\rho}(\mathbf{T}_0|_{E_1}, \mathbf{T}_1|_{E_1}) < 1$ .  $\square$

Note that if  $\hat{\rho}(\mathbf{T}_0|_{E_1}, \mathbf{T}_1|_{E_1}) = \lambda_0 < 1$ , then (2.11) holds for all  $\lambda > \lambda_0$ , but not necessarily for  $\lambda = \lambda_0$ .

*Remarks.* (1) As noted before, the sum rule  $\sum c_{2n} = \sum c_{2n+1}$  is not necessary for a continuous solution to exist. An example is the continuous function  $f(x) = x/2$  for  $0 \leq x \leq 2$ ,  $2 - x/2$  for  $2 \leq x \leq 4$ , zero otherwise, which satisfies  $f(x) = \frac{1}{2}f(2x) + f(2x - 2) + \frac{1}{2}f(2x - 4)$ , with  $\sum c_{2n} = 2$ ,  $\sum c_{2n+1} = 0$ . This example has been obtained by “stretching” an equation that does satisfy (2.1): the function  $\tilde{f}(x) = f(2x)$  satisfies

$$\tilde{f}(x) = \frac{1}{2}\tilde{f}(2x) + \tilde{f}(2x - 1) + \frac{1}{2}\tilde{f}(2x - 2).$$

Consequently, the matrices  $\mathbf{T}_0, \mathbf{T}_1$  still have many interesting properties in this case, even though they do not have a common left eigenvector for the eigenvalue 1.

(2) By the same argument as in the remark following Theorem 2.3, we necessarily have  $\lambda_m \geq \frac{1}{2}$  for any choice of  $m \in \mathbb{N}$ , where  $\lambda_m$  is as defined by (2.14).

(3) For any matrix  $\mathbf{T}$  the spectral radius  $\rho(\mathbf{T})$ ,

$$\rho(\mathbf{T}) = \max \{ |\mu|; \mu \text{ eigenvalue of } \mathbf{T} \},$$

is given by

$$\rho(\mathbf{T}) = \lim_{l \rightarrow \infty} \|\mathbf{T}^l\|^{1/l} = \inf_l \|\mathbf{T}^l\|^{1/l}.$$

If we restrict the choices  $d_1, \dots, d_m$  in (2.11) to either all zero or all 1, it follows, therefore, that

$$(2.17) \quad \rho(\mathbf{T}_0|_{E_1}), \rho(\mathbf{T}_1|_{E_1}) \leq \lambda,$$

i.e., all the other eigenvalues of  $\mathbf{T}_0, \mathbf{T}_1$  (excluding 1) have absolute value smaller than  $\lambda < 1$ . The condition  $\rho(\mathbf{T}_0|_{E_1}), \rho(\mathbf{T}_1|_{E_1}) < 1$  is, however, not sufficient to ensure that (2.11) holds. An example is given by

$$N = 3; \quad c_0 = -.75, \quad c_1 = .2, \\ c_2 = 1.75, \quad c_3 = .8.$$

In this case the spectra of  $\mathbf{T}_0, \mathbf{T}_1$  are  $\{1, -.75, .95\}$  and  $\{1, .8, .95\}$ , respectively. We have, therefore,  $\rho(\mathbf{T}_0|_{E_1}) = \rho(\mathbf{T}_1|_{E_1}) = .95$ . On the other hand, the spectrum of  $\mathbf{T}_0, \mathbf{T}_1$  is  $\{1, .50976\dots, -1.06226\dots\}$ , implying that  $\lambda_m \geq (1.06226\dots)^{1/2} > 1$  for any  $m$ .

(4) Because of the special structure (2.8) of  $\mathbf{T}_0, \mathbf{T}_1$ , it can easily be shown that

$$\begin{aligned} \text{spectrum}(\mathbf{T}_0) &= \{c_0\} \cup \text{spectrum}(\mathbf{M}), \\ \text{spectrum}(\mathbf{T}_1) &= \{c_N\} \cup \text{spectrum}(\mathbf{M}). \end{aligned}$$

In order for (2.11) to be satisfied, it is therefore necessary that  $|c_0|, |c_N| \leq 1$ .

(5) At the end of Deslauriers and Dubuc (1989), which mainly concerns the solutions of an equation of type (1.2) corresponding to a symmetric Lagrangian interpolation scheme, a conjecture is presented concerning generalizations to other, non-Lagrangian, interpolation schemes. Translated in our present terminology, this conjecture reads as follows.

**CONJECTURE 2.5.** *Assume that  $\sum_n c_{2n} = \sum_n c_{2n+1} = 1$ . Define the bounded operator  $\mathbf{A}$  on  $l^2(\mathbb{Z})$  by  $(\mathbf{A}a)_j = \sum_l c_{2j-l} a_l$ . Then there exists a continuous nontrivial solution to (1.2) if 1 is a nondegenerate eigenvalue of  $\mathbf{A}$  and if all the other eigenvalues of  $\mathbf{A}$  have modulus strictly smaller than 1.*

The operator  $\mathbf{A}$  does have a rather bothersome spectrum, however; we find that all the complex numbers with modulus strictly smaller than 1 are in the point spectrum of  $\mathbf{A}$ . For the simple case  $c_0 = 1, c_1 = c_{-1} = \frac{1}{2}, c_n = 0$  for  $|n| > 1$ , e.g., and  $\lambda \in \mathbb{C}, |\lambda| < 1$ , the sequence  $a^\lambda$  defined by

$$\begin{aligned} a^\lambda_{-n} &= 0, & n \in \mathbb{N} \\ a^\lambda_1 &= 0, & a^\lambda_2 = 1, & a^\lambda_{3 \cdot 2^m} = -2\lambda^m, & a^\lambda_{4 \cdot 2^m} = \lambda^m(\lambda + 1), \\ a^\lambda_n &= 0 & \text{if } n \geq 5, & & n \neq 3 \cdot 2^m \text{ or } 4 \cdot 2^m \end{aligned}$$

is clearly in  $l^2(\mathbb{Z})$ , and satisfies  $\mathbf{A}a^\lambda = \lambda a^\lambda$ . It follows that the closed unit disk is part of the spectrum of  $\mathbf{A}$ . In practice, it may be very hard to decide whether 1 is the only element on the unit circle that is not only in the spectrum of  $\mathbf{A}$ , but also a true eigenvalue of  $\mathbf{A}$ . For this reason, this conjecture, even if true, does not seem to give an easily checkable criterion for a given set of  $c_n$  to lead to a continuous solution of (1.2).

We can prove a result analogous to Conjecture 2.5. Define, for any two matrices  $\mathbf{S}_0, \mathbf{S}_1$ , their “generalized spectral radius” by Daubechies and Lagarias (1991)

$$\rho(\mathbf{S}_1, \mathbf{S}_2) = \limsup_{n \rightarrow \infty} [ \max_{\substack{d_j=0, \text{ or } 1 \\ j=1, \dots, n}} \rho(\mathbf{S}_{d_1} \cdots \mathbf{S}_{d_n})^{1/n} ],$$

where  $\rho(\mathbf{S})$  denotes the usual spectral radius. Berger and Wang (1991) prove that

$$\rho(\mathbf{S}_1, \mathbf{S}_2) = \hat{\rho}(\mathbf{S}_1, \mathbf{S}_2).$$

Consequently Theorem 2.2. and Lemma 2.4 give the following.

**THEOREM 2.6.** *Assume that the  $c_n, n = 0, \dots, N$  satisfy  $\sum c_{2n} = 1 = \sum c_{2n+1}$ . Then there exists a continuous nontrivial solution to (1.2) if  $\rho(\mathbf{T}_0|_{E_1}, \mathbf{T}_1|_{E_1}) < 1$ .*

It would be of interest to determine necessary and sufficient conditions on  $\{c_n\}$  for the existence of a continuous solution to (2.1) having  $\sum c_{2n} = 1 = \sum c_{2n+1}$ .

(6) The same analysis can be done for two-scale difference equations having larger (integer) values of  $k$ . In general, there will then be  $k$  different  $(N_0 + 1) \times (N_0 + 1)$ -matrices  $\mathbf{T}_0, \dots, \mathbf{T}_{k-1}$ , where  $N_0$  is the largest integer strictly smaller than  $N/(k - 1)$ . Instead of binary expansions of  $x \in [0, 1]$ , we take the expansion of  $x$  in base  $k$ . Otherwise, the proofs carry over without change.

**3. Higher-order regularity.** If  $L$  additional “sum rules” of the type (2.1) are satisfied, then techniques similar to those that proved continuity and Hölder continuity

in the preceding section can be used to show that  $f \in C^l$ , with  $l > 1$ . The enlarged set of “sum rules” is

$$(3.1) \quad \sum_{n=0}^N (-1)^n n^l c_n = 0 \quad \text{for } l = 0, 1, \dots, L;$$

this is equivalent to requiring that  $p(\xi) = \frac{1}{2} \sum_{n=0}^N c_n e^{in\xi}$  is divisible by  $(1 + e^{i\xi})^{L+1}$ . For  $L = 0$ , (3.1) reduces to (2.1). We define the vectors  $\mathbf{u}_j \in \mathbb{R}^N, j = 1, \dots, L + 1$ , by

$$(3.2) \quad (\mathbf{u}_j)_k = k^{j-1}, \quad k = 1, \dots, N.$$

The vector  $\mathbf{u}_1$  is a common left eigenvector of  $\mathbf{T}_0, \mathbf{T}_1$ , with eigenvalue 1. Similarly the  $\mathbf{u}_j$  lead us to left eigenvectors of  $\mathbf{T}_0$  and  $\mathbf{T}_1$  with eigenvalue  $2^{-j+1}$  (see below). This spectral analysis of  $\mathbf{T}_0, \mathbf{T}_1$  can then be used to prove the following generalization of Theorem 2.2.

**THEOREM 3.1.** *Assume that the  $c_n, n = 0, \dots, N$ , satisfy  $\sum_{n=0}^N c_n = 2$  and*

$$\sum_{n=0}^N (-1)^n n^l c_n = 0 \quad \text{for } l = 0, 1, \dots, L.$$

*For every  $m = 1, \dots, L + 1$ , define  $E_m$  to be the subspace of  $\mathbb{R}^N$  orthogonal to  $U_m = \text{Span}\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$  where  $\mathbf{u}_j = (1^{j-1}, 2^{j-1}, \dots, N^{j-1})$ . Assume that there exist  $\frac{1}{2} \leq \lambda < 1, 0 \leq l \leq L (l \in \mathbb{N})$  and  $C > 0$  such that, for all binary sequences  $(d_j)_{j \in \mathbb{N}}$ , and all  $m \in \mathbb{N}$ ,*

$$(3.3) \quad \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{L+1}}\| \leq C \lambda^m 2^{-m}.$$

*Then*

- (1) *There exists a nontrivial continuous  $L^1$ -solution  $f$  for the two-scale equation (1.2) associated with the  $c_n$ ;*
- (2) *This solution  $f$  is  $l$  times continuously differentiable;*
- (3) *If  $\lambda > \frac{1}{2}$ , then the  $l$ th derivative  $f^{(l)}$  of  $f$  is Hölder continuous, with exponent at least  $|\ln \lambda| / \ln 2$ ; if  $\lambda = \frac{1}{2}$ , then the  $l$ th derivative  $f^{(l)}$  of  $f$  is almost Lipschitz: it satisfies*

$$|f^{(l)}(x + t) - f^{(l)}(t)| \leq C |t| |\ln |t||.$$

*Remarks.* (1) The restriction  $\lambda \leq \frac{1}{2}$  means only that we pick the largest possible integer  $l \leq L$  for which (3.3) holds with  $\lambda < 1$ . If  $l = L$ , then we shall see below that necessarily  $\lambda \geq \frac{1}{2}$ ; if  $l < L$  and  $\lambda < \frac{1}{2}$ , then we could replace  $l$  by  $l + 1$  and  $\lambda$  by  $2\lambda$ , and (3.3) would hold for a larger integer  $l$ .

(2) The formulation of Theorem 3.1 implicitly assumes that  $L + 1 < N$ . If  $L + 1 = N$ , then  $U_{L+1} = \mathbb{R}^N, E_{L+1} = \{0\}$ , and condition (3.3) becomes meaningless. In the case  $L + 1 = N$ , the  $(N + 1)$  coefficients  $c_n$  are completely determined by the  $N$  “sum rules” (3.1) and the requirement  $\sum_{n=0}^N c_n = 2$ . The characteristic determinant of the resulting system of  $N + 1$  linear equations is different from zero (it can be written as a positive linear combination of positive Vandermonde determinants); the unique solution to the system is  $c_n = 2^{-N+1} \binom{N}{n}, n = 0, \dots, N$ . For  $N = 2$ , e.g., we find  $c_0 = \frac{1}{2}, c_1 = 1, c_2 = \frac{1}{2}$ . The corresponding function  $f(x)$  is given by

$$f(x) = \begin{cases} x & 0 \leq x \leq 1, \\ 2 - x & 1 \leq x \leq 2. \end{cases}$$

This function is Lipschitz but not  $C^1$ . This is typical of what happens for larger values of  $N$ , when  $L + 1 = N$ : the corresponding function  $f$  is the  $B$ -spline function of degree  $L$ , which is  $C^{L-1}$ , and its  $(L - 1)$ -th derivative is Lipschitz, but not everywhere differentiable. The points where  $f^{(L-1)}$  fails to be differentiable are  $0, 1, \dots, N$ , where the left

and right  $L$ th derivatives of  $f$  do not coincide. In most of what follows we shall implicitly assume  $L+1 < N$ .

To prove Theorem 3.1, we shall need several technical lemmas. The first one shows how the  $\mathbf{u}_j$  are related to left eigenvectors of  $\mathbf{T}_0, \mathbf{T}_1$  with eigenvalue  $2^{-j+1}$ .

LEMMA 3.2. *Assume that (3.1) holds. Let  $\mathbf{T}_0, \mathbf{T}_1$  be defined as in (2.8),  $\mathbf{u}_j, j = 1, \dots, L+1$  as in (3.2). Define  $U_j = \text{Linear span } \{\mathbf{u}_1, \dots, \mathbf{u}_j\}, j = 1, \dots, L+1$ . For all  $j = 1, \dots, L+1, 2^{-j}$  is an eigenvalue of both  $\mathbf{T}_0, \mathbf{T}_1$ . The corresponding left eigenvectors, denoted by  $\mathbf{e}_j^0, \mathbf{e}_j^1$  respectively, are both in  $U_j$ . We fix their normalization by requiring that  $\mathbf{e}_j^0 - \mathbf{u}_j, \mathbf{e}_j^1 - \mathbf{u}_j \in U_{j-1}$ . Then*

$$(3.4) \quad \mathbf{e}_j^0 = \sum_{k=1}^j (-1)^{j-k} \binom{j-1}{k-1} \mathbf{e}_k^1$$

or, equivalently,

$$(3.5) \quad \mathbf{e}_j^1 = \sum_{k=1}^j \binom{j-1}{k-1} \mathbf{e}_k^0.$$

*Proof.* (1) Define, for  $k=0, \dots, L$ ,

$$C_k = \sum_n (2n)^k c_{2n} = \sum_n (2n+1)^k c_{2n+1}.$$

We then easily check from the definitions (2.8), (3.2) of  $\mathbf{T}_0, \mathbf{T}_1$ , and the  $\mathbf{u}_j$  that

$$\begin{aligned} \mathbf{u}_j \cdot \mathbf{T}_1 &= 2^{-j+1} \sum_{n=1}^j \binom{j-1}{n-1} C_{j-n} \mathbf{u}_n, \\ \mathbf{u}_j \cdot \mathbf{T}_0 &= 2^{-j+1} \sum_{n=1}^j \binom{j-1}{n-1} \sum_{l=0}^{j-n} \binom{j-n}{l} C_l \mathbf{u}_n, \end{aligned}$$

for  $j = 1, \dots, L+1$ . It follows that, for every  $j, 1 \leq j \leq L+1$ , there are left eigenvectors for both  $\mathbf{T}_0, \mathbf{T}_1$ , with eigenvalue  $2^{-j+1}$ , in the linear span  $U_j$  of  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j\}$ . We denote these eigenvectors by  $\mathbf{e}_j^0, \mathbf{e}_j^1$ , respectively. Then

$$(3.6) \quad \mathbf{e}_j^0 = \sum_{n=1}^j a_{j,n} \mathbf{u}_n, \quad \mathbf{e}_j^1 = \sum_{n=1}^j b_{j,n} \mathbf{u}_n,$$

where we take  $a_{j,j} = 1 = b_{j,j}$ .

(2) From (3.6) and  $\mathbf{e}_j^0 \cdot \mathbf{T}_0 = 2^{-j+1} \mathbf{e}_j^0, \mathbf{e}_j^1 \cdot \mathbf{T}_1 = 2^{-j+1} \mathbf{e}_j^1$ , we derive

$$(3.7) \quad \begin{aligned} b_{j,n} &= \sum_{j=n}^j 2^{j-j} \binom{k-1}{n-1} C_{k-n} b_{j,k}, \\ a_{j,n} &= \sum_{k=m}^j \sum_{i=0}^{k-m} \binom{k-1}{m-1} \binom{k-m}{i} C_i 2^{j-k} a_{j,k}. \end{aligned}$$

These equations determine  $a_{j,n}, b_{j,n}$  recursively, starting from  $a_{j,j}, b_{j,j}$ . From (3.7) it follows that

$$a_{j,n} = \frac{(j-1)!}{(n-1)!} \alpha_{j-n}, \quad b_{j,n} = \frac{(j-1)!}{(n-1)!} \beta_{j-n},$$

where  $\alpha_0 = 1 = \beta_0$ , and the  $\alpha_k, \beta_k$  satisfy the recursions

$$(3.8a) \quad \beta_k = \sum_{i=0}^k \frac{1}{(k-i)!} C_{k-i} 2^i \beta_i$$

$$(3.8b) \quad \alpha_k = \sum_{i=0}^k \sum_{n=i}^k 2^i \frac{1}{(k-n)!(n-i)!} C_{k-n} \alpha_i$$

(3) Together with  $\alpha_0 = 1 = \beta_0$ , the recursions (3.8) imply that

$$(3.9) \quad \alpha_k = \sum_{i=0}^k (-1)^i \frac{1}{i!} \beta_{k-i}$$

or, equivalently,

$$\beta_k = \sum_{i=0}^k \frac{1}{i!} \alpha_{k-i}.$$

One way of checking this is to verify that if the  $\beta_i$  satisfy (3.8a), then the right-hand side of (3.9) satisfies (3.8b). It follows that

$$a_{j,n} = \sum_{k=n}^j \binom{j-1}{k-1} (-1)^{j-k} b_{k,n}.$$

This implies (3.4) and (3.5).  $\square$

Instead of the piecewise linear splines in § 2, here we shall use piecewise polynomial splines of degree  $2l + 1$ , where  $l$  is the same as in (3.3). We shall determine the initial spline function  $f_0$  for the iteration by fixing the values of  $f_0^{(k)}(m)$ ,  $0 \leq n \leq N$ ,  $0 \leq k \leq l$ . They will be defined in terms of the right eigenvectors of  $\mathbf{T}_0, \mathbf{T}_1$  for the eigenvalues  $1, \frac{1}{2}, \dots, 2^{-l}$ . We first show that these eigenvalues are all simple.

LEMMA 3.3. *Assume that (3.1) and (3.3) hold. Then the  $l+1$  eigenvalues  $1, \frac{1}{2}, \dots, 2^{-l}$  of  $\mathbf{T}_0, \mathbf{T}_1$  are all simple.*

*Proof.* Since the  $e_j^0, h = 1, \dots, L+1$ , are left eigenvectors of  $\mathbf{T}_0$  with eigenvalue  $2^{-j+1}$ , there exists an appropriate basis transformation  $\mathbf{B}$  such that  $\mathbf{B}\mathbf{T}_0\mathbf{B}^{-1}$  has the form

$$\left( \begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & & \vdots & \vdots & & \vdots \\ \vdots & & & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 2^{-L} & 0 & \cdots & 0 \end{array} \right),$$

$\mathbf{A} \qquad \qquad \qquad \mathbf{C}$

where  $\mathbf{A}, \mathbf{C}$  are  $(N-L-1) \times (L+1)$ - and  $(N-L-1) \times (N-L-1)$ -dimensional matrices, respectively. The roots of the characteristic equation for  $\mathbf{T}_0$  are, therefore,  $1, \frac{1}{2}, \dots, 2^{-L}$  together with the roots of the characteristic equation for  $\mathbf{C}$ . It is easily checked that  $\mathbf{C} = \mathbf{B}|_{E_{L+1}} \mathbf{T}_0|_{E_{L+1}} \mathbf{B}^{-1}|_{\mathbf{B}E_{L+1}}$ ; the spectral radius  $\rho(\mathbf{C})$  of  $\mathbf{C}$  can, therefore, be bounded by

$$\begin{aligned} \rho(\mathbf{C}) &= \limsup_{k \rightarrow \infty} \|\mathbf{C}^k\|^{1/k} \\ &\leq \limsup_{k \rightarrow \infty} [\|\mathbf{B}\|_{E_{L+1}} \|\mathbf{T}_0^k|_{E_{L+1}}\| \|\mathbf{B}^{-1}\|_{\mathbf{B}E_{L+1}}]^{1/k} \\ &\leq \limsup_{k \rightarrow \infty} [\|\mathbf{B}\|_{E_{L+1}} \|\mathbf{B}^{-1}\|_{\mathbf{B}E_{L+1}} C \lambda^k 2^{-lk}]^{1/k} \\ &= \lambda 2^{-l} < 2^{-l}. \end{aligned}$$

It follows that the eigenvalues  $1, \frac{1}{2}, \dots, 2^{-l}$  of  $\mathbf{T}_0$  are also simple. A similar argument applies to  $\mathbf{T}_1$ .  $\square$

Since the eigenvalues  $1, \frac{1}{2}, \dots, 2^{-l}$  of  $\mathbf{T}_0$  are also simple, it makes sense to introduce the corresponding right eigenvectors  $\tilde{\mathbf{e}}_k^0, 1 \leq k \leq l+1$ , which are uniquely determined, up to normalization. We fix their normalization by requiring

$$\mathbf{e}_k^0 \cdot \tilde{\mathbf{e}}_{k'}^0 = \delta_{kk'}, \quad 1 \leq k, k' \leq l+1.$$

We similarly define right eigenvectors  $\tilde{\mathbf{e}}_k^1$  of  $\mathbf{T}_1$  corresponding to the eigenvalues  $1, \frac{1}{2}, \dots, 2^{-l}$ , with normalization determined by

$$\mathbf{e}_k^1 \cdot \mathbf{e}_{k'}^1 = \delta_{kk'}, \quad 1 \leq k, k' \leq l+1.$$

We now define the piecewise polynomial spline for  $f_0$  of degree  $2l+1$  by

$$(3.10) \quad f_0^{(k)}(n) = \begin{cases} (-1)^k k! (\tilde{\mathbf{e}}_{k+1}^0)_{n+1}, & k=0, \dots, l, \\ & n=0, \dots, N-1, \\ 0, & k=0, \dots, l, \\ & n=N. \end{cases}$$

*Remark.* Because  $(\mathbf{T}_0)_{in} = c_0 \delta_{n1}$ , the vector  $(1, 0, \dots, 0)$  is a left eigenvector of  $\mathbf{T}_0$  with eigenvalue  $c_0$ . It follows that  $c_0 \notin \{1, \frac{1}{2}, \dots, 2^{-l}\}$ , since  $c_0 = 2^{-k}$ ,  $0 \leq k \leq l$  would imply the existence of two linearly independent left eigenvectors for  $2^{-k}$ . Consequently,  $(\tilde{\mathbf{e}}_{k+1}^0)_1 = (1, 0, \dots, 0) \cdot \tilde{\mathbf{e}}_{k+1}^0 = 0$  for  $0 \leq k \leq l$ . (We can prove completely analogously that the right eigenvectors  $\tilde{\mathbf{e}}_k^1$  of  $\mathbf{T}_1$  satisfy  $(\tilde{\mathbf{e}}_{k+1}^1)_N = 0$  for  $k=1, \dots, l$ .) For  $n=0$ , (3.10), therefore, specializes to

$$f_0^{(k)}(0) = 0, \quad k=0, \dots, l,$$

so that, despite appearances, there is complete symmetry between the constraints at zero and at  $N$ , the two ends of the support of  $f_0$ .

