10. Daubechies, I., The wavelet transform, time-frequency localization and signal analysis, *IEEE Transactions on Information Theory* **36** (1990), 961-1005.
11. Daubechies, I., A. Grossman, and Y. Meyer, Painless non-orthogonal expansions, J. Math. Phys. **27** (1986), 1271-1283.
12. Delves, L. and J. Mohamed, *Computational Methods for Integral Equations*, Cambridge University Press, 1985.
13. Grossman, A. and J. Morlet, Decomposition of Hardy functions into square integrable wavelets of constant shape, SIAM J. Math. Anal. **15** (1984), 723-736.
14. Hochstadt, H., *Integral Equations*, Wiley, 1989.
15. Meyer, Y., Principe d'incertitude, bases Hilbertiennes et algèbres d'opérateurs, *Séminaire Bourbaki* **662** (1985-1986).
16. Meyer, Y., Ondelettes, functions splines, et analyses graduées, Univ. of Torino, 1986.
17. Mallat, S., Multiresolution approximation and wavelets, Trans. Amer. Math. Soc. **135** (1989), 69-88.
18. Schulz, G., Iterative berechnung der reziproken matrix, *Zeitschrift für Angewandte Mathematik und Mechanik* **13** (1933), 57-59.
19. Strang, G., Wavelets and Dilation Equations: A Brief Introduction, SIAM Review **31** (1989), 614-627.
20. Strang, G. and G. Fix, A Fourier analysis of the finite element variational method, in *Constructive Aspects of Functional Analysis*, Edizioni Cremonese, Rome, 1973.

Supported in part by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy under Contract DE-AC03-76SF-00098.

As mentioned above, similar techniques can be used to obtain higher derivatives and certain other differential and integral operators, including, for instance, the Hilbert transform. This development suggests that the evolution of a variety of differential equations in wavelet bases may become efficient, which would strengthen the arsenal of the numerical analyst attacking problems requiring highly adaptive schemes.

§5. Summary

We have illustrated the construction of wavelets and similar wavelet-like bases, their properties of orthogonality, approximation, compact support, and most distinguishingly, time-frequency localization through dilation invariance (and near-invariance). These bases lead to the sparse representation of integral operators and the rapid solution of integral equations. Differential operators are also represented as sparse matrices in wavelet bases, which permits the construction of adaptive algorithms for time-dependent partial differential equations.

References

1. Abramowitz, M. and I. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D.C., 1972.
2. Alpert, B., *Sparse Representation of Smooth Linear Operators*, Ph.D. thesis, Yale University Department of Computer Science Report #814, August, 1990.
3. Alpert, B., Rapidly-convergent quadratures for integral operators with singular kernels, Lawrence Berkeley Laboratory Report LBL-30091, December, 1990.
4. Alpert, B., Construction of simple multiscale bases for fast matrix operations, *Wavelets and Their Applications*, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, L. Raphael, and M. Ruskai, eds., Jones and Bartlett, 1991.
5. Alpert, B. and V. Rokhlin, A fast algorithm for the evaluation of Legendre expansions, SIAM J. Sci. Statist. Comput. **12** (1991), 158-179.
6. Alpert, B., G. Beylkin, R. Coifman, and V. Rokhlin, Wavelet bases for the fast solution of second-kind integral equations, Yale University Department of Computer Science Report #837, December, 1990. SIAM J. Sci. Statist. Comput. to appear.
7. Beylkin, G., On the representation of operators in compactly supported wavelet bases, preprint, August, 1990.
8. Beylkin, G., R. Coifman, and V. Rokhlin, Fast wavelet transforms and numerical algorithms I, *Communications on Pure and Applied Mathematics* **XLIV** (1991), 141-183.
9. Daubechies, I., Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics* **XLI** (1988), 909-996.

