WAVELETS

RONALD A. DeVORE and BRADLEY J. LUCIER

1. Introduction

The subject of "wavelets" is expanding at such a tremendous rate that it is impossible to give, within these few pages, a complete introduction to all aspects of its theory. We hope, however, to allow the reader to become sufficiently acquainted with the subject to understand, in part, the enthusiasm of its proponents toward its potential application to various numerical problems. Furthermore, we hope that our exposition can guide the reader who wishes to make more serious excursions into the subject. Our viewpoint is biased by our experience in approximation theory and data compression; we warn the reader that there are other viewpoints that are either not represented here or discussed only briefly. For example, orthogonal wavelets were developed primarily in the context of signal processing, an application which we touch on only indirectly. However, there are several good expositions (e.g., [Da1] and [RV]) of this application. A discussion of wavelet decompositions in the context of Littlewood-Paley theory can be found in the monograph of Frazier, Jawerth, and Weiss [FJW]. We shall also not attempt to give a complete discussion of the history of wavelets. Historical accounts can be found in the book of Meyer [Me] and the introduction of the article of Daubechies [Da1]. We shall try to give enough historical commentary in the course of our presentation to provide some feeling for the subject's development.

The term "wavelet" (originally called wavelet of constant shape) was introduced by J. Morlet. It denotes a univariate function ψ (multivariate wavelets exist as well and will be discussed subsequently), defined on $\mathbb{R}$, which, when subjected to the fundamental operations of shifts (i.e., translation by integers) and dyadic dilation, yields an orthogonal basis of $L_2(\mathbb{R})$. That is, the functions $\psi_{j,k} := 2^{k/2}\psi(2^k\cdot - j)$, $j,k \in \mathbb{Z}$, form a complete orthonormal system for $L_2(\mathbb{R})$. In this work, we shall call such a function an orthogonal wavelet, since there are many generalizations of wavelets that drop the requirement of orthogonality. The Haar function $H := \chi_{[0,1/2)} - \chi_{[1/2,1)}$, which will be discussed in more detail in the section that follows, is the simplest example of an orthogonal wavelet. Orthogonal wavelets with higher smoothness (and even compact support) can also be constructed. But before considering that and other questions, we wish first to motivate the desire for such wavelets.

A version of this paper appeared in *Acta Numerica,* A. Iserles, ed., Cambridge University Press, v. 1 (1992), pp. 1–56. This work was supported in part by the National Science Foundation (grants DMS-8922154 and DMS-9006219), the Air Force Office of Scientific Research (contract 89-0455-DEF), the Office of Naval Research (contracts N00014-90-1343, N00014-91-J-1152, and N00014-91-J-1076), the Defense Advanced Research Projects Agency (AFOSR contract 90-0323), and the Army High Performance Computing Research Center at the University of Minnesota.

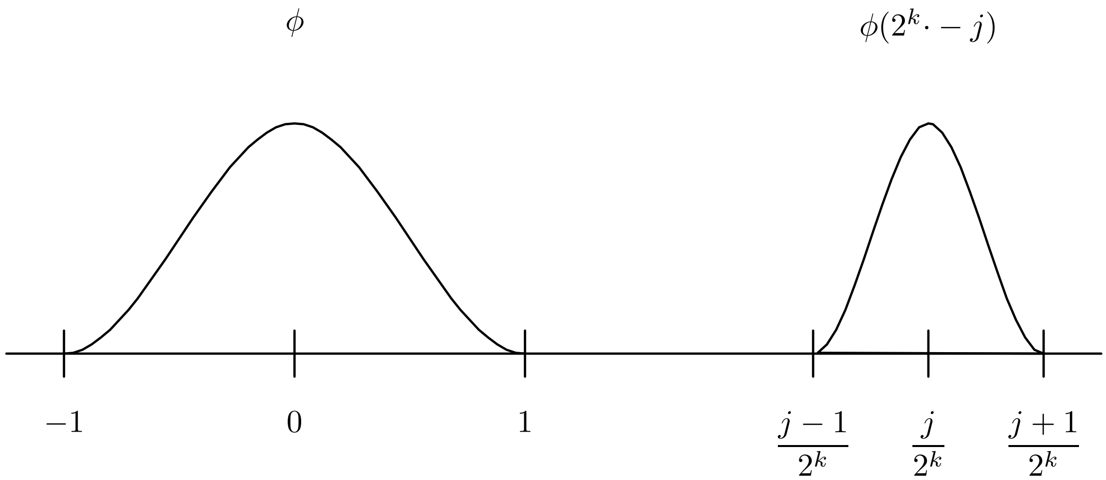

FIGURE 1. An example of functions ϕ and $\phi(2^k\cdot - j)$.

We view a wavelet ψ as a "bump" (and think of it as having compact support, though it need not). Dilation squeezes or expands the bump and translation shifts it (see Figure 1). Thus, $\psi_{j,k}$ is a scaled version of ψ centered at the dyadic integer $j2^{-k}$. If k is large positive, then $\psi_{j,k}$ is a bump with small support; if k is large negative, the support of $\psi_{j,k}$ is large.

The requirement that the set $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ forms an orthonormal system means that any function $f \in L_2(\mathbb{R})$ can be represented as a series

$$f = \sum_{j,k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle\, \psi_{j,k} \tag{1.1}$$

with $\langle f, g\rangle := \int_{\mathbb{R}} f\overline{g}\,dx$ the usual inner product of two $L_2(\mathbb{R})$ functions. We view (1.1) as building up the function f from the bumps $\psi_{j,k}$. Bumps corresponding to small values of k contribute to the broad resolution of f; those corresponding to large values of k give finer detail.

The decomposition (1.1) is analogous to the Fourier decomposition of a function $f \in L_2(\mathbb{T})$ in terms of the exponential functions $e_k := e^{ik\,\cdot}$, but there are important differences. The exponential functions e_k have global support. Thus, all terms in the Fourier decomposition contribute to the value of f at a point x. On the other hand, wavelets are usually either of compact support or fall off exponentially at infinity. Thus, only the terms in (1.1) corresponding to $\psi_{j,k}$ with $j2^{-k}$ near x make a large contribution at x. The representation (1.1) is in this sense local. Of course, exponential functions have other important properties; for example, they are eigenfunctions for differentiation. Many wavelets have a corresponding property captured in the "refinement equation" for the function ϕ from which the wavelet ψ is derived, as discussed in §3.1.

Another important property of wavelet decompositions not present directly in the Fourier decomposition is that the coefficients in wavelet decompositions usually encode all information needed to tell whether f is in a smoothness space, such as the Sobolev and Besov spaces. For example, if ψ is smooth enough, then a function

f is in the Lipschitz space $\mathrm{Lip}(\alpha, L_\infty(\mathbb{R}))$, $0 < \alpha < 1$, if and only if

$$\sup_{j,k} 2^{k(\alpha+\frac{1}{2})} |\langle f, \psi_{j,k}\rangle| \tag{1.2}$$

is finite, and (1.2) is an equivalent seminorm for $\mathrm{Lip}(\alpha, L_\infty(\mathbb{R}))$.

All this would be of little more than theoretical interest if it were not for the fact that one can efficiently compute wavelet coefficients and reconstruct functions from these coefficients. Such algorithms, known as "fast wavelet transforms" are the analogue of the Fast Fourier Transform and follow simply from the refinement equation mentioned above.

In many numerical applications, the orthogonality of the translated dilates $\psi_{j,k}$ is not vital. There are many variants of wavelets, such as the prewavelets proposed by Battle [Ba] and the ϕ-transform of Frazier and Jawerth [FJ], that do not require orthogonality. Typically, for a given function ψ, one wants the translated dilates $\psi_{j,k}$, $j, k \in \mathbb{Z}$, to form a stable basis (also called a Riesz basis) for $L_2(\mathbb{R})$. This means that each $f \in L_2(\mathbb{R})$ has a unique series decomposition in terms of the $\psi_{j,k}$, and that the ℓ_2 norm of the coefficients in this series is equivalent to $\|f\|_{L_2(\mathbb{R})}$ (this will be discussed in more detail in §3.1). In other applications, when approximating in $L_1(\mathbb{R})$, for example, one must abandon the requirement that $\psi_{j,k}$, $j, k \in \mathbb{Z}$, form a stable basis of $L_1(\mathbb{R})$, because none exists. (The Haar system is a Schauder basis for $L_1([0,1])$, for example, but the representation is not $L_1([0,1])$-stable.) For such applications, one can use redundant representations of f, with ψ a box spline, for example.

We have, to this point, restricted our discussion to univariate wavelets. There are several constructions of multivariate wavelets but the final form of this theory is yet to be decided. We shall discuss two methods for constructing multivariate wavelets; one is based on tensor products while the other is truly multivariate.

The plan of the paper is as follows. Section 2 is meant to introduce the topic of wavelets by studying the simplest orthogonal wavelets, which are the Haar functions. We discuss the decomposition of $L_p(\mathbb{R})$ using the Haar expansion, the characterization of certain smoothness spaces in terms of the coefficients in the Haar expansion, the fast Haar transform, and multivariate Haar functions. Section 3 concerns itself with the construction of wavelets. It begins with a discussion of the properties of shift-invariant spaces, and then gives an overview of the construction of univariate wavelets and prewavelets within the framework of multiresolution. Later, mention is made of Daubechies' specific construction of orthonormal wavelets of compact support. We finish with a discussion of wavelets in several dimensions.

Section 4 examines how to calculate the coefficients of wavelet expansions via the so-called Fast Wavelet Transform. Section 5 is concerned with the characterization of functions in certain smoothness classes called Besov spaces in terms of the size of wavelet coefficients. Section 6 turns to numerical applications. We briefly mention some uses of wavelets in nonlinear approximation, data compression (and, more specifically, image compression), and numerical methods for partial differential equations.

2. The Haar Wavelets

2.1. Overview. The Haar functions are the most elementary wavelets. While they have many drawbacks, chiefly their lack of smoothness, they still illustrate in the most direct way some of the main features of wavelet decompositions. For this reason, we shall consider in some detail the properties that make them suitable for numerical applications. We hope that the detail we provide at this stage will render more convincing some of the later statements we make, without proof, about more general wavelets.

We consider first the univariate case. Let $H := \chi_{[0,1/2)} - \chi_{[1/2,1)}$ be the Haar function that takes the value 1 on the left half of $[0,1]$ and the value -1 on the right half. By translation and dilation, we form the functions

$$H_{j,k} := 2^{k/2} H(2^k \cdot - j), \qquad j,k \in \mathbb{Z}. \tag{2.1.1}$$

Then, $H_{j,k}$ is supported on the dyadic interval $I_{j,k} := [j2^{-k}, (j+1)2^{-k})$.

It is easy to see that these functions form an orthonormal system. In fact, given two of these functions $H_{j,k}$, $H_{j',k'}$, $k' \geq k$ and $(j,k) \neq (j',k')$, we have two possibilities. The first is that the dyadic intervals $I_{j,k}$ and $I_{j',k'}$ are disjoint, in which case $\int_{\mathbb{R}} H_{j,k} H_{j',k'} = 0$ (because the integrand is identically zero). The second possibility is that $k' > k$ and $I_{j',k'}$ is contained in one of the halves J of $I_{j,k}$. In this case $H_{j,k}$ is constant on J while $H_{j',k'}$ takes the values ± 1 equally often on its support. Hence, again $\int_{\mathbb{R}} H_{j,k} H_{j',k'} = 0$.

We want next to show that $\{H_{j,k} \mid j,k \in \mathbb{Z}\}$ is complete in $L_2(\mathbb{R})$. The following development gives us a chance to introduce the concept of multiresolution, which is the main vehicle for constructing wavelets and which will be discussed in more detail in the section that follows. Let $\mathcal{S} := \mathcal{S}^0$ denote the subspace of $L_2(\mathbb{R})$ that consists of all piecewise-constant functions with integer breakpoints; i.e., functions in $\mathcal{S}$ are constant on each interval $[j, j+1)$, $j \in \mathbb{Z}$. Then $\mathcal{S}$ is a *shift-invariant space*: if $S \in \mathcal{S}$, each of its *shifts*, $S(\,\cdot\, + k)$, $k \in \mathbb{Z}$, is also in $\mathcal{S}$. A simple orthonormal basis for $\mathcal{S}$ is given by the shifts of the function $\phi := \chi_{[0,1]}$. Namely, each $S \in \mathcal{S}$ has a unique representation

$$S = \sum_{j \in \mathbb{Z}} c(j) \phi(\,\cdot\, - j), \qquad (c(j)) \in \ell_2(\mathbb{Z}). \tag{2.1.2}$$

By dilation, we can form a *scale* of spaces

$$\mathcal{S}^k := \{S(2^k\,\cdot\,) \mid S \in \mathcal{S}\}, \qquad k \in \mathbb{Z}.$$

Thus, $\mathcal{S}^k$ is the space of piecewise-constant $L_2(\mathbb{R})$ functions with breakpoints at the dyadic integers $j2^{-k}$. The normalized dyadic shifts $\phi_{j,k} := 2^{k/2}\phi(2^k \cdot - j) = 2^{k/2}\phi(2^k(\,\cdot\, - j2^{-k}))$ with step $j2^{-k}$, $j \in \mathbb{Z}$, of the function $\phi(2^k\,\cdot\,)$, form an orthonormal basis for $\mathcal{S}^k$. However, to avoid possible confusion, we note that the totality of all such functions $\phi_{j,k}$ is not a basis for the space $L_2(\mathbb{R})$ because there is redundancy. For example, $\phi = (\phi_{0,1} + \phi_{1,1})/\sqrt{2}$.

Clearly, we have $\mathcal{S}^k \subset \mathcal{S}^{k+1}$, $k \in \mathbb{Z}$, so the spaces $\mathcal{S}^k$ get "thicker" as k gets larger and "thinner" as k gets smaller. We are interested in the limiting spaces

$$\mathcal{S}^\infty := \overline{\bigcup \mathcal{S}^k} \text{ and } \mathcal{S}^{-\infty} := \bigcap \mathcal{S}^k, \tag{2.1.3}$$

since these spaces hold the key to showing that the Haar basis is complete. We claim that

$$\mathcal{S}^{\infty} = L_2(\mathbb{R}) \text{ and } \mathcal{S}^{-\infty} = \{0\}. \tag{2.1.4}$$

The first of these claims is equivalent to the fact that any function in $L_2(\mathbb{R})$ can be approximated arbitrarily well (in the $L_2(\mathbb{R})$ norm) by the piecewise-constant functions from $\mathcal{S}^k$ provided k is large enough. For example, it is enough to approximate f by its best $L_2(\mathbb{R})$ approximation from $\mathcal{S}^k$. This best approximation is given by the orthogonal projector P_k from $L_2(\mathbb{R})$ onto $\mathcal{S}^k$. It is easy to see that

$$P_k f(x) = \frac{1}{|I_{j,k}|}\int_{I_{j,k}} f, \qquad x \in I_{j,k},\ j \in \mathbb{Z}.$$

To verify the second claim in (2.1.4), we suppose that $f \in \bigcap \mathcal{S}^k$. Then, f is constant on each of $(-\infty, 0)$ and $[0,\infty)$, and since $f \in L_2(\mathbb{R})$, we must have $f = 0$ a.e. on each of these intervals.

Now, consider again the projector P_k from $L_2(\mathbb{R})$ onto $\mathcal{S}^k$. By (2.1.4), $P_k f \to f$, $k \to \infty$. We also claim that $P_k f \to 0$, $k \to -\infty$. Indeed, if this were false, then we could find a $C > 0$ and a subsequence $m_j \to -\infty$ for which $\|P_{m_j} f\|_{L_2(\mathbb{R})} \geq C$ for all m_j. By the weak-$*$ compactness of $L_2(\mathbb{R})$, we can also assume that $P_{m_j} f \to g$, weak-$*$, for some $g \in L_2(\mathbb{R})$. Now, for any $m \in \mathbb{Z}$, all $P_{m_j} f$ are in $\mathcal{S}^m$ for m_j sufficiently large and negative. Since $\mathcal{S}^m$ is weak-$*$ closed, $g \in \mathcal{S}^m$. Hence $g \in \bigcap \mathcal{S}^m$ implies $g = 0$ a.e. This gives a contradiction because by orthogonality

$$\int_{\mathbb{R}} |P_{m_j} f|^2\, dx = \int_{\mathbb{R}} \overline{f} P_{m_j} f\, dx \to \int_{\mathbb{R}} \overline{f} g\, dx = 0.$$

It follows that each $f \in L_2(\mathbb{R})$ can be represented by the series

$$f = \sum_{k\in\mathbb{Z}} (P_k f - P_{k-1} f) = \sum_{k\in\mathbb{Z}} Q_k f, \qquad Q_{k-1} := P_k - P_{k-1} \tag{2.1.5}$$

because the partial sums, $P_n f - P_{-n} f$, of this series tend to f as $n \to \infty$.

To complete the construction of the Haar wavelets, we need the following simple remarks about projections. If $Y \subset X$ are two closed subspaces of $L_2(\mathbb{R})$ and P_X and P_Y are the orthogonal projectors onto these spaces, then $Q := P_X - P_Y$ is the orthogonal projector from $L_2(\mathbb{R})$ onto $X \ominus Y$, the orthogonal complement of Y in X (this follows from the identity $P_Y P_X = P_Y$). Thus, the operator $Q_{k-1} := P_k - P_{k-1}$ appearing in (2.1.5) is the orthogonal projector onto $W^{k-1} := \mathcal{S}^k \ominus \mathcal{S}^{k-1}$. The spaces W^k are the dilates of the *wavelet* space

$$W := \mathcal{S}^1 \ominus \mathcal{S}^0. \tag{2.1.6}$$

Since the spaces W^k, $k \in \mathbb{Z}$, are mutually orthogonal, we have $W^k \perp W^j$, $j \neq k$, and (2.1.5) shows that $L_2(\mathbb{R})$ is the orthogonal direct sum of the W^k:

$$L_2(\mathbb{R}) = \bigoplus_{k\in\mathbb{Z}} W^k. \tag{2.1.7}$$

How does the Haar function fit into all this? Well, the main point is that H and its translates $H(\cdot - k)$ form an orthonormal basis for W. Indeed, $H = 2\phi(2\,\cdot) - P_0(2\phi(2\,\cdot))$, which shows that H is in $\mathcal{S}^1 \ominus \mathcal{S}^0 = W$. On the other hand, the identities $H + \phi = 2\phi(2\,\cdot)$ and $\phi - H = 2\phi(2\,\cdot - 1)$ show that the shifts of ϕ together with the shifts of H will generate all the half-shifts of $\eta := \phi(2\,\cdot)$. Since the half-shifts of $\sqrt{2}\,\eta$ form an orthonormal basis for $\mathcal{S}^1$, the shifts of H must be complete in W.

By dilation, the functions $H_{j,k}$, $j \in \mathbb{Z}$, form a complete orthonormal system for W^k. Hence, we can represent the orthogonal projector Q_k onto W^k by

$$Q_k f = \sum_{j\in\mathbb{Z}} \langle f, H_{j,k}\rangle\, H_{j,k}.$$

Using this in (2.1.5), we have for any $f \in L_2(\mathbb{R})$ the decomposition

$$f = \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}} \langle f, H_{j,k}\rangle\, H_{j,k}. \tag{2.1.8}$$

In other words, the functions $H_{j,k}$, $j,k \in \mathbb{Z}$, form an orthonormal basis for $L_2(\mathbb{R})$.

2.2. The Haar decomposition in $L_p(\mathbb{R})$. While the Haar decomposition is initially defined only for functions in $L_2(\mathbb{R})$, it is worth noting that Haar decompositions also hold for other spaces of functions. In this section, we shall discuss the Haar representation for functions in $L_p(\mathbb{R})$, $1 \le p < \infty$. A similar analysis can be given when $p = \infty$ if $L_\infty(\mathbb{R})$ is replaced by the space of uniformly continuous functions that vanish at ∞, equipped with the $L_\infty(\mathbb{R})$ norm.

If $f \in L_p(\mathbb{R})$, the Haar coefficients $\langle f, H_{j,k}\rangle$ are well defined and we can ask whether the Haar series (2.1.8) converges in $L_p(\mathbb{R})$ to f. To answer this question, we fix a value of $1 \le p < \infty$ and a $k \in \mathbb{Z}$ and examine the projector P_k, which is initially defined only on $L_2(\mathbb{R})$. For any $f \in L_2(\mathbb{R})$, we have $P_k f = \sum_{I\in D_k} f_I \chi_I$ where D_k denotes the collection of dyadic intervals of length 2^{-k}, and where $f_I := \frac{1}{|I|}\int_I f\,dx$, $I \in D_k$, is the average of f over I. In this form, the projector P_k has a natural extension to $L_p(\mathbb{R})$ and takes values in the space $S^k(\chi, L_p(\mathbb{R}))$ of all functions in $L_p(\mathbb{R})$ that are piecewise-constant functions with breakpoints at the dyadic integers $j2^{-k}$, $j \in \mathbb{Z}$.

In representing P_k on $L_p(\mathbb{R})$, it is useful to change our normalization slightly. We fix a value of p and consider the $L_p(\mathbb{R})$-normalized characteristic functions $\phi_{j,k,p} := 2^{k/p}\phi(2^k\cdot - j)$, $\phi := \chi_{[0,1]}$, which satisfy $\int_{\mathbb{R}} |\phi_{j,k,p}|^p\,dx = 1$. Then,

$$P_k f = \sum_{j\in\mathbb{Z}} \langle f, \phi_{j,k,p'}\rangle\, \phi_{j,k,p}, \qquad \frac{1}{p} + \frac{1}{p'} = 1.$$

From Hölder's inequality, we find $|\langle f, \phi_{j,k,p'}\rangle|^p \le \int_{I_{j,k}} |f|^p\,dx$ and so

$$\|P_k f\|^p_{L_p(\mathbb{R})} = \sum_{j\in\mathbb{Z}} |\langle f, \phi_{j,k,p'}\rangle|^p \le \sum_{j\in\mathbb{Z}} \int_{I_{j,k}} |f|^p\,dx = \int_{\mathbb{R}} |f|^p\,dx.$$

Therefore, P_k is a bounded operator with norm 1 on the space $L_p(\mathbb{R})$.

If $f \in L_p(\mathbb{R})$, then since P_k is a projector of norm one,

$$\|f - P_k f\|_{L_p(\mathbb{R})} = \inf_{S \in \mathcal{S}^k} \|(I - P_k)(f - S)\|_{L_p(\mathbb{R})} \le 2\,\mathrm{dist}(f, \mathcal{S}^k)_{L_p(\mathbb{R})}. \tag{2.2.1}$$

It follows that $P_k f \to f$ in $L_p(\mathbb{R})$ for each $f \in L_p(\mathbb{R})$.