As in § 2, we again use the “folding” process that associates to a function  $g: \mathbb{R} \rightarrow \mathbb{R}$ , with support  $[0, N]$ , the vector valued function  $\mathbf{w}, [0, 1] \rightarrow \mathbb{R}^N$  by means of

$$[\mathbf{w}(x)]_n = g(x+n-1), \quad x \in [0, 1], \quad n=1, \dots, N.$$

The function  $g$  is  $l$  times continuously differentiable if and only if  $\mathbf{w}$  is  $l$  times continuously differentiable on  $[0, 1]$ ,

$$(3.11) \quad [\mathbf{w}^{(k)}(0)]_n = [\mathbf{w}^{(k)}(1)]_{n-1}, \quad 2 \leq n \leq N, \quad 0 \leq k \leq l,$$

$$(3.12) \quad [\mathbf{w}^{(k)}(0)]_1 = 0 = [\mathbf{w}^{(k)}(1)]_N, \quad 0 \leq k \leq l.$$

“Folding”  $f_0$  leads to the vector valued function  $\mathbf{v}_0: [0, 1] \rightarrow \mathbb{R}^N$ , each component of which is a polynomial of degree  $2l+1$ ;  $\mathbf{v}_0$  satisfies (3.11) and (3.12). In particular, for  $n=2, \dots, N$ ,  $k=0, \dots, l$ ,

$$(3.13) \quad \begin{aligned} [\mathbf{v}_0^{(k)}(0)]_n &= [\mathbf{v}_0^{(k)}(1)]_{n-1} = (-1)^k k! [\tilde{\mathbf{e}}_{k+1}^0]_n, \\ [\mathbf{v}_0^{(k)}(0)]_n &= [\mathbf{v}_0^{(k)}(1)]_N = 0. \end{aligned}$$

Iterating the linear operator  $V$  defined by (2.3) on  $f_0$  leads to a sequence  $f_j$  of piecewise polynomial splines of degree  $2l+1$  and finer knot sets; their “folded” versions  $\mathbf{v}_j$  can also be written as

$$\mathbf{v}_j = \mathbf{V}^j \mathbf{v}_0 \quad \text{or} \quad \mathbf{v}_j(x) = \mathbf{T}_{d_1(x)} \mathbf{T}_{d_2(x)} \cdots \mathbf{T}_{d_j(x)} \mathbf{v}_0(\tau^j x)$$

(see § 2). We now have the following lemma.

LEMMA 3.4. For every  $j \in \mathbb{N}$  and  $0 \leq k \leq \min(2l+1, L)$ ,

$$(3.14) \quad \mathbf{e}_{k+1}^0 \cdot \mathbf{v}_j(x) = (-1)^k x^k.$$

*Proof.* (1) We start by proving the assertion for  $j=0$ . By the construction of  $\mathbf{v}_0, \mathbf{e}_k^0 \cdot \mathbf{v}_0(x) = P_k(x)$  is a polynomial on  $[0, 1]$  of degree  $2l+1$ . It is completely determined by the values of the  $P_k^{(m)}(0), P_k^{(m)}(1), 0 \leq m \leq l$ . It, therefore, suffices to prove

$$(3.15) \quad P_k^{(m)}(0) = (-1)^k k! \delta_{km},$$

$$(3.16) \quad P_k^{(m)}(1) = \begin{cases} (-1)^k \frac{k!}{(k-m)!} & \text{if } m \leq k, \\ 0 & \text{if } m > k. \end{cases}$$

We have (use (3.13))

$$\begin{aligned} P_k^{(m)}(0) &= \mathbf{e}_{k+1}^0 \cdot \mathbf{v}_0^{(m)}(0) = (-1)^m m! \mathbf{e}_{k+1}^0 \cdot \tilde{\mathbf{e}}_{m+1}^0 \\ &= (-1)^k k! \delta_{km}, \end{aligned}$$

which proves (3.15), and

$$(3.17) \quad P_k^{(m)}(1) = \mathbf{e}_{k+1}^0 \cdot \mathbf{v}_0^{(m)}(1) = (-1)^m m! \sum_{n=1}^{N-1} (\mathbf{e}_{k+1}^0)_n (\tilde{\mathbf{e}}_{m+1}^0)_{n+1}.$$

Due to the special structure of  $\mathbf{T}_0, \mathbf{T}_1$ , we have

$$(\tilde{\mathbf{e}}_{m+1}^0)_{n+1} = (\tilde{\mathbf{e}}_{m+1}^1)_n.$$

Since  $(\tilde{\mathbf{e}}_{m+1}^1)_N = 0$  (see the remark after (3.10)), (3.17) reduces to

$$P_k^{(m)}(1) = (-1)^m m! \mathbf{e}_{k+1}^0 \cdot \tilde{\mathbf{e}}_{m+1}^1,$$

which can be computed with the help of Lemma 3.2:

$$\begin{aligned} \mathbf{e}_{k+1}^0 \cdot \tilde{\mathbf{e}}_{m+1}^1 &= \sum_{r=0}^k (-1)^{k-r} \binom{k}{r} \mathbf{e}_{r+1}^1 \cdot \tilde{\mathbf{e}}_{m+1}^1 \\ &= \sum_{r=0}^k (-1)^{k-r} \binom{k}{r} \delta_{rm} \\ &= \begin{cases} 0 & \text{if } k < m, \\ (-1)^{k-m} \binom{k}{m} & \text{if } k \geq m. \end{cases} \end{aligned}$$

Hence

$$P_k^{(m)}(1) = \begin{cases} (-1)^k \frac{k!}{(k-m)!} & \text{if } k \geq m \\ 0 & \text{if } k < m, \end{cases}$$

which proves (3.16) and establishes (3.14) for  $j=0$ .

(2) For higher values of  $j$  we proceed by induction. Suppose (3.14) holds for  $\mathbf{v}_j$ . Then

$$(3.18) \quad \mathbf{e}_{k+1}^0 \cdot \mathbf{v}_{j+1}(x) = \mathbf{e}_{k+1}^0 \cdot \mathbf{T}_{d_1(x)} \mathbf{v}_j(\tau x).$$

If  $x \leq \frac{1}{2}$ , then  $d_1(x) = 0$  and (3.18) becomes

$$\begin{aligned} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}_{j+1}(x) &= 2^{-k} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}_j(2x) \\ &= 2^{-k} (-1)^k (2x)^k = (-1)^k x^k, \end{aligned}$$

which is (3.14) again for index  $j + 1$ . If  $x > \frac{1}{2}$ , then we use Lemma 3.2 again:

$$\begin{aligned} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}_{j+1}(x) &= \sum_{r=0}^k (-1)^{k-r} \binom{k}{r} \mathbf{e}_{r+1}^1 \cdot \mathbf{T}_1 \mathbf{v}_j(2x-1) \\ &= \sum_{r=0}^k (-1)^{k-r} \binom{k}{r} 2^{-r} \sum_{s=0}^r \binom{r}{s} \mathbf{e}_{s+1}^0 \cdot \mathbf{v}_j(2x-1) \\ &= \sum_{s=0}^k (2x-1)^2 \binom{k}{s} (-1)^2 2^{-s} \sum_{t=10}^{k-s} \binom{k-s}{t} (-1)^{k-s-t} 2^{-t} \\ &= (-2)^{-k} \sum_{s=0}^k \binom{k}{s} (2x-1)^s \\ &= (-2)^{-k} (2x)^k = (-1)^k x^k, \end{aligned}$$

which is again (3.14). This proves the lemma.  $\square$

We need one more technical lemma before we attack the proof of Theorem 3.1.

LEMMA 3.5. Assume that (3.1) and (3.3) hold. If  $\lambda > \frac{1}{2}$ , then there exists  $C > 0$  so that, for all binary sequences  $(d_j)_{j \in \mathbb{N}}$ , all  $m \in \mathbb{N}$ , and all  $k$  with  $l \leq k \leq L$ ,

$$(3.19) \quad \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{k+1}}\| \leq C\lambda^m 2^{-ml}.$$

If  $\lambda = \frac{1}{2}$ , then (3.19) still holds for  $l + 1 \leq k \leq L$ ; for  $k = l$ , (3.19) is replaced by the slightly weaker bound

$$(3.19') \quad \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{l+1}}\| \leq Cm 2^{-m(l+1)}.$$

Note that we implicitly assume  $l < L$ ; if  $l = L$ , then the lemma is trivial.

*Proof.* (1) We prove (3.19) by induction, working from high to low values of  $k$ . For  $k = L$ , (3.19) is (3.3), and we have nothing to prove.

(2) Because of the existence of the left eigenvectors  $e_m^0, e_m^1, 1 \leq m \leq L + 1$  for  $\mathbf{T}_0, \mathbf{T}_1$ , with eigenvalues  $1, \frac{1}{2}, \dots, 2^{-L}$ , and the relationships (3.4) and (3.5) between these eigenvectors, there exists an appropriate basis transformation  $\mathbf{B}$  so that  $(\mathbf{B}^{-1,t} \mathbf{e}_m^0)_r = \delta_{mr}$ , and  $\mathbf{B} E_m = \{\mathbf{w}; \mathbf{w}_r = 0 \text{ for } r \leq m\}$ . The matrices  $\mathbf{B} \mathbf{T}_0 \mathbf{B}^{-1}, \mathbf{B} \mathbf{T}_1 \mathbf{B}^{-1}$  have the form

$$(3.20) \quad \mathbf{B} \mathbf{T}_0 \mathbf{B}^{-1} = \left( \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & \\ 0 & \frac{1}{2} & & \vdots & 0 \\ \vdots & \vdots & & 0 & \\ 0 & 0 & \cdots & 0 & 2^{-L} \end{array} \right),$$

$$(3.21) \quad \mathbf{B} \mathbf{T}_1 \mathbf{B}^{-1} = \left( \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & \\ t_{2,1} & \frac{1}{2} & & \vdots & 0 \\ \vdots & & & 0 & \\ t_{L+1,1} & \cdots & t_{L+1,L} & 2^{-L} & \end{array} \right).$$

The matrices  $\mathbf{B}|_{E_{k+1}} \mathbf{T}_d|_{E_{k+1}} \mathbf{B}^{-1}|_{\mathbf{B} E_{k+1}}$  are then obtained by deleting the first  $k + 1$  rows and columns of (3.20), (3.21). Let us denote these  $(N - k - 1) \times (N - k - 1)$ -dimensional matrices by  $\mathbf{S}_d^k$ . It then follows that, for  $l \leq k \leq L - 1$

$$(3.22) \quad \mathbf{S}_d^k = \left( \begin{array}{c|ccc} 2^{-k-1} & 0 & \cdots & 0 \\ \mathbf{C}_d^k & & & \mathbf{S}_d^{k+1} \end{array} \right)$$

where  $\mathbf{C}_0^k, \mathbf{C}_1^k$  are  $(N - k - 2)$ -dimensional column vectors.

(3) Let us now assume that (3.19) holds for index  $k + 1 \geq l + 1$ . Then

$$\begin{aligned} \|\mathbf{S}_{d_1}^{k+1} \cdots \mathbf{S}_{d_m}^{k+1}\| &\leq \|\mathbf{B}\|_{E_{k+2}} \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}\|_{E_{k+2}} \|\mathbf{B}^{-1}\|_{\mathbf{B}E_{k+2}} \\ &\leq C \|\mathbf{B}^{-1}\| \|\mathbf{B}\| \lambda^m 2^{-ml} \leq C_{k+1} \lambda^m 2^{-ml}. \end{aligned}$$

From (3.22) it follows that

$$(3.23) \quad \mathbf{S}_{d_1}^k \cdots \mathbf{S}_{d_m}^k = \left( \begin{array}{ccc|ccc} 2^{-(k+1)m} & & & 0 & 0 & \cdots & 0 \\ \hline \sum_{r=1}^m 2^{-(k+1)(m-r)} \mathbf{S}_{d_1}^{k+1} & \cdots & \mathbf{S}_{d_{r-1}}^{k+1} \mathbf{C}_{d_r}^k & \mathbf{S}_{d_1}^{k+1} & \cdots & \mathbf{S}_{d_m}^{k+1} & \end{array} \right),$$

where

$$\begin{aligned} &\left\| \sum_{r=1}^m 2^{-(k+1)(m-r)} \mathbf{S}_{d_1}^{k+1} \cdots \mathbf{S}_{d_{r-1}}^{k-1} \mathbf{C}_{d_r}^k \right\| \\ &\leq C_{k+1} \sum_{r=1}^m 2^{-(k+1)(m-r)} (\lambda 2^{-l})^{r-1} \cdot \max(\|\mathbf{C}_0^k\|, \|\mathbf{C}_1^k\|) \\ &\leq C' 2^{-(k+1)(m-1)} \frac{(\lambda 2^{k+1-l})^m - 1}{\lambda 2^{k+1-l} - 1} \\ &\leq C'' \lambda^m 2^{-lm}, \end{aligned}$$

where we have assumed that  $\lambda 2^{k+1-l} > 1$ , i.e.,  $k \geq l + 1$  or  $\lambda > \frac{1}{2}$ . We come back below to the case  $k = l + 1$ ,  $\lambda = \frac{1}{2}$ .

It follows that the three pieces in (3.23) are all bounded by  $C'' \lambda^m 2^{-lm}$  (since  $k \geq l + 1$  and  $\lambda \geq \frac{1}{2}$ ). There exists, therefore,  $C_k$  (independent of  $m$  or the  $d_j$ ) so that

$$\|\mathbf{S}_{d_1}^k \cdots \mathbf{S}_{d_m}^k\| \leq C_k \lambda^m 2^{-ml}.$$

This proves the induction step if  $k \geq l + 1$  or  $\lambda > \frac{1}{2}$ .

(4) We now treat the case  $k = l + 1$ ,  $\lambda = \frac{1}{2}$ . In this case we find

$$\begin{aligned} \left\| \sum_{r=1}^m 2^{-k(m-r)} \mathbf{S}_{d_1}^{k+1} \cdots \mathbf{S}_{d_{r-1}}^{k+1} \mathbf{C}_{d_r}^k \right\| &\leq C_{k+1} \sum_{r=1}^m 2^{-(l+1)(m-r)} 2^{-(l+1)(r-1)} \\ &= C_{k+1} 2^{-(l+1)(m-1)} (m-1). \end{aligned}$$

Therefore, the three pieces in (3.23) can all be bounded by  $C' 2^{-(l+1)m} m$ , which proves (3.19').  $\square$

We are now ready to attack the proof of Theorem 3.1.

*Proof of Theorem 3.1.* (1) As a consequence of Lemma 3.4 we have, for  $0 \leq k \leq l$ , and all  $j, j' \in \mathbb{N}$ ,

$$\mathbf{e}_{k+1}^0 \cdot [\mathbf{v}_j(x) - \mathbf{v}_{j'}(x)] = 0,$$

hence  $\mathbf{v}_j(x) - \mathbf{v}_{j'}(x) \in E_{l+1}$  for all  $x \in [0, 1]$ , all  $j, j' \in \mathbb{N}$ .

(2) It follows that

$$\begin{aligned} \|\mathbf{v}_{j+1}(x) - \mathbf{v}_j(x)\| &= \|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_j(x)} [\mathbf{v}_1(\tau^j x) - \mathbf{v}_0(\tau^j x)]\| \\ &\leq C \gamma_\lambda(j) \lambda^j 2^{-lj} \sup_{y \in [0,1]} \|\mathbf{v}_1(y) - \mathbf{v}_0(y)\|, \end{aligned}$$

where

$$\gamma_\lambda(j) = \begin{cases} 1 & \text{if } \lambda > \frac{1}{2}, \\ j & \text{if } \lambda = \frac{1}{2}. \end{cases}$$

Consequently,

$$\begin{aligned} \|\mathbf{v}_j(x)\| &\leq \|\mathbf{v}_0(x)\| + \sum_{r=0}^{j-1} \|\mathbf{v}_{r+1}(x) - \mathbf{v}_r(x)\| \\ &\leq C \left[ 1 + \sum_{r=0}^{j-1} \gamma_\lambda(r) \lambda^r 2^{-lr} \right], \end{aligned}$$

which is bounded uniformly in  $j$  and  $x$ .

(3) We now use this uniform bound to prove that the  $\mathbf{v}_j$  constitute a Cauchy sequence in  $L^\infty$ . We have

$$\begin{aligned} \|\mathbf{v}_{j+r}(x) - \mathbf{v}_j(x)\| &= \|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_j(x)}[\mathbf{v}_r(x) - \mathbf{v}_0(x)]\| \\ &\leq C \gamma_\lambda(j) \lambda^j 2^{-lj}, \end{aligned}$$

which can be made arbitrarily small by choosing  $j$  large enough, independent of  $r$ . The  $\mathbf{v}_j$  tend, therefore, to a limit  $\mathbf{v}$ , which satisfies

$$(3.24) \quad \|\mathbf{v}(x) - \mathbf{v}_j(x)\| \leq C \gamma_\lambda(j) \lambda^j 2^{-lj}.$$

(4) Since all the  $\mathbf{v}_j$  satisfy

$$\begin{aligned} [\mathbf{v}_j(0)]_{n+1} &= [\mathbf{v}_j(1)]_n, \quad 1 \leq n \leq N-1, \\ [\mathbf{v}_j(0)]_1 &= 0 = [\mathbf{v}_j(1)]_N, \end{aligned}$$

so does  $\mathbf{v}$ . It follows that  $\mathbf{v}$  can be “unfolded” into a continuous function  $f$ , for which (3.24) translates into

$$(3.25) \quad \|f - f_j\|_{L^\infty} \leq C \phi_\lambda(2^{-j}),$$

where  $\phi_\lambda(t) = t^{l+\alpha}$  if  $\lambda > \frac{1}{2}$ ,  $\phi_\lambda(t) = |\log_2 t| t^{l+1}$  if  $\lambda = \frac{1}{2}$ , with  $\alpha = |\ln \lambda| / \ln 2$ .

(5) Formula (3.25) tells us again how well  $f$  can be approximated by piecewise polynomials, and this can be translated into smoothness estimates on  $f$ . This result is no doubt well-known to spline experts, but we could not find in the literature a full proof of the exact result we needed. For convenience’s sake, we offer the following proof “from scratch.”

Note first that, by Lemma 3.4,

$$\mathbf{e}_{k+1}^0 \cdot \mathbf{v}_j^{(l)}(x) = \begin{cases} 0 & \text{if } k < l, \\ (-1)^l l! & \text{if } k = l. \end{cases}$$

It follows that  $\mathbf{v}_j^{(l)}(x) - \mathbf{v}_{j'}^{(l)}(x') \in E_{l+1}$  for all  $j, j'$ , all  $x, x'$ . On the other hand, the recursive definition of the  $\mathbf{v}_j$  leads to

$$\mathbf{c}_j^{(l)}(x) = 2^l \mathbf{T}_{d_1(x)} \mathbf{v}_{j-1}^{(l)}(\tau x).$$

Together, these two observations imply

$$\begin{aligned} \|\mathbf{v}_{j+1}^{(l)}(x) - \mathbf{v}_j^{(l)}(x)\| &= 2^{lj} \|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_j(x)}[\mathbf{v}_1^{(l)}(\tau^j x) - \mathbf{v}_0^{(l)}(\tau^j x)]\| \\ &\leq C \lambda^j \gamma_\lambda(j) \sup_{y \in [0,1]} [\|\mathbf{v}_1^{(l)}(y)\| + \|\mathbf{v}_0^{(l)}(y)\|], \end{aligned}$$

which can be used, similarly to the argument in (2) above, to prove that the  $\mathbf{v}_j^{(l)}(x)$  are bounded uniformly in  $j$  and  $x$ , and that the  $\mathbf{v}_j^{(l)}$  are a Cauchy sequence in  $L^\infty$ . The limit of the  $\mathbf{v}_j^{(l)}$  is necessarily the  $l$ th derivative of  $\mathbf{v}$ . Moreover,

$$\|\mathbf{v}^{(l)}(x) - \mathbf{v}_j^{(l)}(x)\| \leq C \lambda^j \gamma_\lambda(j),$$

uniformly in  $x$ .