We apply the dilation equation (12) to equation (35) to obtain

$$\begin{aligned} c_k &= \int_{-\infty}^{\infty} \sum_l a_l \, \frac{d\phi}{dx}(2x-l)\, 2 \sum_m a_m \, \phi(2(x-k)-m)\, dx \\ &= \sum_l \sum_m a_l \, a_m \int_{-\infty}^{\infty} \frac{d\phi}{dx}(2x-l)\, \phi(2x-2k-m)\, 2dx \\ &= \sum_l \sum_m a_l \, a_m \, c_{2k+m-l}, \qquad k \in \mathbb{Z}. \end{aligned} \tag{36}$$

The scaling function $\phi = {}_N\phi$ is supported on the finite interval $[0, 2N-1]$. As a result, we see from (35) that $c_k = 0$ for $|k| \geq 2N-1$ (the integrand vanishes). Also, integration by parts yields $c_{-k} = -c_k$. Combining these two observations, (36) becomes a system of $2N-2$ equations in the $2N-2$ unknowns $c_1, \ldots, c_{2N-2}$. Due to its homogeneity, the rank of the system is deficient by one, and another equation is required to determine the scale of the c_k. The supplemental equation is obtained from the fact that the function $f(x) = x$ is a linear combination of translates of ϕ, for $N \geq 2$. In particular,

$$x = \sum_k \mu_k \phi(x-k), \tag{37}$$

where

$$\begin{aligned} \mu_k &= \int_{-\infty}^{\infty} \phi(x-k)\, x\, dx \\ &= \int_{-\infty}^{\infty} \phi(x-k)\, (x-k)\, dx + \int_{-\infty}^{\infty} \phi(x-k)\, k\, dx \\ &= \int_{-\infty}^{\infty} \phi(x)\, x\, dx + k \\ &= \mu_0 + k. \end{aligned}$$

Combining (37) with

$$1 = \sum_k \phi(x-k)$$

and differentiating by x yields

$$1 = \sum_k k \frac{d}{dx}\phi(x-k),$$

which in combination with (35) gives the desired supplemental equation

$$\sum_k k\, c_k = -1. \tag{38}$$

Equations (36) and (38) may be solved directly to obtain the coefficients $c_1, \ldots, c_{2N-2}$.

The technique of basis reordering for sparse LU factorization of integral operators, as described, works very well for one-dimensional problems, even outperforming the Schulz method. For numerical examples, the reader is referred to [4], [6]. It does not appear, however, that the method of factorization cleanly generalizes to higher dimensional problems, where the Schulz method is expected maintain its good performance.

4.5 Representation of Differential Operators

We have seen that integral operators can be expressed in wavelet bases as sparse matrices, to high precision. Certain differential operators, by contrast, are represented exactly by (infinite) sparse matrices. In [7] G. Beylkin determines the representations of various operators using Daubechies wavelets. Here we illustrate his technique for the derivative operator d/dx.

For a function represented as a wavelet expansion

$$f(x) = \sum_{n,k} \alpha_{nk}\, w_{nk}(x),$$

the derivative

$$\frac{df}{dx}(x) = \sum_{n,k} \alpha_{nk}\, \frac{d}{dx} w_{nk}(x)$$

is determined by the derivative of the basic wavelet function w. Through the definition (18) of w, its derivative is given in turn by the equation

$$\frac{dw}{dx}(x) = \sum_{k} b_k \frac{d}{dx}\phi(x-k).$$

Differentiation, therefore, is reduced to the determination of coefficients c_{nk} in the expansion for $d\phi/dx$,

$$\frac{d\phi}{dx}(x) = \sum_{n,k} c_{nk}\, w_{nk}(x),$$

where c_{nk} is given by the inner product

$$c_{nk} = \int_{-\infty}^{\infty} w_{nk}(x)\, \frac{d\phi}{dx}(x)\, dx, \qquad n,k \in \mathbb{Z}.$$

The application of the definitions of ϕ and w are used to reduce the coefficients c_{nk} to the coefficients

$$c_k = \int_{-\infty}^{\infty} \phi(x-k)\, \frac{d\phi}{dx}(x)\, dx, \qquad k \in \mathbb{Z}. \tag{35}$$