On the other hand, consider $P_k f$ as $k \to -\infty$. If f is continuous and of compact support then at most 2 terms in $P_k f$ are nonzero for k large negative and each coefficient is $\le C2^{k/p'}$. Hence $\|P_k f\|_{L_p(\mathbb{R})} \to 0$ provided $p' < \infty$, i.e., $p > 1$. This shows that

$$f = \sum_{k \in \mathbb{Z}} (P_k f - P_{k-1} f) = \sum_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \langle f, H_{j,k} \rangle\, H_{j,k}. \tag{2.2.2}$$

in the sense of $L_p(\mathbb{R})$ convergence. We see that the Haar representation holds for functions in $L_p(\mathbb{R})$ provided $p > 1$.

But what happens when $p = 1$? Well, as is typical for orthogonal decompositions, the expansion (2.2.2) cannot be valid. Indeed, each of the functions appearing on the right in (2.2.2) has mean value zero. If $g \in L_1(\mathbb{R})$ has mean value zero and f is an arbitrary function from $L_1(\mathbb{R})$, then

$$\int_{\mathbb{R}} |f - g|\, dx \ge |\int_{\mathbb{R}} f\, dx - \int_{\mathbb{R}} g\, dx| = |\int_{\mathbb{R}} f\, dx|.$$

This means that the sum in (2.2.2) cannot possibly converge in $L_1(\mathbb{R})$ to f unless f has mean value zero.

The above phenomenon is typical of decompositions for orthogonal wavelets ψ: They cannot represent all functions in $L_1(\mathbb{R})$. However, if ψ is smooth enough, the representation (2.2.2) will hold for the Hardy space $H_1(\mathbb{R})$ used in place of $L_1(\mathbb{R})$, and in fact this representation will then hold for functions in the Hardy spaces $H_p(\mathbb{R})$ for a certain range of $0 < p < 1$ that depends on the smoothness of ψ. We shall not discuss further the behavior of orthogonal wavelets in H_p spaces but the interested reader can consult Frazier and Jawerth [FJ] for a corresponding theory in a slightly different setting.

2.3. Smoothness spaces. We noted earlier the important fact that wavelet decompositions provide a description of smoothness spaces in terms of the wavelet coefficients. We wish to illustrate this point with the Haar wavelets and the Lipschitz spaces in $L_p(\mathbb{R})$, $1 < p < \infty$.

The Lipschitz spaces $\mathrm{Lip}(\alpha, L_p(\mathbb{R}))$ of $L_p(\mathbb{R})$, $0 < \alpha \le 1$, $1 \le p \le \infty$, consist of all functions $f \in L_p(\mathbb{R})$ for which

$$\|f - f(\,\cdot\, + h)\|_{L_p(\mathbb{R})} = \mathcal{O}(h^\alpha), \qquad h \to 0.$$

A seminorm for this space is provided by

$$|f|_{\mathrm{Lip}(\alpha, L_p(\mathbb{R}))} := \sup_{0 < h < \infty} h^{-\alpha} \|f - f(\,\cdot\, + h)\|_{L_p(\mathbb{R})}.$$

The relationship between the smoothness of f and the size of its Haar coefficients rests on three fundamental inequalities. The first of these says that for a fixed $k \in \mathbb{Z}$, the Haar functions $H_{j,k}$, $j \in \mathbb{Z}$, are $L_p(\mathbb{R})$-stable. Because of the disjoint support of the $H_{j,k}$, $j \in \mathbb{Z}$, stability takes the following particularly simple form: for any sequence $(c(j)) \in \ell_p(\mathbb{Z})$ and $S = \sum_{j\in\mathbb{Z}} c(j)H_{j,k}$, we have

$$\|S\|_{L_p(\mathbb{R})} = \Big(\sum_{j\in\mathbb{Z}} |c(j)|^p 2^{kp(1/2-1/p)}\Big)^{1/p}. \tag{2.3.1}$$

This follows by integrating the identity

$$|S|^p = \sum_{j\in\mathbb{Z}} |c(j)|^p |H_{j,k}|^p.$$

The other two inequalities are related to the approximation properties of $\mathcal{S}^k$ and the projectors P_k:

(2.3.2)

$$\text{(J)}\quad \|f - P_k f\|_{L_p(\mathbb{R})} \le 2\cdot 2^{-k\alpha}|f|_{\mathrm{Lip}(\alpha,L_p(\mathbb{R}))},\quad 0<\alpha\le 1,\ 1\le p\le\infty.$$

$$\text{(B)}\quad |S|_{\mathrm{Lip}(1/p,L_p(\mathbb{R}))} \le 2\cdot 2^{k/p}\|S\|_{L_p(\mathbb{R})},\quad S\in\mathcal{S}^k(\chi,L_p(\mathbb{R})),\ 1\le p\le\infty.$$

The first of these, often called a Jackson inequality (after similar inequalities established by D. Jackson for polynomial approximation), tells how well functions from $\mathrm{Lip}(\alpha, L_p(\mathbb{R}))$ can be approximated by the elements of $\mathcal{S}^k$. The second inequality is known as a Bernstein inequality because of its similarity with the classical Bernstein inequalities for polynomials, established by S. Bernstein.

We shall prove (2.3.2) (J) and (B) for $1 \le p < \infty$. If $I \in D_k$ and $h := |I| = 2^{-k}$, then, for all $x \in I$, we obtain

$$\begin{aligned}|f(x) - f_I| &\le \frac{1}{|I|}\int_I |f(x) - f(y)|\,dy\\ &\le \Big(\frac{1}{|I|}\int_I |f(x)-f(y)|^p\,dy\Big)^{1/p} \le \Big(\frac{1}{|I|}\int_{-h}^{h} |f(x)-f(x+s)|^p\,ds\Big)^{1/p}.\end{aligned}$$

If we raise these last inequalities to the power p, integrate over I, and then sum over all $I \in D_k$, we obtain

$$\|f - P_k f\|^p_{L_p(\mathbb{R})} \le \frac{1}{h}\int_{-h}^{h}\int_{\mathbb{R}} |f(x)-f(x+s)|^p\,dx\,ds \le 2|h|^{\alpha p}|f|^p_{\mathrm{Lip}(\alpha,L_p(\mathbb{R}))},$$

which implies the Jackson inequality.

The Jackson inequality can also be proved from more general principles. Since the P_k have norm 1 on $L_p(\mathbb{R})$ and are projectors, we have

$$\|f - P_k f\|_{L_p(\mathbb{R})} \le (1 + \|P_k\|)\operatorname{dist}(f,\mathcal{S}^k)_{L_p(\mathbb{R})} \le 2\operatorname{dist}(f,\mathcal{S}^k)_{L_p(\mathbb{R})}. \tag{2.3.3}$$

Thus, the Jackson inequality follows from the fact that functions in $\mathrm{Lip}(\alpha, L_p(\mathbb{R}))$ can be approximated by the elements of $\mathcal{S}^k$ with an error not exceeding the right side of (2.3.2)(J).

To prove the Bernstein inequality, we note that any $S = \sum_{j\in\mathbb{Z}} c(j)\chi_{I_{j,k}}$ in $\mathcal{S}^k(\chi, L_p(\mathbb{R}))$ has norm:

$$\|S\|^p_{L_p(\mathbb{R})} = \sum_{j\in\mathbb{Z}} |c(j)|^p |I_{j,k}| = \sum_{j\in\mathbb{Z}} |c(j)|^p 2^{-k}. \tag{2.3.4}$$

We fix an $h > 0$. If $h \geq 2^{-k}$ (i.e., $h^{-1/p} \leq 2^{k/p}$), then

$$h^{-1/p}\|S(\,\cdot\,+h) - S\|_{L_p(\mathbb{R})} \leq 2^{k/p}(\|S(\,\cdot\,+h)\|_{L_p(\mathbb{R})} + \|S\|_{L_p(\mathbb{R})}) = 2\cdot 2^{k/p}\|S\|_{L_p(\mathbb{R})}.$$

If $h < 2^{-k}$ then

$$|S(x+h) - S(x)| = \begin{cases} 0, & x \in [j2^{-k}, (j+1)2^{-k} - h), \\ |c(j+1) - c(j)|, & x \in [(j+1)2^{-k} - h, (j+1)2^{-k}). \end{cases}$$

Therefore

$$\begin{aligned} h^{-1/p}\|S(\,\cdot\,+h) - S\|_{L_p(\mathbb{R})} &= h^{-1/p}\Big(\sum_{j\in\mathbb{Z}} |c(j+1) - c(j)|^p h\Big)^{1/p} \\ &= 2^{k/p}\Big(\sum_{j\in\mathbb{Z}} |c(j+1) - c(j)|^p 2^{-k}\Big)^{1/p} \\ &= 2^{k/p}\|S(\,\cdot\,+2^{-k}) - S\|_{L_p(\mathbb{R})} \leq 2\cdot 2^{k/p}\|S\|_{L_p(\mathbb{R})}. \end{aligned}$$

With the Jackson and Bernstein inequalities in hand, it is now easy to show that

$$|f|_{\mathrm{Lip}(\alpha, L_p(\mathbb{R}))} \approx \sup_{k\in\mathbb{Z}} 2^{k(\alpha+1/2-1/p)}\Big(\sum_{j\in\mathbb{Z}} |\langle f, H_{j,k}\rangle|^p\Big)^{1/p}, \quad 0 < \alpha < 1/p. \tag{2.3.5}$$

(It will be convenient to use the notation $A \approx B$ to mean that the two ratios A/B and B/A of the functions A and B are bounded from above independently of the variables; in (2.3.5), independently of f.) First, from the Jackson inequality,

$$\|P_k f - P_{k-1}f\|_{L_p(\mathbb{R})} \leq \|f - P_k f\|_{L_p(\mathbb{R})} + \|f - P_{k-1}f\|_{L_p(\mathbb{R})} \leq C\,2^{-k\alpha}|f|_{\mathrm{Lip}(\alpha, L_p(\mathbb{R}))}.$$

If we write $P_k f - P_{k-1}f = \sum_{j\in\mathbb{Z}}\langle f, H_{j,k-1}\rangle\, H_{j,k-1}$ and replace $\|P_k f - P_{k-1}f\|_{L_p(\mathbb{R})}$ by the sum in (2.3.1) (with $c(j) = \langle f, H_{j,k-1}\rangle$), we obtain that the right side of (2.3.5) does not exceed a multiple of the left.

To reverse this inequality, we fix a value of h and choose $n \in \mathbb{Z}$ so that $2^{-n} \leq h \leq 2^{-n+1}$. We write $f = \sum_{k\in\mathbb{Z}} w_k$ with $w_k := (P_{k+1}f - P_k f)$ and estimate

$$\begin{aligned} &\|f(\,\cdot\,+h) - f\|_{L_p(\mathbb{R})} \\ &\qquad \leq \sum_{k\geq n} \|w_k(\,\cdot\,+h) - w_k\|_{L_p(\mathbb{R})} + \sum_{k<n} \|w_k(\,\cdot\,+h) - w_k\|_{L_p(\mathbb{R})}. \end{aligned} \tag{2.3.6}$$

The first sum does not exceed

$$\Big(\sup_{k\ge n} 2^{k\alpha}\|w_k\|_{L_p(\mathbb{R})}\Big)\Big(\sum_{k\ge n} 2^{-k\alpha}\Big) \le Ch^\alpha \sup_{k\ge n} 2^{k\alpha}\|w_k\|_{L_p(\mathbb{R})}.$$

Similarly, using the Bernstein inequality, the second sum does not exceed

$$\begin{aligned} h^{1/p}\sum_{k<n}|w_k|_{\mathrm{Lip}(1/p,L_p(\mathbb{R}))} &\le C\,h^{1/p}\sum_{k<n}2^{k/p}\|w_k\|_{L_p(\mathbb{R})}\\ &\le C\,h^{1/p}\Big(\sup_{k<n}2^{k\alpha}\|w_k\|_{L_p(\mathbb{R})}\Big)\Big(\sum_{k<n}2^{k(1/p-\alpha)}\Big)\\ &\le C\,h^{\alpha}\sup_{k<n}2^{k\alpha}\|w_k\|_{L_p(\mathbb{R})}.\end{aligned}$$

If we write $w_k = \sum_{j\in\mathbb{Z}}\langle f, H_{j,k}\rangle\, H_{j,k}$ and use (2.3.1) to replace $\|w_k\|_{L_p(\mathbb{R})}$ by

$$2^{k(1/2-1/p)}\Big(\sum_{j\in\mathbb{Z}}|\langle f, H_{j,k}\rangle|^p\Big)^{1/p}$$

in each of these expressions, and then use the resulting expression in (2.3.6), we obtain

$$\|f(\,\cdot\,+h)-f\|_{L_p(\mathbb{R})} \le Ch^\alpha \sup_{k\in\mathbb{Z}} 2^{k(\alpha+1/2-1/p)}\Big(\sum_{j\in\mathbb{Z}}|\langle f, H_{j,k}\rangle|^p\Big)^{1/p},$$

which shows that the left side of (2.3.5) does not exceed a multiple of the right.

The restriction $\alpha < 1/p$ arises because the Haar function is not smooth; for smoother wavelets, the range of α can be increased.

2.4. The fast Haar transform. In numerical applications of the Haar decomposition, one must work with only a finite number of the functions $H_{j,k}$. The choice of which functions to use is often made as follows. Given a function $f\in L_2(\mathbb{R})$, we choose a large value of n, compatible with the accuracy we wish to achieve, and we replace f by P_nf with P_n, as before, the $L_2(\mathbb{R})$ projector onto $\mathcal{S}^n$, the space of piecewise-constant functions in $L_2(\mathbb{R})$ with breakpoints at the dyadic integers $j2^{-n}$, $j\in\mathbb{Z}$. If f has compact support then P_nf is a finite linear combination of the characteristic functions χ_I, $I\in D_n$. If f does not have compact support, it is necessary to truncate this sum (which is justified because $\int_{\mathbb{R}\setminus[-a,a]}|f|^2\,dx\to 0$, $a\to\infty$).

We can now write

$$P_nf = (P_nf-P_{n-1}f)+\cdots+(P_1f-P_0f)+P_0f = P_0f+\sum_{k=0}^{n-1}Q_kf, \tag{2.4.1}$$

which is a finite Haar decomposition. We have started this decomposition with P_0f but we could have equally well started at any other dyadic level.

The fast Haar transform gives an efficient method for finding the coefficients in the expansions

$$Q_k f = \sum_{j\in\mathbb{Z}} d(j,k)H_{j,k}, \quad d(j,k) := \langle f, H_{j,k}, \rangle \tag{2.4.2}$$

and

$$P_k f = \sum_{j\in\mathbb{Z}} c(j,k)\phi_{j,k}, \quad c(j,k) := \langle f, \phi_{j,k}\rangle. \tag{2.4.3}$$

These coefficients are related to the integrals of f over the intervals $I_{j,k} := [j2^{-k}, (j+1)2^{-k})$:

$$\begin{aligned} c(j,k) &= 2^{k/2}\int_{I_{j,k}} f\,dx, \\ d(j,k) &= 2^{k/2}\left[\int_{I_{2j,k+1}} f\,dx - \int_{I_{2j+1,k+1}} f\,dx\right]. \end{aligned}$$

Therefore, if the coefficients $c(j,k+1)$, $j\in\mathbb{Z}$, are known, then

$$\begin{aligned} c(j,k) &= \frac{1}{\sqrt{2}}(c(2j,k+1)+c(2j+1,k+1)), \\ d(j,k) &= \frac{1}{\sqrt{2}}(c(2j,k+1)-c(2j+1,k+1)). \end{aligned} \tag{2.4.4}$$

In other words, starting with the known values of $c(j,n)$ at level n, we can iteratively compute all values $d(j,k)$ and $c(j,k)$ needed for (2.4.1) from (2.4.4). The computation of the $c(j,k)$ at dyadic levels $k\neq 0$ is necessary for the recurrence even though we are in the end not interested in their values.

There is a similar formula for reconstructing a function from its Haar coefficients. Now, suppose that we know the coefficients appearing in (2.4.1), i.e., the values $c(j,0)$, $j\in\mathbb{Z}$, and $d(j,k)$, $j\in\mathbb{Z}$, $k=1,\dots,n$, and we wish to find $c(j,n)$, i.e., to reconstruct f. For this we need only use the recursive formulas

$$\begin{aligned} c(2j,k+1) &= \frac{1}{\sqrt{2}}(c(j,k)+d(j,k)), \\ c(2j+1,k+1) &= \frac{1}{\sqrt{2}}(c(j,k)-d(j,k)). \end{aligned}$$

More information on the structure of the fast Haar transform can be found in §5.

2.5. Multivariate Haar functions. There is a simple method to construct multivariate wavelets from a given univariate wavelet, which, for the Haar wavelets, takes the following form. Let $\phi_0 := \phi = \chi_{[0,1]}$ and $\phi_1 := \psi = H$ and let V denote the set of vertices of the cube $\Omega := [0,1]^d$. For each $v=(v_1,\dots,v_d)$ in V and $x=(x_1,\dots,x_d)$ from $\mathbb{R}^d$, we let $\psi_v(x) := \prod_{i=1}^d \phi_{v_i}(x_i)$. The functions ψ_v are piecewise constant, taking the values ± 1 on the d-tants of Ω. The set

$\Psi := \{\psi_v \mid v \in V,\ v \neq 0\}$ is the set of multidimensional Haar functions; there are $2^d - 1$ of them. They generate by dilation and translation an orthonormal basis for $L_2(\mathbb{R}^d)$. That is, the collection of functions $2^{kd/2}\psi_v(2^k\cdot - j)$, $j \in \mathbb{Z}^d$, $k \in \mathbb{Z}$, $v \in V \setminus \{0\}$, forms a complete orthonormal basis for $L_2(\mathbb{R}^d)$.

Another way to view the multidimensional Haar functions is to consider the shift-invariant space $\mathcal{S}$ of piecewise-constant functions on the dyadic cubes of unit length in $\mathbb{R}^d$. A basis for $\mathcal{S}$ is provided by the shifts of $\chi_{[0,1]^d}$. Note that the space $\mathcal{S}$ is the tensor product of the univariate spaces of piecewise-constant functions with integer breakpoints. The collection of all shifts of the Haar functions $\psi_v \in \Psi$ forms an orthonormal basis for the space $W := \mathcal{S}^1 \ominus \mathcal{S}^0$. Properties of the multivariate Haar wavelets follow from the univariate Haar function. For example, there is a fast Haar transform and a characterization of smoothness spaces in terms of Haar coefficients. We leave the formulation of these properties to the reader.

3. The Construction of Wavelets

3.1. Overview. We turn now to the construction of smoother orthogonal wavelets. Almost all constructions of orthogonal wavelets begin by using multiresolution, which was introduced by Mallat [Ma] (an interesting exception, presented by Strömberg [St], apparently gave the first smooth orthogonal wavelets). We begin with a brief overview of multiresolution that we will expand on in later sections.

Let $\phi \in L_2(\mathbb{R}^d)$ and let $\mathcal{S} := \mathcal{S}(\phi)$ be the shift-invariant subspace of $L_2(\mathbb{R}^d)$ generated by ϕ. That is, $\mathcal{S}(\phi)$ is the $L_2(\mathbb{R}^d)$ closure of finite linear combinations of ϕ and its shifts $\phi(\,\cdot + j)$, $j \in \mathbb{Z}^d$. By dilation, we form the scale of spaces

$$\mathcal{S}^k := \{S(2^k\,\cdot) \mid S \in \mathcal{S}\}. \tag{3.1.1}$$

Then $\mathcal{S}^k$ is invariant under dyadic shifts $j2^{-k}$, $j \in \mathbb{Z}^d$. In the construction of Haar functions, we had $d = 1$, and $\mathcal{S}$ was the space of piecewise-constant functions with integer breakpoints. That is, $\mathcal{S} = \mathcal{S}(\phi)$ with $\phi := \chi := \chi_{[0,1]}$. Other examples for the reader to keep in mind, which result in smoother wavelets, are to take for $\mathcal{S}$ the space of cardinal spline functions of order r in $L_2(\mathbb{R})$. A cardinal spline is a piecewise polynomial function defined on $\mathbb{R}$, of local degree $< r$, that has breakpoints at the integers and has global smoothness C^{r-2}. Then $\mathcal{S} = \mathcal{S}(N_r)$ with N_r the (nonzero) cardinal B-spline that has knots at $0, 1, \ldots, r$. These B-splines are easiest to define recursively: $N_1 := \chi$ and $N_r := N_{r-1} * N_1$, with the usual operation of convolution

$$f * g(x) := \int_{\mathbb{R}} f(x-y)g(y)\,dy.$$

For example, N_2 is a hat function, N_3 a C^1 piecewise quadratic, and so on. In the multivariate case, the primary examples to keep in mind are the tensor product of univariate B-splines: $N(x) := N(x_1, \ldots, x_d) := N(x_1)\cdots N(x_d)$, and the box splines, which will be introduced and discussed later.

Multiresolution begins with certain assumptions on the scale of spaces $\mathcal{S}^k$ and shows under these assumptions how to construct an orthogonal wavelet ψ from the generating function ϕ. The usual assumptions are:

$$\text{(3.1.2)}\qquad \begin{aligned} &\text{(i)} \quad \mathcal{S}^k \subset \mathcal{S}^{k+1}, \quad k \in \mathbb{Z};\\ &\text{(ii)} \quad \overline{\bigcup \mathcal{S}^k} = L_2(\mathbb{R}^d);\\ &\text{(iii)} \quad \bigcap \mathcal{S}^k = \{0\};\\ &\text{(iv)} \quad \{\phi(\,\cdot - j)\}_{j\in\mathbb{Z}^d} \text{ forms an } L_2(\mathbb{R}^d)\text{-stable basis for } \mathcal{S}. \end{aligned}$$

We have already seen the role of (ii) and (iii) in the context of Haar decompositions. The assumption (iv) means that there exist positive constants C_1 and C_2 such that each $S \in \mathcal{S}$ has a unique representation

$$\text{(3.1.3)}\qquad \begin{aligned} &\text{(i)} \quad S = \sum_{j\in\mathbb{Z}^d} c(j)\phi(\,\cdot - j), \text{ and}\\ &\text{(ii)} \quad C_1\|S\|_{L_2(\mathbb{R}^d)} \le \left(\sum_{j\in\mathbb{Z}^d} |c(j)|^2\right)^{1/2} \le C_2\|S\|_{L_2(\mathbb{R}^d)}. \end{aligned}$$

If ϕ has $L_2(\mathbb{R}^d)$-stable shifts then it follows by a change of variables that for each $k \in \mathbb{Z}$, the function $2^{kd/2}\phi(2^k\,\cdot\,)$ has $L_2(\mathbb{R}^d)$-stable $2^{-k}\mathbb{Z}^d$ shifts. We shall mention later how the assumption (3.1.2)(iv) can be weakened.