The remainder of our argument is similar to (7) in the proof of Theorem 2.3. Take  $x \leq y$  in  $[0, N]$  so that  $2^{-(j+1)} \leq y - x \leq 2^{-j}$  (a slight extension of the argument can be used on a neighborhood of  $[0, N]$ , so that the results are also true at zero and  $N$ ). Then there exists  $m \in \mathbb{N}$  so that either  $(m - 1)2^{-j} \leq x \leq y \leq m2^{-j}$  or  $(m - 1)2^{-j} \leq x \leq m2^{-j} \leq y \leq (m + 1)2^{-j}$ . We discuss the second case; the first is similar. Then

$$|f^{(l)}(x) - f^{(l)}(y)| \leq 2C\lambda^j \gamma_\lambda(j) + |f_j^{(l)}(x) - f_j^{(l)}(m2^{-j})| + |f_j^{(l)}(y) - f_j^{(l)}(m2^{-j})|,$$

by the bound on  $\mathbf{v}^{(l)} - \mathbf{v}_j^{(l)}$ . There exists  $n$  so that  $x' = x - n$  and  $m'2^{-j} = m2^{-j} - n$  are both in  $[0, 1]$ ; moreover, we can choose binary expansions for  $x'$  and  $m'2^{-j}$  with the same  $j$  first digits. Consequently,

$$\begin{aligned} |f_j^{(l)}(x) - f_j^{(l)}(m2^{-j})| &\leq \|\mathbf{v}_j^{(l)}(x') - \mathbf{v}_j^{(l)}(m'2^{-j})\| \\ &= 2^{lj} \|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_j(x)} [\mathbf{v}_0^{(l)}(\tau^j x') - \mathbf{v}_0^{(l)}(\tau^j (m'2^{-j}))]\| \\ &\leq C\lambda^j \gamma_\lambda(j), \end{aligned}$$

where we have used (3.3), the uniform boundedness of  $\mathbf{v}_0^{(l)}(y)$ , and  $\mathbf{v}_0^{(l)}(y) - \mathbf{v}_0^{(l)}(y') \in E_{l+1}$  for all  $y, y'$ . A similar bound holds for  $|f_j^{(l)}(y) - f_j^{(l)}(m2^{-j})|$ . Putting it all together, we obtain, for  $2^{-(j+1)} \leq |x - y| \leq 2^{-j}$ ,

$$|f^{(l)}(x) - f^{(l)}(y)| \leq C' \lambda^j \gamma_\lambda(j),$$

which translates to

$$|f^{(l)}(x) - f^{(l)}(y)| \leq C|x - y|^{\|\log_2 \lambda\|}$$

if  $\lambda > \frac{1}{2}$ , and to

$$|f^{(l)}(x) - f^{(l)}(y)| \leq C|\log_2 |x - y|| |x - y|$$

if  $\lambda = \frac{1}{2}$ .  $\square$

*Remarks.* (1) If  $l = L$ , then we would not need Lemma 3.5, so that the assumption  $\lambda \geq \frac{1}{2}$  would never be used. The argument of (7) would work if (3.3) held for  $\lambda < \frac{1}{2}$ , but it would lead to a Hölder exponent  $\alpha$  larger than 1 for  $f^{(l)}$ , hence to  $f^{(l)} \equiv 0$ . This is incompatible with the fact that  $f$  is compactly supported, except if  $f \equiv 0$ . It follows that for matrices  $\mathbf{T}_0, \mathbf{T}_1$  constructed as in (2.8), the constant  $\lambda$  in (3.3) is necessarily  $\geq \frac{1}{2}$  if  $l = L$ .

(2) We chose  $f_0$  so that  $\mathbf{v}_0^{(k)}(0) = (-1)^k k! \tilde{\mathbf{e}}_{k+1}^0$  (see (3.13)). Since, for  $x < 2^{-j}$ ,

$$\mathbf{v}_j^{(k)}(x) = 2^{kj} \mathbf{T}_0^j \mathbf{v}_0^{(k)}(2^j x),$$

it follows that

$$\begin{aligned} \mathbf{v}_j^{(k)}(0) &= (-1)^k k! 2^{kj} \mathbf{T}_0^j \tilde{\mathbf{e}}_{k+1}^0 \\ &= (-1)^k k! \tilde{\mathbf{e}}_{k+1}^0, \end{aligned}$$

where we have used  $\mathbf{T}_0 \tilde{\mathbf{e}}_{k+1}^0 = 2^{-k} \tilde{\mathbf{e}}_{k+1}^0$ . Hence the  $f_j^{(k)}(n)$ ,  $0 \leq k \leq l$ , are independent of  $j$ . Because of the bounds (3.25), this implies also, for  $0 \leq k \leq l$ ,

$$\begin{aligned} f^{(k)}(n) &= f_j^{(k)}(n) = (-1)^k k! [\tilde{\mathbf{e}}_{k+1}^0]_{n+1}, \quad 0 \leq n \leq N - 1, \\ f^{(k)}(N) &= 0. \end{aligned} \tag{3.26}$$

In § I.5 we had already seen that the  $f^{(k)}(n)$  were linked to the right eigenvectors of  $\mathbf{M}$  for the eigenvalues  $2^{-k}$ , but it wasn't clear how to choose their normalization, explicit in (3.26). Note that  $f_0^{(k)}(n) = f^{(k)}(n)$  for  $0 \leq k \leq l$  implies that  $f_j^{(k)}$  converges to  $f^{(k)}$  for

$0 \leq k \leq l$  (see Theorem 4.2 in part I): not only do the  $f_j$  themselves converge, their derivatives up to  $l$ th order do as well.

(3) By continuity, Lemma 3.4 carries over to  $\mathbf{v}(x)$ : for  $0 \leq k \leq \min(2l+1, L)$ ,

$$(3.27) \quad \mathbf{e}_{k+1}^0 \cdot \mathbf{v}(x) = (-1)^k x^k.$$

The following lemma states that (3.27) holds for all  $k \leq L$ , even if  $2l+1 < L$ .

LEMMA 3.6. For all  $0 \leq k \leq L$ ,  $\mathbf{e}_{k+1}^0 \cdot \mathbf{v}(x) = (-1)^k x^k$ .

*Proof.* (1) If  $2l+1 \geq L$ , then we have nothing to prove. We assume  $2l+1 < L$  and restrict ourselves to  $k > 2l+1$ .

(2) We only need to prove the lemma for dyadic rationals  $x$ ; by continuity it then holds for all  $x \in [0, 1]$ . Take, therefore,  $x = n2^{-j}$ . The proof works by induction on  $j$ .

(3) If  $j = 0$ , then only  $n = 0, 1$  lead to  $x \in [0, 1]$ . Since  $\mathbf{v}(x) = \mathbf{T}_{d_1(x)}\mathbf{v}(\tau x)$ , we have  $\mathbf{v}(0) = \mathbf{T}_0\mathbf{v}(0)$ ,  $\mathbf{v}(1) = \mathbf{T}_1\mathbf{v}(1)$ . Consequently,

$$\mathbf{e}_{k+1}^0 \cdot \mathbf{v}(0) = \mathbf{e}_{k+1}^0 \cdot \mathbf{T}_0\mathbf{v}(0) = 2^{-k} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}(0);$$

hence  $\mathbf{e}_{k+1}^0 \cdot \mathbf{v}(0) = 0$ . Similarly,  $\mathbf{e}_{k+1}^1 \cdot \mathbf{v}(1) = 0$ , which implies (use Lemma 3.2)

$$\begin{aligned} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}(1) &= \sum_{m=0}^k (-1)^{k-m} \binom{k}{m} \mathbf{e}_{m+1}^1 \cdot \mathbf{v}(1) \\ &= (-1)^k \mathbf{e}_1^1 \cdot \mathbf{v}(1) = (-1)^k, \end{aligned}$$

where we have used (3.27) for  $k = 0$ . This proves  $\mathbf{e}_{k+1}^0 \cdot \mathbf{v}(x) = (-1)^k x^k$  for all dyadic rationals  $x = n2^{-j}$  with  $j = 0$ .

(4) Suppose that the lemma holds for all  $x = n2^{-j}$ ,  $j$  fixed,  $0 \leq n \leq 2^j$ . Take  $y$  of the form  $r2^{-j-1}$ ,  $r \in \mathbb{N}$ ,  $0 \leq r \leq 2^{j+1}$ . If  $y \leq \frac{1}{2}$ , then

$$\begin{aligned} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}(y) &= \mathbf{e}_{k+1}^0 \cdot \mathbf{T}_0\mathbf{v}(2y) \\ &= 2^{-k} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}(2y) = (-1)^k 2^{-k} (2y)^k \quad (\text{by induction}) \\ &= (-1)^k y^k. \end{aligned}$$

If  $y \geq \frac{1}{2}$ , then

$$\mathbf{e}_{k+1}^0 \cdot \mathbf{v}(y) = \mathbf{e}_{k+1}^0 \cdot \mathbf{T}_1\mathbf{v}(2y-1).$$

As in the proof of Lemma 3.4, we have

$$\mathbf{e}_{k+1}^0 \mathbf{T}_1 = 2^{-k} \sum_{s=1}^k \binom{k}{s} (-1)^{k-s} \mathbf{e}_{s+1}^0.$$

Hence

$$\begin{aligned} \mathbf{e}_{k+1}^0 \cdot \mathbf{v}(y) &= 2^{-k} \sum_{s=0}^k \binom{k}{s} (-1)^{k-s} (1-2y)^s \quad (\text{by induction}) \\ &= 2^{-k} (-2y)^k = (-y)^k. \end{aligned}$$

This proves the lemma.  $\square$

As in § 2, the condition (3.3) can be reduced to a condition involving only a finite number of matrix norms or interpreted as a ‘‘spectral’’ constraint:

$$(3.3) \Leftrightarrow \hat{\rho}(\mathbf{T}_0|_{E_{L+1}}, \mathbf{T}_1|_{E_{L+1}}) < 2^{-l},$$

$$(3.28) \quad \Leftrightarrow \exists m \in \mathbb{N} \text{ so that}$$

$$\gamma_m^L = \max_{\substack{d_j=0 \text{ or } 1 \\ j=1, \dots, m}} \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{L+1}}\|^{1/m} < 2^{-l}.$$

If (3.28) holds for some  $m$ , then (3.3) holds with  $\lambda = 2^l \gamma_m^L$ .

Even the condition  $\gamma_m^L < 2^{-l}$  for some  $m \in \mathbb{N}$  may be hard to check in practice: for any fixed  $m$ , it is necessary to verify the norms of  $2^m$  matrices. Some simplification may occur if the  $c_n$  are symmetric,  $c_j = c_{N-j}$ , but even then the number of matrices is huge if  $m$  is large. In the case of the Lagrangian interpolation scheme with dilation factor 2 and 5 nodes (see § 5), we find that  $\gamma_m^1 < \frac{1}{2}$  only if  $m \geq 7$ . Because the  $c_n$  are symmetric in this case, the number of matrices to check is only  $2^6$  instead of  $2^7$ , but that still amounts to a large number of computations for a fairly simple example. We can reduce the number of computations significantly by some simple tricks listed in the following proposition.

PROPOSITION 3.7. *The following are all equivalent to (3.3).*

(1<sup>0</sup>) *For some  $m \in \mathbb{N}$ ,*

$$(3.29) \quad \gamma_m^L < 2^{-l}.$$

(2<sup>0</sup>) *For some  $N \times N$ -matrix  $\mathbf{B}$  with  $\det \mathbf{B} \neq 0$ , and for some  $m \in \mathbb{N}$ ,*

$$(3.30) \quad \gamma_{m;\mathbf{B}}^L < 2^{-l},$$

where

$$\gamma_{m;\mathbf{B}}^k = \max_{\substack{d_j=0 \text{ or } 1 \\ j=1, \dots, m}} \|\bar{\mathbf{T}}_{d_1} \cdots \bar{\mathbf{T}}_{d_m}|_{\mathbf{B}E_{k+1}}\|^{1/m},$$

with  $\bar{\mathbf{T}}_d = \mathbf{B}\mathbf{T}_d\mathbf{B}^{-1}$ .

(3<sup>0</sup>) *There exists a finite collection of “building blocks”  $D^j, j = 1, \dots, J$ , with  $D^j = \{d_1^j, d_2^j, \dots, d_{k_j}^j\}$ , each  $d_n^j = 0$  or 1, which is complete, in the sense that every dyadic sequence can be written as a sequence of blocks  $D^j$  and for which*

$$(3.31) \quad \max_{j=1, \dots, J} t_{D^j}^L < 2^{-l},$$

where  $t_{D^j}^m = \|\mathbf{T}_{d_1^j} \cdots \mathbf{T}_{d_{k_j}^j}|_{E_{m+1}}\|^{1/k_j}$ .

*Proof.* (1) The proof of the equivalence (3.3)  $\Leftrightarrow$  (3.29) is similar to the proof of Lemma 2.3.

(2) We prove (3.29)  $\Leftrightarrow$  (3.30). Since

$$\|\bar{\mathbf{T}}_{d_1} \cdots \bar{\mathbf{T}}_{d_n}|_{\mathbf{B}E_{L+1}}\| \leq \|\mathbf{B}\| \|\mathbf{B}^{-1}\| \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_n}|_{\mathbf{B}E_{L+1}}\|,$$

we have  $\gamma_{n;\mathbf{B}}^L \leq [\|\mathbf{B}\| \|\mathbf{B}^{-1}\|]^{1/n} \gamma_n^L$ . By the same argument used in the proof of Lemma 2.7,  $\gamma_n^L \leq C^{1/n} \gamma_m^L$  for some  $C > 0$ . If  $\gamma_m^L < 2^{-l}$ , it follows, therefore, that  $\gamma_{n;\mathbf{B}}^L < 2^{-l}$  for large enough  $n$ . This proves (3.29)  $\Rightarrow$  (3.30). The converse implication is proved in the same way.

(3) The implication (3.29)  $\Rightarrow$  (3.31) is obvious: it suffices to take the  $2^m$  “building blocks,” each consisting of  $m$  entries, with every entry zero or 1.

(4) Finally, we prove (3.31)  $\Rightarrow$  (3.29), which completes the proof. Suppose (3.31) is satisfied. Define  $K = \max \{k_j; j = 1, \dots, J\}$ ,  $C = \max \{\|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_n}|_{E_{L+1}}\|; n \leq K, d_j = 0$  or 1 for each  $j\}$ . Take now any fixed sequence  $(d_j)_{j \in \mathbb{N}}$ , and  $n \in \mathbb{N}$ . The  $n$ -tuple  $(d_1 \cdots d_n)$  can be written as a sequence of building blocks  $D^j$ , followed by a stretch of entries that has at most length  $K$ ,

$$(d_1 \cdots d_n) = (D^{j_1} \cdots D^{j_n} \tilde{D}).$$

If  $\alpha = \max t_{D^j}^L$ , and  $|D^j|$  denotes the number of entries in  $D^j$ , then we have, therefore,

$$\begin{aligned} \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_n}|_{E_{L+1}}\| &\leq 2^{-l(|D^{j_1}| + \dots + |D^{j_n}|)} \alpha^{|D^{j_1}| + \dots + |D^{j_n}|} C \\ &\leq 2^{-lm} \alpha^m C 2^{lK} \alpha^{-K}, \end{aligned}$$

which amounts to (3.3).  $\square$

Note that the bounds on  $\gamma_m^L$ , on  $\gamma_{m;\mathbf{B}}^L$ , or on the  $t_{D^j}^L$  all lead to lower bounds for the Hölder exponent of  $f^{(L)}$ , since (3.29)–(3.31) are equivalent with (3.3), with  $2^l \gamma_m^L$ ,  $2^l \gamma_{m;\mathbf{B}}^L$ , or  $2^l \max t_{D^j}^L$  playing the role of  $\lambda$ .

It is, of course, possible to combine (3.30) and (3.31) and to define  $\bar{t}_{D_j}^L$  (for the matrices  $\bar{\mathbf{T}}_d = \mathbf{B}\mathbf{T}_d\mathbf{B}^{-1}$ ). A bound on the  $\bar{t}_{D_j}^L$ , similar to (3.31), is again equivalent with (3.3).

*Remark.* The proof of Theorem 3.1 presented here relies on arguments which translate convergence rate to  $f$  by spline functions into regularity estimates for  $f$ . There exist generalizations of lattice two-scale difference equations to higher dimensions for which it does not seem possible to find appropriate spline functions playing the role of the  $f_j$  here (one such generalization is outlined in the Appendix). For this reason we present a sketch of a different proof in the Appendix, a fully detailed version which would be longer and more complicated than the proof in this section, but which does not use spline functions.

**4. Local regularity and fractal sets.** We assume again that the  $c_n, n=0, \dots, N$  satisfy the  $L+1$  sum rules (3.2) (with  $L+1 < N$ ), as well as  $\sum_{n=0}^N c_n = 2$ . If (3.3) is satisfied, then Theorem 3.1 tells us that  $f$  is  $l$  times continuously differentiable, and that  $f^{(l)}$  is Hölder continuous with exponent  $|\ln \lambda|/\ln 2$ . The proof for the uniform Hölder continuity of  $f^{(l-1)}$  uses that, for any point  $x$ ,

$$(4.1) \quad \|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_m(x)}|_{E_{L+1}}\| \leq 2^{-ml} \lambda^m.$$

In some cases, more accurate bounds on the left-hand side of (4.1) can be obtained for some values of  $x$ , depending on the relative frequency of the digits 0 and 1 in the binary expansion of  $x$ . This can then be used to compute local Hölder exponents, which may be larger than the uniformly valid Hölder exponents for  $f^{(l-1)}$ . More precisely, we have the following theorem.

**THEOREM 4.2.** *Assume that the  $c_n, n=0, \dots, N$ , satisfy the  $L+1$  sum rules (3.2), and assume that  $\sum_{n=0}^N c_n = 2$ , with  $L < N - 1$ . Assume that there exist  $m \in \mathbb{N}$ , and  $\mu_0, \mu_1 < 1$  such that*

$$(4.2) \quad \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{L+1}}\| \leq 2^{-\lambda} \mu_1^{s_m} \mu_0^{m-s_m}$$

for all possible combinations of  $d_j, d_j = 0$  or  $1, j = 1, \dots, m$ , with  $s_m = \sum_{j=1}^m d_j$ .

Let  $f$  be the solution  $f$  of the two-scale difference equation (1.2) associated with the  $c_n$ , and let  $f^{(l)}$  be its  $l$ th derivative, which exists and is Hölder continuous, with exponent  $\min(|\ln \mu_1|, |\ln \mu_2|)/\ln 2$ , by Theorem 3.1. Take  $x \in [0, N]$ . Assume that the decimal part of the binary expansion of  $x$  satisfies

$$(4.3) \quad r_m(x) = \frac{1}{m} \sum_{j=1}^m d_j(x) \text{ tends to a limit } r(x) \text{ as } m \rightarrow \infty,$$

$$(4.4) \quad 0 < r(x) < 1.$$

Then the following holds:

- (1) For all  $\epsilon > 0$ , there exist  $\delta > 0$  and  $C < \infty$  such that

$$(4.5) \quad |f^{(L)}(x) - f^{(L)}(x+t)| \leq C|t|^{\alpha(x)-\epsilon} \quad \text{if } |t| < \delta,$$

where  $\alpha(x) = -\min\{1, [(1-r(x))|\ln \mu_0| + r(x)|\ln \mu_1|]/\ln 2\}$ , with  $r(x) = r(x - \lfloor x \rfloor)$  as defined by (4.4).

- (2) If  $\mu_0 \geq \frac{1}{2} > \mu_1$  and  $r(x) > (\ln 2 - |\ln \mu_0|)/|\ln \mu_1| - |\ln \mu_0|$ , or if  $\mu_1 \geq \frac{1}{2} > \mu_0$  and  $r(x) < (|\ln \mu_0| - \ln 2)/(|\ln \mu_0| - |\ln \mu_1|)$ , then  $f^{(l)}$  is differentiable in  $x$ .

To prove this theorem we need the following lemma.

LEMMA 4.2. Take  $x \in [0, 1]$ . Assume that (4.3) and (4.4) are satisfied. Then, for all  $\varepsilon > 0$ , there exists an  $N_\varepsilon$  so that, for all  $m \geq N_\varepsilon$ , the  $m$  first digits of the binary expansions of  $x$  and  $x + t$  are identical whenever  $|t| < 2^{-m(1+\varepsilon)-1}$ .

Proof. (1) Fix  $\varepsilon > 0$ . Choose  $\delta$  small enough so that

$$(4.6) \quad [1 - r(x) - \delta]^{-1}2\delta \leq \varepsilon \quad \text{and} \quad [r(x) - \delta]^{-1}2\delta \leq \varepsilon.$$

There exists  $N_\varepsilon$  so that  $|r_m(x) - r(x)| \leq \delta$  for  $m \geq N_\varepsilon$ .

(2) Now choose  $m \geq N_\varepsilon$  and  $0 \leq t < 2^{-m(1+\varepsilon)-1}$ . If  $d_{m+1}(x) = 0$ , then it follows from  $t < 2^{-m-1}$  that  $d_j(x) = d_j(x+t)$  for all  $j \leq m$ , and we are done. Suppose that  $d_{m+1}(x) = 1 = d_{m+2}(x) = \dots = d_{m+s}(x)$ , and  $d_{m+s+1}(x) = 0$ . Then the condition (4.6) implies an upper bound on  $s$ , since  $r_{m+s}(x) \leq r(x) + \delta$ , while  $r_m(x) \geq r(x) - \delta$ . Using  $r_{m+s}(x) = (m+s)^{-1}[mr_m(x) + s]$  together with these restrictions leads to  $s \leq m2\delta[1 - r(x) - \delta]^{-1} \leq m\varepsilon$ . Therefore,  $t < 2^{-m(1+\varepsilon)-1} \leq 2^{-(m+s)-1}$ . Since  $d_{m+s+1}(x) = 0$ , this implies  $d_j(x) = d_j(x+t)$  for all  $j \leq m+s$ .