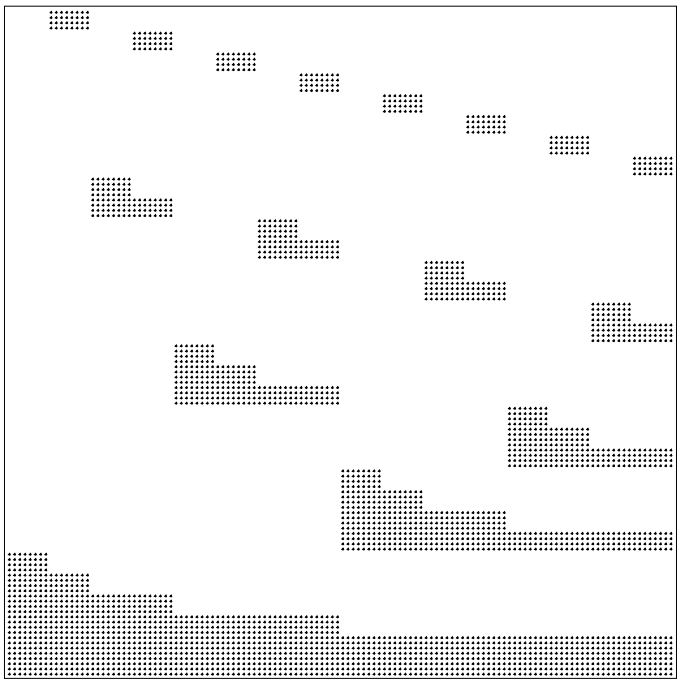

Figure 8. The matrix represents a reordering of the rows (basis vectors) of the matrix in Figure 4. In this order the basis is used to transform $\tilde{K}^n$ into a matrix supporting sparse LU factorization (shown below).

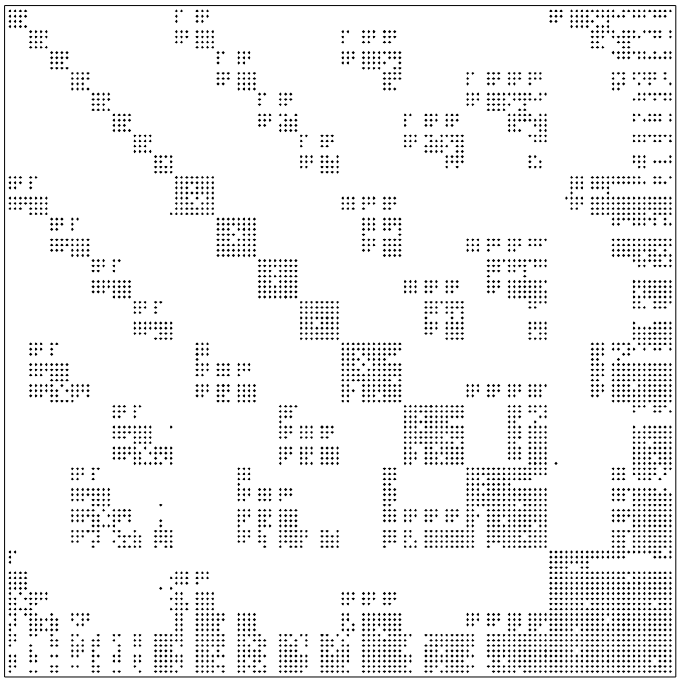

Figure 9. The non-negligible matrix elements ($\epsilon = 10^{-3}$) are shown for the integral operator with kernel $\log|x-y|$ expressed in the basis shown above. The matrix can be factored into lower- and upper-triangular matrices with no increase in the number of elements.