Assumption (3.1.2)(i) is a severe restriction on the underlying function ϕ. Because each space $\mathcal{S}^k$ is obtained from $\mathcal{S}$ by dilation, we see that (3.1.2)(i) is satisfied if and only if $\mathcal{S} \subset \mathcal{S}^1$, or, equivalently, if ϕ is in the space $\mathcal{S}^1$. From the $L_2(\mathbb{R}^d)$-stability of the set $\{\phi(\,\cdot - j)\}_{j\in\mathbb{Z}^d}$, this is equivalent to

$$\phi(x) = \sum_{j\in\mathbb{Z}^d} a(j)\phi(2x - j) \tag{3.1.4}$$

for some sequence $(a(j)) \in \ell_2(\mathbb{Z}^d)$. Equation (3.1.4) is called the *refinement equation* for ϕ, since it says that ϕ can be expressed as a linear combination of the scaled functions $\phi(2\,\cdot - j)$, which are at the finer dyadic level. We shall discuss the refinement equation in more detail later and for now only point out that this equation is well known for the B-spline of order r, for which it takes the form

$$N_r(x) = 2^{-r+1}\sum_{j=0}^{r}\binom{r}{j}N_r(2x - j). \tag{3.1.5}$$

Because of (3.1.2)(i), the *wavelet space*

$$W := \mathcal{S}^1 \ominus \mathcal{S}^0$$

is a subspace of $\mathcal{S}^1$. By dilation, we obtain the scaled wavelet spaces W^k, $k \in \mathbb{Z}$. Then, W^k is orthogonal to $\mathcal{S}^k$ and

$$\mathcal{S}^{k+1} = \mathcal{S}^k \oplus W^k. \tag{3.1.6}$$

Since $W^j \subset \mathcal{S}^k$ for $j < k$, it follows that W_j and W_k are orthogonal. From this and (3.1.2)(ii) and (iii), we obtain

$$L_2(\mathbb{R}^d) = \bigoplus_{k\in\mathbb{Z}} W^k. \tag{3.1.7}$$

We find wavelets by showing that W is shift invariant and finding its generators. For example, when $d = 1$, W is a principal shift-invariant space, that is, it can be generated by one element ψ, i.e. $W = \mathcal{S}(\psi)$. Of course, there are many such generators ψ for W. In the multivariate case, the space W will be generated by $2^d - 1$ such functions.

We find an orthogonal wavelet in one dimension by determining a ψ whose shifts form an orthonormal basis for W. Indeed, once such a function ψ is found, the scaled functions $\psi_{j,k} := 2^{k/2}\psi(2^k\cdot - j)$ will then form an orthonormal basis for $L_2(\mathbb{R})$.

Generators ψ for W whose shifts are not orthogonal are nonorthogonal wavelets. For example, if ψ has shifts that are $L_2(\mathbb{R})$-stable (but not orthonormal), the functions $\psi_{j,k} := 2^{k/2}\psi(2^k\cdot - j)$ form an $L_2(\mathbb{R})$-stable basis for $L_2(\mathbb{R})$. While they do not form an orthonormal system, they still possess orthogonality between levels,

$$\int_{\mathbb{R}} \psi_{j,k}\psi_{j',k'}\,dx = 0, \quad k \neq k',$$

which is enough for most applications. After Battle [Ba], we call such functions ψ *prewavelets*.

The construction of (univariate) orthogonal wavelets introduced by Mallat [Ma], begins with a function ϕ that has orthonormal shifts (rather than just $L_2(\mathbb{R})$-stability). Mallat shows that the function

$$\psi := \sum_{j\in\mathbb{Z}} (-1)^j \overline{a(1-j)}\phi(2\cdot - j), \tag{3.1.8}$$

with $(a(j))$ the refinement coefficients of (3.1.4), is an orthogonal wavelet. (It is easy to check that ψ is orthogonal to the shifts of ϕ by using the refinement equation (3.1.4)). A construction similar to that of Mallat was used by Chui and Wang [CW1] and Micchelli [Mi] to produce prewavelets. In the construction of prewavelets, they begin with a function ϕ that has $L_2(\mathbb{R})$-stable shifts (but not necessarily orthonormal shifts). Then a formula similar to (3.1.8) gives a prewavelet ψ (see (3.4.15)).

To find generators for the wavelet space W, we shall follow the construction of de Boor, DeVore, and Ron [BDR1], which is somewhat different from that of Mallat. We simply take suitable functions η in the space $\mathcal{S}^1$ and consider their orthogonal projections $P\eta$ onto the space $\mathcal{S}$. The error function $w := \eta - P\eta$ is then an element of W. By choosing appropriate functions η, we obtain a set of generators for W. In one dimension, only one function is needed to generate W and any reasonable choice for η results in such a generator. The most obvious choices, $\eta := \phi(2\,\cdot)$ or $\eta := \phi(2\cdot - 1)$, lead to the wavelet (3.1.8) or its prewavelet analogue.

If we begin with the orthonormalized shifts of the B-spline $\phi = N_r$ as the basis for $\mathcal{S}$, the construction of Mallat gives the spline wavelets ψ of Battle-Lemarié (see, e.g., [Ba]), which have smoothness C^{r-2}. The support of ψ is all of $\mathbb{R}$, although ψ does decay exponentially at infinity. More details are given in §3.4. If we do not orthonormalize the shifts, we obtain the spline prewavelets of Chui and Wang [CW], which have compact support (in fact minimal support among all functions in W).

It is a more substantial problem to construct smooth orthogonal wavelets of compact support and this was an outstanding achievement of Daubechies [Da] (see §3.5). Daubechies' construction depends on finding a *compactly supported* function $\phi \in C^r$ that satisfies the assumptions of multiresolution and has *orthonormal* shifts. In this way, she is able to apply Mallat's construction to obtain a compactly supported orthogonal wavelet ψ in C^r.

The construction of multivariate wavelets by multiresolution is based on similar ideas. We want now to find a set of generators $\Psi = \{\psi\}$ for the wavelet space W. There are typically $2^d - 1$ functions in Ψ. This is an orthogonal wavelet set if the totality of functions $\psi_{j,k}$, $j \in \mathbb{Z}^d$, $k \in \mathbb{Z}$, $\psi \in \Psi$, forms an orthonormal basis for $L_2(\mathbb{R}^d)$. For this to hold, it is sufficient to have orthogonality between $\psi(\cdot - j)$ and $\tilde{\psi}(\cdot - j')$, $(j, \psi) \neq (j', \tilde{\psi})$. If the shifts of the functions $\psi \in \Psi$ form an $L_2(\mathbb{R}^d)$-stable basis for W, we say this is a prewavelet set. In this case, we shall still have the orthogonality between levels: $\psi_{j,k} \perp \tilde{\psi}_{j',k'}$ if $k \neq k'$. Sometimes we also require orthogonality between $\tilde{\psi} \in \Psi$ and all of the $\psi(\cdot - j)$, $j \in \mathbb{Z}^d$, $\psi \neq \tilde{\psi}$. Because the construction of multivariate wavelets is significantly more complicated and more poorly understood than the construction of wavelets of one variable, we shall postpone the discussion of multivariate wavelets until §3.6.

In the following sections, we shall show how to construct wavelets and prewavelets in the setting of multiresolution. These constructions depend on a good description of the space $\mathcal{S} := \mathcal{S}(\phi)$ in terms of Fourier transforms, which is the topic of the next section.

3.2. Shift-invariant spaces. Because multiresolution is based on a family of shift-invariant spaces, it is useful to have in mind the structure of these spaces before proceeding with the construction of wavelets and prewavelets. The structure of shift-invariant spaces and their application to approximation and wavelet construction were developed in a series of papers by de Boor, DeVore, and Ron [BDR], [BDR1], [BDR2]; much of the material in our presentation is taken from these references.

We recall that a closed subspace $\mathcal{S}$ of $L_2(\mathbb{R}^d)$ is shift invariant if $S(\cdot + j)$, $j \in \mathbb{Z}^d$, is in $\mathcal{S}$ whenever $S \in \mathcal{S}$. We have already encountered the space $\mathcal{S}(\phi)$, which is the $L_2(\mathbb{R}^d)$-closure of finite linear combinations of the shifts of ϕ. We say that such a space is a *principal shift-invariant space* (in analogy with principal ideals). More generally, if Φ is a finite set of $L_2(\mathbb{R}^d)$ functions, then the space $\mathcal{S}(\Phi)$ is the $L_2(\mathbb{R}^d)$-closure of finite linear combinations of the shifts of the functions $\phi \in \Phi$. Of course, a general shift-invariant subspace of $L_2(\mathbb{R}^d)$ need not be finitely generated.

We are interested in describing the space $\mathcal{S}(\Phi)$ in terms of its Fourier transforms.

We let

$$\hat{f}(x) := \int_{\mathbb{R}^d} f(y)e^{-ix\cdot y}\,dy$$

denote the Fourier transform of an $L_1(\mathbb{R}^d)$ function f. The Fourier transform has a natural extension from $L_1(\mathbb{R}^d)\cap L_2(\mathbb{R}^d)$ to $L_2(\mathbb{R}^d)$ and more generally to tempered distributions. We assume that the reader is familiar with the rudiments of Fourier transform theory.

The Fourier transform of $f(\,\cdot\,+t)$, $t\in\mathbb{R}^d$, is $e_t\hat{f}$; we shall use the abbreviated notation

$$e_t(x) := e^{ix\cdot t}$$

for the exponential functions. Now, suppose that the shifts of ϕ form an $L_2(\mathbb{R}^d)$-stable basis for $\mathcal{S}(\phi)$. Then from (3.1.3), each $S\in\mathcal{S}(\phi)$ can be written as $S=\sum_{j\in\mathbb{Z}^d} c(j)\phi(\,\cdot\,-j)$ with $(c(j))\in\ell_2(\mathbb{Z}^d)$. Therefore,

$$\hat{S}(y)=\sum_{j\in\mathbb{Z}^d} c(j)e_{-j}(y)\hat{\phi}(y)=\tau(y)\hat{\phi}(y),\quad \tau(y):=\sum_{j\in\mathbb{Z}^d} c(j)e_{-j}(y). \tag{3.2.1}$$

Here τ is an $L_2(\mathbb{T}^d)$ function (i.e., of period 2π in each of the variables $y_1,\dots,y_d$). The $L_2(\mathbb{R}^d)$-stability of the shifts of ϕ can easily be seen to be equivalent to the statement

$$\|\tau\|_{L_2(\mathbb{T}^d)} \approx \|S\|_{L_2(\mathbb{R}^d)}. \tag{3.2.2}$$

The characterization (3.2.1) allows one to readily decide when a function is in $\mathcal{S}(\phi)$. Even when the shifts of ϕ are not $L_2(\mathbb{R}^d)$-stable, one can characterize $\mathcal{S}(\phi)$ by (see [BDR])

$$\widehat{\mathcal{S}(\phi)} = \{\hat{S}=\tau\hat{\phi}\in L_2(\mathbb{R}^d)\mid \tau \text{ is } 2\pi\text{-periodic}\}. \tag{3.2.3}$$

By dilation, (3.2.3) gives a characterization of the scaled spaces $\mathcal{S}^k$, $\mathcal{S}=\mathcal{S}(\phi)$. For example, the functions in $\mathcal{S}^1$ are characterized by $\hat{S}=\tau\hat{\eta}\in L_2(\mathbb{R}^d)$, $\eta:=\phi(2\,\cdot)$, with τ a 4π-periodic function.

A similar characterization holds for a finite set Φ of generators for a shift-invariant space $\mathcal{S}(\Phi)$. We say that this set provides $L_2(\mathbb{R}^d)$-stable shifts if the totality of all functions $\phi(\,\cdot\,-j)$, $j\in\mathbb{Z}^d$, $\phi\in\Phi$, forms an $L_2(\mathbb{R}^d)$-stable basis for $\mathcal{S}(\Phi)$. In this case, a function $S\in\mathcal{S}(\Phi)$ is described by its Fourier transform

$$\hat{S}=\sum_{\phi\in\Phi}\tau_\phi\hat{\phi},$$

where the functions τ_ϕ, $\phi\in\Phi$, are in $L_2(\mathbb{T}^d)$ and

$$\|S\|_{L_2(\mathbb{R}^d)} \approx \sum_{\phi\in\Phi}\|\tau_\phi\|_{L_2(\mathbb{T}^d)}.$$

It is clear that the values at points congruent modulo 2π of the Fourier transform of a function S in $\mathcal{S}(\phi)$ are related. If we know $\hat{\phi}(x)$ and $\hat{S}(x)$, then, because τ has period 2π, all other values $\hat{S}(x+\alpha)$, $\alpha \in 2\pi\mathbb{Z}^d$, are determined. It is natural to try to remove this redundancy. The following *bracket product* is useful for this purpose. If f and g are in $L_2(\mathbb{R}^d)$, we define

$$[f,g] := \sum_{\beta \in 2\pi\mathbb{Z}^d} f(\cdot+\beta)\overline{g(\cdot+\beta)}. \tag{3.2.4}$$

Then $[f,g]$ is a function in $L_1(\mathbb{T}^d)$.

One particular use of the bracket product is to relate inner products on $\mathbb{R}^d$ to inner products on $\mathbb{T}^d$. For example, if $f,g \in L_2(\mathbb{R}^d)$ and $j \in \mathbb{Z}^d$, then

$$(2\pi)^d \int_{\mathbb{R}^d} f(x)\overline{g(x-j)}\,dx = \int_{\mathbb{R}^d} e_j(y)\hat{f}(y)\overline{\hat{g}(y)}\,dy = \int_{\mathbb{T}^d} e_j(\theta)[\hat{f},\hat{g}](\theta)\,d\theta. \tag{3.2.5}$$

Thus, these inner products are the Fourier coefficients of $[\hat{f},\hat{g}]$. In particular, a function f is orthogonal to all the shifts of g if and only if $[\hat{f},\hat{g}]=0$ a.e., in which case one obtains that all shifts of f are orthogonal to all the shifts of g (which also follows directly by a simple change of variables).

Another application of the bracket product is to relate integrals over $\mathbb{R}^d$ to integrals over $\mathbb{T}^d$. For example, if $\hat{S}=\tau\hat{\phi}$ with τ of period 2π, then

$$(2\pi)^d \|S\|_{L_2(\mathbb{R}^d)} = \|\tau[\hat{\phi},\hat{\phi}]^{1/2}\|_{L_2(\mathbb{T}^d)}. \tag{3.2.6}$$

Returning to $L_2(\mathbb{R}^d)$-stability for a moment, it follows from (3.2.6) and (3.2.2) that the shifts of ϕ are $L_2(\mathbb{R}^d)$-stable if and only if $C_1 \le [\hat{\phi},\hat{\phi}] \le C_2$, a.e., for constants $C_1, C_2 > 0$. Also, the shifts of ϕ are orthonormal if and only if $[\hat{\phi},\hat{\phi}]=1$ a.e. For example, if we begin with a function ϕ with $L_2(\mathbb{R}^d)$-stable shifts, then the function ϕ_* with Fourier transform

$$\hat{\phi}_* := \frac{\hat{\phi}}{[\hat{\phi},\hat{\phi}]^{1/2}} \tag{3.2.7}$$

has orthonormal shifts (this is the standard way to orthogonalize the shifts of ϕ). Incidentally, this orthogonalization procedure applies whenever $[\hat{\phi},\hat{\phi}]$ vanishes only on a set of measure zero in $\mathbb{T}^d$, in particular for any compactly supported ϕ. That is, it is not necessary to assume that ϕ has $L_2(\mathbb{R}^d)$-stable shifts in order to orthonormalize its shifts

The bracket product is useful in describing projections onto shift-invariant spaces. Let ϕ be an arbitrary $L_2(\mathbb{R}^d)$ function and let $P := P_\phi$ denote the $L_2(\mathbb{R}^d)$ projector onto the space $\mathcal{S}(\phi)$. Then for each $f \in L_2(\mathbb{R}^d)$, Pf is the best $L_2(\mathbb{R}^d)$ approximation to f from $\mathcal{S}(\phi)$. It was shown in [BDR] that

$$\widehat{Pf} = \frac{[\hat{f},\hat{\phi}]}{[\hat{\phi},\hat{\phi}]}\hat{\phi}. \tag{3.2.8}$$

Here and later, we use the convention that $0/0=0$. We note some properties of (3.2.8). First, $[\hat{f},\hat{\phi}]$ is 2π-periodic and therefore the form of $\widehat{Pf}$ matches that required by (3.2.3). If ϕ has orthonormal shifts, then $[\hat\phi, \hat\phi] = 1$ a.e., and in view of (3.2.5), the formula (3.2.8) is the usual one for the $L_2(\mathbb{R}^d)$ projector. If $\hat\phi / [\hat\phi, \hat\phi]$ is the Fourier transform of an $L_2(\mathbb{R}^d)$ function γ (this holds, for example, if ϕ has $L_2(\mathbb{R}^d)$-stable shifts), then

$$Pf = \sum_{j\in\mathbb{Z}^d} \gamma_j(f)\phi(\,\cdot - j), \quad \gamma_j(f) := \int_{\mathbb{R}^d} f(x)\overline{\gamma(x-j)}\,dx, \tag{3.2.9}$$

as follows from (3.2.5). Whenever ϕ has compact support and $L_2(\mathbb{R}^d)$-stable shifts, the function γ decays exponentially.

The bracket product is also useful in describing properties of shift-invariant spaces $\mathcal{S}(\Phi)$ that are generated by a finite set Φ of functions from $L_2(\mathbb{R}^d)$. The properties of the generating set Φ are contained in its Gramian

$$G(\Phi) := \Big([\hat\phi, \hat\psi]\Big)_{\phi,\psi\in\Phi}. \tag{3.2.10}$$

This is a matrix of 2π-periodic functions from $L_1(\mathbb{T}^d)$. For example, the shifts of the functions in Φ form an orthonormal basis for $\mathcal{S}(\Phi)$ if and only if $G(\Phi)$ is the identity matrix a.e. on $\mathbb{T}^d$. The generating set Φ provides an $L_2(\mathbb{R}^d)$-stable basis for $\mathcal{S}(\Phi)$ if and only if $G(\Phi)$ and $G(\Phi)^{-1}$ exist and are a.e. bounded on $\mathbb{T}^d$ with respect to some (and then every) matrix norm. For proofs, see [BDR2].

3.3. The conditions of multiresolution. The question arises as to when the conditions (3.1.2) of multiresolution are satisfied for a function $\phi \in L_2(\mathbb{R}^d)$. We mention, without proof, two sufficient conditions on ϕ for (3.1.2)(ii) and (iii) to hold. Jia and Micchelli [JM] have shown that if the shifts of ϕ are $L_2(\mathbb{R}^d)$-stable, if ϕ satisfies the refinement equation (3.1.4) with coefficients $(a(j))$ in $\ell_1(\mathbb{Z}^d)$, and if $\sum_{j\in\mathbb{Z}^d} |\phi(x+j)|$ is in $L_2(\mathbb{T}^d)$, then (3.1.2)(ii) and (iii) are satisfied.

On the other hand, in [BDR1] it is shown that (3.1.2)(ii) and (iii) are satisfied whenever $\phi \in L_2(\mathbb{R}^d)$ satisfies the refinement condition (3.1.2)(i) and, in addition, $\operatorname{supp}[\hat\phi, \hat\phi] = \mathbb{T}^d$; by this we mean that $[\hat\phi, \hat\phi]$ vanishes only on a set of measure zero. In particular, these conditions are satisfied whenever ϕ has compact support and satisfies the refinement condition. Since these conditions are satisfied for all functions ϕ that we shall encounter (in fact for all functions ϕ that have been considered in wavelet construction by multiresolution), it is not necessary to verify separately (3.1.2)(ii) and (iii)—they automatically hold. We also note that in the construction of wavelets and prewavelets in [BDR1] it is not necessary to assume that ϕ has $L_2(\mathbb{R}^d)$-stable shifts.

We now discuss the refinement condition (3.1.2)(i). In view of the characterization (3.2.3) of shift-invariant spaces, this condition is equivalent to

$$\hat\phi = A\hat\eta, \quad \eta := \phi(2\,\cdot\,) \tag{3.3.1}$$

for some 4π-periodic function A. If ϕ has $L_2(\mathbb{R}^d)$-stable shifts then this condition becomes the refinement equation (3.1.4) and $A(y) = \sum_{j\in\mathbb{Z}^d} a(j)e_{-j/2}$ in the sense of $L_2(2\mathbb{T}^d)$ convergence.

It was shown in [BDR1] that one can construct generators for the wavelet space W even when (3.1.2)(iv) does not hold. For example, this condition can be replaced by the assumption that $\operatorname{supp}\hat{\phi} = \mathbb{R}^d$. We also note that if $[\hat{\phi}, \hat{\phi}]$ is nonzero a.e., then we can always find a generator ϕ_* for $\mathcal{S}$ with orthonormal shifts, so the condition (3.1.2)(ii) is satisfied for this generator (and the other conditions of multiresolution remain the same). However, the generator ϕ_* does not satisfy the same refinement equation as ϕ (for example, the refinement equation for ϕ_* may be an infinite sum even if the equation for ϕ is a finite sum) and ϕ_* may not have compact support even if ϕ has compact support, so the construction that gives ϕ_* is not completely satisfactory. Furthermore, we would like to describe the wavelets and prewavelets directly in terms of the original ϕ. This is especially the case when ϕ does not have $L_2(\mathbb{R}^d)$-stable shifts, since then we can say nothing about the decay of ϕ_* even when ϕ has compact support. In the remainder of this presentation, we shall assume that ϕ has $L_2(\mathbb{R}^d)$-stable shifts.

3.4. Constructions of univariate wavelets. In this section we restrict our attention to wavelets in one variable, because multiresolution is simpler and better understood for a single variable than for several variables. We suppose that ϕ satisfies the assumptions (3.1.2) of multiresolution and follow the ideas presented in [BDR1].

Fundamentally, the approach of [BDR1] is quite simple. We take a function $\eta \in \mathcal{S}^1$ and consider its error $w := \eta - P\eta$ of best $L_2(\mathbb{R})$ approximation by the elements of $\mathcal{S}^0 = \mathcal{S} = \mathcal{S}(\phi)$. Here P is the $L_2(\mathbb{R})$ projector onto $\mathcal{S}(\phi)$ given by (3.2.8). Clearly $w \in W$ and we shall show that with any reasonable choice for η, the function w is a generator of W, i.e., $W = \mathcal{S}(w)$. Thus, because of the characterization (3.2.3) of principal shift-invariant spaces, we can obtain all other generators for w by operations on the Fourier transform side. Here are the details.

We take $\eta := \phi(2\,\cdot)$, which is clearly in $\mathcal{S}^1$. Then, $w := \eta - P\eta$ is in W and by virtue of (3.2.8) has Fourier transform

$$\hat{w} = \hat{\eta} - \frac{[\hat{\eta}, \hat{\phi}]}{[\hat{\phi}, \hat{\phi}]}\hat{\phi}. \tag{3.4.1}$$

It is convenient to introduce (for a function $f \in L_2(\mathbb{R}^d)$) the abbreviated notation

$$\tilde{f} := [\hat{f}, \hat{f}]^{1/2}, \tag{3.4.2}$$

since this expression occurs often in wavelet constructions. Another description of $\tilde{f}$ is

$$\tilde{f} = \Big(\sum_{j \in 2\pi\mathbb{Z}^d} |\hat{f}(\cdot + j)|^2\Big)^{1/2}.$$

We see that $\tilde{f}$ is a 2π-periodic function, and if f has compact support then $\tilde{f}^2$ is a trigonometric polynomial, because of (3.2.5). The analogue of this function for half-shifts is

$$\tilde{\tilde{f}} := \Big(\sum_{j \in 4\pi\mathbb{Z}^d} |\hat{f}(\cdot + j)|^2\Big)^{1/2}, \tag{3.4.3}$$

which is now a 4π-periodic function. In particular f has orthonormal half-shifts if and only if $\tilde{\tilde{f}} = 2^{-d/2}$ a.e., and $L_2(\mathbb{R}^d)$-stable half-shifts if and only if $C_1 \le \tilde{\tilde{f}} \le C_2$, a.e., for constants $C_1, C_2 > 0$.