(3) The argument for  $-2^{-m(1-\varepsilon)-1} < t \leq 0$  is similar. If  $d_{m+1}(x) = 1$ , then  $d_j(x) = d_j(x+t)$  for all  $j < m$  because  $|t| < 2^{-m-1}$ . If  $d_{m+1}(x) = 0 = d_{m+2}(x) = \dots = d_{m+s}(x)$ , and  $d_{m+s+1}(x) = 1$ , then  $s \leq m2\delta[r(x) - \delta]^{-1} \leq m\varepsilon$ . Hence  $t > -2^{-(m+s)-1}$ , which implies  $d_j(x) = d_j(x+t)$  for all  $j \leq m+s$ .  $\square$

We now proceed with the proof of Theorem 4.1.

Proof of Theorem 4.1. (1) Since  $\mu_0, \mu_1 < 1$ , (4.2) implies (3.29), hence (3.3), so that  $f$  is  $l$  times continuously differentiable. By Lemma 3.6 its “folded” version  $\mathbf{v}$  satisfies

$$\mathbf{e}_{k+1}^0 \cdot \mathbf{v}^{(l)}(y) = (-1)^l l! \delta_{lk} \quad \text{for } 0 \leq k \leq l,$$

which implies

$$\mathbf{v}^{(l)}(y) - \mathbf{v}^{(l)}(y') \in E_{l+1} \quad \text{for all } y, y' \in [0, 1].$$

(2) It follows from (4.2) that there exists  $C > 0$  so that, for all  $p \in \mathbb{N}$ ,

$$(4.7) \quad \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_p}|_{E_{L+1}}\| \leq C2^{-lp} \mu_1^s \mu_0^{p-s_p}$$

(by the same argument as in the proof of Lemma 2.3). The same arguments used in the proof of Lemma 3.5 can then be used to derive that, for  $p \geq 1$ ,

$$(4.8) \quad \|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_p}|_{E_{l+1}}\| \leq Cp2^{-lp} \mu_1^s \mu_0^{p-s_p},$$

where the constant  $C$  may be different from that in (4.7), and where we have introduced the extra  $p$  for the case where  $\mu_1^s \mu_0^{n-s_n} = 2^{-n}$  for some  $n$ .

(3) Fix  $x \in [0, N]$ , such that  $r(x)$  is well defined and  $0 < r(x) < 1$ . Note that this excludes all dyadic rationals  $x$ , since these have  $r(x) = 0$  or  $1$ , depending on which binary expansion is chosen (for dyadic rationals, there are two binary expansions). Fix  $\varepsilon > 0$ . There exists  $\gamma > 0$  so that

$$(4.9) \quad \begin{aligned} & [(1 - r(x) - \gamma/2)|\ln \mu_0| + (r(x) - \gamma/2)|\ln \mu_1|]/(1 + \gamma) \\ & \geq (1 - r(x))|\ln \mu_0| + r(x)|\ln \mu_1| - \varepsilon/2 = \alpha(x) - \varepsilon/2. \end{aligned}$$

Define  $\delta_1 = 2^{-N_\gamma(1+\gamma)-1}$ , where  $N_\gamma$  is chosen as in the proof of Lemma 4.2.

(4) We have  $x = n + y$ , with  $0 < y < 1$ ,  $n = 0, 1, \dots$ , or  $N - 1$ . Take  $|t| < \delta = \min(\delta_1, x - n, n + 1 - x)$ . Then  $\lfloor x + t \rfloor = \lfloor x \rfloor = n$ . It follows that

$$\begin{aligned} |f^{(l)}(x) - f^{(l)}(x+t)| &= |\mathbf{v}_{n+1}^{(l)}(x-n) - \mathbf{v}_{n+1}^{(l)}(x+t-n)| \\ &\leq \|\mathbf{v}^{(l)}(y) - \mathbf{v}^{(l)}(y+t)\|. \end{aligned}$$

(5) Since  $|t| < \delta$ , there exist  $p \geq N_\gamma$  so that

$$(4.10) \quad 2^{-(p+1)(1+\gamma)-1} \leq |t| < 2^{-p(1+\gamma)-1}.$$

By Lemma 4.2,  $d_j(y) = d_j(y+t)$  for  $j \leq p$ . Hence

$$\mathbf{v}^{(l)}(y) - \mathbf{v}^{(l)}(y+t) = 2^{pl} \mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)} [\mathbf{v}^{(l)}(\tau^p y) - \mathbf{v}^{(l)}(\tau^p y + 2^p t)].$$

It follows from (4.8) that

$$(4.11) \quad \|\mathbf{v}^{(l)}(y) - \mathbf{v}^{(l)}(y+t)\| \leq 2^{pl} C 2^{-pl} p \mu_1^{pr_p(y)} \mu_0^{p(1-r_p(y))}.$$

(6) From the proof of Lemma 4.2 we have

$$|r_p(y) - r(y)| \leq \delta,$$

with  $[r(y) - \delta]^{-1} 2\delta \leq \gamma$ , hence  $\delta \leq r(y)\gamma/2 \leq \gamma/2$ . Substituting this in (4.11) we find

$$\begin{aligned} \|\mathbf{v}^{(l)}(y) - \mathbf{v}^{(l)}(y+t)\| &\leq C p \mu_1^{p[r(y)-\gamma/2]} \mu_0^{p[1-r(y)-\gamma/2]} \\ &\leq C' p 2^{-[(p+1)(1+\gamma)][r(y)-\gamma/2] \ln \mu_1 + [1-r(y)-\gamma/2] \ln \mu_0} / \ln 2 \\ &\leq C'' (1 + |\ln |t||) |t|^{\alpha(x) - \varepsilon/2} \\ &\leq C''' |t|^{\alpha(x) - \varepsilon}, \end{aligned}$$

where we have used (4.9). This proves the first part of the theorem.

(7) Suppose now that  $\mu_0 \geq \frac{1}{2} > \mu_1$ , and that  $r(x) > (\ln 2 - |\ln \mu_0| / |\ln \mu_1| - |\ln \mu_0|)$ , i.e.,  $\mu_1^{r(x)} \mu_0^{1-r(x)} < \frac{1}{2}$ . Choose  $\gamma$  so that

$$\mu_1^{r(x)-\gamma/2} \mu_0^{1-r(x)-\gamma/2} \leq \frac{1}{2} (\frac{1}{2} + \mu_1^{r(x)} \mu_0^{1-r(x)}) = \lambda(x)/2 < \frac{1}{2}$$

and

$$2^{-2\gamma} \geq \lambda(x).$$

Choose  $N_\gamma$  corresponding to  $\gamma$ , as in Lemma 4.2. Then, as before, for  $p \geq N_\gamma$ ,  $|r_p(y) - r(y)| \leq \gamma/2$ ; hence

$$\mu_1^{r_p(x)} \mu_0^{1-r_p(x)} \leq \lambda(x)/2 < \frac{1}{2}.$$

It follows then from (4.2) that there exists  $C > 0$  so that, for  $p \geq N_\gamma$ ,

$$\|\mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)}|_{E_{l+1}}\| \leq 2^{-p(l+1)} \lambda(x)^p.$$

The same arguments used in the proof of Lemma 3.5 lead to

$$(4.12) \quad \|\mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)}|_{E_{l+2}}\| \leq C 2^{-p(l+1)} \lambda(x)^p.$$

(We assume  $\lambda(x) > \frac{1}{2}$ ; if necessary we replace  $\lambda(x)$  by  $\frac{1}{2} + \varepsilon$ .) Since  $\lambda(x) < 1$ , this means that the eigenvalue  $2^{-(l+1)p}$  of  $\mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)}$  is simple. It follows that  $\mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)}$  has a left eigenvector  $\mathbf{e}_{l+2}(p; y)$  and a right eigenvector  $\tilde{\mathbf{e}}_{l+2}(p; y)$  for this eigenvalue, both uniquely determined, except for their normalization. By the structure of  $\mathbf{T}_0$  and  $\mathbf{T}_1$ ,  $\mathbf{e}_{l+2}(p; y)$  is necessarily a linear combination of the  $\mathbf{u}_{k+1}$ ,  $0 \leq k \leq l+1$ , or equivalently, of the  $\mathbf{e}_{k+1}^0$ ,  $0 \leq k \leq l+1$ ; we fix its normalization by requiring that the coefficient of  $\mathbf{e}_{l+2}^0$  (which is necessarily nonzero) is equal to 1. We also fix the normalization of  $\tilde{\mathbf{e}}_{l+2}(p; y)$  by the requirement

$$\mathbf{e}_{l+2}(p; y) \cdot \tilde{\mathbf{e}}_{l+2}(p; y) = 1,$$

or equivalently, since  $\tilde{\mathbf{e}}_{l+2}(p; y) \in E_{l+1}$ ,  $\mathbf{e}_{l+2}^0 \cdot \tilde{\mathbf{e}}_{l+2}(p; y) = 1$ .

(8) The vectors  $\tilde{\mathbf{e}}_{l+2}(p; y)$  are uniformly bounded in  $p$  and converge to a limit as  $p \rightarrow \infty$ , as shown by the following argument. We have

$$\begin{aligned} &\tilde{\mathbf{e}}_{l+2}(p; y) - \tilde{\mathbf{e}}_{l+2}(p+1; y) \\ &= 2^{(l+1)p} \mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)} [\tilde{\mathbf{e}}_{l+2}(p; y) - 2^{l+1} \mathbf{T}_{d_{p+1}(y)} \tilde{\mathbf{e}}_{l+2}(p+1; y)]. \end{aligned}$$

Since  $\mathbf{e}_{l+2}^0 \cdot \mathbf{T}_d = 2^{-(l+1)}\mathbf{e}_{l+2}^0 +$  linear combination of  $\mathbf{e}_{k+1}^0, 0 \leq k \leq l$ , we find  $\mathbf{e}_{l+2}^0 \cdot [\tilde{\mathbf{e}}_{l+2}(p+y) - 2^{l+1}\mathbf{T}_{d_{p+1}(y)}\tilde{\mathbf{e}}_{l+2}(p+1; y)] = 0$ , i.e.,  $\tilde{\mathbf{e}}_{l+2}(p; y) - 2^{l+1}\mathbf{T}_{d_{p+1}(y)}\tilde{\mathbf{e}}_{l+2}(p+1; y) \in E_{l+2}$ . Hence, by (4.12),

$$(4.13) \quad \|\tilde{\mathbf{e}}_{l+2}(p; y) - \tilde{\mathbf{e}}_{l+2}(p+1; y)\| \leq C\lambda(x)^p [\|\tilde{\mathbf{e}}_{l+2}(p; y)\| + \|\tilde{\mathbf{e}}_{l+2}(p+1; y)\|],$$

which implies, for large enough  $p$ ,

$$\begin{aligned} \|\tilde{\mathbf{e}}_{l+2}(p+1; y)\| &\leq [1 + C\lambda(x)^p][1 - C\lambda(x)^p]^{-1}\|\tilde{\mathbf{e}}_{l+2}(p; y)\| \\ &\leq \exp[3C\lambda(x)^p]\|\tilde{\mathbf{e}}_{l+2}(p; y)\|. \end{aligned}$$

Hence, for all  $p > p_0$ , where  $p_0$  is chosen large enough,

$$\|\tilde{\mathbf{e}}_{l+2}(p; y)\| \leq \exp[3C\lambda(x)^{p_0}(1 - \lambda(x))^{-1}]\|\tilde{\mathbf{e}}_{l+2}(p_0; y)\| \leq C',$$

and, by (4.13),

$$\|\tilde{\mathbf{e}}_{l+2}(p+r; y) - \tilde{\mathbf{e}}_{l+2}(p; y)\| \leq 2CC'\lambda(x)^p(1 - \lambda(x))^{-1},$$

so that the  $\tilde{\mathbf{e}}_{l+2}(p; y)$  form a Cauchy sequence in  $p$ , with limit  $\tilde{\mathbf{e}}_{l+2}(y)$ .

(9) Any  $\mathbf{u}$  in  $E_{l+1}$  can be written as

$$\mathbf{u} = \tilde{\mathbf{e}}_{l+2}(p; y)(\mathbf{e}_{l+2}^0 \cdot \mathbf{u}) + \mathbf{u}',$$

where  $\mathbf{u}' \in E_{l+2}$ , since  $\mathbf{e}_{l+2}^0 \cdot \mathbf{u}' = 0$  and  $\|\mathbf{u}'\| \leq C\|\mathbf{u}\|$ , with  $C$  independent of  $p$ , because the  $\tilde{\mathbf{e}}_{l+2}(p; y)$  are bounded uniformly in  $p$ .

Choose now  $p_1 = \max(p_0, N_\gamma)$ , with  $p_0, N_\gamma$  as in (7) and (6). For  $|t| < 2^{-p_1(1+\gamma)-1}$  there exists  $p \geq p_1$  so that

$$2^{-(p+1)(1+\gamma)-1} \leq |t| < 2^{-p(1+\gamma)-1},$$

which implies that  $x$  and  $x+t$ , or  $y$  and  $y+t$ , have the same  $p$  first digits in their binary expansion. Then

$$\mathbf{v}^{(l)}(y+t) - \mathbf{v}^{(l)}(y) = 2^{lp}\mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)}\mathbf{u},$$

with  $\mathbf{u} = \mathbf{v}^{(l)}(\tau^p y + 2^p t) - \mathbf{v}^{(l)}(\tau^p y) \in E_{l+1}$ . By Lemma 3.6,  $\mathbf{e}_{l+2}^0 \cdot \mathbf{u} = (-1)^{l+1}(l+1)!2^p t$ , so that

$$\begin{aligned} \mathbf{v}^{(l)}(y) - \mathbf{v}^{(l)}(y+t) &= 2^{lp}\mathbf{T}_{d_1(y)} \cdots \mathbf{T}_{d_p(y)}[(-1)^{l+1}(l+1)!2^p t\tilde{\mathbf{e}}_{l+2}(p; y) + \mathbf{u}'] \\ &= (-1)^{l+1}(l+1)!t\tilde{\mathbf{e}}_{l+2}(p; y) + \mathbf{u}'', \end{aligned}$$

with  $\|\mathbf{u}''\| \leq 2^{lp}C2^{-p(l+1)}\lambda(x)^p$  by (4.12), by the  $p$ -incident bound on  $\mathbf{u}$ , and the boundedness of  $\mathbf{v}^{(l)}$ . Because  $\lambda(x) \leq 2^{-\gamma}$ , it follows that,

$$(4.14) \quad \begin{aligned} &\left\| \frac{\mathbf{v}^{(l)}(y+t) - \mathbf{v}^{(l)}(y)}{t} - (-1)^{l+1}!\tilde{\mathbf{e}}_{l+2}(y) \right\| \\ &\leq (l+1)!\|\tilde{\mathbf{e}}_{l+2}(p; y) - \tilde{\mathbf{e}}_{l+2}(y)\| + C2^{\gamma+2}2^{-\gamma p}. \end{aligned}$$

As  $|t| \rightarrow 0, p \rightarrow \infty$ , and the right-hand side of (4.14) tends to zero, so that  $\mathbf{v}$  is  $(l+1)$  times differentiable in  $y$  (with  $(l+1)$ -th derivative  $(-1)^{l+1}(l+1)!\tilde{\mathbf{e}}_{l+2}(y)$ ). Hence,  $f$  is  $(l+1)$  times differentiable in  $x = [x] + y$ , and the theorem is proved.  $\square$

Under the conditions of Theorem 4.1, we find that different Hölder exponents  $\alpha_r$  correspond to the sets  $V_r$ ,

$$(4.15) \quad V_r = \{x \in [0, N]; r(x) = r(x - [x]) = r\}.$$

These sets are fractal sets. Their Hausdorff dimension can be computed explicitly; it is given by the following theorem, conjectured in I. J. Good (1941) and proved in H. G. EGGLESTON (1949).

**THEOREM 4.3.** For  $x \in [0, 1]$ ,  $k \in \mathbb{N}$ ,  $k \geq 2$ , and  $0 \leq l \leq k - 1$ , define  $r_n(x; l, k)$  to be the number of times, divided by  $n$ , that the digit 1 occurs among the first  $n$  digits of the expansion of  $x$  in basis  $k$ . For any  $k$ -tuple  $(r_0, r_1, \dots, r_{k-1}) = r$  with  $0 \leq r_l \leq 1$  for all  $l \leq k - 1$  and  $\sum_{l=0}^{k-1} r_l = 1$ , define

$$V_r(k) = \{x[0, 1]; \lim_{n \rightarrow \infty} r_n(x; l, k) = r_l \text{ for } l = 0, \dots, k - 1\}.$$

Then  $V_r$  has fractal dimension  $\delta$  defined by

$$k^{-\delta} = \prod_{l=0}^{k-1} r_l^{r_l}.$$

*Remark.* Note that the choice  $r_l = k^{-1}$ ,  $l = 0, \dots, k - 1$  leads to  $\delta = 1$ . This was to be expected, since in this case  $V_r$  contains all the normal numbers and, therefore, has Hausdorff dimension 1.

Specializing to  $k = 2$ , we find therefore that  $V_r$ , as defined by (4.15), has fractal dimension

$$\delta = \frac{r|\ln r| + (1 - r)|\ln(1 - r)|}{\ln 2}.$$

**COROLLARY 4.4.** We assume the same conditions as in Theorem 4.1, with  $\frac{1}{2} \leq \mu_0, \mu_1$ . Then, for any  $\alpha$ ,  $\min(|\ln \mu_0|, |\ln \mu_1|) < \alpha < \max(|\ln \mu_0|, |\ln \mu_1|)$ , the set  $S_\alpha$  of  $x \in [0, N]$  on which  $f$  is Hölder continuous with exponent at least  $\alpha$  is a fractal set with Hausdorff dimension  $d_\alpha$ , with

$$d_\alpha = 1 \quad \text{if } \alpha \leq \frac{|\ln \mu_0| + |\ln \mu_1|}{2 \ln 2},$$

$$d_\alpha \cong \frac{r|\ln r| + (1 - r)|\ln(1 - r)|}{\ln 2} \quad \text{if } \alpha \geq \frac{|\ln \mu_0| + |\ln \mu_1|}{2 \ln 2},$$

with

$$r = \frac{\alpha \ln 2 - |\ln \mu_0|}{|\ln \mu_1| - |\ln \mu_0|}.$$

*Proof.* The proof follows immediately from

$$\dim(V_r) = \frac{r|\ln r| + (1 - r)|\ln(1 - r)|}{\ln 2}. \quad \square$$

As announced in the introduction, we thus find a hierarchy of fractal sets  $S_\alpha$ , with decreasing fractal dimension  $d_\alpha$  for increasing Hölder exponents  $\alpha$ . Similar fractal sets can be introduced if  $\mu_0 < \frac{1}{2} \leq \mu_1$  or  $\mu_1 < \frac{1}{2} \leq \mu_0$ .

*Remarks.* (1) The different Hölder exponents, and the associated hierarchy of fractal sets do not occur as obviously if the coefficients  $c_n$ ,  $n = 0, \dots, N$  are symmetric, i.e., if  $c_n = c_{N-n}$ ,  $n = 0, \dots, N$ . In this case we find indeed that

$$\mathbf{OT}_0\mathbf{O}^{-1} = \mathbf{T}_1,$$

where  $\mathbf{O}$  is the orthogonal matrix defined by

$$\mathbf{O}_{ij} = \delta_{N+1-i-j}.$$

We easily check that the subspaces  $U_l$  are invariant under  $\mathbf{O}$ ,

$$\mathbf{O}u_l = \sum_{n=1}^l \binom{l-1}{n-1} (N+1)^{l-n} (-1)^{n+1} u_n.$$

It follows that the  $E_k$  are invariant under  $\mathbf{O}' = \mathbf{O}^{-1} = \mathbf{O}$ . ( $\mathbf{O}$  is an involution.) Consequently,

$$\mathbf{T}_l|_{V_k} = \mathbf{O}|_{V_k} \mathbf{T}_0|_{V_k} \mathbf{O}^{-1}|_{V_k}$$

for all  $l = 1, \dots, L+1$ , and

$$\|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{V_l}\| = \|\mathbf{T}_{1-d_1} \cdots \mathbf{T}_{1-d_m}|_{V_l}\|.$$

If an inequality of type (4.2) were to hold for the  $\mathbf{T}_d$ , this symmetry would immediately imply that the same inequality would hold if  $\mu_0, \mu_1$  were both replaced by  $\min(\mu_0, \mu_1)$ . The different Hölder exponents  $\alpha(x)$  then all collapse to the uniformly valid exponent  $|\ln(\min(\mu_0, \mu_1))|/\ln 2 - \varepsilon$ . Some fractal structure can still persist, however. An example of this is given in § 5.3.

(2) The “tricks” in Proposition 3.7 also apply here: in order to verify bounds of the type (4.2) it suffices, e.g., to check a similar bound for the matrices  $\bar{\mathbf{T}}_d = \mathbf{B}\mathbf{T}_d\mathbf{B}^{-1}$ , restricted to  $\mathbf{B}E_{L+1}$ , where  $\mathbf{B}$  is any invertible matrix. A convenient choice of  $\mathbf{B}$  may greatly simplify computations.