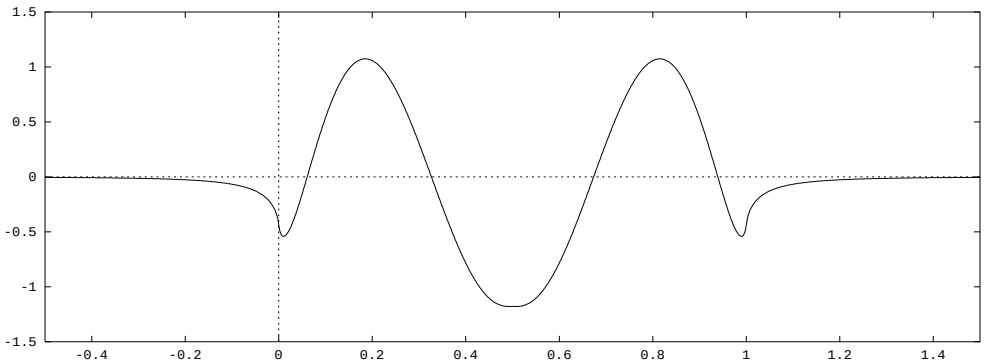

Figure 7. One of the multi-wavelet basis functions for $N = 4$ is convolved with $\log|x|$ and the image graphed. The function decays as the Nth power of the distance from its center, so effectively vanishes beyond the neighbors of the interval of support of the basis function.

From equation (32) we obtain $I - X_m A = (I - X_0 A)^{2^m}$, which in combination with equation (34) and simple manipulation yields bound (33). ■

The Schulz method is a notably simple scheme for matrix inversion and its convergence is extremely rapid. It is rarely used, however, because it involves matrix-matrix multiplications on each iteration; for most problem formulations, this process requires order $O(n^3)$ operations for an $n \times n$-matrix. As we have seen above, on the other hand, an integral operator A represented in a discrete wavelet-like basis has only $O(n)$ elements (to finite precision). In addition, $A^T A$ and $(A^T A)^m$ are similarly sparse. This property enables us to employ the Schulz algorithm to compute A^{-1} in order $O(n)$ operations.

Sparse LU Factorization. For dense matrices, computation of the inverse is almost never desirable. The decomposition into lower-triangular and upper-triangular (LU) factors requires roughly one third as many operations, and is equally useful. One might suppose that this advantage would also hold for sparse matrices with sparse inverses: perhaps it is possible to factor the sparse matrix $A = U\,\tilde{K}^n\,U^T$ into LU factors which are themselves sparse.

Unfortunately, direct factorization of A produces substantial fill-in of zero elements, and lower and upper triangular factors that are not sparse. This fill-in results from a "smearing" of the near-diagonal blocks. These blocks represent the interactions ${u_i}^T\,\tilde{K}^n\,u_j$ of basis elements u_i and u_j that are supported on adjacent intervals (see Figure 7).

Reordering the basis elements of U, we can construct a basis for which the elements are sorted into "levels" such that the basis elements on different intervals, but on one level, are separated from each other, and only interact with the elements of a single interval on each higher level. This ordering is illustrated in Figure 8. The reordered basis can then be used to transform the Nyström discretization $\tilde{K}^n$ of the integral operator into a sparse matrix lacking subdiagonal and superdiagonal blocks. In this form, shown in Figure 9, direct Gaussian elimination produces sparse lower- and upper-triangular factors.

4.4 Solution of Integral Equations

The representation of integral operators as sparse matrices, via transformation to wavelet coordinates, leads to new methods for the solution of integral equations. The integral equation (26), written in operator notation as

$$(\gamma + \mathcal{K})f = g,$$

has (formal) solution

$$f = (\gamma + \mathcal{K})^{-1} g.$$

The operator $(\gamma+\mathcal{K})^{-1}$ can be applied to g with the conjugate gradient method (conjugate residual if $A = \gamma + \mathcal{K}$ is nonsymmetric). This well-established method for sparse matrices is very fast if A is well conditioned. The number of iterations, which grows as the square root of the condition number (linearly in the condition number for conjugate residual), becomes rather large for poorly-conditioned problems.

Alternatively, one can directly invert A, obtaining a sparse inverse, or compute a sparse LU-factorization of A.

Schulz Method of Matrix Inversion. Schulz's method [18] is an iterative, quadratically convergent algorithm for computing the inverse of a matrix. Its performance is characterized as follows.