We return now to the construction of wavelets. We can multiply $\hat{w}$ by any 2π-periodic function, and as long as the resulting function is in $L_2(\mathbb{R})$, it will be the Fourier transform of a function in W. We multiply (3.4.1) by $\tilde{\phi}^2$, which clears the denominator. The result is the function w_0 with Fourier transform

$$\hat{w}_0 := \tilde{\phi}^2\hat{\eta} - [\hat{\eta}, \hat{\phi}]\hat{\phi}. \tag{3.4.4}$$

We note that w_0 has compact support whenever ϕ does.

We can calculate the bracket products appearing in (3.4.4) by using the refinement relation $\hat{\phi} = A\hat{\eta}$ (see (3.3.1)) with A a 4π-periodic function. For example, to calculate $\tilde{\phi}$,

$$\begin{aligned}
\tilde{\phi}^2 &= \sum_{j\in 2\pi\mathbb{Z}} |\hat{\phi}(\,\cdot\,+j)|^2 \\
&= \sum_{j\in 4\pi\mathbb{Z}} (|A(\,\cdot\,+j)|^2|\hat{\eta}(\,\cdot\,+j)|^2 + |A(\,\cdot\,+j+2\pi)|^2|\hat{\eta}(\,\cdot\,+j+2\pi)|^2) \\
&= |A|^2\tilde{\tilde{\eta}}^2 + |A(\,\cdot\,+2\pi)|^2\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi).
\end{aligned}$$

Similarly, $[\hat{\eta}, \hat{\phi}] = \overline{A}\tilde{\tilde{\eta}}^2 + \overline{A(\,\cdot\,+2\pi)}\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi)$. Therefore,

$$\begin{aligned}
\hat{w}_0 &= \Big\{|A|^2\tilde{\tilde{\eta}}^2 + |A(\,\cdot\,+2\pi)|^2\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi) - \overline{A}A\tilde{\tilde{\eta}}^2 - \overline{A(\,\cdot\,+2\pi)}A\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi)\Big\}\,\hat{\eta} \\
&= \{A(\,\cdot\,+2\pi) - A\}\,\overline{A(\,\cdot\,+2\pi)}\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi)\hat{\eta}.
\end{aligned}$$

We can make one more simplification in the last representation for $\hat{w}_0$. The function $\frac{1}{2}e_{1/2}\{A - A(\,\cdot\,+2\pi)\}$ is 2π-periodic. Therefore, dividing by this function, we obtain the function

$$\hat{\psi} := 2e_{-1/2}\overline{A(\,\cdot\,+2\pi)}\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi)\hat{\eta}. \tag{3.4.5}$$

It is easy to see (and is shown in (3.4.14)) that ψ is in $L_2(\mathbb{R})$. It follows, therefore, that ψ is in W and $\mathcal{S}(\psi) \subset W$. The following argument shows that we really have $\mathcal{S}(\psi) = W$.

If we replace η by $\eta_1 := \eta(\,\cdot\,-1/2)$ (which is also in $\mathcal{S}^1$) and repeat the above construction, in place of w_0 we obtain the function w_1 whose Fourier transform is

$$\hat{w}_1 = e_{-1/2}\,\{A(\,\cdot\,+2\pi) + A\}\,\overline{A(\,\cdot\,+2\pi)}\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi)\hat{\eta}.$$

Hence, dividing by $A(\,\cdot\,+2\pi) + A$ (which is 2π-periodic), we arrive at the same function ψ. The importance of this fact is that we can reverse these two processes. In other words, we can multiply $\hat{\psi}$ by a 2π-periodic function and obtain $\hat{\eta} - \widehat{P\eta}$ (respectively $\hat{\eta}_1 - \widehat{P\eta_1}$). Hence, both of these functions are in $\mathcal{S}(\psi)$. Since $P\eta$ is in $\mathcal{S}(\phi)$, $\eta = P\eta + (\eta - P\eta)$ is in $\mathcal{S}(\phi) + \mathcal{S}(\psi)$. Similarly, η_1 is in this space. Since the full shifts of η and η_1 generate $\mathcal{S}^1(\phi)$, we must have $W = \mathcal{S}(\psi)$. This confirms our earlier remark that W is a principal shift-invariant space. Since we can obtain ψ from w and w_0 by multiplying by 2π-periodic functions, both w and w_0 are also generators of W.

We consider some examples that show that ψ is the (pre)wavelet constructed by various authors.

Orthogonal wavelets. To obtain orthogonal wavelets, Mallat [Ma] begins with a function ϕ that satisfies the assumptions (3.1.2) of multiresolution and whose shifts are orthonormal. This is equivalent to $\tilde{\phi} = 1$ a.e., and (by a change of variables) to the half-shifts of $\sqrt{2}\,\eta$ being orthonormal, i.e., to $\tilde{\tilde{\eta}} = 1/2$ a.e. When this is used in (3.4.5), we obtain

$$\hat{\psi} = e_{-1/2}\overline{A(\,\cdot\, + 2\pi)}\hat{\eta}, \tag{3.4.6}$$

which is the orthogonal wavelet of Mallat. To see that the shifts of ψ are orthonormal, one simply computes

$$\tilde{\psi}^2 = |A(\,\cdot\, + 2\pi)|^2\tilde{\tilde{\eta}}^2 + |A|^2\tilde{\tilde{\eta}}^2(\,\cdot\, + 2\pi) = \frac{1}{4}\{|A(\,\cdot\, + 2\pi)|^2 + |A|^2\} = 1, \tag{3.4.7}$$

where the last equality follows from the identity

$$\begin{aligned} 1 = \tilde{\phi}^2 = \tilde{\tilde{\phi}}^2 + \tilde{\tilde{\phi}}^2(\,\cdot\, + 2\pi) &= |A|^2\tilde{\tilde{\eta}}^2 + |A(\,\cdot\, + 2\pi)|^2\tilde{\tilde{\eta}}^2(\,\cdot\, + 2\pi) \\ &= \frac{1}{4}\{|A(\,\cdot\, + 2\pi)|^2 + |A|^2\}. \end{aligned} \tag{3.4.8}$$

The Fourier transform identity (3.4.6) is equivalent to the identity (3.1.8).

We note that from the orthogonal wavelet ψ of (3.1.8) (respectively (3.4.6)), we obtain all other orthogonal wavelets in W by multiplying $\hat{\psi}$ by a 2π-periodic function τ of unit modulus. Indeed, we know that any element $w \in W$ satisfies $\hat{w} = \tau\hat{\psi}$ with $\tau \in L_2(\mathbb{T})$. To have $[\hat{w}, \hat{w}] = 1$ a.e., the function τ must satisfy $|\tau(y)| = 1$ a.e. in $\mathbb{T}$.

As an example, we consider the cardinal B-spline N_r of order r. To obtain orthogonal wavelets by Mallat's construction, one need only manipulate various Laurent series. First, one orthogonalizes the shifts of N_r. This gives the spline $\phi = \mathcal{N}_r$ whose Fourier transform is

$$\hat{\phi} = \hat{\mathcal{N}}_r := \frac{\hat{N}_r}{\tilde{N}_r}. \tag{3.4.9}$$

It is easy to compute the coefficients in the expansion

$$\tilde{N}_r^2 = \sum_{j\in\mathbb{Z}} \alpha(j)e_{-j}. \tag{3.4.10}$$

In fact, we know from (3.2.5) that this is a trigonometric polynomial whose coefficients are

$$\begin{aligned} \alpha(j) = \int_{\mathbb{R}} N_r(x-j)N_r(x)\,dx &= \int_{\mathbb{R}} N_r(r+j-x)N_r(x)\,dx \\ = [N_r * N_r](j+r) &= N_{2r}(j+r), \quad j \in \mathbb{Z}, \end{aligned}$$

because N_r is symmetric about its midpoint.

The polynomial $\rho_{2r}(z) := z^r \sum_{j\in\mathbb{Z}} \alpha(j)z^{-j}$ is the Euler-Frobenius polynomial of order $2r$, which plays a prominent role in cardinal spline interpolation (see [Sch1]). It is well known that ρ_{2r} has no zeros on $|z| = 1$. Hence, the reciprocal $1/\rho_{2r}$ is analytic in a nontrivial annulus that contains the unit circle in its interior. One can easily find the coefficients of reciprocals and square roots of Laurent series inductively. By finding the coefficients of $\rho_{2r}^{-1/2}$, we obtain the coefficients $\beta(j)$ appearing in the expansion

$$\phi(x) = \mathcal{N}_r(x) = \sum_{j\in\mathbb{Z}} \beta(j) N_r(x-j). \tag{3.4.11}$$

Because ρ_{2r} has no zeros on $|z| = 1$, we conclude that the coefficients $\beta(j)$ decrease exponentially. The spline $\mathcal{N}_r$ together with its shifts form an orthonormal basis for the cardinal spline space $\mathcal{S}(N_r)$. They are sometimes referred to as the Franklin basis for $\mathcal{S}(N_r)$.

Now that we have the spline $\phi := \mathcal{N}_r$ in hand, we can obtain the spline wavelet $\psi = \mathcal{N}_r^*$ of Battle-Lemarié [Ba] from formula (3.1.8). For this, we need to find the refinement equation for ϕ. We begin with the refinement equation (3.1.5) for the B-spline N_r, which we write in terms of Fourier transforms as $\hat{N}_r = A_0\hat{\eta}_0$ with $\eta_0 := N_r(2\,\cdot\,)$ and $A_0 = 2^{-r+1}\sum_{j=0}^{r}\binom{r}{j}e_{-j/2}$ a 4π-periodic trigonometric polynomial. It follows that

$$\hat{\phi} = A\hat{\eta}, \quad \eta := \mathcal{N}_r(2\,\cdot\,), \quad A(y) = \frac{\hat{\phi}(y)}{\frac{1}{2}\hat{\phi}(y/2)} = \tilde{N}_r(y/2)\tilde{N}_r^{-1}(y)A_0(y). \tag{3.4.12}$$

In terms of the B-spline N_r, this gives

$$\hat{\psi}(y) = e_{-1/2}(y)\overline{A(y+2\pi)}\hat{\eta}(y) = \frac{1}{2}e_{-1/2}(y)\overline{A(y+2\pi)}\tilde{N}_r^{-1}(y/2)\hat{N}_r(y/2). \tag{3.4.13}$$

In other words, to find the orthogonal spline wavelet ψ of (3.4.13), we need to multiply out the various Laurent expansions making up $\overline{A(\,\cdot\,+2\pi)}\tilde{N}_r^{-1}(\,\cdot\,/2)$. This gives the coefficients $\gamma(j)$, $j \in \mathbb{Z}$, in the representation

$$\psi(x) = \sum_{j\in\mathbb{Z}} \gamma(j) N_r(2x-j).$$

We emphasize that each of the Laurent series converges in an annulus containing the unit circle. This means that the coefficients $\gamma(j)$ converge exponentially to zero when $j \to \pm\infty$.

Prewavelets. For the construction of prewavelets, we do not assume that the shifts of ϕ are orthonormal, but only that they are $L_2(\mathbb{R})$-stable, i.e., we assume (3.1.2)(iv). Then, it is easy to see that the function ψ defined by (3.4.5) is a prewavelet. Indeed, we already know that ψ is a generator for W and it is enough to check that it has $L_2(\mathbb{R})$-stable shifts. For this, we compute

$$\frac{\tilde{\psi}^2}{4} = |A(\,\cdot\,+2\pi)|^2\tilde{\tilde{\eta}}^4(\,\cdot\,+2\pi)\tilde{\tilde{\eta}}^2 + |A|^2\tilde{\tilde{\eta}}^4\tilde{\tilde{\eta}}^2(\,\cdot\,+2\pi). \tag{3.4.14}$$

Since the shifts of ϕ are $L_2(\mathbb{R})$-stable, so are the half-shifts of η. This means that $C_1 \le \tilde{\tilde{\eta}} \le C_2$ for constants $C_1, C_2 > 0$. Moreover, the formula

$$\tilde{\phi}^2 = |A|^2 \tilde{\tilde{\eta}}^2 + |A(\,\cdot\, + 2\pi)|^2 \tilde{\tilde{\eta}}^2(\,\cdot\, + 2\pi)$$

shows that $C_1 \le |A|^2 + |A(\,\cdot\, + 2\pi)|^2 \le C_2$, again for positive constants C_1, C_2. Combining this information with (3.4.14) shows that $\tilde{\psi}$ is bounded above and below by positive constants, so that ψ has $L_2(\mathbb{R})$-stable shifts. This also shows that ψ is in $L_2(\mathbb{R})$. The prewavelet ψ was introduced by Chui and Wang [CW] and independently by Micchelli [M].

We can also find a direct representation for ψ in terms of the shifts of $\phi(2\,\cdot\,)$. For this we need the Fourier coefficients $\mu(j)$ (of $e_{-j/2}$) for the 4π-periodic function $2\overline{A}\tilde{\tilde{\eta}}$:

$$\mu(j) := \frac{1}{4\pi} \int_{-2\pi}^{2\pi} 2\overline{A}\tilde{\tilde{\eta}}^2 e_{j/2} = \frac{1}{4\pi} \int_{\mathbb{R}} 2\overline{\hat{\phi}}\hat{\eta} e_{j/2} = \int_{\mathbb{R}} \overline{\phi}\eta(\,\cdot + j/2) = \int_{\mathbb{R}} \overline{\phi(x)}\phi(2x+j)\,dx.$$

Using this in (3.4.5), we find that

$$\psi = \sum_{j\in\mathbb{Z}} (-1)^{j+1}\mu(j-1)\phi(2\,\cdot - j), \quad \mu(j) := \int_{\mathbb{R}} \overline{\phi(x)}\phi(2x+j)\,dx. \tag{3.4.15}$$

If ϕ has compact support, then clearly ψ also has compact support. Chui and Wang [CW1] posed the interesting question whether ψ has the smallest support among all the elements in W, to which they gave the following answer. We assume that A is a polynomial, i.e., that ϕ satisfies a finite refinement equation. Next, we note that because $W \subset \mathcal{S}^1$, any $w \in W$ is represented as

$$\hat{w}(y) = e_{-1/2}(y)\overline{B(y + 2\pi)}\hat{\eta}(y) \tag{3.4.16}$$

with B of period 4π. If $B = \sum_{j=m}^{M} b(j)e_{-j/2}$ is a Laurent polynomial with $b(m)b(M) \ne 0$, then w has compact support, and we define the length of B to be $M - m$. We know that there are nonzero polynomials B that satisfy (3.4.16) for some w because $\tilde{\tilde{\eta}}^2$ is a polynomial (since η has compact support) and (3.4.5) implies that for $B_0 := A\tilde{\tilde{\eta}}^2$, w is the prewavelet $\psi \in W$.

B_0 may not have minimal length among all such polynomials, however, because it may be possible to cancel certain symmetric factors from B_0. To see this, we write $B_0(y) = e_M(y/2)P(e^{-iy/2})$ with P an algebraic polynomial, and we let $Q(z^2) := \prod_\lambda (z - \lambda)$, with the product taken over all λ with λ and $-\lambda$ both zeros of P. Then, the factorization $P(z) = Q(z^2)P_*(z)$ gives the factorization $B_0(y) = \tau(y)B_*(y)$ with τ a trigonometric polynomial of period 2π that does not vanish. Therefore, the function ψ_* with Fourier transform

$$\hat{\psi}_*(y) = e_{-1/2}(y)\overline{B_*(y + 2\pi)}\hat{\eta}(y), \quad B_*(y) := \tau^{-1}(y)B_0(y), \tag{3.4.17}$$

is in W and has smaller length than B_0. A simple argument (which we do not give) shows that B_* has smallest length. For most prewavelets of interest, $B_* = B_0$.

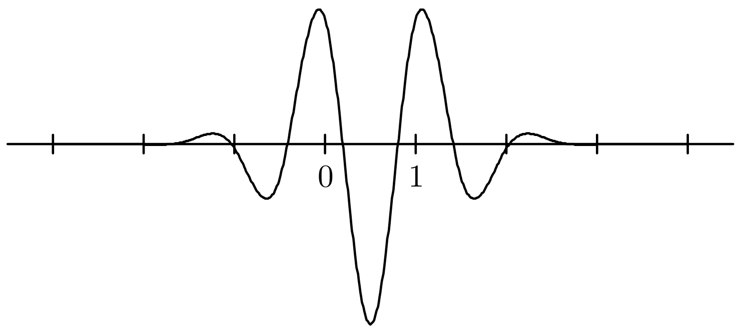

FIGURE 2. The Chui-Wang spline prewavelet for $r = 4$, which has support $[-3, 4]$. The vertical scale is stretched by a factor of eight.

The problem of finding a wavelet w in the form (3.4.16) with B a polynomial of minimal length, which is solved by $w = \psi_*$, is not always equivalent to finding the wavelet with minimal support; here the word "support" means the interval of smallest length outside of which w vanishes identically. In general, there may be wavelets w of compact support that can be represented by (3.4.16) with B not a polynomial. However, Ben-Artzi and Ron [B-AR] show that this is impossible whenever the following property holds: The linear combination $\sum_{j\in\mathbb{Z}} \gamma(j)\phi(\,\cdot\, - j)$ (which converges pointwise, since ϕ has compact support) is identically zero if and only if all the coefficients $\gamma(j)$ are 0. Under these assumptions, the wavelet ψ_* has minimal support (see [BDR1] for details).

For a prewavelet ψ, we have the wavelet decomposition

$$f = \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}^d} c_{j,k}(f)\psi_{j,k}, \quad c_{j,k}(f) := \int_{\mathbb{R}} f\overline{\gamma_{j,k}} \tag{3.4.18}$$

where γ has Fourier transform $\hat\gamma = \hat\psi/[\hat\psi, \hat\psi]$. This follows from the representation (3.2.9) for the projector P from $L_2(\mathbb{R})$ onto W. It is useful to note that when ψ has compact support, the function γ will generally not have compact support because of the division by the bracket product $[\hat\psi, \hat\psi]$. Thus, there is in some sense a trade-off between the simplicity of the prewavelet and the complexity of the coefficient functional.

We consider the following important example. Let $\phi := N_r$ be the cardinal B-spline of order r, which is known to have linearly independent shifts. Then, the function ψ in (3.4.15) is a spline function with compact support. It is easy to see that $A\tilde{\tilde\eta}^2$ has no symmetric zeros so that ψ has minimal support. We note also that it is shown in [BDR1] that the shifts of ψ are themselves linearly independent. From formula (3.4.15), we see that ψ is supported on $[1 - r, r]$. Up to a shift, the spline ψ is the minimally supported spline prewavelet of Chui and Wang [CW]; see Figure 2.

3.5. Daubechies' compactly supported wavelets. The orthogonal spline wavelets of §3.4, which decay exponentially at infinity, can be chosen to have any specified finite order of smoothness. It is natural to ask whether orthogonal wavelets can be constructed that have both any specified finite order of smoothness and compact support. A celebrated construction of Daubechies [Da] leads to such wavelets, which are frequently used in numerical applications. Space prohibits us from giving all the details of Daubechies' construction, but the following discussion will outline the basic ideas.

To construct a compactly supported wavelet with a prescribed smoothness order r and compact support, one finds a special finite sequence $(a(j))$ such that the refinement equation (3.1.4) has a solution $\phi \in C^r$ with orthogonal shifts. The orthogonal wavelet ψ of (3.4.5) will then obviously have compact support and the same smoothness. Before we begin, it is necessary to understand which properties of the sequence $(a(j))$ guarantee the existence of a function ϕ with the desired properties, i.e., we need to understand the nature of solutions to the refinement equation (3.1.4). This has been studied in another context, namely, in subdivision algorithms for computer aided geometric design (see, for example, the paper of Cavaretta, Dahmen, and Micchelli [CDM] for a discussion of subdivision). As was pointed out by Dahmen and Micchelli [DM], it is possible to derive part of Daubechies' construction from the subdivision approach. However, we shall describe Daubechies' original construction.

Let r be a nonnegative integer that corresponds to the desired order of smoothness, and let $(a(j))$ with $a(j) = 0$, $|j| > m$, and $a(m) \neq 0$, be the sequence of the refinement equation (3.1.4) for the function ϕ we want to construct. The sequence $(a(j))$ and the Fourier transform of ϕ are related by

$$\hat{\phi}(y) = \mathcal{A}(y/2)\hat{\phi}(y/2), \quad \mathcal{A}(y) := \frac{1}{2}\sum_{j=-m}^{m} a(j)e^{-ijy}. \tag{3.5.1}$$

Here we use a slightly different normalization for the refinement function ($\mathcal{A}(y) = \frac{1}{2}A(2y)$). If $\hat{\phi}$ is continuous at 0 and $\hat{\phi}(0) = 1$, we can, at least in a formal sense, write

$$\hat{\phi}(y) = \lim_{k\to\infty} A_k(y) \tag{3.5.2}$$

where

$$A_k(y) := \prod_{j=1}^{k} \mathcal{A}(y/2^j). \tag{3.5.3}$$

We note that $A_k^*(y) := A_k(2^k y)$ is a trigonometric polynomial of degree $(2^k - 1)m$. The key to Daubechies' construction is to impose conditions on $\mathcal{A}$ (which are therefore conditions on the sequence $(a(j))$) that not only make (3.5.2) rigorous but also guarantee that the function ϕ defined by (3.5.2) has the desired smoothness and has orthonormal shifts.

We first note that if the shifts of ϕ are orthonormal then, as was shown in (3.4.8),

$$|\mathcal{A}(y)|^2 + |\mathcal{A}(y+\pi)|^2 = 1, \qquad y \in \mathbb{T}. \tag{3.5.4}$$

The converse to this is almost true. Namely, Daubechies' construction shows that (3.5.4) together with some mild assumptions (related to the convergence in (3.5.3)) imply the orthonormality of the shifts of ϕ. For this, the following identities, which follow from (3.5.4) by induction, are useful:

$$\sum_{j=0}^{2^k-1} |A_k^*(y+j2^{-k}2\pi)|^2 = 1, \quad k = 1, 2 \dots . \tag{3.5.5}$$

We next want to see what properties of $\mathcal{A}$ guarantee smoothness for ϕ. The starting point is the following observation. If $\int_{\mathbb{R}} \phi(x)\,dx \neq 0$, then integrating the refinement equation (3.1.4) gives $\sum_{j\in\mathbb{Z}} a(j) = 2$. Hence, $\mathcal{A}(0) = 1$ and $\mathcal{A}(\pi) = 0$. We can therefore write

$$\mathcal{A}(y) = (1+e^{iy})^N \alpha(y), \quad \|\alpha\|_{L_\infty(\mathbb{T})} = 2^{-\theta},\ \alpha(0) = 2^{-N}, \tag{3.5.6}$$

for a suitable integer $N > 0$, a real number θ, and a function α.