(3) In establishing local Hölder exponents and local differentiability we have restricted ourselves to  $x$  such that  $r(x) = \lim_{n \rightarrow \infty} r_n(x)$  is well defined and  $0 < r(x) < 1$ . In fact, it is possible to handle more general  $x$  as well. In fact, Lemma 4.2 tells us that for those  $x$  for which  $r(x) = \lim_{m \rightarrow \infty} r_m(x)$  exists and  $0 < r(x) < 1$ , there can be no “abnormally long” stretches of  $0-s$  or  $1-s$ . If  $x$  is not of this type (it is easy to construct such  $x$ ; they constitute a set of zero Lebesgue measure, however), then we need to control these stretches in some other way. We show here how this can be done when  $1 > \mu_0 > \mu_1 > \frac{1}{2}$ . For all  $x \in [0, 1]$ ,  $n \in \mathbb{N}$ , we define

$$\omega_n^1(x) = \min \{k \in \mathbb{N}; d_{n-k}(x) = 0\} / n.$$

Then  $n\omega_n^1(x)$  is exactly the length of the stretch of digits 1 ending at  $n$ ; in particular, if  $d_n(x) = 0$ , then  $\omega_n^1(x) = 0$ . A detailed analysis shows that, for  $0 \leq t < 2^{-n}$ ,

$$\|\mathbf{v}^{(l)}(x) - \mathbf{v}^{(l)}(x+t)\| \leq C[\mu_1^{nr_n(x)} \mu_0^{n(1-r_n(x))} + \mu_1^{n(n_x(x) - \omega_n^1(x)) + 1} \mu_0^{n(1-r_n(x) + \omega_n^1(x))}].$$

For  $-2^{-n} < t \leq 0$ , we find

$$\|\mathbf{v}^{(l)}(x) - \mathbf{v}^{(l)}(x+t)\| \leq C\mu_1^{nr_n(x)} \mu_0^{n(1-r_n(x))},$$

where  $\omega_n^1$  doesn't enter because  $\mu_0 > \mu_1$ . Now define

$$\bar{r}(x) = \limsup_{n \rightarrow \infty} r_n(x),$$

$$\underline{r}(x) = \liminf_{n \rightarrow \infty} r_n(x),$$

$$\bar{r}^1(x) = \limsup_{n \rightarrow \infty} [r_n(x) - \omega_n^1(x)],$$

$$\underline{r}_1(x) = \liminf_{n \rightarrow \infty} [r_n(x) - \omega_n^1(x)].$$

Then it follows that, for small enough  $t$

$$\begin{aligned} \|\mathbf{v}^{(l)}(x) - \mathbf{v}^{(l)}(x+t)\| &\leq C|t|^{\{r_1(x)|\ln \mu_1| + (1-\bar{r}^1(x))|\ln \mu_0|\}/\ln 2} && \text{if } t \geq 0, \\ &\leq C|t|^{\{r(x)|\ln \mu_1| + \bar{r}(x)|\ln \mu_0|\}/\ln 2} && \text{if } t \leq 0. \end{aligned}$$

This effectively defines two local Hölder exponents  $\alpha_{\pm}(x)$ , one from above and one from below. Similar bounds can be derived if  $\mu_0 < \mu_1$ ; in this case we need to introduce  $\omega_n^0(x)$ , measuring the stretch of zeros preceding the digit  $d_n(x)$ .

The situation  $\mu_0 \geq \frac{1}{2} > \mu_1$  or  $\mu_1 \geq \frac{1}{2} > \mu_0$  is a bit more tricky, but can be dealt with in the same way. Note that if  $r(x)$  is well defined, then  $\bar{r}(x) = \underline{r}(x) = \bar{r}^1(x) = \underline{r}^1(x) = r(x)$ .

(4) The two Hölder exponents  $\alpha_{\pm}(x)$  introduced in the preceding remark, one from above and one from below, are particularly interesting in dyadic rational points  $x$ . For such points, we have two possible binary expansions, one ending in all zeros, the expansion “from above,” which we denote by  $d^+(x)$ , and one ending in all ones, the expansion “from below,” denoted by  $d^-(x)$ . For the Hölder exponent from above, we have to start from  $d^+(x)$ . We then have

$$\bar{r}^1(x) = \underline{r}_1(x) = 0$$

leading to  $\alpha_+(x) = |\ln \mu_0| / \ln 2$ . Similarly, for the Hölder exponent from below, starting from  $d^-(x)$ , we find

$$\bar{r}(x) = \underline{r}(x) = 1;$$

hence,  $\alpha_-(x) = |\ln \mu_1| / \ln 2$ .

If  $\mu_0 > \frac{1}{2} > \mu_1$ , we similarly find that  $f$  is left but not right differentiable in every dyadic rational point.

**5. Examples.** We treat three classes of examples, namely orthonormal wavelets, the de Rham function and generalizations, and the interpolation functions arising in Deslauriers and Dubuc’s Lagrangian interpolation scheme or studied in detail by Dyn, Gregory, and Levin.

In many of these examples we shall use the tricks proposed in Proposition 3.7. In particular, we shall try to determine matrices  $\mathbf{B}$  so that the  $\mathbf{B}\mathbf{T}_d\mathbf{B}^{-1}$  are easier to study than the  $\mathbf{T}_d$  themselves. If the  $c_n$  satisfy  $L + 1$  sum rules, then a very particular choice for  $\mathbf{B}$  is the following:

$$\mathbf{B}_{ij} = \begin{cases} (i-1)! \binom{j-1}{i-1} & \text{for } i \leq L+1, \\ L! \binom{j-i+L}{L} & \text{for } i > L+1, \end{cases}$$

where we use the standard convention that the binomial coefficient  $\binom{n_1}{n_2}$  vanishes if  $n_2 > n_1$ . This is a triangular matrix; for  $i \leq L + 1$ , the  $i$ th row can be written as  $\mathbf{u}_i$  – linear combination of  $\mathbf{u}_k, k < i$ . It is easy to check that the inverse matrix  $\mathbf{B}^{-1}$  is again a triangular matrix, given by

$$(\mathbf{B}^{-1})_{ij} = \begin{cases} (-1)^{i+j} \binom{j-1}{i-1} [(j-1)!]^{-1} & \text{for } j \leq L+1, \\ (-1)^{i+j} \binom{L+1}{i-j+L+1} (L!)^{-1} & \text{for } j > L+1. \end{cases}$$

Note that for  $i \leq L + 1$ , the  $i$ th row of  $\mathbf{B}$  is a linear combination of  $\mathbf{u}_1, \dots, \mathbf{u}_i$ . Because of the special choice of the first  $L + 1$  rows of  $\mathbf{B}$ , and because of the spectral properties of the  $\mathbf{T}_d$ , with their nested left invariant subspaces spanned by the  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}, k \leq L + 1$ , we find, therefore, that

$$\begin{aligned} (\mathbf{B}\mathbf{T}_d\mathbf{B}^{-1})_{ij} &= 0 && \text{if } i \leq L+1, j > i, \\ &= 2^{-i+1} && \text{if } i = j \leq L+1. \end{aligned}$$

This means we are in the situation described in the proofs of Lemmas 3.3, 3.5; in order to prove bounds on products of the restrictions  $\mathbf{T}_d|_{E_{L+1}}$ , it suffices to consider the smaller matrices obtained by deleting the first  $L + 1$  rows and columns of the  $\mathbf{B}\mathbf{T}_d\mathbf{B}^{-1}$ . These submatrices are completely determined by  $\mathbf{T}_d$  and by the last  $(N - L - 1)$  rows of  $\mathbf{B}$  and the last  $(N - L - 1)$  columns of  $\mathbf{B}^{-1}$ . The matrix elements in these rows and columns depend only on the difference  $j - i$  between column and row index. Since the  $(\mathbf{T}_d)_{i,j} = c_{2i-j-1+d}$  depend only on  $2i - j$ , this property will, therefore, be shared by the submatrices representing  $\mathbf{B}|_{E_{L+1}}\mathbf{T}_d|_{E_{L+1}}\mathbf{B}^{-1}|_{E_{L+1}}$ . The entries of these submatrices will, in fact, be given exactly by the coefficients of the quotient of  $\sum c_n e^{in\xi}$  by  $(1 + e^{i\xi})^{L+1}$ . That is to say, if

$$\frac{1}{2} \sum_n c_n e^{in\xi} = \left(\frac{1 + e^{i\xi}}{2}\right)^{L+1} \frac{1}{2} \sum_n q_n e^{in\xi},$$

then the last  $(N - L - 1)$  rows and columns of the  $\mathbf{B}\mathbf{T}_d\mathbf{B}^{-1}$  will be given by

$$(\mathbf{B}\mathbf{T}_d\mathbf{B}^{-1})_{ij} = 2^l q_{2i-j-(N-L)+d}.$$

This observation reduces the study to much smaller matrices, obtained by ‘‘peeling off’’ the sum rules; this is the analog, in our matrix notation, of the reduction from an interpolation subdivision scheme  $S$  to the ‘‘smaller’’ subdivision schemes  $S^{(n)}$  in Dyn and Levin (1989). A similar observation can also be found in Deslauriers and Dubuc (1989).

**5.A. Orthonormal wavelets with compact support.** An orthonormal basis of wavelets is a family of functions generated from one single function by dilations and translations,

$$(5.1) \quad \psi_{jk}(x) = 2^{-j/2} \psi(2^{-j}x - k), \quad j, k \in \mathbb{Z},$$

such that the resulting  $\psi_{jk}$  constitute an orthonormal basis of  $L^2(\mathbb{R})$ . A typical construction of such a basis involves the construction of an auxiliary function  $\phi$  such that

$$(5.2) \quad \phi(x) = \sum c_n \phi(2x - n)$$

for some family of coefficients  $c_n$ . In order to lead to an orthonormal wavelet basis, the  $c_n$  should satisfy the condition

$$(5.3) \quad |p(\xi)|^2 + |p(\xi + \pi)|^2 = 1,$$

where  $p(\xi) = \frac{1}{2} \sum_n c_n e^{in\xi}$ . If this condition is satisfied, then there exists an  $L^2$ -solution to (5.2), and the associated  $\psi$  is given by

$$(5.4) \quad \psi(x) = \sum_n (-1)^n c_{n+1} \phi(2x + n).$$

For a thorough discussion of this construction, with proofs, see Mallat (1989) and Meyer (1990). Equation (5.2) is a two-scale difference equation. If there is only a finite number of nonvanishing  $c_n$ , then (modulo some convergence conditions)  $\phi$  has finite support, and so has  $\psi$ , by (5.4). In Daubechies (1988) the structure of finite families  $\{c_n; n = 0, \dots, N\}$  satisfying (5.3) was analyzed. It turns out that these  $c_n$  are *always* nonsymmetric if we also require  $\phi$  to be continuous. If this continuity requirement is dropped, then there exists one and only one symmetric solution:  $c_0 = c_1 = 1$ , all other  $c_n = 0$ , which correspond to  $\phi(x) = 1$  if  $0 \leq x < 1$ , 0 otherwise, or  $\psi(x) = 1$  if  $0 \leq x < \frac{1}{2}$ ,  $-1$  if  $\frac{1}{2} \leq x < 1$ , zero otherwise; the  $\psi_{jk}$  then constitute the Haar basis. Apart from this example, the interesting finite families of  $c_n$  are asymmetric, and the corresponding  $\phi$ ,

therefore, have local regularity properties associated with fractal sets, as described in § 4. The simplest examples are

$$(5.5) \quad N = 3, \quad c_0 = \frac{1 + \sqrt{3}}{4}, \quad c_2 = \frac{3 - \sqrt{3}}{4},$$

$$c_1 = \frac{3 + \sqrt{3}}{4}, \quad c_3 = \frac{1 - \sqrt{3}}{4},$$

$$(5.6) \quad N = 5, \quad c_0 = [1 + \sqrt{10} + \sqrt{5 + 2\sqrt{10}}]/16,$$

$$c_1 = [5 + \sqrt{10} + 3\sqrt{5 + 2\sqrt{10}}]/16,$$

$$c_2 = [10 - 2\sqrt{10} + 2\sqrt{5 + 2\sqrt{10}}]/16,$$

$$c_3 = [10 - 2\sqrt{10} - 2\sqrt{5 + 2\sqrt{10}}]/16,$$

$$c_4 = [5 + \sqrt{10} - 3\sqrt{5 + 2\sqrt{10}}]/16,$$

$$c_5 = [1 + \sqrt{10} - \sqrt{5 + 2\sqrt{10}}]/16.$$

Higher-order examples (always corresponding to odd  $N$ ) cannot be written in closed form; a table with numerical values of the  $c_n$ , for  $N$  up to 19, is given in Daubechies (1988), as well as the recipe for their computation. For any  $N$ , we denote the associated solution of (1.2) by  $\phi_N$ .

For  $N = 7$ , we have for instance

$$(5.7) \quad \begin{aligned} c_0 &= .325803428051, & c_4 &= -.264507167369, \\ c_1 &= 1.01094571509, & c_5 &= .043616300474, \\ c_2 &= .892200138247, & c_6 &= .046503601071, \\ c_3 &= -.039575026235, & c_7 &= -.014986989330. \end{aligned}$$

(Note that these  $c_n$  are larger, by a factor  $\sqrt{2}$ , than the coefficients  $h(n)$  in Daubechies ((1988), Table 1); the  $h(n)$  are normalized by  $\sum_n h(n) = \sqrt{2}$ ).

Let us discuss in detail what the analysis of §§ 2-4 leads to when applied to the concrete examples (5.5)-(5.7).

**A.  $N = 3$ .** The  $c_n$  defined in (5.5) satisfy two “sum rules” of type (3.1), corresponding to  $l = 0, 1$  or  $L = 1$ . It follows that  $N - (L + 1) = 3 - 2 = 1$ , so that  $E_{L+1} = E_2$  is one-dimensional. The matrices  $\mathbf{T}_0, \mathbf{T}_1$  are given by

$$\mathbf{T}_0 = \begin{pmatrix} \frac{1 + \sqrt{3}}{4} & 0 & 0 \\ \frac{3 - \sqrt{3}}{4} & \frac{3 + \sqrt{3}}{4} & \frac{1 + \sqrt{3}}{4} \\ 0 & \frac{1 - \sqrt{3}}{4} & \frac{3 - \sqrt{3}}{4} \end{pmatrix},$$

$$\mathbf{T}_1 = \begin{pmatrix} \frac{3 + \sqrt{3}}{4} & \frac{1 + \sqrt{3}}{4} & 0 \\ \frac{1 - \sqrt{3}}{4} & \frac{3 - \sqrt{3}}{4} & \frac{3 + \sqrt{3}}{4} \\ 0 & 0 & \frac{1 - \sqrt{3}}{4} \end{pmatrix}.$$

Their eigenvalues are  $1, \frac{1}{2}, (1 + \sqrt{3})/4$  for  $\mathbf{T}_0$ , and  $1, \frac{1}{2}, (1 - \sqrt{3})/4$  for  $\mathbf{T}_1$ . Since  $1 > (1 + \sqrt{3})/4 > \frac{1}{2} > |(1 - \sqrt{3})/4|$  we are in the situation described in Theorem 4.1, with

$\mu_0 > \frac{1}{2} > \mu_1$ . Because  $E_{L+1}$  is one-dimensional,  $\mathbf{T}_0|_{E_{L+1}}$  and  $\mathbf{T}_1|_{E_{L+1}}$  commute and

$$\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{L+1}} = \text{multiplication by } 2^{-m} \left(\frac{1+\sqrt{3}}{2}\right)^{m-s_m} \left(\frac{1-\sqrt{3}}{2}\right)^{s_m} \text{ on } E_2,$$

for all  $m \in \mathbb{N}$ . Because we have sharp bounds for the  $\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{L+1}}$ , the lower bounds for the Hölder exponents given in Theorem 4.5 are also sharp, except for the extra  $\epsilon$  which we always have to introduce in local Hölder exponents.

Since (3.3) is satisfied with  $l = 0$ ,  $\lambda = (1 + \sqrt{3})/4$ , it follows from Theorem 3.1 that  $\phi_3$  is continuous and that  $\phi_3$  is Hölder continuous with exponent  $1 - \ln[(1 + \sqrt{3})/2] \ln 2 \cong .5500 \cdots$ . The following example shows that this estimate is sharp. Applying (1.2) to  $x = 2^{-m}$  leads to  $\phi_3(2^{-m}) = ((1 + \sqrt{3})/4)^m \phi_3(1)$ . The values of  $\phi_3(1)$ ,  $\phi_3(2)$  are determined by the right eigenvector of  $\mathbf{T}_1$  for the eigenvalue 1, which leads to  $\phi_3(1) = (1 + \sqrt{3})/2$ . Hence, for  $x_m = 2^{-m} \rightarrow 0$ ,

$$|\phi_3(0) - \phi_3(x_m)| = \left(\frac{1+\sqrt{3}}{2}\right)^m |x_m|^{-\ln[(1+\sqrt{3})/4]/\ln 2} = \left(\frac{1+\sqrt{3}}{2}\right)^m |x_m|^{.5500 \cdots}.$$

By Theorem 4.1,  $\phi_3$  is differentiable in all  $x$  such that  $r(x)$  is well defined and

$$1 > r(x) > \frac{\ln\left(\frac{1+2\sqrt{3}}{2}\right)}{\ln\left(\frac{1+\sqrt{3}}{2}\right) + \left|\ln\left(\frac{\sqrt{3}}{2}\right)\right|} \cong .2368 \cdots.$$

In particular,  $\phi_3$  is differentiable in all normal numbers, since  $r(x) = \frac{1}{2}$  for  $x$  normal. Consequently,  $\phi_3$  is differentiable almost everywhere, as announced earlier. For values of  $r$  between .2368 and zero, there exists a corresponding fractal set (see Theorem 4.3) on which  $\phi_3$  is Hölder continuous with a Hölder exponent, determined by  $r$ , between .5500  $\cdots$  and 1.

In dyadic rationals  $x$ , we find that  $\phi_3$  is differentiable from below, but is only Hölder continuous from above, with exponent .5500  $\cdots$  (see Remark 4 at the end of § 4). This is already illustrated by the behavior of  $\phi_3$  near  $x = 0$  (see above); it also accounts for the “jaggedness” of the graph of  $\phi_3$  at dyadic points (see Fig. 1).

**B.  $N = 5$ .** We now turn to the properties of  $\phi_5$ . The corresponding  $c_n$ , given in (5.6), satisfy three sum rules of type (3.1), corresponding to  $L = 2$ . In this case  $N - (L + 1) = 5 - 3 = 2$ , i.e.,  $E_{L+1} = E_3$  is two-dimensional, and obtaining good estimates for  $\|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_{L+1}}\|$  is not as straightforward as in the previous case. Explicit computation shows that

$$\begin{aligned} \text{spectrum } \mathbf{T}_0 &= \{1, \frac{1}{2}, \frac{1}{4}, (\sqrt{5+2\sqrt{10}} + \sqrt{10} + 1)/16, (1 - \sqrt{10})/8\} \\ &= \{1, \frac{1}{2}, \frac{1}{4}, .470467 \cdots, -.270284 \cdots\}, \\ \text{spectrum } \mathbf{T}_1 &= \{1, \frac{1}{2}, \frac{1}{4}, (1 - \sqrt{10})/8, (1 + \sqrt{10} - \sqrt{5+2\sqrt{10}})/16\} \\ &= \{1, \frac{1}{2}, \frac{1}{4}, -.270274 \cdots, .049817 \cdots\}. \end{aligned} \tag{5.8}$$

The spectral radii of both  $\mathbf{T}_0|_{V_3}$  and  $\mathbf{T}_1|_{V_3}$  are, therefore, strictly larger than  $\frac{1}{4}$ , which means that we cannot expect better than a  $C^1$ -result of  $\phi_5$ , with some Hölder continuity for  $\phi'_5$ . In order to obtain this much, we need to prove that (3.3) holds for  $l = 1$ , i.e., that  $\|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_3}\| \leq C 2^{-m} \lambda^m$  for some  $\lambda \leq 1$ . Straightforward estimates of  $\|\mathbf{T}_0|_{E_3}\|$  and  $\|\mathbf{T}_1|_{E_3}\|$  are much too large for our goal. To make computations easier, we apply Proposition 3.7 and work with  $\mathbf{B}\mathbf{T}_0\mathbf{B}^{-1}$ ,  $\mathbf{B}\mathbf{T}_1\mathbf{B}^{-1}$  instead, for suitably chosen  $\mathbf{B}$ .

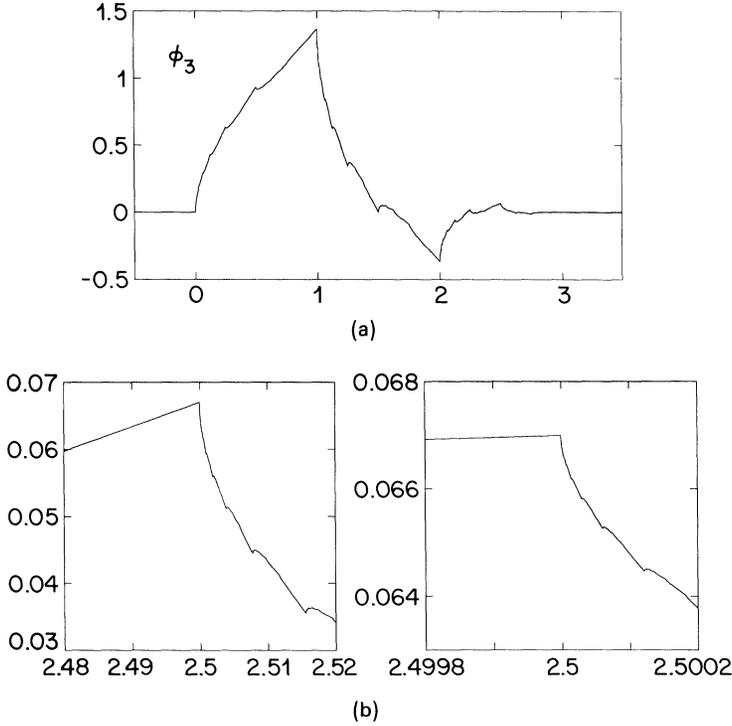


FIG. 1. (a) The  $L^1$ -solution  $\phi_3$  to the two-scale difference equation  $f(x) = [(1 + \sqrt{3})f(2x) + (3 + \sqrt{3})f(2x - 1) + (3 - \sqrt{3})f(2x - 2) + (1 - \sqrt{3})f(2x - 3)]/4$ . The function  $\phi_3$  is Hölder continuous with exponent  $2 - \log_2(1 + \sqrt{3}) = .5500\dots$ . Moreover,  $\phi_3$  is almost everywhere differentiable; in dyadic rational points in  $[0, 3]$   $\phi_3$  is left but not right differentiable (see text). (b) Two successive blow-ups of  $\phi_3$  near  $x = 2.5$ , illustrating the self-similar behavior of  $\phi_3$ .