Lemma 2. *Suppose that A is an invertible matrix, X_0 is the matrix given by $X_0 = A^T/\|A^T A\|$, and for $m = 0, 1, 2, \ldots$ the matrix X_{m+1} is defined by the recursion*

$$X_{m+1} = 2X_m - X_m A X_m.$$

Then X_{m+1} satisfies the formula

$$I - X_{m+1}A = (I - X_m A)^2. \tag{32}$$

Furthermore, $X_m \to A^{-1}$ as $m \to \infty$ and for any $\epsilon > 0$ we have

$$\|I - X_m A\| < \epsilon \quad \textit{provided} \quad m \geq 2\log_2 \kappa(A) + \log_2 \log(1/\epsilon), \tag{33}$$

where $\kappa(A) = \|A\| \cdot \|A^{-1}\|$ is the condition number of A and the norm is given by $\|A\| =$ (largest eigenvalue of $A^T A)^{1/2}$.

Proof: Equation (32) is obtained directly from the definition of X_{m+1}. Bound (33) is equally straightforward. Noting that $A^T A$ is symmetric positive-definite and letting λ_0 denote the smallest and λ_1 the largest eigenvalue of $A^T A$ we have

$$\begin{aligned} \|I - X_0 A\| &= \left\| I - \frac{A^T A}{\|A^T A\|} \right\| \\ &= 1 - \lambda_0/\lambda_1 \\ &= 1 - \kappa(A)^{-2}. \end{aligned} \tag{34}$$

Use of the Nyström method was the primary motivation behind the development of the discrete wavelet-like bases.

4.2 Sparsity in Wavelet Bases

The example at the beginning of the section suggests that the number of basis functions required to represent the function $f_a(x) = \log|x-a|$ to precision ϵ is of order $O(\log(1/\epsilon))$. This is indeed the case for this function, as well as other functions analytic except at separated, integrable singularities. Since the Nyström discretization $\tilde{K}^n$ consists of columns with elements described by functions like f_a, one might expect that its transformation $U\,\tilde{K}^n$ would be sparse (to high precision), containing only $O(n \log n)$ non-negligible elements. This is the case. Furthermore, the complete similarity transformation $U\,\tilde{K}^n\,U^T$, which exploits the smoothness in the rows of $\tilde{K}^n$, as well as the columns, is yet more sparse, containing only $O(n)$ non-negligible elements.

The story for the matrix $\bar{K}^n$ is similar, but here the matrix contains $O(n \log n)$ non-negligible elements. This sparsity is proved for several specific examples in [2].

4.3 Multiplication of Integral Operators

The product of two integral operators with smooth kernels itself has a smooth kernel, and it can be represented as a sparse matrix in wavelet coordinates.

We define integral operators $\mathcal{K}_1$ and $\mathcal{K}_2$ by the formula

$$(\mathcal{K}_i f)(x) = \int_a^b K_i(x,y)\,f(y)\,dy, \qquad i = 1,2.$$

The product operator $\mathcal{K}_3 = \mathcal{K}_1\mathcal{K}_2$ is given by the formula

$$\begin{aligned}(\mathcal{K}_1\mathcal{K}_2 f)(x) &= \int_a^b K_1(x,t) \int_a^b K_2(t,y)\,f(y)\,dy\,dt \\ &= \int_a^b \left(\int_a^b K_1(x,t)\,K_2(t,y)\,dt \right) f(y)\,dy \\ &= \int_a^b K_3(x,y)\,f(y)\,dy,\end{aligned}$$

where the kernel $\mathcal{K}_3$ of the product has the form

$$K_3(x,y) = \int_a^b K_1(x,t)\,K_2(t,y)\,dt.$$

If kernels K_1 and K_2 are analytic except along the diagonal $x = t$, where they have integrable singularities, then the same is true of the product kernel K_3. As a result, the product operator $\mathcal{K}_3$ also has a sparse representation in a wavelet basis.

cretization of integral equations, the Nyström method approximates the integral operator

$$(\mathcal{K}f)(x) = \int_a^b K(x,y)\, f(y)\, dy$$

by a quadrature for selected values of x. We define the operator R by the formula

$$(Rf)(x) = \sum_{j=1}^{n} \omega_j\, K(x, x_j)\, f(x_j),$$

which approximates $\mathcal{K}$ for appropriate choice of quadrature weights ω_j and nodes x_j. The values of x are chosen to coincide with the quadrature nodes $x_1, \ldots, x_n$, and the original integral equation (26) is approximated by the system of equations

$$f(x_i) \cdot \gamma + \sum_{j=1}^{n} \omega_j\, K(x_i, x_j)\, f(x_j) = g(x_i), \qquad i = 1, \ldots, n. \tag{31}$$