By carefully estimating the partial products A_k, it can be shown that whenever $\mathcal{A}$ satisfies (3.5.6) for some $\theta > 1/2$, the product (3.5.2) converges to a function in $L_2(\mathbb{R})$ that decays like $|x|^{-\theta}$ as $|x| \to \infty$. The limit function is the Fourier transform of the solution ϕ to the refinement equation (3.5.1). We see that the larger we can make θ in (3.5.6), the smoother ϕ is. For example, if $\theta > r+1$, then ϕ is in C^r.

What is the role of the integer N in (3.5.6)? Practically, one must increase N to find a function $\alpha(y)$ that satisfies (3.5.6) for large θ. In addition, the local approximation properties of the spaces $\mathcal{S}^k(\phi)$ are determined by N; see §5.

Once it is shown that there is a function ϕ that satisfies the refinement equation for the given sequence $(a(j))$, it remains to show that ϕ has compact support and orthonormal shifts. Here the arguments have the same character as those used to analyze subdivision algorithms for the graphical display of curves and surfaces. Assume that $\mathcal{A}$ satisfies (3.5.4) and (3.5.6) for some $\theta > 1/2$ and let χ denote the characteristic function of $[-1/2, 1/2]$. Then $\hat{\chi}(y) = (\sin y/2)/(y/2)$. We define ϕ_k to be the function whose Fourier transform is $\hat{\phi}_k(y) := A_k(y)\hat{\chi}(2^{-k}y)$. It can then be shown that

$$\int_{\mathbb{R}} |\phi(y) - \phi_k(y)|^2\,dy \to 0, \quad k \to \infty. \tag{3.5.7}$$

If $A_k^* = \sum_j a^*(j,k)e_{-j}$, then $\sum_j a^*(j,k)\hat{\chi}(x-j)$ has Fourier transform $A_k^*(y)\hat{\chi}(y)$ and $\hat{\phi}_k(y) = A_k^*(2^{-k}y)\hat{\chi}(2^{-k}y)$. Therefore,

$$\phi_k(x) = \sum_{j\in\mathbb{Z}} a^*(j,k) 2^k \chi(2^k x - j). \tag{3.5.8}$$

Since the coefficients $a^*(j,k)$ of A_k^* are 0 for $|j| > (2^k-1)m$, we obtain that ϕ_k is supported in $[-m, m]$. Letting $k \to \infty$, we obtain from (3.5.7) that ϕ is also supported on $[-m, m]$. From (3.5.8), (3.5.5), and the orthonormality of the functions $2^{k/2}\chi(2^k\cdot - j)$, $j \in \mathbb{Z}$, we have

$$\int_{\mathbb{R}} \phi_k(x)\phi_k(x-\ell)\,dx = 2^k \sum_{\mu-\nu=2^k\ell} a^*(\mu,k)a^*(\nu,k) = \delta(\ell), \quad \ell \in \mathbb{Z}.$$

Here the last equality follows by expanding the identity (3.5.5). Letting $k \to \infty$, we obtain that $\{\phi(\cdot - j)\}_{j\in\mathbb{Z}}$ is an orthonormal system.

The above outline shows that a C^r, compactly supported function ϕ with orthonormal shifts exists if (3.5.4) and (3.5.6) hold for a sequence $(a(j))$ and two numbers N and $\theta > r+1$. The following arguments show that such sequences exist.

We look for an $\mathcal{A}$ of the form (3.5.6) with α a trigonometric polynomial with real coefficients. Then, $|\alpha(y)|^2 = \alpha(y)\overline{\alpha(y)} = \alpha(y)\alpha(-y)$ is an even trigonometric polynomial, and

$$|\alpha(y)|^2 = T(\cos y) = T(1 - 2\sin^2 y/2) = R(\sin^2 y/2)$$

with R an algebraic polynomial. The identity (3.5.4) now becomes

$$(\cos^2 y/2)^N R(\sin^2 y/2) + (\sin^2 y/2)^N R(\cos^2 y/2) = 2^{-2N}.$$

With $t := \sin^2 y/2$, we have

$$(1-t)^N R(t) + t^N R(1-t) = 2^{-2N}. \tag{3.5.9}$$

Therefore, to find $\mathcal{A}$, we must find an algebraic polynomial R that satisfies (3.5.9). It is easy to see that the degree of R must be at least $N-1$. We can find R of this degree by writing R in the Bernstein form

$$R(t) = \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} t^k (1-t)^{N-k-1}.$$

Then, (3.5.9) becomes

(3.5.10)
$$(1-t)^N \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} t^k (1-t)^{N-k-1} + t^N \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} (1-t)^k t^{N-k-1}$$
$$= 2^{-2N} = 2^{-2N} \sum_{k=0}^{2N-1} \binom{2N-1}{k} t^k (1-t)^{2N-1-k}.$$

We see that

$$\lambda_k := 2^{-2N} \frac{\binom{2N-1}{k}}{\binom{N-1}{k}}, \qquad k = 0, 1, \dots, N-1,$$

satisfies (3.5.10), and we denote the polynomial with these coefficients by R_N.

It is important to observe that $R_N(t)$ is nonnegative for $0 \le t \le 1$, because we wish to show that there is a trigonometric polynomials $\alpha(y)$ such that $R_N(\sin^2 y/2) = |\alpha(y)|^2$, i.e., we somehow have to take a "square root" of R_N. For this, we use the classical theorem of Fejer-Riesz (see for example Karlin and Studden [KS, pg. 185]) that says that if R is nonnegative on $[0,1]$, then $R(\sin^2 y/2) = |\alpha(y)|^2$

for some trigonometric polynomial α with real coefficients and of the same degree as R. We let α_N be the trigonometric polynomial corresponding to R_N.

We now set $\mathcal{A}_N(y) := (1+e^{iy})^N\alpha_N(y)$ and note that $\mathcal{A}_N$ satisfies (3.5.4) because R_N satisfies (3.5.9). Therefore, the function ϕ defined via the limit process (3.5.2) has compact support and orthonormal shifts. The corresponding orthogonal wavelet $\psi =: \mathcal{D}_{2N}$ defined by (3.1.8) for the refinement coefficients $(a(j))$ is an orthogonal wavelet with compact support. It is easy to show that $\mathcal{D}_{2N}$ is supported in $[-(N-1), N]$.

The question now is what is the smoothness of $\mathcal{D}_{2N}$. Here the matter can become somewhat technical (see Daubechies [Da] and Meyer [Me1]). However, the following "poor man's" argument based on Stirling's formula at least shows that given any integer r, if we choose N sufficiently large, the orthogonal wavelet $\mathcal{D}_{2N}$ will have smoothness C^r.

Because the Bernstein coefficients of R_N are monotone, it follows that R_N is increasing on $[0,1]$. Therefore, $\max_{0\le t\le 1} R_N(t) = R_N(1) = \lambda_{N-1} = 2^{-2N}\binom{2N-1}{N}$. Therefore, $\|\alpha_N\|^2_{L_\infty(\mathbb{T})}$ is bounded by

$$2^{-2N}\binom{2N-1}{N} \le \frac{2^{-2N}\sqrt{2\pi(2N-1)}(2N-1)^{2N-1}e^{-(2N-1)}}{\sqrt{2\pi N}N^N e^{-N}\sqrt{2\pi(N-1)}(N-1)^{N-1}e^{-(N-1)}} \le C_0 N^{-1/2},$$

by Stirling's formula. We see that given any value of $\theta > 0$, we can choose N large enough so that (3.5.6) is satisfied for that θ, and the function $\hat\phi$ satisfies $|\hat\phi(x)| \le C(1+|x|)^{-\theta}$. Hence, for any $r < \theta - 1$, ϕ, and hence $\mathcal{D}_{2N}$, is in C^r.

For $N = 1$, the Daubechies construction gives $\phi = \chi_{[0,1]}$ and $\mathcal{D}_2$ is the Haar function. For $N = 2$, the polynomial $R_2(t) = (1+2t)/16$ and

$$\mathcal{A}_2(y) = (1+e^{iy})^2\left(\frac{\sqrt{3}+1}{8} - \frac{\sqrt{3}-1}{8}e^{-iy}\right) = (1+e^{iy})^2\alpha_2(y).$$

Then $\alpha_2(y)$ satisfies $|\alpha_2(y)| \le \sqrt{3}/4 < 2^{-1}$. Therefore, the function ϕ and the wavelet $\mathcal{D}_4 := \psi$ corresponding to this choice is continuous. (See Figure 3 for a graph of ϕ and ψ.) A finer argument shows that $\mathcal{D}_4$ is in $\mathrm{Lip}(.55, L_\infty(\mathbb{R}))$. The reader can consult Daubechies [Da] for a table of the refinement coefficients of $\mathcal{D}_{2N}$ for other values of N and a more precise discussion of the smoothness of $\mathcal{D}_{2N}$ in $L_\infty(\mathbb{R})$.

3.6. Multivariate wavelets. There are two approaches to the construction of multivariate wavelets for $L_2(\mathbb{R}^d)$. The first, the tensor product approach, we now briefly describe. In this section, V will denote the set of vertices of the cube $[0,1]^d$ and $V' := V\setminus\{0\}$. Let ϕ be a univariate function satisfying the conditions (3.1.2) of multiresolution and let ψ be an orthogonal wavelet obtained from ϕ. For $\phi_0 := \phi$, $\phi_1 := \psi$, the collection Ψ of functions

$$\psi_v(x_1,\dots,x_d) := \phi_{v_1}(x_1)\cdots\phi_{v_d}(x_d), \quad v \in V', \tag{3.6.1}$$

generates, by dilation and translation, a complete orthonormal system for $L_2(\mathbb{R}^d)$. More precisely, the collection of functions $\psi_{j,k,v} := 2^{kd/2}\psi_v(2^k\cdot - j)$, $j \in \mathbb{Z}^d$, $k \in \mathbb{Z}$,

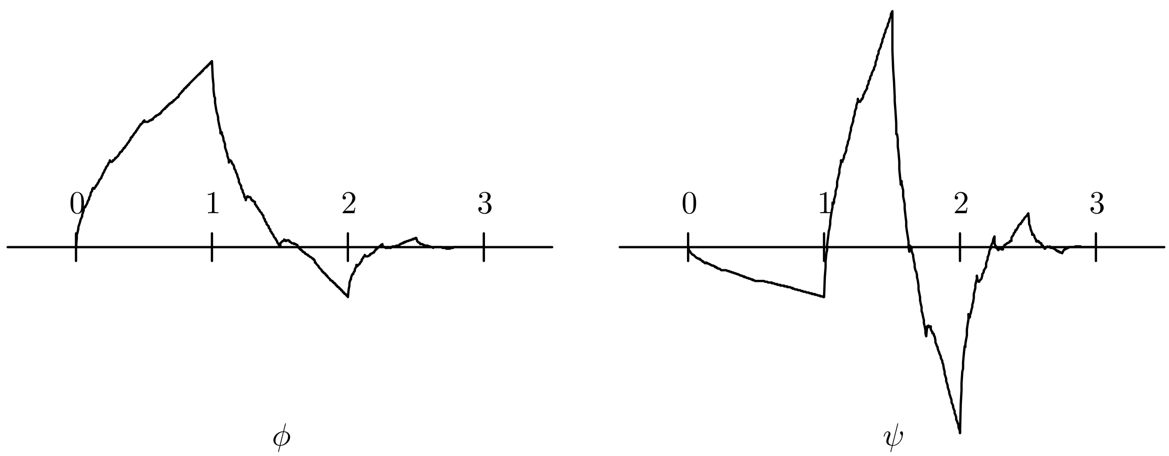

FIGURE 3. The function ϕ and the Daubechies wavelet $\psi = \mathcal{D}_{2N}$ when $N = 2$.

$v \in V'$, forms a complete orthonormal system for $L_2(\mathbb{R}^d)$: each $f \in L_2(\mathbb{R}^d)$ has the series representation

$$f = \sum_{v \in V'} \sum_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}^d} \langle f, \psi_{j,k,v} \rangle \, \psi_{j,k,v} \tag{3.6.2}$$

in the sense of convergence in $L_2(\mathbb{R}^d)$. This construction also applies to prewavelets, thereby yielding a stable basis for $L_2(\mathbb{R}^d)$.

Another view of the tensor product wavelets is the following. We let $\mathcal{S}$ be the space generated by the shifts of the function $x \longmapsto \phi(x_1) \cdots \phi(x_d)$. Then, the wavelets ψ_v are generators for the wavelet space $W := \mathcal{S}^1 \ominus \mathcal{S}^0$.

The second way to construct multivariate wavelets uses multiresolution in several dimensions. We let ϕ be a function in $L_2(\mathbb{R}^d)$ that satisfies the conditions (3.1.2) of multiresolution for $\mathcal{S} := \mathcal{S}(\phi)$, and we seek a set Ψ of generators for the wavelet space $W := \mathcal{S}^1 \ominus \mathcal{S}^0$. For example, if we want an orthonormal wavelet basis for $L_2(\mathbb{R}^d)$, we would seek Ψ such that the totality of functions $\psi(\cdot - j)$, $j \in \mathbb{Z}^d$, $\psi \in \Psi$, forms an orthonormal basis for W. By dilation and translation, we would obtain the collection of functions $\psi_{j,k} := 2^{kd/2}\psi(2^k \cdot - j)$, $\psi \in \Psi$, $j \in \mathbb{Z}^d$, $k \in \mathbb{Z}$, which together form an orthonormal basis for $L_2(\mathbb{R}^d)$. Each function f in $L_2(\mathbb{R}^d)$ has the representation

$$f = \sum_{\psi \in \Psi} \sum_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}^d} \langle f, \psi_{j,k} \rangle \, \psi_{j,k}. \tag{3.6.3}$$

For a prewavelet set Ψ, we would require $L_2(\mathbb{R}^d)$-stability in place of orthogonality. Sometimes, we might require additionally that the shifts of ψ and those of $\tilde{\psi}$ are orthogonal whenever ψ and $\tilde{\psi}$ are different functions in Ψ.

Constructing orthogonal wavelets and prewavelets by this second approach is complicated by the fact that there does not seem to be a straightforward way to choose a canonical orthogonal wavelet set from the many possible wavelet sets Ψ.

The book of Meyer [Me] contains first results on the construction of multivariate wavelet sets by the second approach. This was expanded upon in the paper of Jia and Micchelli [JM]. These treatments are not always constructive; for example, the latter paper employs in some contexts the Quillen-Suslin theorem from commutative algebra. Several papers [RS], [RS1], [CSW], and [LM] treat special cases, such as the construction of orthogonal wavelet and prewavelet sets when ϕ is taken to be a box spline. The paper of Riemenschneider and Shen [RS] is particularly important, since it gives a constructive approach that applies in two and three dimensions to a wide class of functions ϕ.

We shall follow the approach of [BDR1], which is based on the structure of shift-invariant spaces. This approach immediately gives a generating set for W, which can then be exploited to find orthogonal wavelet and prewavelet sets. de Boor, DeVore, and Ron start with a function ϕ that satisfies the refinement relation (3.1.2)(i) and whose Fourier transform satisfies $\operatorname{supp}\hat{\phi} = \mathbb{R}^d$. It is not necessary in this approach to assume (3.1.2) (ii) and (iii)—they follow automatically. It is also not necessary to assume (3.1.2) (iv). In particular, this approach applies to any compactly supported ϕ. To simplify our discussion, we shall assume in addition to (3.1.2)(i) that ϕ has compact support and that the shifts of ϕ are $L_2(\mathbb{R}^d)$-stable; we refer the reader to [BDR1] for a discussion of the more general theory.

The usual starting point for the construction of multivariate wavelets is the fact that the dilated space $\mathcal{S}^1$ of $\mathcal{S} := \mathcal{S}(\phi)$ is generated by the half-shifts of $\eta := \phi(2\,\cdot\,)$, and therefore also by the full shifts of the functions $\eta_v := \eta(\,\cdot\, - v/2)$, $v \in V$. The assumption that $\operatorname{supp}\hat{\phi} = \mathbb{R}^d$ is important because it implies that the set $\Phi := \{\phi_v := \phi(x - v/2) \mid v \in V\}$ is also a generating set for $\mathcal{S}^1$, i.e., $\mathcal{S}^1 = \mathcal{S}(\Phi)$. Φ is more useful than $\{\eta_v \mid v \in V\}$ as a generating set because Φ contains a function that is in $\mathcal{S}^0$, namely ϕ. In analogy with the univariate construction, we see that with P the $L_2(\mathbb{R}^d)$ projector onto $\mathcal{S}$, the functions $\phi_v - P\phi_v$, $v \in V'$, form a generating set for W. From (3.2.8), we calculate the Fourier transforms of these functions and multiply them by $[\hat{\phi}, \hat{\phi}]$ to obtain the functions w_v, with Fourier transform

$$\hat{w}_v := [\hat{\phi}, \hat{\phi}]\hat{\phi}_v - [\hat{\phi}_v, \hat{\phi}]\hat{\phi}, \quad v \in V'. \tag{3.6.4}$$

The set $\mathcal{W} := \{w_v \mid v \in V'\}$ is another generating set for W. We note that because we assume ϕ has compact support, the two bracket products appearing in (3.6.4) are trigonometric polynomials, and hence the functions w_v also have compact support.

The set $T\mathcal{W}$, with $T = (\tau_{v,v'})_{v,v' \in V'}$ a matrix of 2π-periodic functions, is another generating set for W if $\det(T) \neq 0$ a.e.

It is easy to find an orthogonal wavelet set by this approach. Because the functions in $\mathcal{W}$ have compact support, the Gramian matrix $([\hat{w}_v, \hat{w}_{v'}])_{v,v' \in V'}$ has trigonometric polynomials as its entries. Since this matrix is symmetric and positive semidefinite, its determinant is nonzero a.e. We can use Gauss elimination (Cholesky factorization) without division or pivoting to diagonalize $G(\mathcal{W})$. That is, we can find a (symmetric) matrix $T = (\tau_{v,v'})_{v,v' \in V'}$ of trigonometric polynomials such that $\mathcal{W}^* := T\mathcal{W}$ has Gramian $G(T\mathcal{W}^*) = TG(\mathcal{W})T^*$ that is a diagonal matrix with trigonometric polynomial entries. If w_v^* are the functions in $\mathcal{W}^*$, then the functions w_v^{**} with Fourier transforms $\hat{w}_v^{**} := \hat{w}_v^* / [\hat{w}_v^*, \hat{w}_v^*]^{1/2}$, $v \in V'$, have shifts that form an orthonormal basis for W. Indeed,

$$[\hat{w}_v^{**}, \hat{w}_{v'}^{**}] = \frac{[\hat{w}_v^*, \hat{w}_{v'}^*]}{[\hat{w}_v^*, \hat{w}_v^*]^{1/2}[\hat{w}_{v'}^*, \hat{w}_{v'}^*]^{1/2}} \ ,$$

which shows that the new set of generators $\mathcal{W}^{**}$ has the identity matrix as its Gramian.

The disadvantage of the orthogonal wavelet set $\mathcal{W}^{**}$ is that usually we can say nothing about the decay of the functions w_v^{**}, since the above construction may involve division by trigonometric polynomials that have zeros. However, when ϕ has $L_2(\mathbb{R}^d)$-stable half-shifts, the above construction can be modified to give an orthogonal wavelet set whose elements decay exponentially (see [BDR1]).

While the assumption that the half-shifts of ϕ are $L_2(\mathbb{R}^d)$-stable is often not realistic, we shall assume it a little longer in order to introduce some new ideas that can later be modified to drop the stability assumption. Under the half-shift stability assumption, we have that $\tilde{\tilde{\phi}}$ is a trigonometric polynomial of period 4π that has no zeros. Therefore, $\hat{\phi}_* := \hat{\phi}/\tilde{\tilde{\phi}}$ serves to define an $L_2(\mathbb{R}^d)$ function in $\mathcal{S}^1(\phi)$ that decays exponentially and has orthogonal half-shifts. Moreover, the function w with Fourier transform $\hat{w} := \hat{\phi}/\tilde{\tilde{\phi}}^2$ is also in $\mathcal{S}^1(\phi)$ and decays exponentially. Therefore, with $[\,\cdot\,,\,\cdot\,]_{1/2}$ the bracket product for half-shifts (which is defined as in (3.2.4) except that the sum is taken over $4\pi\mathbb{Z}^d$), we have

$$[\hat{\phi}, \hat{w}]_{1/2} = [\hat{\phi}_*, \hat{\phi}_*]_{1/2} = 1 \text{ a.e.} \tag{3.6.5}$$

The Fourier coefficients (with respect to the $e_{-j/2}$, $j \in \mathbb{Z}^d$) of $[\hat{\phi}, \hat{w}]_{1/2}$ are the inner products of ϕ with half-shifts of w. Hence, all nontrivial half-shifts of w are orthogonal to ϕ. This means that the functions in $\mathcal{W}_0 := \{w(\,\cdot\, + v/2) \mid v \in V'\}$ are all in W. It is easy to see that they generate W, that is, $W = \mathcal{S}(\mathcal{W}_0)$.

Thus, in the special case we are considering, W is generated by the nontrivial half-shifts of a single function w. It is natural to ask whether this holds in general (i.e., when we do not assume stability of half-shifts). To see that this is indeed true, we modify the argument in (3.6.5). If we multiply $\hat{w}$ by the 2π-periodic function $\prod_{\lambda\in 2\pi V} \tilde{\tilde{\phi}}(\,\cdot\, + \lambda)^2$, the result is a compactly supported function $w_* \in L_2(\mathbb{R}^d)$, with Fourier transform

$$\hat{w}_* := \hat{\phi} \prod_{\lambda \in 2\pi V'} \tilde{\tilde{\phi}}(\,\cdot\, + \lambda)^2. \tag{3.6.6}$$

We find that

$$[\hat{\phi}, \hat{w}_*]_{1/2} = \prod_{\lambda \in 2\pi V} \tilde{\tilde{\phi}}(\,\cdot\, + \lambda)^2. \tag{3.6.7}$$

Because the right side is 2π-periodic, we deduce that the inner product of ϕ with $w_*(\,\cdot\, - j/2)$ is zero whenever $j = v + 2k$ with $v \in V'$ and $k \in \mathbb{Z}^d$. Hence, the functions $w_*(\,\cdot\, + v/2)$, $v \in V'$, are all in W and it is easy to see that they are also a generating set for W.