As a first step, we choose  $\mathbf{B}_1$  as outlined at the start of § 5, which in this case means

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 6 & 12 \\ 0 & 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

As explained above, the matrices  $\mathbf{B}_1 \mathbf{T}_d \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}$  ( $d = 0$  or  $1$ ) are simply obtained by stripping  $\mathbf{B}_1 \mathbf{T}_d \mathbf{B}_1^{-1}$  of their first three columns and rows. This results in

$$(5.9) \quad \begin{aligned} \mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3} &= \begin{pmatrix} .470467 \dots & 0 \\ .049817 \dots & -.270284 \dots \end{pmatrix}, \\ \mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3} &= \begin{pmatrix} -.270284 \dots & .4770467 \dots \\ 0 & .049817 \dots \end{pmatrix}; \end{aligned}$$

as announced earlier, these reduced matrices can also be obtained directly from  $(1 + e^{i\xi})^{-3} \sum_n c_n e^{in\xi}$ . The norms of these matrices are still larger than we would like

them. We apply, therefore, the same trick again and multiply (5.9) on the left by  $\mathbf{B}_2$ , on the right by  $\mathbf{B}_2^{-1}$ , where  $\mathbf{B}_2$  is a conveniently chosen  $2 \times 2$ -matrix. The choice

$$\mathbf{B}_2 = \begin{pmatrix} 1 & 0 \\ .33188 \dots & 4.92450 \dots \end{pmatrix}$$

diagonalizes  $\mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}$  and reduces  $\mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}$  to a symmetric matrix,

$$\mathbf{B}_2 (\mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}) \mathbf{B}_2^{-1} = \begin{pmatrix} .470467 \dots & 0 \\ 0 & -.270284 \dots \end{pmatrix}$$

and

$$\mathbf{B}_2 (\mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}) \mathbf{B}_2^{-1} = \begin{pmatrix} -.238644 \dots & -.095535 \dots \\ -.095535 \dots & .018177 \dots \end{pmatrix}.$$

Consequently, the norms of these matrices are given by their spectral radii, which implies

$$\begin{aligned} \|\mathbf{B}_2 (\mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}) \mathbf{B}_2^{-1}\| &= \rho[\mathbf{B}_2 (\mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}) \mathbf{B}_2^{-1}] \\ &= \rho(\mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}) = \rho(\mathbf{T}_1|_{E_3}), \end{aligned}$$

and similarly

$$\|\mathbf{B}_2 (\mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_3}) \mathbf{B}_2^{-1}\| = \rho(\mathbf{T}_0|_{E_3}).$$

It follows that, for all binary sequences  $d$ , and all  $m$ ,

$$\|\mathbf{T}_{d_1} \dots \mathbf{T}_{d_m}|_{E_3}\| \leq C 2^{-n} \mu_1^{s_n} \mu_0^{n-s_n},$$

where  $s_n = \sum_{j=1}^n d_j$ , and  $\mu_1 = [2\rho(\mathbf{T}_1|_{E_3})] = .540569 \dots$ ,  $\mu_0 = [2\rho(\mathbf{T}_0|_{E_3})] = .940934 \dots$ . On the other hand,  $\mu_d \geq 2\|\mathbf{T}_d|_{E_3}\|^{1/n} \geq 2\rho(\mathbf{T}_d|_{E_3})$  ( $d = 0$  or  $1$ ). Since  $\rho(\mathbf{T}_d|_{E_2}) = \rho(\mathbf{T}_d|_{E_3})$  in this case (see (5.8)), the above estimates are, therefore, the sharpest possible for  $\mu_0, \mu_1$ .

Clearly  $\frac{1}{2} < \mu_1 < \mu_0 < 1$ . It follows that  $\phi_5$  is continuously differentiable, and that  $\phi'_5$  is Hölder continuous with exponent

$$(5.10) \quad |\log \mu_0| / \log 2 - \varepsilon = .087833 \dots$$

Since the sharpest estimate for  $\mu_1$  is strictly larger than  $\frac{1}{2}$ ,  $\phi_5$  is nowhere twice differentiable; because  $\mu_1 < \mu_0$ , there exists a hierarchy of fractal sets on which  $\phi'_5$  has a larger Hölder exponent than (5.10). In particular, the Hölder exponent of  $\phi'_5$  on the full set of normal numbers is

$$[|\log \mu_0| + |\log \mu_1|] / 2 \log 2 - \varepsilon = .487641 \dots - \varepsilon.$$

The function  $\phi_5$  is plotted in Fig. 2. At first sight, we have the impression that  $\phi_5$  has a sharp peak at  $x = 1$ , contradicting its differentiability at this point. Successive blowups of  $\phi_5$  around  $x = 1$  show that this peak is not really “sharp” (see Fig. 2(b)).

C.  $N = 7$ . The  $c_n$  corresponding to  $\phi_7$  are given by (5.7); they satisfy (by construction; see Daubechies (1988)) four sum rules of type (3.1), corresponding to  $L = 3$ . It follows that  $E_{L+1} = E_4$  is three-dimensional. Explicit computation leads to

$$\text{spectrum } \mathbf{T}_0 = \{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, .325803 \dots, -.279620 \dots, .093804 \dots\},$$

$$\text{spectrum } \mathbf{T}_1 = \{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, -.279620 \dots, .093804 \dots, -.014986 \dots\}.$$

Since  $\rho(\mathbf{T}_0|_{V_4}), \rho(\mathbf{T}_1|_{V_4}) > \frac{1}{4}$ ,  $\phi_7$  is at most once continuously differentiable. In order to prove that  $\phi_7$  is indeed  $C^1$  and to estimate the Hölder exponent of  $\phi'_7$ , we need to

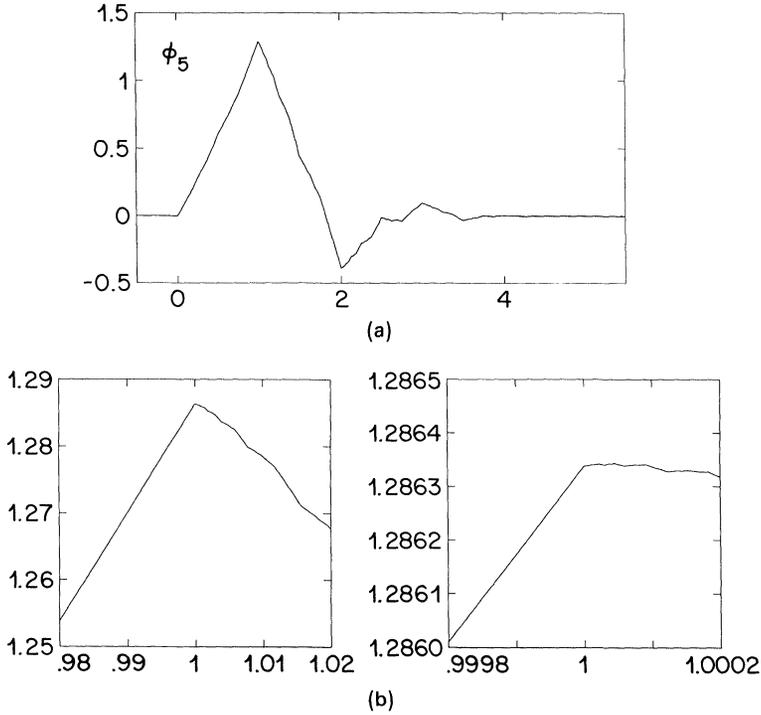


FIG. 2. (a) The  $L^1$ -solution  $\phi_5$  to the two-scale difference equation  $f(x) = \sum_{n=0}^5 c_n f(2x - n)$ , with  $c_n$  as given by (5.6). The function  $\phi_5$  is continuously differentiable, despite appearances. (b) Two blow-ups of  $\phi_5(x)$  around  $x = 1$ , showing that the peak of  $\phi_5$  at  $x = 1$  is not really “sharp.” We find  $\phi_5'(1) = -(\tilde{e}_2^0)_2 = 1.63845 \dots$ .

find bounds of the type (3.3), with  $l = 1$ ,  $\lambda < 1$ . In order to derive such bounds, we shall again study  $\|\mathbf{B}\mathbf{T}_\nu\mathbf{B}^{-1}|_{\mathbf{B}E_4}\|$  for suitably chosen, invertible  $\mathbf{B}$ . We choose  $\mathbf{B}_1$  according to the recipe at the start of this section,

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 2 & 6 & 12 & 20 & 30 \\ 0 & 0 & 0 & 6 & 24 & 60 & 120 \\ 0 & 0 & 0 & 0 & 6 & 24 & 60 \\ 0 & 0 & 0 & 0 & 0 & 6 & 24 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 \end{pmatrix}.$$

The restriction to  $\mathbf{B}_1E_4$  of the matrices  $\mathbf{B}_1\mathbf{T}_d\mathbf{B}_1^{-1}$  then simply consists of omitting the first four rows and columns, leading to

$$\mathbf{B}_1\mathbf{T}_0\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4} = \begin{pmatrix} .325803 \dots & 0 & 0 \\ .106451 \dots & -.292267 \dots & .325803 \\ 0 & -.014986 \dots & .106451 \end{pmatrix},$$

$$\mathbf{B}_1\mathbf{T}_1\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4} = \begin{pmatrix} -.292267 \dots & .325803 \dots & 0 \\ -.014986 \dots & .106451 \dots & -.292267 \\ 0 & 0 & -.014986 \end{pmatrix}.$$

We recognize again the same characteristic structure as in the matrices  $\mathbf{T}_d$  themselves (the  $ij$ -element depends only on  $2i - j$ ); the entries are the coefficients of the

trigonometric polynomial  $2^{-4}(1 + e^{i\xi})^{-4} \sum_n c_n e^{in\xi}$ . The matrix

$$\mathbf{B}_2 = \begin{pmatrix} 1 & 0 & 0 \\ -.144737 \cdots & 0.315439 \cdots & -8.125871 \cdots \\ -.466989 \cdots & 2.655918 \cdots & -2.241310 \cdots \end{pmatrix}$$

diagonalizes  $\mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_4}$ , and reduces  $\mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 V_4}$  to

$$\mathbf{B}_2 \mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1} \mathbf{B}_2^{-1}|_{\mathbf{B}_1 E_4} = \begin{pmatrix} -.238104 \cdots & -.034981 \cdots & .126825 \cdots \\ .034981 \cdots & -.002306 \cdots & -.004837 \cdots \\ .126825 \cdots & .080468 \cdots & .039608 \cdots \end{pmatrix}.$$

The norm of this last matrix (computed by  $\|A\| = [\rho(A^t A)]^{1/2}$ ) is  $.296060 \cdots$ ; since  $\mathbf{B}_2 \mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1} \mathbf{B}_2^{-1}|_{\mathbf{B}_1 E_4}$  is diagonal, its norm is given by the largest eigenvalue of  $\mathbf{T}_0|_{E_4}$ , which is  $.325803 \cdots$ . We have thus proved that

$$\|\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_m}|_{E_4}\| \leq C 2^{-n} (.651606 \cdots)^{n-s_n} (.592120 \cdots)^{s_n},$$

which implies that  $\phi_7 \in C^1$  and that  $\phi_7'$  is Hölder continuous with exponent

$$\frac{|\log (.651606 \cdots)|}{\log 2} = .617926 \cdots.$$

Since our estimate of  $\mu_0$  is sharp, this Hölder exponent is sharp as well. Since  $\mu_1 < \mu_0$ , we again have a hierarchy of fractal sets corresponding to larger Hölder exponents. Note, however, that our estimate for  $\mu_1$  is very likely not sharp (the norm of  $\mathbf{B}_2 \mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1} \mathbf{B}_2^{-1}|_{\mathbf{B}_1 E_4}$  is larger than its spectral radius, and it might be possible to sharpen this estimate by other choices of  $\mathbf{B}_2$ , without losing the sharp estimate on  $\mathbf{T}_0$ ). The range of Hölder exponents might, therefore, be even larger than suggested by our estimates.

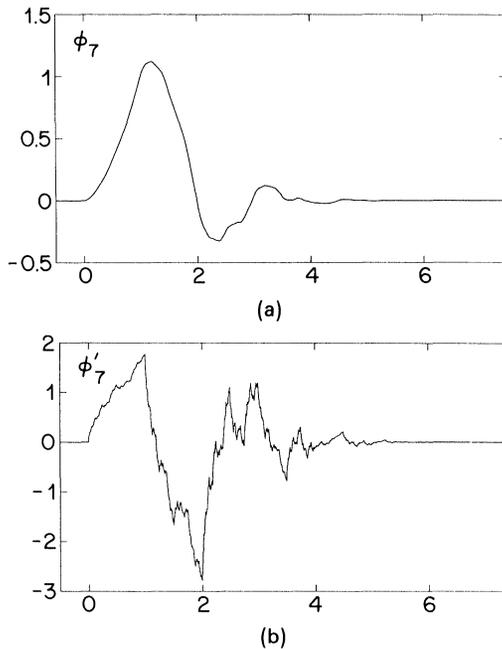


FIG. 3. (a) The  $L^1$ -solution  $\phi_7$  to the two-scale difference equation  $f(x) = \sum_{n=0}^7 c_n f(2x - r)$ , with  $c_n$  as given in (5.7). The function  $\phi_7$  is continuously differentiable. (b) The derivative  $\phi_7'$  of  $\phi_7$ ;  $\phi_7'$  is Hölder continuous with exponent  $.6179 \cdots$ .

Figures 3(a) and 3(b) give plots of  $\phi_7$  and  $\phi'_7$ , respectively. Both plots have been realized via the cascade algorithm (see part I, § 4).

*Remark.* Estimates for the Hölder exponents of  $\phi_3, \phi_5,$  and  $\phi_7$  were already computed in Daubechies (1988) via a different method, involving the Fourier transform. The table below compares the estimates found here with these earlier results. With the notation  $\phi \in C^{n+\alpha}$  for  $\phi \in C^n, \phi^{(n)}$  Hölder continuous with exponent  $\alpha$  ( $\alpha \in [0, 1[$ ), the estimates for  $n + \alpha$  compares as follows:

	best estimate in (Daubechies (1988))	as computed here
$\phi_3$	$.5 - \varepsilon$	$.5500 \dots$
$\phi_5$	$.915 \dots - \varepsilon$	$1.0878 \dots$
$\phi_7$	$1.275 \dots - \varepsilon$	$1.6179 \dots$

Our present results are significantly better; moreover, they are optimal.

**5.B. The de Rham function and generalizations.** The de Rham function, as defined in Part I, § 6, is the normalized  $L^1$ -solution to the two-scale difference equation

$$\phi(x) = \phi(3x) + \frac{1}{3}[\phi(3x - 1) + \phi(3x + 1)] + \frac{2}{3}[\phi(3x - 2) + \phi(3x + 2)].$$

Since the scaling factor is three instead of two, we are in a slightly different case from before. This example will illustrate, however, that our techniques can be used for all integer scaling factors  $k$ , modulo minor adaptations. In general, we have  $(k - 1)$ -matrices  $\mathbf{T}_0, \mathbf{T}_1, \dots, \mathbf{T}_{k-1}$ , with matrix elements given by

$$(\mathbf{T}_l)_{ij} = c_{ki-j-N_1-k+1,l}, \quad 1 \leq i, j \leq \frac{N_2 + N_1}{k - 1},$$

where we assume that the index  $n$  of the  $c_n$  ranges from  $-N_1$  to  $N_2$ . In this case, with  $k = 3, c_0 = 1, c_1 = c_{-1} = \frac{1}{3}, c_2 = c_{-2} = \frac{2}{3}$ , we have, therefore, three  $2 \times 2$  matrices,

$$\mathbf{T}_0 = \begin{pmatrix} \frac{2}{3} & 0 \\ \frac{1}{3} & 1 \end{pmatrix} \quad \mathbf{T}_1 = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} \quad \mathbf{T}_2 = \begin{pmatrix} 1 & \frac{1}{3} \\ 0 & \frac{2}{3} \end{pmatrix}$$

We have  $\sum_{n=-2}^2 c_n = 3$ , and the  $c_n$  satisfy one sum rule:

$$(5.11) \quad \sum c_{3n} = \sum c_{3n+1} = \sum c_{3n+2} = 1.$$

The vector  $(1, 1)$  is a common left eigenvector for  $\mathbf{T}_0, \mathbf{T}_1, \mathbf{T}_2$ , with eigenvalue 1. The remaining eigenvalues of  $\mathbf{T}_0, \mathbf{T}_1, \mathbf{T}_2$  are, respectively,  $\frac{2}{3}, -\frac{1}{3}$ , and  $\frac{2}{3}$ . For  $x \in [0, 1]$  we shall now use the ternary expansion of  $x$  (since  $k = 3$ ); as in Theorem 4.3, we define

$$r_n(x, i) = \frac{1}{n} \# \{d_j(x) = i; j \leq n\} \quad i = 0, 1 \text{ or } 2,$$

where  $d_j(x)$  are the digits of the ternary expansion of  $x$  (each equal to zero, 1 or 2). Then

$$(5.12) \quad \mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_n(x)}|_{E_1} = \text{multiplication by } 3^{-n} 2^{n[r_n(x;0)+r_n(x;2)]} (-1)^{nr_n(x;1)}.$$

By the extension of Theorem 2.2 to scale factors  $k \neq 2$ , it follows that  $\phi$  is continuous. Since

$$\|\mathbf{T}(n; d(x))|_{E_1}\| \leq \left(\frac{2}{3}\right)^n,$$

$\phi$  is Hölder continuous, with exponent  $|\ln(\frac{2}{3})| \ln 3 = .36907 \dots$ . In fact, this example is so easy (partly due to  $c_n \geq 0$  for all  $n$ ) that the Hölder continuity can be established without recourse to the matrix analysis presented in this paper; we then find that the Hölder exponent of  $\phi$  is indeed exactly  $|\ln(2\frac{2}{3})| / \ln 3$ .

From (5.12) it also follows that the local Hölder exponent of  $\phi$  at  $y \in [-1, 1]$  is often larger than the uniformly valid exponent  $\alpha_{\text{uniform}} = .36907 \dots$ . For  $y = n + x$ ,  $n = -1$  or  $0$ ,  $x \in [0, 1]$ , we find that

$$|\phi(y + t) - \phi(y)| \leq C|t|^{\alpha(y) - \epsilon},$$

with

$$\alpha(y) = \frac{\ln 3 - (r_0 + r_2) \ln 2}{\ln 3} = \alpha_{\text{uniform}} + r_1(x) \frac{\ln 2}{\ln 3},$$

where we have assumed that  $r_1(x) = \lim_{n \rightarrow \infty} [r_n(x; 1)]$  exists. Only where  $r_1(x) = 0$  is  $\alpha(y) = \alpha_{\text{uniform}}$ . On the other hand, there exists  $y$  such that  $r_1(x) = 1$ , namely  $y \in [0, 2]$  of the type  $3^{-j}(l + \frac{1}{2})$ ; in these points  $\phi$  is Lipschitz. As usual, there exists a hierarchy of fractal sets corresponding to the Hölder exponents between and  $\alpha_{\text{uniform}}$ ; their Hausdorff dimension is given by Theorem 4.3. In particular, for the full set of normal numbers, we find

$$\alpha_{\text{normal}} = \alpha_{\text{uniform}} + \frac{1}{3} \frac{\ln 2}{\ln 3} = 1 - \frac{2}{3} \frac{\ln 2}{\ln 3} = .57938 \dots$$

A straightforward generalization of the de Rham function is obtained by choosing

$$c_0 = 1, c_1 = c_{-1} = \frac{1}{2} - \gamma, c_2 = c_{-2} = \frac{1}{2} + \gamma,$$

corresponding to the two-scale difference equation

$$\begin{aligned} \phi^\gamma(x) &= \phi^\gamma(3x) + (\frac{1}{2} - \gamma)[\phi^\gamma(3x - 1) + \phi^\gamma(3x + 1)] \\ &\quad + (\frac{1}{2} + \gamma)[\phi^\gamma(3x - 2) + \phi^\gamma(3x + 2)]. \end{aligned}$$

In this case the matrices  $\mathbf{T}_0, \mathbf{T}_1, \mathbf{T}_2$  are

$$\mathbf{T}_0 = \begin{pmatrix} \frac{1}{2} + \gamma & 0 \\ \frac{1}{2} - \gamma & 1 \end{pmatrix}, \mathbf{T}_1 = \begin{pmatrix} \frac{1}{2} - \gamma & \frac{1}{2} + \gamma \\ \frac{1}{2} + \gamma & \frac{1}{2} - \gamma \end{pmatrix}, \mathbf{T}_2 = \begin{pmatrix} 1 & \frac{1}{2} - \gamma \\ 0 & \frac{1}{2} + \gamma \end{pmatrix}.$$

We have again  $\sum c_{3n} = \sum c_{3n+1} = \sum c_{3n+2} = 1$ , causing  $\mathbf{T}_0, \mathbf{T}_1, \mathbf{T}_2$  to have a common left eigenvector for the eigenvalue 1. The remaining eigenvalues are  $\frac{1}{2} + \gamma, -2\gamma, \frac{1}{2} - \gamma$ , respectively. It follows that  $\phi^\gamma$  is continuous only if  $|\frac{1}{2} + \gamma| < 1$  and  $|2\gamma| < 1$ , i.e., if  $|\gamma| < \frac{1}{2}$ . For  $\gamma = \frac{1}{6}$ , we obviously revert to the de Rham case; for  $\gamma = -\frac{1}{6}$  the function  $\phi^\gamma$  is piecewise linear,  $\phi^{-1/6}(x) = 1 + x$  for  $-1 \leq x \leq 0, 1 - x$ , for  $0 \leq x \leq 1$ , and zero otherwise; for  $\gamma = 0$  the resulting  $\phi^0$  consists of a copy and its mirror image of the Cantor-Lebesgue function on  $[0, 1]$ .