(Compare to the Galerkin discretization (30).) For the trapezoidal rule, the quadrature weights ω_j are equal, except at the ends $j = 1$ and $j = n$. For kernels with singularities, however, the trapezoidal rule does not provide a good approximation of the integral. Quadratures have been developed in which the weights near the ends and near the singularities are altered to yield rapidly-convergent schemes [3]. In this case each weight depends on the argument x_i as well as the quadrature node x_j, so it becomes ω_{ij}, for $i, j = 1, \ldots, n$.

Even with these adjustments to the trapezoidal rule, most of the quadrature weights have constant value and the smoothness of the matrix $\tilde{K}^n = \{\omega_{ij} K(x_i, x_j)\}_{i,j=1,\ldots,n}$ depends primarily on the smoothness of the kernel K. For a kernel which is smooth except for diagonal singularities, the matrix $\tilde{K}^n$ can be transformed by a change of basis to a sparse matrix, to high precision. In particular, the wavelet-like basis matrix U defined in §3.5 can be used to obtain the similarity-transformed matrix $U\, \tilde{K}^n\, U^T$, which has the desired sparse structure. In fact, a picture of $U\, \tilde{K}^n\, U^T$ is nearly indistinguishable from Figure 6.

Remark. The Galerkin method and Nyström method are two techniques for the discretization of integral operators: which is preferred? Both are conceptually straightforward and an error analysis has been developed for each method (see, *e.g.*, [12]); the Nyström method offers, however, some computational benefits. Using the Nyström method with the trapezoidal rule, or high-order corrected trapezoidal rule, the kernel is evaluated just once for each element in the computed matrix. With the Galerkin method, on the other hand, a matrix element corresponds to an integral of the kernel with the basis elements in both coordinates. An appropriate quadrature must be applied to compute each of these elements, generally requiring many kernel evaluations. This complication usually makes the Galerkin method uncompetitive with the Nyström method.

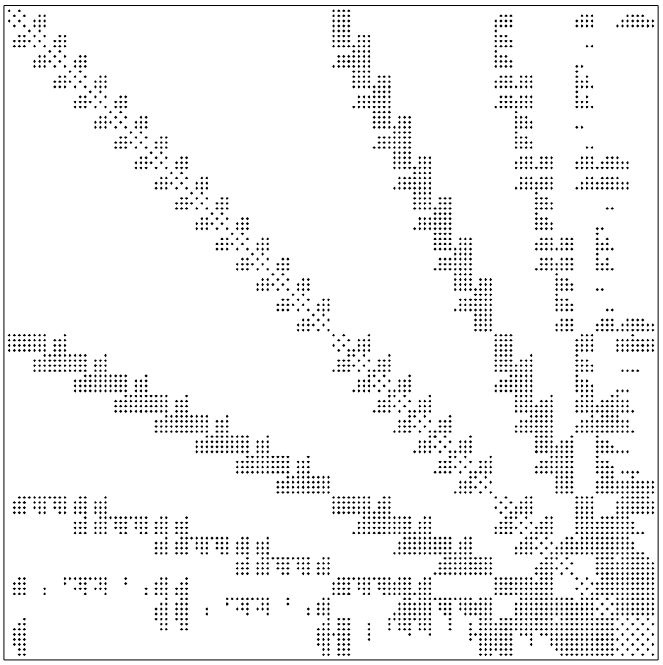

Figure 6. The integral operator with kernel $K(x,y) = \log|x-y|$ is discretized by the Galerkin method, with the multi-wavelet basis of order $N = 4$. The truncation is set at $n = 128$ basis functions. The dots represent elements above a threshold, determined so that the relative error ϵ is bounded at 10^{-3}.