While the nontrivial half-shifts of w_* are a generating set for W, they have the drawback that they usually do not provide an $L_2(\mathbb{R}^d)$-stable basis. The usual approach towards constructing an $L_2(\mathbb{R}^d)$-stable basis for W is to begin with the generating set $\{\eta_v \mid v \in V\}$, $\eta := \phi(2\,\cdot\,)$, for $\mathcal{S}^1(\phi)$. With this as a starting point, Meyer [Me, III, §6] and Jia and Micchelli [JM] have shown the existence of a set of generators for W consisting of compactly supported functions whose shifts provide an $L_2(\mathbb{R}^d)$-stable basis for W. However, their proofs are not constructive. In one, two, or three dimensions, and with an additional assumption on the symmetry of ϕ, Riemenschneider and Shen [RS], [RS1] have given a constructive proof for the existence of such a generating set, which we now describe.

We begin again with the function $\hat{w} := \hat{\phi}/\tilde{\tilde{\phi}}^2$. If $\tau = \sum_{j\in\mathbb{Z}^d} c(j)e_{-j/2}$ is a 4π-periodic function whose Fourier coefficients $c(j) = 0$ whenever $j \in 2\mathbb{Z}^d$, then the function with Fourier transform $\tau\hat{w}$ is in W provided it is in $L_2(\mathbb{R}^d)$. The condition on these Fourier coefficients is equivalent to requiring that

$$\sum_{v\in 2\pi V} \tau(\,\cdot\, + v) = 0 \text{ a.e.} \tag{3.6.8}$$

In particular, we can use this method to produce functions in W as follows.

We assume that ϕ is real valued, has $L_2(\mathbb{R}^d)$-stable shifts, and satisfies the refinement equation (3.3.1) for a *real valued function* A. This assumption on A is a new ingredient; it will be fulfilled for example if $\phi(-x) = \phi(x)$. (More generally, one only needs symmetry about the center of the support of ϕ.) If $\alpha \in V'$ and $v_\alpha \in 2\pi V$, then we claim that the function ψ_α with Fourier transform

$$\hat{\psi}_\alpha := 2e_{\alpha/2}\,(A\tilde{\tilde{\eta}}^2)(\,\cdot\, + v_\alpha)\hat{\eta} \tag{3.6.9}$$

is in W, provided $e_{\alpha/2}(v_\alpha) = -1$. Indeed, the refinement equation says that $\hat{w} = \hat{\eta}/(A\tilde{\tilde{\eta}}^2)$. We obtain $\hat{\psi}_\alpha$ from $\hat{w}$ by multiplying by $\tau_\alpha := 2e_{\alpha/2}A\tilde{\tilde{\eta}}^2 A(\,\cdot\, + v_\alpha)\tilde{\tilde{\eta}}^2(\,\cdot\, + v_\alpha)$. The vertices v and $v + v_\alpha$ (modulo 2π) contribute values in (3.6.8) that are negatives of one another. Hence, (3.6.8) is satisfied for τ_α and $\psi_\alpha \in W$. The functions ψ_α have compact support when A is a polynomial and ϕ has compact support.

We are allowed to make any assignment of $\alpha \longmapsto v_\alpha$ with $e_{\alpha/2}(v_\alpha) = -1$. To obtain an $L_2(\mathbb{R}^d)$-stable basis for W, we need a special assignment with the property that $\alpha - \beta$ (modulo 2) is assigned $v_\alpha - v_\beta$ (modulo 2π). If such a special assignment can be made, then a simple computation shows that with $\mu := 2A\tilde{\tilde{\eta}}^2$,

$$[\hat{\psi}_\alpha, \hat{\psi}_\beta] = \sum_{v\in 2\pi V} e_{(\alpha-\beta)/2}(\,\cdot\, + v)\,\mu(\,\cdot\, + v + v_\alpha)\,\mu(\,\cdot\, + v + v_\beta)\,\tilde{\tilde{\eta}}^2(\,\cdot\, + v). \tag{3.6.10}$$

For example, if the shifts of ϕ are orthonormal, then $\tilde{\tilde{\eta}}^2 = 1/2$. In this case, for $\alpha \neq \beta$, the terms of the sum in (3.6.10) are negatives of one another for the two values v and $v+(v_\alpha+v_\beta)$ (this is where we need to assume that a special assignment exists), and hence the sum in (3.6.10) is 0. When $\alpha = \beta$, this sum is

$$\sum_{v\in 2\pi V} A^2(\,\cdot\, + v)\tilde{\tilde{\eta}}^2(\,\cdot\, + v) = \sum_{v\in 2\pi V} \tilde{\tilde{\phi}}^2(\,\cdot\, + v) = \tilde{\phi}^2 = 1 \text{ a.e.}$$

Hence, the Gramian of $\Psi := \{\psi_\alpha \mid \alpha \in V'\}$ is the identity matrix. We obtain in this way an orthonormal basis for W.

If we begin with a ϕ that has $L_2(\mathbb{R}^d)$-stable shifts then a slightly more complicated argument shows that the functions ψ_α, $\alpha \in V'$, are an $L_2(\mathbb{R}^d)$-stable basis for W.

This leaves the question of when we can make an assignment $\alpha \longmapsto v_\alpha$ of this special type. Such assignments are possible for $d = 1, 2, 3$ but not for $d > 3$. For example, when $d = 2$, we can make the assignments as follows:

$$\begin{aligned} (0,0) &\longmapsto 2\pi(0,0), \quad (0,1) \longmapsto 2\pi(0,1), \\ (1,0) &\longmapsto 2\pi(1,1), \quad (1,1) \longmapsto 2\pi(1,0). \end{aligned}$$

The above construction will give orthogonal wavelet and prewavelet sets for box splines. To illustrate this, we consider briefly the following special box splines on $\mathbb{R}^2$. Let Δ be the triangulation of $\mathbb{R}^2$ obtained from grid lines $x_1 = k$, $x_2 = k$, and $x_2 - x_1 = k$, $k \in \mathbb{Z}$. Let M be the Courant element for this partition. Thus, M is the piecewise linear function for this triangulation that takes the value 1 at $(0,0)$ and the value 0 at all other vertices. The Fourier transform of M is

$$\hat{M}(y_1, y_2) = \left(\frac{\sin(y_1/2)}{y_1/2}\right)\left(\frac{\sin(y_2/2)}{y_2/2}\right)\left(\frac{\sin((y_1+y_2)/2)}{(y_1+y_2)/2}\right).$$

By convolving M with itself, we obtain higher order box splines defined recursively by $M_1 := M$ and $M_r := M * M_{r-1}$. Then M_r is a compactly supported piecewise polynomial (with respect to Δ) of degree $3r - 2$ and smoothness C^{2r-2}. Since $\hat{M}$ is real, the box spline M_r satisfies the refinement identity (3.3.1) with A real. Therefore, if we take $\phi := M_r$ and $\mathcal{S} = \mathcal{S}(M_r)$, the construction of Riemenschneider and Shen applies to give a prewavelet set Ψ consisting of three compactly supported piecewise polynomials for the partition $\Delta/2$. The set Ψ provides an $L_2(\mathbb{R}^d)$-stable basis for the wavelet space W.

4. Fast Wavelet Transforms

It is easy to compute the coefficients in wavelet decompositions iteratively with a technique similar to the fast Haar transform. We shall limit our discussion to Daubechies' orthogonal wavelets with compact support in one dimension. However, the ideas presented here apply equally well to other orthogonal wavelets and to prewavelets. We let ϕ be a real-valued, compactly-supported function with orthonormal shifts that satisfies the conditions of multiresolution and in particular the refinement equation (3.1.4). The function ϕ is real and the refinement coefficients are real and finite in number. The orthogonal wavelet ψ is then obtained from ϕ by (3.1.8).

A numerical application begins with a finite representation of a function f as a wavelet sum. This is accomplished by choosing a large value of n, commensurate with the numerical accuracy desired, and taking an approximation to f of the form

$$S_n = \sum_{j \in \mathbb{Z}} f_j \phi_{j,n}, \tag{4.1}$$

with only a finite number of nonzero coefficients f_j. The coefficients f_j are obtained from f in some suitable way. For many applications, it suffices to take $f_j := f(j2^{-n})$. The point $j2^{-n}$ corresponds to the support of $\phi_{j,n}$.

Since $S_n \in \mathcal{S}^n$, we have

$$S_n = P_n S_n = P_0 S_n + \sum_{k=1}^{n} (P_k S_n - P_{k-1} S_n) = P_0 S_n + \sum_{k=0}^{n-1} \sum_{j\in\mathbb{Z}} c(j,k)\psi(2^k x - j). \tag{4.2}$$

We will present an efficient algorithm for computing the coefficients $c(j,k)$ from the f_j and an efficient method to recover S_n from the coefficients $c(j,k)$.

The algorithm presented below has two main features. First, it computes the coefficients $c(j,k)$ using only f_j and the coefficients $(a(j))$ of the refinement equation (3.1.4) for ϕ. In other words, one never needs to find a concrete realization of the functions ϕ and ψ. Second, the iterative computations are particularly simple to program and run very quickly—the complexity of the fast wavelet transform of 2^n coefficients is $\mathcal{O}(2^n)$; in contrast, the complexity of the Fast Fourier Transform is $\mathcal{O}(n2^n)$.

During one step of our algorithm, we wish to find the coefficients of $P_{k-1}S$ when $S = \sum_{j\in\mathbb{Z}} s(j)\phi_{j,k}$ is in $\mathcal{S}^k$. The coefficients of $P_{k-1}S = \sum_{i\in\mathbb{Z}} s'(i)\phi_{i,k-1}$ are the inner product of S with the $\phi_{i,k-1}$. We therefore compute, using (3.1.4),

$$\begin{aligned} s'(i) = \int_{\mathbb{R}} \Big[\sum_{j\in\mathbb{Z}} s(j)\phi_{j,k}\Big]\phi_{i,k-1} &= \int_{\mathbb{R}} \Big[\sum_{j\in\mathbb{Z}} s(j)\phi_{j,k}\Big]\Big[\frac{1}{\sqrt{2}}\sum_{\ell\in\mathbb{Z}} a(\ell)\phi_{2i+\ell,k}\Big] \\ &= \frac{1}{\sqrt{2}}\sum_{j\in\mathbb{Z}} a(j-2i)s(j). \end{aligned}$$

In other words, the sequence $\boldsymbol{s}' := (s'(i))$ is obtained from $\boldsymbol{s} := (s(i))$ by matrix multiplication:

$$\boldsymbol{s}' = \boldsymbol{A}\boldsymbol{s}, \quad \boldsymbol{A} := (\alpha_{i,j}), \quad \alpha_{i,j} := \frac{1}{\sqrt{2}} a(j-2i),\ i,j\in\mathbb{Z}. \tag{4.3}$$

Let Q_k be the projector onto the wavelet space W^k. A similar calculation tells us how to compute the coefficients $\boldsymbol{t} = (t(i))$ of the projection $Q_{k-1}S = \sum_{i\in\mathbb{Z}} t(i)\psi_{i,k-1}$ of $S \in \mathcal{S}^k$ onto W_{k-1}:

$$\boldsymbol{t} = \boldsymbol{B}\boldsymbol{s}, \quad \boldsymbol{B} := (\beta_{i,j}), \quad \beta_{i,j} := \frac{1}{\sqrt{2}} b(j-2i),\ i,j\in\mathbb{Z}, \tag{4.4}$$

where $b_j := (-1)^j a(1-j)$ are the coefficients of the wavelet ψ given in (3.1.8).

This gives the following schematic diagram for computing the wavelet coefficients $(c(j,k))$:

$$\begin{array}{ccccccc} & \boldsymbol{A} & & \boldsymbol{A} & & \boldsymbol{A} & \\ S_n & \rightarrow & P_{n-1}S_n & \rightarrow & \dots & \rightarrow & P_0 S_n \\ & \searrow \boldsymbol{B} & & \searrow \boldsymbol{B} & & \searrow \boldsymbol{B} & \\ & & Q_{n-1}S_n & & \dots & & Q_0 S_n \end{array} \tag{4.5}$$

In other words, to compute the coefficients of $P_{k-1}S_n$ we apply the matrix $\boldsymbol{A}$ to the coefficients of P_kS_n, while to compute those of $Q_{k-1}S_n$ we apply the matrix $\boldsymbol{B}$ to the coefficients of P_kS_n. The coefficients $c(j,k)$, $j \in \mathbb{Z}$, are the coefficients of Q_kS_n. This is valid because $Q_{k-1}P_kS_n = Q_{k-1}S_n$ and $P_{k-1}P_kS_n = P_{k-1}S_n$.

Now suppose that we know the coefficients of P_0S_n and Q_kS_n, $k = 0, \dots, n-1$. How do we reconstruct S_n? We need to rewrite an element $S \in \mathcal{S}^k$, $S = \sum_{j\in\mathbb{Z}} s(j)\phi_{j,k}$ as an element of $\mathcal{S}^{k+1}$, $S = \sum_{i\in\mathbb{Z}} s'(i)\phi_{i,k+1}$. From the refinement equation (3.1.4), we find

$$S = \sum_{j\in\mathbb{Z}} s(j)\left[\frac{1}{\sqrt{2}}\sum_{\ell\in\mathbb{Z}} a(\ell)\phi_{\ell+2j,k+1}\right] = \sum_{i\in\mathbb{Z}}\left[\frac{1}{\sqrt{2}}\sum_{j\in\mathbb{Z}} a(i-2j)s(j)\right]\phi_{i,k+1}.$$

Therefore, we can express the computation of $\boldsymbol{s}'$ from $\boldsymbol{s}$ as multiplication by the transpose $\boldsymbol{A}^*$ of $\boldsymbol{A}$:

$$\boldsymbol{s}' = \boldsymbol{A}^*\boldsymbol{s}, \quad \boldsymbol{A}^* := (\alpha^*_{i,j}), \quad \alpha^*_{i,j} := \frac{1}{\sqrt{2}}a(i-2j),\ i,j \in \mathbb{Z}. \tag{4.6}$$

A similar calculation tells us how to rewrite a sum $S = \sum_{i\in\mathbb{Z}} t(i)\psi_{i,k}$ as a sum $S = \sum_{i\in\mathbb{Z}} t'(i)\phi_{i,k+1}$:

$$\boldsymbol{t}' = \boldsymbol{B}^*\boldsymbol{t}, \quad \boldsymbol{B}^* := (\beta^*_{i,j}), \quad \beta^*_{i,j} := \frac{1}{\sqrt{2}}b(i-2j),\ i,j \in \mathbb{Z}. \tag{4.7}$$

The reconstruction of S_n from Q_kS_n, $k = 0, \dots, n-1$, and P_0S_n is then given schematically by:

$$\begin{array}{ccccccc} & \boldsymbol{A}^* & & \boldsymbol{A}^* & & \boldsymbol{A}^* & \\ P_0S_n & \to & P_1S_n & \to & \dots & \to & S_n \\ & \nearrow \boldsymbol{B}^* & & & & \nearrow \boldsymbol{B}^* & \\ Q_0S_n & & \dots & & Q_{n-1}S_n & & \end{array} \tag{4.8}$$

The matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{A}^*$, and $\boldsymbol{B}^*$ have a small finite number of nonzero elements in each row, so each of the operations in (4.5) and (4.6) has computational complexity proportional to the number of unknowns.

The reconstruction algorithm can be used to display graphically a finite wavelet sum S. We choose a large value of n, and use the reconstruction algorithm to write $S = \sum_{j\in\mathbb{Z}} s(j,n)\phi_{j,n}$. The piecewise linear function with values $s(j,n)$ at $j2^{-n}$ is an approximation to the graph of S. Such procedures for graphical displays are known as subdivision algorithms in computer aided geometric design.

The matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have many remarkable properties summarized by:

$$\begin{gathered} \boldsymbol{B}\boldsymbol{A}^* = \boldsymbol{0}, \\ \boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{I}, \qquad \boldsymbol{B}\boldsymbol{B}^* = \boldsymbol{I}, \\ \boldsymbol{A}^*\boldsymbol{A} + \boldsymbol{B}^*\boldsymbol{B} = \boldsymbol{I}. \end{gathered} \tag{4.9}$$

The first equation represents the orthogonality between W and $\mathcal{S}^0$, the second the fact that the shifts of ϕ and ψ are orthonormal, and the third the orthogonal decomposition $\mathcal{S}^1 = \mathcal{S}^0 \oplus W$.

5. Smoothness Spaces and Wavelet Coefficients

We have seen in §2 that one can determine when a function $f \in L_p(\mathbb{R})$ is in a Lipschitz space $\mathrm{Lip}(\alpha, L_p(\mathbb{R}))$, $0 < \alpha < 1/p$, by examining the coefficients of the Haar expansion of f. In fact, one can often characterize membership in general smoothness spaces in terms of the size of coefficients in general wavelet or prewavelet expansions. We do not have the space to explain in detail how such characterizations are proved, but we shall outline one approach, based on approximation, that parallels the arguments in §2 about Haar wavelets. A more complete presentation, much along the lines given here, can be found in the book of Meyer [Me]. The article of Frazier and Jawerth [FJ] gives a much more general and extensive treatment of wavelet-like decompositions from the viewpoint of Littlewood-Paley theory.

We shall suppose that ϕ satisfies the conditions (3.1.2) of multiresolution. We also suppose that ϕ has compact support. This is not a necessary assumption for the characterizations given below (it can be replaced by suitable polynomial decay at infinity) but it will simplify our discussion. We shall also assume that $1 < p < \infty$. The arguments that follow can be modified simply to apply when $p = \infty$; the analysis for $p \le 1$ can also be developed as below, but then it must be carried out in the setting of the Hardy spaces $H_p(\mathbb{R}^d)$.

We fix a value of p and let $\mathcal{S} := \mathcal{S}(\phi, L_p(\mathbb{R}^d))$ be the $L_p(\mathbb{R}^d)$ closure of the finite linear combination of shifts of ϕ.

We assume that the shifts of ϕ form an $L_p(\mathbb{R}^d)$-stable basis for $\mathcal{S}$. For functions with compact support, this holds whenever the shifts of ϕ form an $L_2(\mathbb{R}^d)$-stable basis for $\mathcal{S}(\phi, L_2(\mathbb{R}^d))$ (see Jia and Micchelli [JM]). It follows that the dilated functions $\phi_{j,k,p} := 2^{k/p}\phi(2^k\cdot - j)$, $j \in \mathbb{Z}^d$, form an $L_p(\mathbb{R}^d)$-stable basis of $\mathcal{S}^k$, for each $k \in \mathbb{Z}$. That is, there are constants $C_1, C_2 > 0$ such that each $S \in \mathcal{S}^k$ can be represented as $S = \sum_{j\in\mathbb{Z}^d} c(j,k,p)(S)\phi_{j,k,p}$ with

$$C_1\Big(\sum_{j\in\mathbb{Z}^d} |c(j,k,p)|^p\Big)^{1/p} \le \|S\|_{L_p(\mathbb{R}^d)} \le C_2\Big(\sum_{j\in\mathbb{Z}^d} |c(j,k,p)|^p\Big)^{1/p}. \tag{5.1}$$

Next, we show that the orthogonal projector P from $L_2(\mathbb{R}^d)$ onto $\mathcal{S}(\phi, L_2(\mathbb{R}^d))$ has a natural extension to a bounded operator from $L_p(\mathbb{R}^d)$ onto $\mathcal{S}$. We can represent P as in (3.2.9):

$$Pf = \sum_{j\in\mathbb{Z}} \gamma_j(f)\phi(\,\cdot - j), \quad \gamma_j(f) := \int_{\mathbb{R}^d} f(x)\overline{\gamma(x-j)}\,dx. \tag{5.2}$$

The function $\gamma \in L_\infty(\mathbb{R}^d)$ decays exponentially at infinity and hence is in $L_q(\mathbb{R}^d)$ for $1 \le q \le \infty$. In particular, $\gamma \in L_{p'}(\mathbb{R}^d)$, and (5.2) serves to define P on $L_p(\mathbb{R}^d)$. The compact support of ϕ and the exponential decay of γ then combine to show that P is bounded on $L_p(\mathbb{R}^d)$. By dilation, we find that the projectors P_k (which map $L_p(\mathbb{R}^d)$ onto $\mathcal{S}^k$) are bounded independently of k.

The projector Q from $L_2(\mathbb{R}^d)$ onto the wavelet space W is also represented in the form (5.2) and has an extension to a bounded operator on $L_p(\mathbb{R}^d)$ for the same reasons as above. We can also derive the boundedness of Q from the formula $Q = P_1 - P_0$.

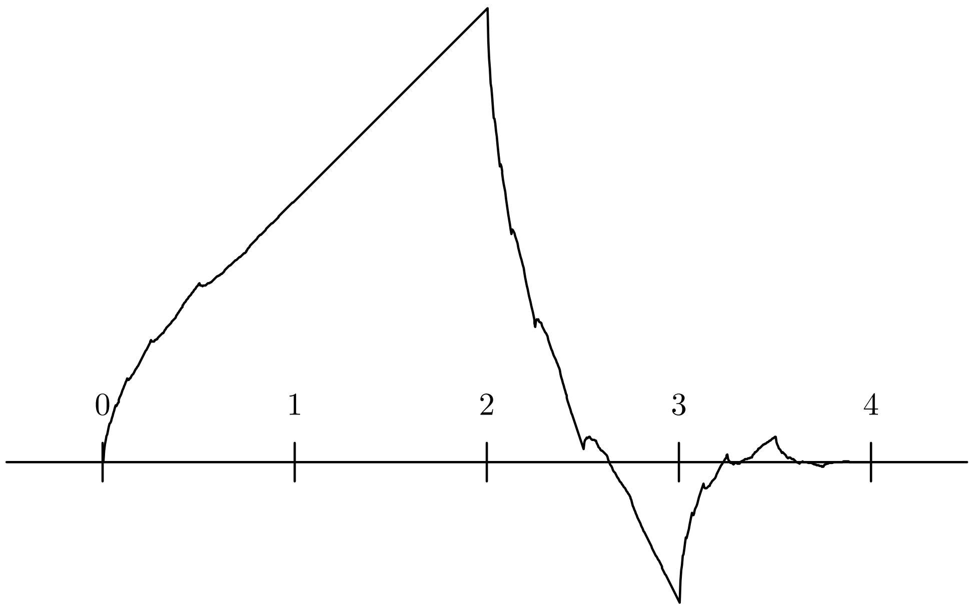

FIGURE 4. The function $0 \cdot \phi(x+1) + 1 \cdot \phi(x) + 2 \cdot \phi(x-1)$, with ϕ given by Daubechies' formula (3.5.1) with $N = 2$; see Figure 3. Note the linear segment between $x = 1$ and $x = 2$.

Since the P_k are bounded projectors onto $\mathcal{S}^k$, their approximation properties are determined by the approximation properties of the spaces $\mathcal{S}^k$. Consequently, we want to bound the error of approximation by elements in $\mathcal{S}^k$ of functions in certain smoothness classes. In particular, we are interested in determining for which spaces $\mathcal{S}^k$ it is true that

$$\operatorname{dist}(f, \mathcal{S}^k)_{L_p(\mathbb{R}^d)} \le C 2^{-kr} |f|_{W^r(L_p(\mathbb{R}^d))}; \tag{5.3}$$

here $W^r(L_p(\mathbb{R}^d))$ is the Sobolev space of functions with r (weak) derivatives in $L_p(\mathbb{R}^d)$ with its usual norm and seminorm. This well-studied problem originated with the work of Schoenberg [Sch], and was later developed by Strang and Fix [SF] for application to finite elements. Strang and Fix show that when ϕ has compact support, a sufficient condition for (5.3) to hold is that

$$\hat{\phi}(0) \ne 0 \quad \text{and} \quad D^\nu \hat{\phi}(2\pi\alpha) = 0,\ |\nu| < r,\ \alpha \in \mathbb{Z}^d,\ \alpha \ne 0. \tag{5.4}$$

This condition is also necessary in a certain context; see [BDR] and [BR] for a history of the Strang-Fix conditions.