The same analysis as before results in Hölder continuity for  $\phi^\gamma$ , with exponent  $\alpha_\gamma = (\min [|\ln(\frac{1}{2} + \gamma)|, |\ln(2\gamma)|]) / \ln 3$ . For  $y = n + x, n = -1$  or  $0, x \in [0, 1]$ , such that  $r_1(x) = \lim_{n \rightarrow \infty} [r_n(x; 1)]$  exists, we find a larger local Hölder exponent

$$\begin{aligned} (5.13) \quad \alpha_\gamma(y) &= \frac{|\ln(\frac{1}{2} + \gamma)|}{\ln 3} + r_1(x) \frac{\ln \frac{\frac{1}{2} + \gamma}{|2\gamma|}}{\ln 3} \quad \text{if } \frac{1}{2} > \gamma \geq -\frac{1}{6}, \\ &\frac{|\ln |2\gamma||}{\ln 3} + [1 - r_1(x)] \frac{\ln \frac{|2\gamma|}{\frac{1}{2} + \gamma}}{\ln 3} \quad \text{if } -\frac{1}{6} \geq \gamma > -\frac{1}{2}. \end{aligned}$$

For  $\gamma = -\frac{1}{6}$ , all these Hölder exponents (local and uniform) collapse to 1, which was to be expected since  $\phi^{-1/6}$  is piecewise linear. If, for some  $\gamma$  and some  $y, \alpha_\gamma(y)$  as

given by (5.13) is larger than 1, then this means that  $\phi^\gamma$  is differentiable in  $y$ . This happens, for instance, if  $(\frac{1}{2} + \gamma)^2 |2\gamma| < \frac{1}{27}$ , and if  $r_1(x) = \frac{1}{3}$ . In particular,  $\phi^\gamma$  is differentiable almost everywhere if  $-\frac{1}{2} < \gamma < .05921689 \dots$ .

Fig. 4 gives the graphs of  $\phi^\gamma$  for the values  $\gamma = \frac{1}{6}$  (de Rham case),  $\gamma = \frac{1}{12}$ ,  $\gamma = .05$ ,  $\gamma = 0$  and  $\gamma = -\frac{1}{12}$ .

**5.C. Lagrangian interpolation functions and generalizations.** In Deslauriers and Dubuc (1989), general symmetric Lagrangian interpolation schemes are defined, for arbitrary integer scaling factor  $k > 1$ , and an arbitrary number of nodes. Deslauriers and Dubuc characterize the regularity of the associated “fundamental functions,” which are solutions to two-scale difference equations with scaling factor  $k$ ,

$$f(x) = f(kx) + \sum_{n=1}^{kN-1} c_n [f(kx - n) + f(kx + n)],$$

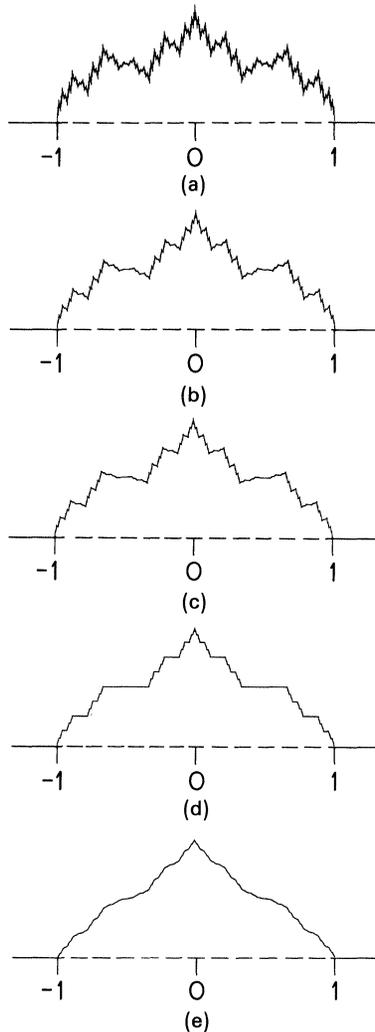


FIG. 4. The generalized de Rham function  $\phi^\gamma$  for different values of  $\gamma$ ; (a)  $\gamma = \frac{1}{6}$  (the de Rham case), (b)  $\gamma = \frac{1}{12}$ , (c)  $\gamma = .05$ , (d)  $\gamma = 0$  (characteristic function of Cantor set), (e)  $\gamma = -\frac{1}{12}$ . In the cases (c), (d), and (e),  $\phi^\gamma$  is almost everywhere differentiable.

where the  $c_n$  have the peculiarity that  $c_{km} = \delta_{m0}$  (see Part I). The  $c_n$  in a Lagrangian symmetric interpolation scheme are completely determined by the requirement that  $p(\xi) = k^{-1}[1 + 2 \sum_{n=1}^{kN-1} c_n \cos n\xi]$  be divisible by  $(1 + e^{i\xi} + \dots + e^{i(k-1)\xi})^L$ , with  $L$  as large as possible for the given  $N$  (this leads to  $L = 2N$ ). The resulting  $p(\xi)$  is positive for real  $\xi$ , which is of great help in the analysis of the regularity of  $f$ . We have indeed (see part I or Deslauriers and Dubuc (1989))  $\hat{f}(\xi) = \prod_{j=1}^{\infty} p(k^{-j}\xi)$ ; since  $p(\xi) = ((1 - e^{ik\xi})/k(1 - e^{i\xi}))^L q(\xi)$ , this leads to  $\hat{f}(\xi) = ((1 - e^{i\xi})/\xi)^L \prod_{j=1}^{\infty} q(k^{-j}\xi)$ . The quotient  $q(\xi)$  is a trigonometric polynomial,  $q(\xi) = \sum_n q_n e^{in\xi}$ . The regularity of  $f$  is given by the largest  $\lambda$  such that  $|\xi|^\lambda \hat{f}(\xi)$  is in  $L^1(\mathbb{R})$ ; it turns out that these  $L^1$ -norms can be estimated in terms of  $N$  and the spectral radius of a finite matrix, constructed from the  $q_n$  in the same way as  $\mathbf{T}_0, \mathbf{T}_1$  are constructed from the  $c_n$  (see § 7, 8 in Deslauriers and Dubuc (1989)). A similar technique was used by one of us, independently, to obtain estimates in the  $L^1$ -norm of  $|\xi|^\lambda |\hat{\phi}_N(\xi)|$ , where  $\phi_N$  are the functions associated to orthonormal wavelet bases (see § 5.A). In that case, however, the function  $p$  was not positive, and it was necessary to use the Cauchy-Schwarz inequality to reduce everything to more tractable  $L^2$ -estimates, which, however, led to less sharp estimates (see the table at the end of § 5.A).

Let us see now how the present methods, which avoid Fourier transforms, perform on these interpolation functions. We shall restrict ourselves to one of the simpler examples, with  $k = 2$  and  $N = 2$ . In this case the  $c_n$  are

$$(5.14) \quad c_0 = 1, \quad c_1 = c_{-1} = \frac{9}{16}, \quad c_3 = c_{-3} = -\frac{1}{16}.$$

We have two matrices:  $\mathbf{T}_0, \mathbf{T}_1$ ,

$$\mathbf{T}_0 = \begin{pmatrix} -\frac{1}{16} & 0 & 0 & 0 & 0 & 0 \\ \frac{9}{16} & 0 & -\frac{1}{16} & 0 & 0 & 0 \\ \frac{9}{16} & 1 & \frac{9}{16} & 0 & -\frac{1}{16} & 0 \\ -\frac{1}{16} & 0 & \frac{9}{16} & 1 & \frac{9}{16} & 0 \\ 0 & 0 & -\frac{1}{16} & 0 & \frac{9}{16} & 1 \\ 0 & 0 & 0 & 0 & -\frac{1}{16} & 0 \end{pmatrix},$$

$$\mathbf{T}_1 = \begin{pmatrix} 0 & -\frac{1}{16} & 0 & 0 & 0 & 0 \\ 1 & \frac{9}{16} & 0 & -\frac{1}{16} & 0 & 0 \\ 0 & \frac{9}{16} & 1 & \frac{9}{16} & 0 & -\frac{1}{16} \\ 0 & -\frac{1}{16} & 0 & \frac{9}{16} & 1 & \frac{9}{16} \\ 0 & 0 & 0 & -\frac{1}{16} & 0 & \frac{9}{16} \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{16} \end{pmatrix}.$$

The  $c_n$  in (5.14) satisfy four sum rules. Because of the symmetry  $c_n = c_{-n}$ , we have  $\mathbf{T}_1 = \mathbf{O}\mathbf{T}_0\mathbf{O}$ , where  $\mathbf{O}$  is the involution  $\mathbf{O}_{ij} = \delta_{i+j-7}$ . Consequently,  $\mathbf{T}_1, \mathbf{T}_0$  have the same eigenvalues; we find

$$\text{spectrum}(\mathbf{T}_0) = \text{spectrum}(\mathbf{T}_1) = \{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, -\frac{1}{16}\},$$

where the root  $\frac{1}{4}$  has multiplicity two in the characteristic polynomial of  $\mathbf{T}_0, \mathbf{T}_1$ . It follows that  $\|\mathbf{T}_0|_{E_3}\| = \|\mathbf{T}_1|_{E_3}\| \cong \frac{1}{4}$ , so that the best we can hope for is a bound of type (3.3) with  $l = 1$  and  $\lambda = \frac{1}{2}$ , corresponding to  $|f'(x+t) - f'(x)| \leq C|t| |\log |t||$ . It turns out that the restriction of (say)  $\mathbf{T}_0$  (the same will be true for  $\mathbf{T}_1$ ) to the two-dimensional space associated with the eigenvalue  $\frac{1}{4}$  can be brought to the normal Jordan form, but cannot be diagonalized. There are, therefore, only one left and one right eigenvector of  $\mathbf{T}_0$  with eigenvalue  $\frac{1}{4}$ , and they are orthogonal. This is what causes the extra factor  $|\log |t||$  in the ‘‘almost Lipschitz’’ bound on  $f'$ . The degeneracy of  $\frac{1}{4}$  is lifted, however,

as soon as  $\mathbf{T}_0, \mathbf{T}_1$  are mixed: if  $(d_1, \dots, d_n)$  does not consist of only  $0-s$  or only  $1-s$ , then the eigenvalue  $4^{-n}$  of  $\mathbf{T}_{d_1} \cdots \mathbf{T}_{d_n}$  is nondegenerate. If, however, the sequence  $d$  has a tail consisting of only  $0-s$  or only  $1-s$  (corresponding to a dyadic rational  $x$ ), then the gap between  $4^{-n}$  and the closest eigenvalue tends to zero as  $n$  tends to infinity, and an attempt at the construction of  $\tilde{\mathbf{e}}_3(n; d(x))$  (see the proof of Theorem 4.1) diverges for  $n \rightarrow \infty$ . If  $f$  were twice continuously differentiable, then its second derivative  $f''$  would be given by these limits  $\tilde{\mathbf{e}}_3(\infty; d(x))$  (see § 4); the divergence of  $\tilde{\mathbf{e}}_3(n; d(x))$  at dyadic rational  $x$  shows, therefore, that  $f$  cannot be twice differentiable at dyadic rationals.

To check that (3.3) holds for  $l=1, \lambda=\frac{1}{2}$ , we shall again compute estimates for  $\mathbf{B}\mathbf{T}_0\mathbf{B}^{-1}, \mathbf{B}\mathbf{T}_1\mathbf{B}^{-1}$ , with a conveniently chosen  $\mathbf{B}$ , rather than work with  $\mathbf{T}_0|_{E_4}, \mathbf{T}_1|_{E_4}$  themselves. The existence of four sum rules suggests that we use for  $\mathbf{B}_1$

$$(5.15) \quad \mathbf{B}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 2 & 6 & 12 & 20 \\ 0 & 0 & 0 & 6 & 24 & 60 \\ 0 & 0 & 0 & 0 & 6 & 24 \\ 0 & 0 & 0 & 0 & 0 & 6 \end{pmatrix}.$$

The matrices  $\mathbf{B}_1\mathbf{T}_0\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4}, \mathbf{B}_1\mathbf{T}_1\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4}$  then reduce to  $2 \times 2$ -matrices,

$$\begin{aligned} \mathbf{B}_1\mathbf{T}_0\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4} &= \begin{pmatrix} -\frac{1}{16} & 0 \\ -\frac{1}{16} & \frac{1}{4} \end{pmatrix}, \\ \mathbf{B}_1\mathbf{T}_1\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4} &= \begin{pmatrix} \frac{1}{4} & -\frac{1}{16} \\ 0 & -\frac{1}{16} \end{pmatrix}. \end{aligned}$$

The matrix

$$(5.16) \quad \mathbf{B}_2 = \begin{pmatrix} 1 & 0 \\ \frac{\sqrt{6}}{2} & \frac{-5\sqrt{6}}{2} \end{pmatrix}$$

then diagonalizes  $\mathbf{B}_1\mathbf{T}_0\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4}$  and reduces  $\mathbf{B}_1\mathbf{T}_1\mathbf{B}_1^{-1}|_{\mathbf{B}_1E_4}$  to a symmetric matrix. It follows that, for  $d=0$  or  $1$ ,

$$\begin{aligned} \|\mathbf{B}_2\mathbf{B}_1\mathbf{T}_d\mathbf{B}_1^{-1}\mathbf{B}_2^{-1}|_{\mathbf{B}_1E_4}\| &= \rho(\mathbf{B}_2\mathbf{B}_1\mathbf{T}_d\mathbf{B}_1^{-1}\mathbf{B}_2^{-1}|_{\mathbf{B}_1E_4}) \\ &= \rho(\mathbf{T}_d|_{E_4}) = \frac{1}{4}. \end{aligned}$$

This implies (3.3) with  $l=1, \lambda=\frac{1}{2}$ . Theorem 3.1 then implies that  $f$  is ‘‘almost’’  $C^2$ , in the sense that  $f \in C^1$  and that  $f'$  is Hölder continuous with exponent  $1 - \varepsilon$ , with  $\varepsilon > 0$  arbitrarily small; in fact  $|f'(x+t) - f'(x)| \leq Ct|\ln |t||$  for sufficiently small  $t$ , a result which was first proved in Dubuc (1986).

Since  $\mathbf{T}_1 = \mathbf{O}\mathbf{T}_0\mathbf{O}^{-1}$ , where  $\mathbf{O}$  is an orthogonal matrix (see above), we necessarily have  $\mu_0 = \mu_1$ , where  $\mu_0, \mu_1$  are as defined in § 4. Nevertheless, we can still improve on the estimate  $\|\mathbf{T}_{d_1(x)} \cdots \mathbf{T}_{d_n(x)}|_{E_2}\| < Cn2^{-2n}$  for special, nondyadic  $x$ . Group the digits in the binary expansion of  $x$  together in pairs,  $p_n(x) = d_{2n-1}(x)d_{2n}(x)$ ,  $n = 1, 2, \dots$ . Assume that the pairs  $01, 10$  occur with an asymptotic frequency larger than zero,

$$p(x) = \liminf_{n \rightarrow \infty} \frac{1}{n} \#\{j < n; p_j(x) = 10 \text{ or } p_j(x) = 01\} > 0.$$

Then it can be shown that, for large enough  $n$ ,

$$(5.17) \quad \|\mathbf{T}(n; d(x))|_{E_3}\| \leq C2^{-2n}\lambda^n,$$

where  $\lambda < 1$ . Moreover,  $\tilde{e}_3(x) = \lim_{n \rightarrow \infty} \tilde{e}_3(n; d(x))$  is well defined in these points  $x$ ; repeating some of the arguments in § 4 shows, therefore, that  $f$  is twice differentiable in  $x$ . The estimate (5.17) is a consequence of the fact that  $\mathbf{B}_2\mathbf{B}_1\mathbf{T}_0\mathbf{B}_1^{-1}\mathbf{B}_2^{-1}|_{\mathbf{B}_1E_3}$  and  $\mathbf{B}_2\mathbf{B}_1\mathbf{T}_1\mathbf{B}_1^{-1}\mathbf{B}_2^{-1}|_{\mathbf{B}_1E_3}$  have a common left eigenvector with eigenvalue  $\frac{1}{8}$ , while also

$$\|\mathbf{B}_2\mathbf{B}_1\mathbf{T}_0\mathbf{T}_1\mathbf{B}_1^{-1}\mathbf{B}_2^{-1}\|_{\mathbf{B}_1E_4}^{1/2} = \|\mathbf{B}_2\mathbf{B}_1\mathbf{T}_1\mathbf{T}_0\mathbf{B}_1^{-1}\mathbf{B}_2^{-1}\|_{\mathbf{B}_1E_4}^{1/2} = .150156 \dots$$

Here  $\mathbf{B}_1, \mathbf{B}_2$  are the matrices defined by (5.15), (5.16), respectively. This implies

$$(5.18) \quad \|\mathbf{B}_2\mathbf{B}_1\mathbf{T}_{d_1} \dots \mathbf{T}_{d_n}\mathbf{B}_2^{-1}\|_{\mathbf{B}_1E_4} \leq C \max(8^{-n}, \|\mathbf{B}_2\mathbf{B}_1\mathbf{T}_{d_1} \dots \mathbf{T}_{d_n}\mathbf{B}_1^{-1}\mathbf{B}_2^{-1}\|_{\mathbf{B}_1E_4}),$$

where  $C$  depends on  $\mathbf{B}_1, \mathbf{B}_2$ , but not on  $n$ . For large  $n$ , the right-hand side of (5.18) is bounded by

$$C[(.150156 \dots)^{p(x)}(\frac{1}{4})^{1-p(x)}]^n \leq C4^{-n}(.600624 \dots)^{np(x)}.$$

Together with (5.18), this implies (5.17), with  $\lambda = (.600624 \dots)^{p(x)}$ .

Note that  $p(x) = \frac{1}{2}$  in every normal point  $x$ . The above argument proves therefore that  $f$  is almost everywhere twice differentiable. Figure 5 gives graphs of both  $f$  and  $f'$  (both were also plotted in Dubuc (1986)). Note that  $f$  looks “smooth,” but that  $f'$  again exhibits “bumps” which repeat themselves at different scales.

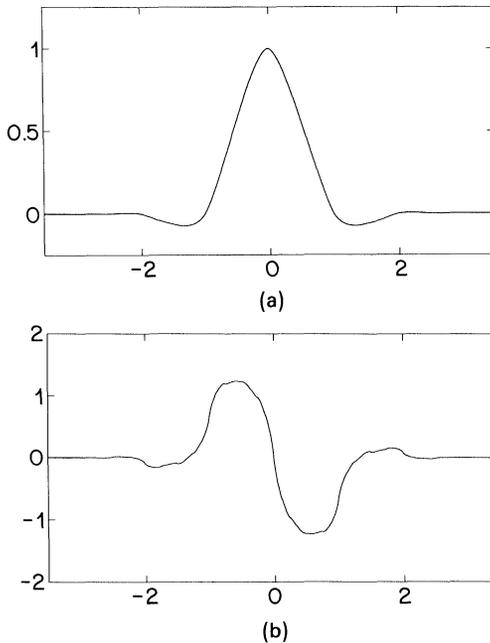


FIG. 5. (a) The  $L^1$ -solution to the two-scale difference equation  $f(x) = f(2x) + \frac{9}{16}[f(2x-1) + f(2x+1)] - \frac{1}{16}[f(2x-3) + f(2x+3)]$ . Thus function is “almost”  $C^2$  (see text). (b) The derivative  $f'$  of the function plotted in (a).

The  $c_n$  in (5.14) were chosen so that, with the restrictions  $c_n = c_{-n}, c_2 = c_{-2} = 0$ , the polynomial  $p(\xi) = \sum_{n=-3}^3 c_n e^{in\xi}$  was divisible by the maximum possible number of factors  $(1 + e^{i\xi})$ . Other examples, still satisfying  $c_n = c_{-n}, c_2 = c_{-2}, c_n = 0$  for  $|n| > 3$ , were mentioned in Deslauriers and Dubuc (1987). These examples satisfy fewer sum rules, ( $p(\xi)$  is divisible by fewer factors  $(1 + e^{i\xi})$ ), and the corresponding  $p(\xi)$  is not positive any more. In order to evaluate the regularity of the function  $f$ , Deslauriers and Dubuc had to resort to other, less optimal techniques. They prove, e.g., that for

$c_1 = c_{-1} = \frac{1}{2} - a$ ,  $c_3 = c_{-3} = a$ , the function  $f$  is continuous if  $a \in ]-\frac{3}{16}, \frac{1}{16}[$ . Our present technique leads to better results. For the case  $a = -\frac{3}{16}$ , e.g., we show that  $f$  is continuously differentiable. To do this, we shall use Proposition 3.7 again. Since the  $c_n$  satisfy only two sum rules in this case, we choose  $\mathbf{B}_1$

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

$\mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_2}$ ,  $\mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_2}$  reduce to  $4 \times 4$ -matrices, with eigenvalues  $\frac{3}{8}$ ,  $-\frac{3}{16}$ , and  $(1 \pm i\sqrt{2})/4$ . The matrices

$$\mathbf{B}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2y/9 & y & 0 & -y \\ \frac{5-26\sqrt{2}i}{81}z & z & \frac{-1+2\sqrt{2}i}{3}z & z \\ \frac{5+26\sqrt{2}i}{81}z' & z' & -\frac{1+2\sqrt{2}i}{3}z' & z' \end{pmatrix},$$

where  $y, z, z'$  are arbitrary complex parameters, diagonalize  $\mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1}|_{\mathbf{B}_1 E_2}$ , so that

$$\|\mathbf{B}_2 \mathbf{B}_1 \mathbf{T}_0 \mathbf{B}_1^{-1} \mathbf{B}_2^{-1}|_{\mathbf{B}_1 E_2}\| = \rho(\mathbf{T}_0|_{E_2}) = |(1 \pm i\sqrt{2})/4| = .433012 \dots$$

For the choice  $y = .3$ ,  $z = .27 + .08i$ ,  $z' = .27 - .08i$ , the norm corresponding to  $\mathbf{T}_1$  is

$$\|\mathbf{B}_2 \mathbf{B}_1 \mathbf{T}_1 \mathbf{B}_1^{-1} \mathbf{B}_2^{-1}|_{\mathbf{B}_2 E_2}\| = .604342 \dots$$

This is still larger than  $\frac{1}{2}$ , and, therefore, not sufficient to show that  $f \in C^1$ . However, if we define

$$t_{d_1 \dots d_m} = \|\mathbf{T}_{d_1, \text{red}} \dots \mathbf{T}_{d_m, \text{red}}\|^{1/m},$$

where  $\mathbf{T}_{d_m, \text{red}} = \mathbf{B}_2 \mathbf{B}_1 \mathbf{T}_d \mathbf{B}_1^{-1} \mathbf{B}_2^{-2}|_{\mathbf{B}_1 E_2}$ , with the parameters in  $\mathbf{B}_2$  fixed as above, then we check that

$$\begin{aligned} & t_0, t_{10}, t_{1100}, t_{110100}, t_{1101010}, t_{11010110}, t_{11010111}, \\ & t_{1101100}, t_{1101101}, t_{110111}, t_{111000}, t_{111001}, \\ (5.19) \quad & t_{1110100}, t_{111010100}, t_{1110101010}, t_{11101010110}, \\ & t_{11101010111}, t_{1110101100}, t_{1110101101}, t_{111010111}, \\ & t_{11101100}, t_{11101101}, t_{1110111}, t_{111100}, \\ & t_{1111010}, t_{11110110}, t_{11110111}, t_{111110} \quad \text{and} \quad t_{111111}. \end{aligned}$$

are all  $< \frac{1}{2}$ .