Schoenberg [Sch] showed that (5.4) guarantees that algebraic polynomials of (total) degree $< r$ are contained locally in the space $\mathcal{S}^k$. This means that for any compact set Ω and any polynomial R with $\deg(R) < r$, there is an $S \in \mathcal{S}$ that agrees with R on Ω.

In summary, the approximation properties of P_k are determined by the largest value of r for which (5.4) is valid. Because we usually know a lot about the Fourier transform of ϕ, the best value of r is easy to determine. For this r, we have (5.3). For example, for $\phi = N_r$, the B-spline of order r, $\hat{\phi}(y) = (1 - e^{-iy})^r/(iy)^r$ satisfies (5.4) for this value of r. Similarly, the Daubechies wavelets satisfy (5.4) for $r = N$, with N the integer appearing in the representation (3.5.6); see Figure 4.

Assume now that there are positive integers r and s such that the following Jackson and Bernstein inequalities hold:

$$\begin{aligned} &\text{(J)} \quad \|f - P_k f\|_{L_p(\mathbb{R}^d)} \le C \cdot 2^{-kr} |f|_{W^r(L_p(\mathbb{R}^d))}, \\ &\text{(B)} \quad |S|_{W^s(L_p(\mathbb{R}^d))} \le C \cdot 2^{ks} \|S\|_{L_p(\mathbb{R}^d)}, \quad S \in \mathcal{S}^k(\phi, L_p(\mathbb{R}^d)),\ 1 \le p \le \infty. \end{aligned} \tag{5.5}$$

(Actually, s need not be an integer.) The Jackson inequality is just a reformulation of (5.3), and the largest value of r for which (J) holds is determined by (5.4). The Bernstein inequality holds if $\phi \in W^s(L_p(\mathbb{R}^d))$ and in particular (since ϕ has compact support) whenever ϕ is in C^s. It is enough to verify (B) for $k = 0$, since (B) would then follow for general k by rescaling. The left seminorm in (B) for $S = \sum_{j \in \mathbb{Z}^d} c(j)\phi(\,\cdot - j)$ is bounded by the $\ell_p(\mathbb{Z}^d)$ norm of the coefficients $(c(j))_{j \in \mathbb{Z}^d}$, which is bounded in turn by the right side of (B) by using the $L_p(\mathbb{R}^d)$-stability of the shifts of ϕ.

Once the Jackson and Bernstein inequalities have been established, we can invoke a general procedure to characterize smoothness spaces in terms of wavelet coefficients. To describe these results, we introduce the Besov spaces, which are a family of smoothness spaces that depend on three parameters. We introduce the Besov spaces for only one reason: They are the spaces that are needed to describe precisely the smoothness of functions that can be approximated to a given order by wavelets. The following discussion is meant as a gentle introduction to Besov spaces for the reader who instinctively dislikes any space that depends on so many parameters.

The Besov space $B^\alpha_q(L_p(\mathbb{R}^d))$, $\alpha > 0$ and $0 < q, p \le \infty$, is a smoothness subspace of $L_p(\mathbb{R}^d)$. The parameter α gives the order of smoothness in $L_p(\mathbb{R}^d)$. The second parameter q gives a finer scaling, which allows us to make subtle distinctions in smoothness of fixed order α. This second parameter is necessary in many embedding and approximation theorems.

To define these spaces, we introduce, for $h \in \mathbb{R}^d$, the r-th difference in the direction h:

$$\Delta^r_h(f, x) := \sum_{j=0}^{r} (-1)^{r+j} \binom{r}{j} f(x + jh).$$

Thus, $\Delta_h(f, x) := f(x+h) - f(x)$ is the first difference of f and the other differences are obtained inductively by a repeated application of Δ_h. With these differences, we can define the moduli of smoothness

$$\omega_r(f, t)_p := \sup_{|h| \le t} \|\Delta^r_h(f, \cdot\,)\|_{L_p(\mathbb{R}^d)}, \quad t > 0,$$

for each $r = 1, 2, \dots$. The rate at which $\omega_r(f, t)_p$ tends to zero gives information about the smoothness of f in $L_p(\mathbb{R}^d)$. For example, the spaces $\mathrm{Lip}(\alpha, L_p(\mathbb{R}^d))$, which we have discussed earlier, are characterized by the condition $\omega_1(f, t)_p = \mathcal{O}(t^\alpha)$, $0 < \alpha \le 1$.

The Besov spaces are defined for $0 < \alpha < r$ and $0 < p, q \le \infty$ as the set of all functions $f \in L_p(\mathbb{R}^d)$ for which

$$(5.6) \qquad |f|_{B_q^\alpha(L_p(\mathbb{R}^d))} := \begin{cases} \left(\int_0^\infty \left[t^{-\alpha}\omega_r(f,t)_p \right]^q \dfrac{dt}{t} \right)^{\frac{1}{q}}, & 0 < q < \infty, \\ \sup_{t \geq 0} t^{-\alpha}\omega_r(f,t)_p, & q = \infty, \end{cases}$$

is finite. We define the following "norm" for $B_q^\alpha(L_p(\mathbb{R}^d))$:

$$\|f\|_{B_q^\alpha(L_p(\mathbb{R}^d))} := \|f\|_{L_p(\mathbb{R}^d)} + |f|_{B_q^\alpha(L_p(\mathbb{R}^d))}.$$

Because we allow values of p and q less than one, this "norm" does not always satisfy the triangle inequality, but it is always a quasi-norm: There exists a constant C such that for all f and g in $B_q^\alpha(L_p(\mathbb{R}^d))$,

$$\|f+g\|_{B_q^\alpha(L_p(\mathbb{R}^d))} \leq C(\|f\|_{B_q^\alpha(L_p(\mathbb{R}^d))} + \|g\|_{B_q^\alpha(L_p(\mathbb{R}^d))}).$$

Even though the definition of the $B_q^\alpha(L_p(\mathbb{R}^d))$ norm depends on r through the modulus of smoothness, we have not parametrized the spaces by r, for two reasons. First, *no one* can stand spaces that are parametrized by more than three parameters. Second, it can be shown that all values of r greater than α give rise to equivalent norms, so the set of functions in $B_q^\alpha(L_p(\mathbb{R}^d))$ does not depend on r as long as $r > \alpha$.

We note that the family of Besov spaces contains both the Lipschitz spaces $\mathrm{Lip}(\alpha, L_p(\mathbb{R}^d)) = B_\infty^\alpha(L_p(\mathbb{R}^d))$, $0 < \alpha < 1$, and the Sobolev spaces $W^\alpha(L_2(\mathbb{R}^d)) = B_2^\alpha(L_2(\mathbb{R}^d))$, which are frequently denoted by $H^\alpha(\mathbb{R}^d)$.

We have mentioned that once Jackson and Bernstein inequalities have been established, there is a general theory for characterizing membership in Besov spaces by the decay of wavelet coefficients. This is based on results from approximation theory described (among other places) in the articles [DJP], [DP], and the forthcoming book of DeVore and Lorentz [DeLo]. Among other things, this theory states that whenever (5.5) holds, we have that

$$(5.7) \qquad |f|_{B_q^\alpha(L_p(\mathbb{R}^d))} \approx \left(\sum_{k\in\mathbb{Z}} [2^{k\alpha}\|Q_k(f)\|_{L_p(\mathbb{R}^d)}]^q \right)^{1/q}$$

for $0 < \alpha < \min(r,s)$, $1 < p < \infty$, and $0 < q < \infty$.

If Ψ is a wavelet set associated with ϕ, then

$$Q_k(f) = \sum_{\psi\in\Psi} \sum_{j\in\mathbb{Z}^d} \gamma_{j,k,\psi,p}(f)\psi_{j,k,p},$$

with $\psi_{j,k,p} := 2^{kd/p}\psi(2^k \cdot - j)$ the $L_p(\mathbb{R}^d)$-normalized (pre)wavelets, and the $\gamma_{j,k,\psi,p}$ the associated dual functionals. Using the $L_p(\mathbb{R}^d)$-stability of Ψ, we can replace $\|Q_k(f)\|_{L_p(\mathbb{R}^d)}^p$ by $\sum_{j\in\mathbb{Z}^d} |\gamma_{j,k,\psi,p}(f)|^p$ and obtain

$$(5.8) \qquad |f|_{B_q^\alpha(L_p(\mathbb{R}^d))} \approx \left(\sum_{k\in\mathbb{Z}} [2^{k\alpha p} \sum_{\psi\in\Psi} \sum_{j\in\mathbb{Z}^d} |\gamma_{j,k,\psi,p}(f)|^p]^{q/p} \right)^{1/q},$$

with the usual change to a supremum when $q = \infty$. This is the characterization of the Besov space in terms of wavelet coefficients. When $q = p$, (5.8) takes an especially simple form:

$$|f|^p_{B^\alpha_p(L_p(\mathbb{R}^d))} \approx \sum_{k\in\mathbb{Z}} [2^{k\alpha} \sum_{\psi\in\Psi} \sum_{j\in\mathbb{Z}^d} |\gamma_{j,k,\psi,p}(f)|^p]. \tag{5.9}$$

In particular, (5.9) gives an equivalent seminorm for the Sobolev space $H^\alpha(\mathbb{R}^d)$ by taking $p = 2$.

6. Applications

6.1. Wavelet compression. We shall present a few examples that indicate how wavelets can be used in numerical applications. Wavelet techniques have had a particularly significant impact on data compression. We begin by discussing a problem in nonlinear approximation that is at the heart of compression algorithms.

Suppose that Ψ is a (pre)wavelet set and $f \in L_p(\mathbb{R}^d)$, $1 \le p \le \infty$, has the wavelet representation

$$f = \sum_{k\in\mathbb{Z}} \sum_{\psi\in\Psi} \sum_{j\in\mathbb{Z}^d} c_{j,k,\psi,p}(f)\psi_{j,k,p}. \tag{6.1.1}$$

with respect to the $L_p(\mathbb{R}^d)$-normalized functions $\psi_{j,k,p} := 2^{kd/p}\psi(2^k\cdot - j)$. In numerical applications, we must replace the sum in (6.1.1) by a finite sum, and the question arises as to the most efficient way to accomplish this. To make this into a well defined mathematical problem, we fix an integer n, which represents the number of terms we shall allow in the finite sum. Thus, we want to approximate f in the $L_p(\mathbb{R}^d)$ norm by an element from the set

$$\Sigma_n := \Big\{ S = \sum_{(j,k,\psi)\in\Lambda} d_{j,k,\psi}\psi_{j,k,p} \mid |\Lambda| \le n \Big\}, \tag{6.1.2}$$

where $d_{j,k,\psi}$ are arbitrary complex numbers. We have the error of approximation

$$\sigma_n(f)_p := \inf_{S\in\Sigma_n} \|f - S\|_{L_p(\mathbb{R}^d)}. \tag{6.1.3}$$

In contrast to the usual problems in approximation, the set Σ_n is not a linear space since adding two elements of Σ_n results in an element of Σ_{2n}, but not generally an element of Σ_n.

The approximation problem (6.1.3) has a particularly simple solution when $p = 2$ and Ψ is an orthogonal wavelet set. We order the coefficients $c_{j,k,\psi,2}(f)$ by their absolute value. If Λ_n is the set of indices (j, k, ψ) corresponding to n largest values (this set is not necessarily unique), then $S_n = \sum_{(j,k,\psi)\in\Lambda_n} c_{j,k,\psi,2}(f)\psi_{j,k,2}$ attains the infimum in (6.1.3). For prewavelet sets, this selection is optimal within constants, as follows from the $L_2(\mathbb{R}^d)$-stability of the basis $\psi_{j,k,2}$.

It is somewhat surprising that the strategy of the preceding paragraph also is optimal (in a sense to be made clear below) for approximation in $L_p(\mathbb{R}^d)$, $p \ne 2$.

To describe this, we fix a value of $1 < p < \infty$ (slightly weaker results than those stated below are known when $p = \infty$) and for $f \in L_p(\mathbb{R}^d)$, we let Λ_n denote a set of indices corresponding to n largest values of $|c_{j,k,\psi,p}(f)|$. We define $S_n := \sum_{(j,k,\psi)\in\Lambda_n} c_{j,k,\psi,p}(f)\psi_{j,k,p}$ and $\tilde{\sigma}_n(f)_p := \|f - S_n\|_{L_p(\mathbb{R}^d)}$. DeVore, Jawerth, and Popov [DJP] have established various results that relate $\sigma_n(f)_p$ with $\tilde{\sigma}_n(f)_p$ under certain conditions on ψ and the generating function ϕ. For example, it follows from their results that

$$\tilde{\sigma}_n(f)_p = \mathcal{O}(n^{-\alpha/d}) \Longleftrightarrow \sigma_n(f)_p = \mathcal{O}(n^{-\alpha/d}) \tag{6.1.4}$$

for $0 < \alpha < r$. Here, the integer r is related to properties of ϕ. Namely, the generating function ϕ should satisfy the Strang-Fix conditions (5.4) of this order and ϕ and ψ should have sufficient smoothness (for example, C^r is enough). It is also necessary to assume decay for the functions ϕ and ψ; sufficiently fast polynomial decay is enough. We caution the reader that the results in [DJP] are formulated for one wavelet ψ and not a wavelet set Ψ. However, the same proofs apply in the more general setting.

It is also of interest to characterize the functions f that satisfy (6.1.4). That is, we would like to know when we can expect the order of approximation (6.1.4). This has not been accomplished in exactly the form of (6.1.4), but the following variant has been shown in [DJP]. The following are equivalent for $\tau := \tau(\alpha, p) := (\alpha/d+1/p)^{-1}$:

$$\begin{aligned} &\text{(i)} \quad \sum_{n=1}^{\infty} [n^{\alpha/d}\sigma_n(f)_p]^\tau \frac{1}{n} < \infty, \\ &\text{(ii)} \quad \sum_{n=1}^{\infty} [n^{\alpha/d}\tilde{\sigma}_n(f)_p]^\tau \frac{1}{n} < \infty, \\ &\text{(iii)} \quad f \in B_\tau^\alpha(L_\tau(\mathbb{R}^d)). \end{aligned} \tag{6.1.5}$$

Several words of explanation regarding (6.1.5) are in order. First, (6.1.5)(i) is very close to the condition in (6.1.4). For example, (6.1.5)(i) implies (6.1.4), and if (6.1.4) holds for some $\beta > \alpha$ then (6.1.5)(i) is valid. So, roughly speaking, it is the functions $f \in B_\tau^\alpha(L_\tau(\mathbb{R}^d))$ for which the order of approximation in (6.1.4) holds. Secondly, the characterization (6.1.5) says that it is those functions with smoothness of order α in $L_\tau(\mathbb{R}^d)$, $\tau = (\alpha/d + 1/p)^{-1}$, that are approximated with order $\mathcal{O}(n^{-\alpha/d})$ in $L_p(\mathbb{R}^d)$. This should be contrasted with the usual results for approximation from linear spaces (such as finite element methods), which characterize functions with this approximation order as having smoothness of order α in $L_p(\mathbb{R}^d)$. Since $\tau < p$, the nonlinear approximation problem (6.1.3) provides the approximation order (6.1.4) for functions with less smoothness than required by linear methods.

The fact that functions with less smoothness can be approximated well by (6.1.3) is at the essence of wavelet compression. This means that functions with singularities can be handled numerically. Intuitively this is accomplished by retaining in the sum for S_n terms corresponding to functions $\psi_{j,k,p}$ that make a large contribution to f near the singularity. Here the situation is similar to adaptive methods for piecewise polynomial (finite element) approximation that have refined triangulations near a singularity. However, we want to stress that in wavelet compression, it is simple (almost trivial) to approximate optimaly without encountering problems of triangulation. An overview of this approach to data compression using wavelets can be found in [DJL2].

6.2. Image compression. We explain next how a typical algorithm for compression is implemented from the theoretical results of the previous section. This has been accomplished by DeVore, Jawerth, and Lucier for surface compression in [DJL] and image compression in [DJL1]. We shall discuss only image compression.

A digitized grey-scale image consists of an array of picture elements (pixels) represented by numbers that correspond to the brightness (grey scale) of each pixel, with $0 \equiv$ black, say, and $255 \equiv$ white. A grey-scale image has, say, 1024×1024 such numbers taking integer values between 0 and 255. Thus, the image is given by a matrix $(p_j)_{j \in \{0,\dots,1023\}^2}$ with $p_j \in \{0, \dots, 255\}$. It would take a data file of about one million bytes to encode such an image. For purposes of transmission or storage, it is desirable to compress this file.

To use wavelets for image compression, we proceed as follows. We think of the pixel values as associated with the points $j2^{-m}$, $j \in [0, 2^m)^2$, $m = 10$, of the unit square $[0, 1]^2$. In this way, we can think of the image as a discretization of a function f defined on this square.

We choose a function ϕ satisfying the assumptions of multiresolution, and a corresponding wavelet set Ψ that provides a stable basis for $L_2(\mathbb{R}^2)$. Thus, Ψ would consist of three functions, which we shall assume are of compact support.

We would like to represent the image as a wavelet sum. For this purpose, we select coefficients γ_j and consider the function

$$f = \sum_{j \in \Omega} \gamma_j \phi_{j,m} \tag{6.2.1}$$

with Ω the set of indices for which $\phi_{j,m}$ does not vanish identically on $[0, 1]^2$. We think of f as the image and apply the results of the preceding section to compress f.

The coefficients γ_j are to be determined numerically from the pixel values; choosing good values of γ_j is a nontrivial problem, which we do not discuss. A typical choice is to take $\gamma_j = p_j$ for $j2^{-m} \in [0, 1]^2$ and some extension of these values for other j.

Once the coefficients (γ_j) have been determined, we use a fast wavelet transform to write f in its wavelet decomposition

$$f = P_0 f + \sum_{j,k,\psi} c_{j,k,\psi}(f) \psi_{j,k} \tag{6.2.2}$$

with respect to the $L_2(\mathbb{R}^2)$-normalized $\psi_{j,k} := 2^k \psi(2^k \cdot - j)$. We can find the coefficients of f with respect to the $L_p(\mathbb{R}^2)$-normalized ψ's by the relation $c_{j,k,\psi,p} = 2^{2k(1/p-1/2)} c_{j,k,\psi}$. The projection $P_0 f$ has very few terms, which we take intact into the compressed representation at little cost.

To apply the compression algorithm of the previous section, we need to decide on a suitable norm in which to measure the error. The $L_2(\mathbb{R}^2)$ norm is most commonly used, but we argue in [DJL1] that the $L_1(\mathbb{R}^2)$ is a better model for the human eye for error with high spatial frequency.

If one decides to use the $L_p(\mathbb{R}^2)$ norm to measure compression error, then the algorithm of the previous section orders the $L_p(\mathbb{R}^2)$-normalized wavelet coefficients $c_{j,k,\psi,p}$ and chooses the largest of these to keep. Optimally, one would send coefficients in decreasing order of size. Thus, we find a (small) set Λ of ordered triples $\{(j,k,\psi)\}$ that index the largest values of $|c_{j,k,\psi,p}|$ and use for our compressed image

$$g := \sum_{(j,k,\psi)\in\Lambda} c_{j,k,\psi,p}\psi_{j,k,p}.$$

This sum has $|\Lambda|$ terms.

This method of sending coefficients sequentially across a communications link to allow gradual reconstruction of an image by the receiver is known as *progressive transmission.* Our criterion provides a new ordering for the coefficients to be transmitted that depends on the value of p. However, sorting the coefficients requires $\mathcal{O}(m2^{2m})$ operations, while the fast wavelet transform itself takes but $\mathcal{O}(2^{2m})$ operations. Thus, a faster compression method that does not rely on sorting is to be preferred; we proceed to give one.

We discuss compression in $L_2(\mathbb{R}^2)$ for a moment. As noted before, the optimal algorithm is to keep the largest $L_2(\mathbb{R}^2)$ coefficients and to discard the other coefficients. The coefficients to keep can be determined by fixing any the following quantities:

$$\begin{aligned} &\text{(i)} \quad N := |\Lambda|, \\ &\text{(ii)} \quad \|f-g\|_{L_2(\mathbb{R}^2)} = \Bigg(\sum_{(j,k,\psi)\notin\Lambda} |c_{j,k,\psi}|^2\Bigg)^{1/2}, \\ &\text{(iii)} \quad \epsilon := \inf_{(j,k,\psi)\in\Lambda} |c_{j,k,\psi}|. \end{aligned}$$

Setting any one of these quantities determines the other two, and by extension the set Λ, for any function f. In other words, we can prescribe either the number of nonzero coefficients N, the total error $\|f-g\|_{L_2(\mathbb{R}^2)}$, or ϵ, which we consider to be a measure of the local error. If we determine which triples (j,k,ψ) to include in Λ by the third criterion, then we do not need to sort the coefficients, for we can sequentially examine each coefficient and put (j,k,ψ) into Λ whenever $|c_{j,k,\psi}| \geq \epsilon$. This is known as *threshold coding* to the engineers, because one keeps only those coefficients that exceed a specified threshold.

Even more compression can be achieved by noting that we should keep only the most significant bits of the coefficients $c_{j,k,\phi,p}$. Thus, we choose a tolerance $\epsilon > 0$ and we take in the compressed approximation for each $\psi_{j,k,p}$ a coefficient $\tilde{c}_{j,k,\psi,p}$ such that

$$|\tilde{c}_{j,k,\psi,p} - c_{j,k,\psi,p}| < \epsilon. \tag{6.2.3}$$

with the proviso that $\tilde{c}_{j,k,\psi,p} = 0$ whenever $|c_{j,k,\psi,p}| < \epsilon$. Then $\tilde{S} := \sum_{j,k,\psi} \tilde{c}_{j,k,\psi,p}\psi_{j,k,p}$ represents our compressed image, and (6.1.5) holds for this approximation.

This process of keeping only the most significant bits of $c_{j,k,\psi,p}$ is known in the engineering literature as *scalar quantization.* The dependence on the dyadic level k and the space $L_p([0,1]^2)$ in which the error is measured is brought out more clearly when using $L_\infty(\mathbb{R}^2)$-normalized wavelets, i.e., when $\psi_{j,k} = \psi(2^k\cdot - j)$. For these wavelets, as the dyadic level increases, the number of bits of $c_{j,k,\psi}$ taken in $\tilde{c}_{j,k,\psi}$ decreases. For example, if $p = 2$, we would take one less bit at each increment of the dyadic level. On the other hand, if the compression is done in the $L_1([0,1]^2)$ norm, than we would take two fewer bits as we change dyadic levels. See [DJL1].