Since these groups of indices constitute a complete set of building blocks (any binary sequence can be decomposed into them), it follows that

$$\|\mathbf{T}_{d_1} \dots \mathbf{T}_{d_m}|_{E_2}\| \leq 2^{-m} \lambda^m$$

for some  $\lambda < 1$ , if  $m$  is large enough (see Proposition 3.7). Note that in order to derive (5.19) only 56 matrix norms were calculated (this includes candidates that failed, such

as  $t_{1101, \dots}$ ), even though the longest sequences of digits has 11 elements. Checking all  $2^{11}$  sequences of this length would have been much more cumbersome, and would, in fact, not have been sufficient in order to conclude that (3.3) holds, since some of them still lead to  $t_D > \frac{1}{2}$ .

From all this it follows that the function  $f$  corresponding to the two-scale difference equation with  $k = 2, c_0 = 1, c_1 = c_2 = \frac{11}{16}, c_3 = c_{-3} = -\frac{3}{16}$ , all other  $c_n = 0$ , is continuously differentiable. It is plotted in Fig. 6, together with its derivative.

For general  $a, c_0 = 1, c_1 = c_{-1} = \frac{1}{2} - a, c_3 = c_{-3} = a$ , we find that

$$\text{spectrum } (T_0) = \text{spectrum } (T_1) = \left\{ 1, \frac{1}{2}, a, -2a, \frac{1 \pm \sqrt{1+16a}}{4} \right\}.$$

We conjecture that

- the associated function  $f$  is in  $C^1$  if  $-\frac{1}{4} < a < 0$ ,
- the associated function  $f$  is continuous if  $-\frac{1}{2} < a < \frac{1}{2}$ .

**Appendix.** In our analysis of lattice two-scale equations of the type

$$(A.1) \quad f(x) = \sum_{n=0}^N c_n f(2x - n),$$

the interval  $[0, 1]$  and binary expansions play a special role. This is because  $[0, 1]$  has the following two properties:  $[0, 1]$  and its integer translates tile the real line, and  $[0, 1] = \Lambda^{-1}[0, 1] \cup \Lambda^{-1}([0, 1] + 1)$ , where  $\Lambda$  is multiplication by 2. The interval  $[0, 1]$  is, moreover, the unique bounded subset of  $\mathbb{R}$  satisfying these two conditions. If the scale factor in (A.1) were  $k$  instead of 2, then we would define  $\Lambda$  to be multiplication

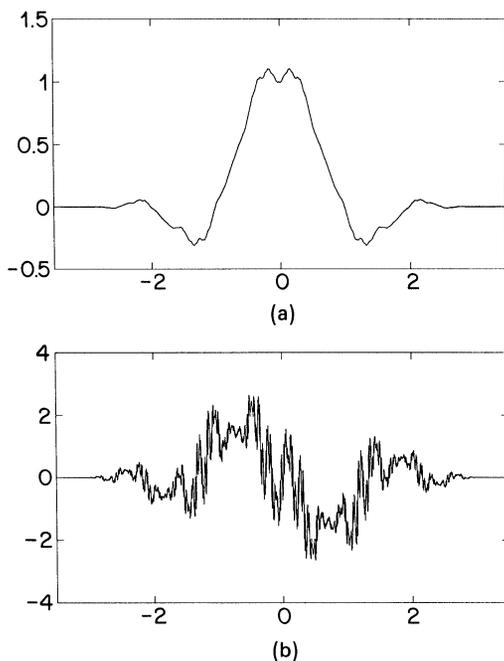


FIG. 6. (a) The  $L^1$ -solution to the two-scale difference equation  $f(x) = f(2x) + \frac{11}{16}[f(2x-1) + f(2x+1)] - \frac{3}{16}[f(2x-3) + f(2x+3)]$ . This function is continuously differentiable (see text). (b) The derivative of the function plotted in (a).

by  $k$ , and  $[0, 1]$  would be tiled by the  $\Lambda^{-1}([0, 1] + m)$ ,  $0 \leq m \leq k - 1$ , leading to  $k$ -adic expansions. All this can be generalized to higher dimensions. We can then write

$$(A.2) \quad f(x) = \sum_{n \in I} c_n f(\Lambda x - n),$$

where  $x \in \mathbb{R}^n$ ,  $I$  is a finite subset of  $\mathbb{Z}^d$ , and  $\Lambda$  is a linear operator on  $\mathbb{R}^d$  that preserves  $\mathbb{Z}^d$  and has all its singular values strictly larger than 1. For  $d = 2$ , examples of such  $\Lambda$  which have been proposed for applications are

$$\Lambda_0 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \Lambda_1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \text{and} \quad \Lambda_2 = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}.$$

$\Lambda_0$  is used in many two-dimensional subdivision schemes; since it consists of a simple uniform dilation, the one-dimensional approach used earlier can be transposed without any problems. The role played before by  $[0, 1]$  will now be played by  $[0, 1]^2$ , and binary expansions still do the trick. For the matrices  $\Lambda_1$  and  $\Lambda_2$  the situation is more complicated.  $\Lambda_1$  was first proposed by M. Vetterli (1984) for a subband coding scheme with exact reconstruction for two-dimensional images; presently, several groups are working on the corresponding orthonormal wavelet bases (Gröchenig and Madych (1992), Lawton and Resnikoff (1991)). An interpolation subdivision scheme using  $\Lambda_2$  is studied in Mongeau (1990); strictly speaking, Mongeau uses

$$\Lambda'_2 = \begin{pmatrix} \frac{3}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{3}{2} \end{pmatrix}$$

for which the triangular lattice  $\{(m + n/2, n\sqrt{3}/2); m, n \in \mathbb{Z}\}$ , not  $\mathbb{Z}^2$ , is invariant. The matrix

$$\mathbf{B} = \begin{pmatrix} 1 & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{3}} \end{pmatrix}$$

maps the triangular lattice to  $\mathbb{Z}^2$  and can be used to translate Mongeau's results to results for lattice two-scale equations involving  $\Lambda_2$ ; we have  $\Lambda_2 = \mathbf{B}\Lambda'_2\mathbf{B}^{-1}$ .

A large part of our analysis can still be carried out, even for nonuniform dilation matrices such as  $\Lambda_1$  and  $\Lambda_2$ . Since  $\Lambda$  has only integer entries, and its singular values are all larger than 1,  $\det \Lambda = k \in \mathbb{Z}$ , with  $k > 1$ . The role played by  $[0, 1]$  for (A.1) will now be played by the unique set  $\Gamma$  defined by

- the collection  $\{\Gamma + n; n \in \mathbb{Z}^d\}$  tiles  $\mathbb{R}^d$ ,
- $\Gamma$  itself is tiled by the  $k$  elements of  $\{\Lambda^{-1}(\Gamma + m); m \in \mathbb{Z}^d \cap \Lambda([0, 1]^d)\}$ .

For both  $\Lambda_1$  and  $\Lambda_2$ , the corresponding set  $\Gamma$  is a set with fractal boundary;  $\Gamma_1$  is the so-called *twin dragon set*. The set  $\Gamma$  will be the “elementary building block” for the support of compactly supported solutions to (A.2), just like  $[0, 1]$  was for (A.1). The number  $N$  of such building blocks constituting support ( $f$ ) will be determined by the nonvanishing coefficients  $c_n$  in (A.2). We can then “fold”  $f$  as follows:

$$[\mathbf{v}(x)]_n = f(x + z_n), \quad x \in \Gamma, z_n \in \mathbb{Z}^d, \quad n = 1, \dots, N,$$

where  $\cup_{n=1}^N (\Gamma + z_n) \supset \text{support}(f)$ . Points in  $\mathbb{R}^d$  can then be represented by  $\Gamma$ -adic expansions; in particular, for  $x \in \Gamma$ , we define

$$d_1(x) = \{m \in \mathbb{Z}^d \cap \Lambda([0, 1[{}^d); x \in \Lambda^{-1}(\Gamma + m)\},$$

$$d_{j+1}(x) = d_j(\Lambda x - d_1(x)).$$

These expansions are unique, except for  $x$  in a set of measure zero (corresponding to the dyadic rationals in the one-dimensional case with  $\Lambda = 2$ ). We can again define a shift operator  $\tau$  on  $\Gamma$  by

$$d_j(\tau x) = d_{j+1}(x) \quad j = 1, 2, \dots,$$

and the equation (A.2) can be rewritten in vector notation as

$$v(x) = T_{d_1(x)}v(\tau x),$$

where the  $k$  matrices  $T_l$  are all determined by the coefficients  $c_n$ . We can then derive existence, continuity, smoothness, etc., of  $v$  (hence of  $f$ ) from the spectral properties of the  $T_l$ . A first extension to higher dimensions can be found in Mongeau (1990), extending our original approach, which did not use spline functions. It seems hard to use spline functions, which are piecewise polynomial and have to be fitted together correctly on the boundaries of the building blocks, when the building blocks have fractal boundaries as in these examples. Mongeau’s work, by using appropriate powers of  $\Lambda$ , reduces everything to pure dilations, and thereby avoids the fractal boundaries, at the price of having to deal with much larger matrices  $T_d$ . It is also possible to deal directly with the fractal sets by extending our original, longer, and more complicated proof without spline functions (as shown in Cohen and Daubechies (1991)). Therefore, we would like to outline here the major steps of the proofs of Theorems 2.2 and 3.1 without spline functions. For simplicity, we return for this outline to the one-dimensional case, with  $\Lambda = \text{multiplication by } 2$ .

What follows is an outline of the proof of Theorem 2.1. The starting point is the equation

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

$$= T(m; d(x))v(\tau^m x).$$

From  $v(0), v(1)$  (which can be determined from the right eigenvector of  $M$  for the eigenvalue 1—see § I.5), we can thus derive  $v(x)$  for all dyadic rationals  $x$ . Existence and continuity of  $v$  will be proved if this definition of  $v$  on a dense set can be continuously extended. This is done in the following steps:

- $e_1$ , the common left eigenvector of  $T_0, T_1$  for the eigenvalue 1, is a left eigenvector of every  $T(m; x)$ , with eigenvalue 1;
- consequently,  $e_1 \cdot v(x) = 1$  for every dyadic rational  $x$ ;
- $\|v(x)\|$  is uniformly bounded, for all dyadic rationals (this uses the condition (2.11), together with  $v(x) - v(x') \in E_1$ );
- we then use that, for small enough  $t$ ,

$$v(x + t) - v(x) = T(m; d(x))[v(\tau^m x + 2^m t) - v(\tau^m x)],$$

together with (2.16), to show that the restriction of  $v$  to the dyadic rationals is continuous and satisfies

$$\|v(x) - v(y)\| \leq C\lambda^m$$

if  $|x - y| \leq 2^{-m}$ . This suffices to show that  $v$  has a continuous extension to all of  $[0, 1]$ , and to prove the Hölder continuity of this extension.

In the details of the proof, we have to be careful, occasionally, because dyadic rationals have two binary extensions. This does not cause any real problem.

Next we sketch the proof of Theorem 3.1.

- Every  $\mathbf{T}(m; d)$  has a left eigenvector  $\mathbf{e}_k(m; d)$  for the eigenvalue  $2^{-(k-1)m}$ ,  $1 \leq k \leq L+1$ , which can be written as  $\mathbf{e}_k(m; d) = \mathbf{e}_k^0 +$  combination of  $\mathbf{e}_{k'}^0$ ,  $k' < k$ . These  $\mathbf{e}_k(m; d)$  are bounded uniformly in  $m \in \mathbb{N}$  and  $d \in \{0, 1\}^{\mathbb{N}}$ . (This is a purely combinatorial fact, and we do not need (3.3) to prove it.)

- For  $k \leq l+1$ , the eigenvalue  $2^{-(k-1)m}$  of  $\mathbf{T}(m; d)$  is simple; there exists a corresponding right eigenvector  $\tilde{\mathbf{e}}_k(m; d)$ , uniquely normalized by  $\mathbf{e}_k(m; d) \cdot \tilde{\mathbf{e}}_k(m; d) = \mathbf{e}_k^0 \cdot \tilde{\mathbf{e}}_k(m; d) = 1$ . These  $\tilde{\mathbf{e}}_k(m; d)$  are bounded uniformly in  $m$  and  $d$ . (To prove this we need (3.3); the proof is analogous to (7) in the proof of Theorem 4.1.)

$$\begin{aligned} \|\mathbf{T}(m; d)|_{E_{k+1}}\| &\leq C\lambda^m 2^{-ml}, & l+1 \leq k \leq L, \\ \|\mathbf{T}(m; d)|_{E_{l+1}}\| &\leq C\lambda^m 2^{-ml} \phi_\lambda(m), \end{aligned}$$

where

$$\phi_\lambda(m) = 1 \quad \text{if } \lambda > \frac{1}{2}, \quad \phi_{1/2}(m) = m,$$

and

$$\|\mathbf{T}(m; d)|_{E_{k+1}}\| \leq C2^{-mk}, \quad 0 \leq k \leq l-1.$$

(Part of this is Lemma 3.5; the third bound is proved analogously.) By taking  $k=0$ , we see that (3.3) implies (2.15), so that  $\mathbf{v}$  can be constructed and is continuous by Theorem 2.3.

- The  $\tilde{\mathbf{e}}_k(n; d)$  converge to a limit for  $n \rightarrow \infty$ . Moreover, if  $x$  is a dyadic rational, then  $\lim_{n \rightarrow \infty} \tilde{\mathbf{e}}_k(n; d^+(x)) = \lim_{n \rightarrow \infty} \tilde{\mathbf{e}}_k(n; d^-(x))$ . We can, therefore, define  $\tilde{\mathbf{e}}_k(x) = \lim_{n \rightarrow \infty} \tilde{\mathbf{e}}_k(n; d(x))$  for all  $x \in [0, 1]$ ,  $l \leq k \leq l+1$ . Finally  $\tilde{\mathbf{e}}_k(x)$  is continuous in  $x$ .

- For all  $x \in [0, 1]$ ,  $k = 0, \dots, L$ ,

$$\mathbf{e}_{k+1}^0 \cdot \mathbf{v}(x) = (-1)^k x^k$$

(see the proof of Lemma 3.6).

- With all these ingredients, we prove, e.g., differentiability of  $\mathbf{v}$  as follows: for  $t$  small enough,

$$\begin{aligned} \mathbf{v}(x+t) - \mathbf{v}(x) &= \mathbf{T}(m; d(x))[\mathbf{v}(\tau^m x + 2^m t) - \mathbf{v}(\tau^m x)] \\ &= 2^{-m} \tilde{\mathbf{e}}_2(m; d(x)) \{ \mathbf{e}_2(m; d(x)) \cdot [\mathbf{v}(\tau^m x + 2^m t) - \mathbf{v}(\tau^m x)] \} \\ &\quad + \mathbf{T}(m; d(x)) \mathbf{r}, \end{aligned}$$

where we have used  $\mathbf{e}_1(m; d(x)) \cdot [\mathbf{v}(y) - \mathbf{v}(z)] = 0$  and where  $\mathbf{r} \in E_2$ ,  $\|\mathbf{r}\|$  bounded independently of  $x, t$ , and  $m$ . Consequently,

$$\frac{\mathbf{v}(x+t) - \mathbf{v}(x)}{t} = \tilde{\mathbf{e}}_2(m; d(x)) + o(1) \rightarrow \tilde{\mathbf{e}}_2(x),$$

and  $\mathbf{v}$  is differentiable in  $x$  if  $l \geq 1$ .

- Higher-order differentiability is established analogously. We find

$$\frac{d^k}{dx^k} \mathbf{v}(x) = (-1)^k k! \tilde{\mathbf{e}}_k(x), \quad j = 0, \dots, l.$$

(There is no need to “guess” the ansatz (3.10) with this approach.)

• All these results for  $\mathbf{v}$  must then be “unfolded” to state results for  $f$ ; in principle, problems could occur at the integers. Since  $[\tilde{\mathbf{e}}_k(0)]_1 = 0 = [\tilde{\mathbf{e}}_k(1)]_N$  and  $[\tilde{\mathbf{e}}_k(0)]_{m+1} = [\tilde{\mathbf{e}}_k(1)]_m$ ,  $m = 1, \dots, N-1$ , unfolding works without problems.

Here ends our outline of the proof without spline functions.

## REFERENCES

- M. BERGER AND Y. WANG (1991), *Bounded semi-groups of matrices*, Linear Algebra Appl., to appear.
- A. CAVARETTA, W. DAHMEN, AND C. MICCHELLI (1989), *Stationary subdivision*, Mem. Amer. Math. Soc., 453, pp. 1–186.
- A. COHEN AND I. DAUBECHIES (1991), *Non-separable two dimensional wavelet bases*, Rev. Mat. Iberoamericana, submitted.
- I. DAUBECHIES (1988), *Orthonormal bases of wavelets with compact support*, Comm. Pure Appl. Math., 41, pp. 909–996.
- I. DAUBECHIES AND J. C. LAGARIAS (1991), *Sets of matrices all infinite products of which converge*, Linear Algebra Appl., 161, pp. 227–263.
- (1991), *Two-scale difference equations. I. Global regularity of solutions*, SIAM J. Math. Anal., 22, pp. 1388–1410.
- G. DE RHAM (1956), *Sur une courbe plane*, J. Math. Pures Appl., 35, pp. 25–42.
- (1957), *Sur un exemple de fonction continue sans dérivée*, Enseign. Math., 3, pp. 71–72.
- (1959), *Sur les courbes limites de polygones obtenue par trisection*, Enseign. Math., 5, pp. 29–43.
- G. DESLAURIERS AND S. DUBUC (1987), *Interpolation dyadique*, in Fractals, dimensions non entières et applications, G. Cherbit, ed., Masson, Paris.
- (1989), *Symmetric iterative interpolation processes*, Constr. Approx., 5, pp. 49–68.
- S. DUBUC (1986), *Interpolation through an iterative scheme*, J. Math. Anal. Appl., 114, pp. 185–204.
- N. DYN, J. A. GREGORY, AND D. LEVIN (1987), *A 4-point interpolatory subdivision scheme for curve design*, Comput. Aided Geom. Design, 4, pp. 257–268.
- (1991), *Analysis of uniform binary subdivision schemes for curve design*, Constr. Approx. 7, pp. 127–147.
- N. DYN AND D. LEVIN (1989), *Interpolating subdivision schemes for the generation of curves and surfaces*, to appear.
- (1990), *Uniform subdivision algorithms for curves and surfaces*, Constr. Approx., to appear.
- H. G. EGGLESTON (1949), *The fractional dimension of a set defined by decimal properties*, Quart. J. Math. Oxford Ser., 20, pp. 31–36.
- I. J. GOOD (1941), Proc. Cambridge Phil. Soc., 37, p. 200.
- K. GRÖCHENIG AND W. R. MADYCH (1992), *Multiscale analysis, Haar bases and self-similar tilings of  $\mathbb{R}^n$* , IEEE Trans. Inform. Theory, 38, pp. 556–568.
- W. LAWTON AND H. RESNIKOFF (1991), *Multidimensional wavelet bases*, SIAM J. Math. Anal., submitted.
- S. MALLAT (1989), *Multiresolution approximation and wavelet orthonormal bases of  $L^2$* , Trans. Amer. Math. Soc., 315, pp. 69–88.
- Y. MEYER (1990), *Ondelettes*, Vol. 1, in Ondelettes et Opérateurs, Hermann, Paris.
- C. A. MICCHELLI (1986), *Subdivision algorithms for curves and surfaces*, Proc. SIGGRAPH, 1986, Dallas, TX.
- C. A. MICCHELLI AND H. PRAUTZSCH (1987), *Refinement and subdivision for spaces of integer translates of compactly supported functions* in Numerical Analysis, D. F. Griffith, G. A. Watson, eds., Academic, New York, pp. 192–222.
- (1987), *Computing curves invariant under halving*, Comput. Aided Geom. Design, 4, pp. 133–140.
- (1989), *Uniform refinement of curves*, Linear Algebra Appl., 114/115, pp. 841–870.
- J. P. MONGEAU (1990), *Propriétés de l’interpolation itérative*, Ph.D. thesis, Ecole Polytechnique, Montréal.
- G.-C. ROTA AND W. G. STRANG (1960), *A note on the joint spectral radius*, Indag. Math., 22, pp. 379–381.
- L. L. SCHUMAKER (1981), *Spline Functions: Basic Theory*, John Wiley, New York.
- M. VETTERLI (1984), *Multi-dimensional subband coding: some theory and algorithms*, Signal Process., 6, pp. 97–112.