6.3. The numerical solution of partial differential equations. Wavelets are currently being investigated for the numerical solution of differential and integral equations (see, e.g., the papers of Beylkin, Coifman, and Rokhlin [BCR] and Jaffard [Ja]). While these applications are only now being developed, we shall consider a couple of simple examples to illustrate the potential of wavelets in this direction.

Elliptic equations. The Galerkin method applied to elliptic partial differential equations gives rise to a matrix problem that involves a so-called stiffness matrix. We present a simple example that illustrates the perhaps surprising fact that the stiffness matrix derived from the wavelet basis can be preconditioned trivially to have a uniformly bounded condition number. In general, this property allows one to use iterative methods, such as the conjugate gradient method, to solve linear systems with great efficiency. The linear systems that arise by discretizing elliptic PDEs have a lot of structure and can in no way be considered general linear systems, and there are many very efficient numerical methods, such as multigrid, that exploit the special structure of these linear systems to solve these systems to high accuracy with very low operation counts. We do not yet know of a complete analysis that shows that computations with wavelets can be more efficient than existing multigrid methods when applied to the linear systems that arise by discretizing elliptic PDEs in the usual way.

Rather than consider Dirichlet and Neumann boundary value problems in several space dimensions, as discussed in [Ja], we shall present only a simple, periodic, second order ODE that illustrates the main points. We shall consider functions defined on the one-dimensional torus $\mathbb{T}$, which is equivalent to $[0,1]$ with the end-points identified, and search for $u = u(x)$, $x \in \mathbb{T}$, that satisfies the equation

$$-u''(x) + u(x) = f(x), \qquad x \in \mathbb{T}, \tag{6.3.1}$$

with $f \in L_2(\mathbb{T})$. In variational form, the solution $u \in W^1(L_2(\mathbb{T}))$ of (6.3.1) satisfies

$$\int_{\mathbb{T}} (u'v' + uv) = \int_{\mathbb{T}} fv, \tag{6.3.2}$$

for all $v \in W^1(L_2(\mathbb{T}))$. We remark that we can take

$$\|u\|^2_{W^1(L_2(\mathbb{T}))} := \int_{\mathbb{T}} ([u']^2 + u^2).$$

To approximate u by Galerkin's method, we must choose a finite-dimensional subspace of $W^1(L_2(\mathbb{T}))$, which we shall choose to be a space spanned by wavelets defined on the circle $\mathbb{T}$. We indicate briefly how to construct periodic wavelets on $\mathbb{T}$ from wavelets on $\mathbb{R}$.

For any f with compact support defined on $\mathbb{R}$, the function

$$f^\circ := \sum_{j\in\mathbb{Z}} f(\cdot + j)$$

is a function of period one, which we call the periodization of f. To obtain wavelets on $\mathbb{T}$, we apply this periodization to wavelets on $\mathbb{R}$. To be specific, we consider only the Daubechies wavelets $\psi := \mathcal{D}_{2N}$ with $N > 2$ because they are contained in $W^1(L_2(\mathbb{T}))$. (The first nontrivial Daubechies wavelet, $\mathcal{D}_4$, is in $W^\alpha(L_2(\mathbb{T}))$ for all $\alpha < 1$ (see [E]), but we do not know if it is in $W^1(L_2(\mathbb{T}))$.) Let ϕ be the function of §3.5 that gives rise to ψ. For each $k \in \mathbb{Z}$, we let $\phi^\circ_{j,k}$ and $\psi^\circ_{j,k}$ be the periodization of the functions $\phi_{j,k}$ and $\psi_{j,k}$ respectively. We define $\tilde{\mathcal{S}}^k$ to be the linear span of the functions $\phi^\circ_{j,k}$, $j = 0, \ldots, 2^k - 1$. The functions in this space are clearly of period one. We also note that the $\phi^\circ_{j,k}$ are orthogonal. Indeed, because they are periodic,

$$\begin{aligned}\int_0^1 \phi^\circ_{j,k}\phi^\circ_{j',k} &= \sum_{\ell\in\mathbb{Z}} \int_0^1 \phi_{j,k}(\,\cdot\, + \ell)\phi^\circ_{j',k} = \sum_{\ell\in\mathbb{Z}} \int_0^1 \phi_{j,k}(\,\cdot\, + \ell)\phi^\circ_{j',k}(\,\cdot\, + \ell)\\ &= \int_{\mathbb{R}} \phi_{j,k}\phi^\circ_{j',k} = \sum_{\ell\in\mathbb{Z}} \int_{\mathbb{R}} \phi_{j,k}\phi_{j',k}(\,\cdot\, + \ell) = \sum_{\ell\in\mathbb{Z}} \int_{\mathbb{R}} \phi_{j,k}\phi_{j'+\ell 2^k,k}.\end{aligned}$$

If $j \neq j'$, each integral in the last sum is zero, since we never have $j = j' + 2^k\ell$. If $j = j'$, then exactly one integral in the last sum is nonzero and its value is one. Similarly, we find that $\psi^\circ_{j,k}$, $j = 0, \ldots, 2^k - 1$, $k \geq 0$, is an orthonormal system for $L_2(\mathbb{T})$. It is easy to check that by adjoining $\phi^\circ_{0,0}$, which is identically one on $\mathbb{T}$, this orthonormal system is complete.

Returning to our construction of periodic wavelet spaces, we define $\tilde{W}^k$ to be the linear span of the functions $\psi^\circ_{j,k}$, $j = 0, \ldots, 2^k - 1$. Then $\tilde{\mathcal{S}}^{k+1} = \tilde{\mathcal{S}}^k \oplus \tilde{W}^k$. To simplify our notation in this section, in which we refer to periodic wavelets only, we drop the superscripts and tildes, and denote by $\phi_{j,k}$, $\psi_{j,k}$, $\mathcal{S}^k$, and W^k, the periodic wavelet bases and spaces.

Returning to the numerical solution of (6.3.1), we choose a positive value of m and approximate u by an element $u_m \in \mathcal{S}^m$ that satisfies:

$$\int_{\mathbb{T}} (u_m' v' + u_m v) = \int_{\mathbb{T}} f v, \quad v \in \mathcal{S}^m. \tag{6.3.3}$$

We can write u_m and v as in (6.2.2). For example,

$$u_m = P_0 u_m + \sum_{k=0}^{m-1} Q_k u_m = \gamma\phi_{0,0} + \sum_{k=0}^{m-1}\sum_{j=0}^{2^k-1} \beta(j,k)\psi_{j,k}. \tag{6.3.4}$$

Because $\phi_{0,0}$ is identically one on the circle $\mathbb{T}$, γ is just the average of u_m on $\mathbb{T}$. From (6.3.3) we see that with $v \equiv 1$, $\gamma = \int_{\mathbb{T}} f$. Thus, we need to determine $\beta(j,k)$.

If we replace u_m in (6.3.3) by its representation (6.3.4), we arrive at a system of equations

$$\sum_{k=0}^{m-1}\sum_{j=0}^{2^k-1}\beta(j,k)\int_{\mathbb{T}}(\psi'_{j,k}\psi'_{j',k'}+\psi_{j,k}\psi_{j',k'})=\int_{\mathbb{T}}f\psi_{j',k'},$$

for $j'=0,\dots,2^{k'}-1$ and $k'=0,\dots,m-1$, or, more succinctly,

$$T\boldsymbol{\beta}=\boldsymbol{f} \tag{6.3.5}$$

where the typical entry in T is $\int_{\mathbb{T}}(\psi'_{j,k}\psi'_{j',k'}+\psi_{j,k}\psi_{j',k'})$ and $\boldsymbol{f}$ a vector with components $\int_{\mathbb{T}}f\psi_{j',k'}$; $\boldsymbol{\beta}=(\beta(j,k))_{j=0,\dots,2^k-1,\ k=0,\dots,m-1}$ is the coefficient vector of the unknown function.

The convergence rate of the conjugate gradient method, a popular iterative method for the solution of systems like (6.3.5), depends on the condition number, $\kappa(T)=\|T\|\|T^{-1}\|$, of the symmetric, positive definite, stiffness matrix T. Now,

$$\|T\|=\sup_{\|\boldsymbol{\alpha}\|_{\ell^2}=1}\boldsymbol{\alpha}^*T\boldsymbol{\alpha}\text{ and }\|T^{-1}\|^{-1}=\inf_{\|\boldsymbol{\alpha}\|_{\ell^2}=1}\boldsymbol{\alpha}^*T\boldsymbol{\alpha},$$

where $\boldsymbol{\alpha}:=(\alpha(j,k))_{j=0,\dots,2^k-1,\ k=0,\dots,m-1}$. For any vector $\boldsymbol{\alpha}$, we form the function $S:=S(\boldsymbol{\alpha})\in\mathcal{S}^m$ by

$$S=\sum_{k=0}^{m-1}\sum_{j=0}^{2^k-1}\alpha(j,k)\psi_{j,k}.$$

For example, $u_m=\gamma\phi_{0,0}+S(\boldsymbol{\beta})$. It follows easily that

$$|\boldsymbol{\alpha}^*T\boldsymbol{\alpha}|=\int_0^1([S']^2+S^2)=\|S\|^2_{W^1(L_2(\mathbb{R}^d))}\approx\sum_{k=0}^{m-1}\sum_{j=0}^{2^k-1}|2^k\alpha(j,k)|^2. \tag{6.3.6}$$

The last equivalence in (6.3.6) is a variant of (5.9) (for $q=p=2$). In (5.9), we considered wavelet representations beginning at the dyadic level $k=-\infty$; we could just as easily have begun at the level $k=0$ (by including P_0) and arrived at the equivalence in (6.3.6).

Equation (6.3.6) shows that the matrix DTD, where D has entries $2^{-k}\delta^j_{j'}\delta^k_{k'}$, satisfies

$$|\boldsymbol{\alpha}^*DTD\boldsymbol{\alpha}|\approx\sum_{k=0}^{m-1}\sum_{j=0}^{2^k-1}|\alpha(j,k)|^2;$$

i.e., the stiffness matrix T can be trivially preconditioned to have a condition number $\kappa(DTD):=\|DTD\|\|D^{-1}T^{-1}D^{-1}\|$ that is bounded independently of m.

Because the subspace $\mathcal{S}^m$ of $L_2(\mathbb{T})$ generated by the Daubechies wavelet D_{2N} locally contains all polynomials of degree $<N$, we have that

$$\inf_{v\in\mathcal{S}^m}\|u-v\|_{W^1(L_2(\mathbb{T}))}\le C2^{-m(N-1)}|u|_{W^N(L_2(\mathbb{T}))}.$$

Because of the special form of (6.3.1), u_m is in fact the $W^1(L_2(\mathbb{T}))$ projection of u onto S^m, so we have

$$\|u - u_m\|_{W^1(L_2(\mathbb{T}))} \leq C 2^{-m(N-1)} |u|_{W^N(L_2(\mathbb{T}))}.$$

One can treat in an almost identical way the equation (6.3.1) defined on the torus $\mathbb{T}^d$ with the left side replaced by $-\nabla \cdot (a \nabla u) + bu$ with a and b bounded, smooth, positive functions on $\mathbb{T}^d$, and obtain the same uniform bound on the condition number. Of course, ultimately, one would like to handle elliptic boundary value problems for a domain Ω. First results in this direction have been obtained by Jaffard [Ja] and in a slightly different (nonwavelet) setting by Oswald (see for example [O]). For example, Jaffard's approach to elliptic equations with Dirichlet boundary conditions is to transform the equation to one with zero boundary conditions by extending the boundary function into the interior of the domain Ω. He then employs in a Galerkin method wavelets whose support is contained strictly inside Ω. However, there has not yet been an analysis of the desired relationship between the extension and the wavelets employed. Another approach is to develop wavelets for the given domain (which do not vanish on the boundary) (see [JaMe]).

Employing wavelets for elliptic problems as outlined above is similar to the use of hierarchical bases in the context of multigrid. However, two points suggest that wavelet bases may be more useful than hierarchical bases. First, the wavelet bases are $L_2(\mathbb{R}^d)$-stable, while the hierarchical basis are not. Second, one can choose from a much greater variety of wavelet bases with various approximation properties (values of N).

Finally, we mention the great potential for compression to be used in conjunction with wavelets for elliptic problems in a similar way to adaptive finite elements (see also [Ja]).

6.4. Time dependent problems. Wavelets have potential application for the numerical solution of time-dependent parabolic and hyperbolic problems. We mention one particular application where the potential of wavelets has at least theoretical foundation.

We consider the solution $u(x,t)$ of the scalar hyperbolic conservation law

$$\begin{aligned} &u_t + f(u)_x = 0, \qquad x \in \mathbb{R}, \quad t > 0, \\ &u(x,0) = u_0(x), \qquad x \in \mathbb{R}, \end{aligned} \tag{6.4.1}$$

in one space dimension. It is well known that the solution to (6.4.1) develops discontinuities (called "shocks") even when the initial condition is smooth. This makes the numerical treatment of (6.4.1), and even more so its analogue in several space dimensions, somewhat subtle. The appearance of shocks calls for adaptive or nonlinear methods.

Considering the appearance of discontinuities in the solution u to (6.4.1), the following regularity result of the authors [DL], [DL1], is quite surprising. If the flux f in (6.4.1) is strictly convex and suitably smooth, and u_0 has bounded variation, then it has been shown that for any $\alpha > 0$ and $\tau := \tau(\alpha) := (\alpha+1)^{-1}$,

$$u_0 \in B^\alpha_\tau(L_\tau(\mathbb{R})) \Longrightarrow u(\,\cdot\,,t) \in B^\alpha_\tau(L_\tau(\mathbb{R})), \tag{6.4.2}$$

for all later time $t > 0$. That is, if u_0 has smoothness of order α in $L_\tau(\mathbb{R})$ then so will $u(\cdot, t)$ for all later time $t > 0$.

The regularity result (6.4.2) has an interpretation in terms of wavelet decompositions. We cannot use orthogonal wavelets in these decompositions because, as we pointed out earlier, they do not provide stable representations of functions in $L_1(\mathbb{R})$ because they have mean value zero. However, the characterization of Besov spaces can be carried out using other, nonorthogonal, wavelets, such as B-splines. With this caveat, (6.4.2) says that whenever u_0 has a wavelet decomposition $\sum_{j,k\in\mathbb{Z}} \gamma_{j,k}(u_0)\psi_{j,k,1}$, with certain $L_1(\mathbb{R})$-normalized wavelets ψ, whose coefficients satisfy

$$\sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} |\gamma_{j,k}(u_0)|^\tau < \infty, \tag{6.4.3}$$

then $u(\cdot, t)$ has a similar wavelet decomposition with the same control (6.4.3) on the coefficients. We want to stress that the results in [DL] and [DL1] do not describe directly how to determine wavelet coefficients at later time $t > 0$ from those of the initial function u_0. That is, there is no direct, theoretically correct, numerical method known to us that describes how to update coefficients with time so that (6.4.3) holds. The regularity result (6.4.2) is proved by showing that whenever u_0 can be approximated well in $L_1(\mathbb{R})$ by piecewise polynomials with free (variable) knots, then $u(\cdot, t)$ can be approximated in the same norm by piecewise algebraic functions (of a certain type) with free knots.

Finally, we mention that the authors have also shown that the analogue of regularity result (6.4.3) does not hold in more than one space dimension. From the point of view of wavelet decompositions, our results seem to indicate that the wavelets described in this presentation are too symmetric to effectively handle the diverse types of singularities that arise in the solution of conservation laws in several space dimensions.

References

[Ba] G. Battle, *A block spin construction of ondelettes, Part I: Lemarié functions*, Comm. Math. Phys. **110** (1987), 601–615.

[B-AR] A. Ben Artzi and A. Ron, *On the integer translates of a compactly supported function: Dual bases and linear projectors*, SIAM J. Math. Anal. **21** (1990), 1550–1562.

[BCR] G. Beylkin, R. Coifman, and V. Rokhlin, *Fast wavelet transforms and numerical algorithms I*, Comm. Pure and Appl. Math. **XLIV** (1991), 141–183.

[BDR] C. de Boor, R. DeVore and A. Ron, *Approximation from shift invariant spaces*, preprint.

[BDR1] C. de Boor, R. DeVore and A. Ron, *Wavelets and prewavelets*, preprint.

[BDR2] C. de Boor, R. DeVore and A. Ron, *The structure of finitely generated shift invariant spaces in $L_2(\mathbb{R}^d)$*, preprint.

[BR] C. de Boor and A. Ron, *Fourier analysis of the approximation power of principal shift-invariant spaces*, Constructive Approximation **8** (1992), 427–462.

[CDM] A. Cavaretta, W. Dahmen and C. Micchelli, *Subdivision for Computer Aided Geometric Design*, Memoirs Amer. Math. Soc. **93** (1991).

[CSW] C. K. Chui, J. Stöckler and J. D. Ward, *Compactly supported box-spline wavelets*, CAT Report, Texas A& M University **230** (1991).

[CW] C. K. Chui and J. Z. Wang, *On compactly supported spline wavelets and a duality principle*, Trans. Amer. Math. Soc. (1991).

[CW1] C. K. Chui and J. Z. Wang, *A general framework for compactly supported splines and wavelets*, CAT Report, Texas A&M University **219** (1990).

[DM] W. Dahmen and C. Micchelli, *On stationary subdivision and the construction of compactly supported wavelets*, Multivariate Approximation and Interpolation (W. Haussmann and K. Jetter, eds.), ISNM vol. 94, Birkhauser Verlag, 1990, pp. 69–89.

[Da] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. on Pure and Appl. Math. **XLI** (1988), 909–996.

[Da1] I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Trans. on Information Theory **36** (1990), 961–1005.

[DJL] R. DeVore, B. Jawerth, and B. Lucier, *Surface compression*, Computer Aided Geometric Design **9** (1992), 219–239.

[DJL1] R. DeVore, B. Jawerth, and B. Lucier, *Image compression through wavelet transform coding*, IEEE Transactions on Information Theory **38** (1992), 719–746, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis.

[DJL2] R. DeVore, B. Jawerth, and B. Lucier, *Data compression using wavelets: Error, smoothness, and quantization*, DCC-91, Data Compression Conference (J. A. Storer and J. H. Reif, eds.), IEEE Computer Society Press, Los Alamitos, CA, 1991, pp. 186–195.

[DJP] R. DeVore, B. Jawerth, and V. Popov, *Compression of wavelet decompositions*, Amer. J. Math. (to appear).

[DeLo] R. DeVore and G. Lorentz, *Constructive Approximation*, Springer Grundlehren, Heidelberg, 1992.

[DL] R. DeVore and B. Lucier, *High order regularity for conservation laws*, Indiana Univ. Math. J. **39** (1990), 413–430.

[DL1] R. DeVore and B. Lucier, *High order regularity for solutions of the inviscid Burgers equation*, Nonlinear Hyperbolic Problems (Proceedings of an Advanced Research Workshop, Bordeaux, France, June 1988, Springer Lecture Notes in Mathematics, vol. 1402) (C. Carasso, P. Charrier, B. Hanouzet, J.-L. Joly, eds.), Springer-Verlag, New York, 1989, pp. 147–154.

[DP] R. DeVore and V. Popov, *Interpolation spaces and non-linear approximation*, Function Spaces and Applications (M. Cwikel, J. Peetre, Y. Sagher, and H. Wallin, eds.), Springer Lecture Notes in Mathematics, vol. 1302, Springer, New York, 1988, pp. 191–205.

[E] T. Eirola, *Sobolev characterization of solutions of dilation equations*, SIAM J. Math. Anal. (to appear).

[FJ] M. Frazier and B. Jawerth, *A discrete transform and decompositions of distribution spaces*, J. of Functional Anal. **93** (1990), 34–170.

[FJW] M. Frazier, B. Jawerth, and G. Weiss, *Littlewood-Paley Theory and the Study of Function Spaces*, Conference Board of the Mathematical Sciences, Regional Conference Series in Mathematics, Number 79, American Mathematical Society, Providence, RI, 1991.

[Ja] S. Jaffard, *Wavelet methods for fast resolution of elliptic problems*, preprint.

[JaMe] S. Jaffard and Y. Meyer, *Bases ondelettes dans des ouverts de* $\mathbb{R}^n$, Jour. Math. Pures et Appl. **68** (1989), 95–108.

[JM] R-Q. Jia and C. Micchelli, *Using the refinement equation for the construction of pre-wavelets II: power of two*, Curves and Surfaces (P. J. Laurent, A. LeMéhauté, and L. Schumaker, eds.), Academic Press, New York, 1991, pp. 209–246.

[KS] S. Karlin and W. Studden, *Tchebycheff Systems with Application to Analysis and Statistics*, Wiley Interscience, New York, 1966.

[LM] R. A. H. Lorentz and W. R. Madych, *Wavelets and generalized box splines*, preprint.

[Ma] S. G. Mallat, *Multiresolution approximations and wavelet orthonormal bases of* $L^2(\mathbb{R})$, Trans. Amer. Math. Soc **315** (1989), 69–87.

[Me] Y. Meyer, *Ondelettes et Opérateurs I: Ondelettes*, Hermann, 1990.

[Me1] Y. Meyer, *Wavelets with compact support*, Zygmund Lectures (University of Chicago), 1987.

[Mi] C. Micchelli, *Using the refinement equation for the construction of pre-wavelets*, Numerical Algorithms **1** (1991), 75–116.

[0] P. Oswald, *On discrete norm estimates related to multilevel preconditioners in the finite elment method*, preprint.

[RS] S. D. Riemenschneider and Z. W. Shen, *Box splines, cardinal splines, and wavelets*, Approximation Theory and Functional Analysis (C. K. Chui, ed.), Academic Press, New York, 1991, pp. 133–149.

[RS1] S. D. Riemenschneider and Z. W. Shen, *Wavelets and prewavelets in low dimensions*, J. Approx. Theory (to appear).

[RV] O. Rioul and M. Vetterli, *Wavelets and signal processing*, IEEE Signal Processing Magazine **8** (1991), no. 4, 14–38.

[Sch] I. Schoenberg, *Contributions to the problem of approximation of equidistant data by analytic functions, Parts A & B*, Quarterly Appl. Math. **IV** (1946), 45–99, 112–141.

[Sch1] I. Schoenberg, *Cardinal Spline Interpolation*, Regional Conference Series in Applied Mathematics, vol. 12, SIAM, Philadelphia, 1973.

[SF] G. Strang and G. Fix, *A Fourier analysis of the finite element variational method*, C.I.M.E. II Ciclo 1971, Constructive Aspects of Functional Analysis (G. Geymonat, ed.), 1973, pp. 793–840.

[St] J. O. Strömberg, *A modified Franklin system and higher-order spline systems on* $\mathbb{R}^n$ *as unconditional basis for Hardy spaces*, Conference in Harmonic Analysis in Honor of Antoni Zygmund (W. Beckner et al., ed., eds.), vol. Vol II, Wadsworth, 1981, pp. 475–493.

Department of Mathematics, University of South Carolina, Columbia, SC 29208

Department of Mathematics, Purdue University, West Lafayette, IN 47907