For most applications, with classical bases (*e.g.*, Fourier or orthogonal polynomials) the matrix $\bar{K}^n = \{K_{ij}\}_{i,j=1,\ldots,n}$ is dense (nearly all of its elements are nonzero). There may be substantial cost in computing all elements K_{ij}. Even after the elements are computed, the cost of the application of the $n \times n$-matrix $\bar{K}^n$ to a vector is of order $O(n^2)$. The solution of (30), if obtained by a direct scheme such as Gaussian elimination, requires order $O(n^3)$ operations. If (30) is solved by an iterative method, which requires an order $O(n^2)$ matrix-vector product on each iteration, the number of iterations may be large, depending on the conditioning of the original integral equation. For large-scale problems, these costs are often prohibitive.

The denseness of $\bar{K}^n$ depends on both the kernel K and the basis. For a wide variety of kernels arising in problems of potential theory, the matrix $\bar{K}^n$ is sparse, to high precision, if the basis is chosen to be a wavelet basis. This theme was developed by G. Beylkin, R. Coifman, and V. Rokhlin [8] for Daubechies wavelets. For these problems, the kernel $K(x,y) = \log|x-y|$ serves as a good model. The value of the kernel varies smoothly as a function of x and y away from the diagonal $x = y$, where the kernel is singular. In the example at the start of this section we saw what happens when $f_a(x) = \log|x-a|$ is expanded in the x-coordinate in a multi-wavelet basis. Now we propose to expand the kernel $K(x,y) = \log|x-y|$ in both coordinates in a multi-wavelet basis; one example of the matrices which result is shown in Figure 6.

Nyström Method. An alternative to the Galerkin method for the dis-

for the existence and uniqueness of solutions to first- and second-kind integral equations, developed by Fredholm (see, *e.g.*, [12] or [14]). Integral equations form a powerful tool for the mathematical formulation of a wide range of physical problems; their relative neglect, compared with differential equations, is due in part to the historical lack of efficient solution techniques. Recently there has been substantial progress toward eliminating this deficiency.

Any equation to be solved numerically must be reduced to a finite dimensional problem, or discretized. There are two basic approaches to the discretization of integral equations. In one, often called the Galerkin method, expansions of f, K, and g are made in some basis, the expansions are truncated, and the resulting finite system of linear equations is solved numerically. In the second, developed by Nyström, the integral is approximated, at each of n points, by an n-point quadrature, yielding again a system of equations that is solved numerically.

Galerkin Method. Suppose that $\{b_1, b_2, \ldots\}$ is an orthonormal basis for $L^2[a,b]$. The expansions of $f, g \in L^2[a,b]$ in the basis are given by

$$f(x) = \sum_{j=1}^{\infty} f_j \, b_j(x), \qquad g(x) = \sum_{j=1}^{\infty} g_j \, b_j(x), \tag{27}$$

where the coefficients f_j and g_j are given by

$$f_j = \int_a^b b_j(x) \, f(x) \, dx, \qquad g_j = \int_a^b b_j(x) \, g(x) \, dx, \qquad j \in \mathbb{N}.$$

Similarly, the expansion for $K \in L^2([a,b] \times [a,b])$ is the integral in both coordinates

$$K(x,y) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} K_{ij} \, b_i(x) \, b_j(y), \tag{28}$$

where the coefficient K_{ij} is the double integral

$$K_{ij} = \int_a^b \int_a^b b_i(x) \, b_j(y) \, K(x,y) \, dx \, dy, \qquad i,j \in \mathbb{N}.$$

Substitution of equations (27) and (28) into integral equation (26) yields an infinite system of equations in the coefficients f_j, g_j, and K_{ij}, namely,

$$f_i \cdot \gamma + \sum_{j=1}^{\infty} K_{ij} \, f_j = g_i, \qquad i \in \mathbb{N}. \tag{29}$$

The expansion for K may be truncated at a finite number of terms, producing the finite system of equations

$$f_i \cdot \gamma + \sum_{j=1}^{n} K_{ij} \, f_j = g_i, \qquad i = 1, \ldots, n. \tag{30}$$

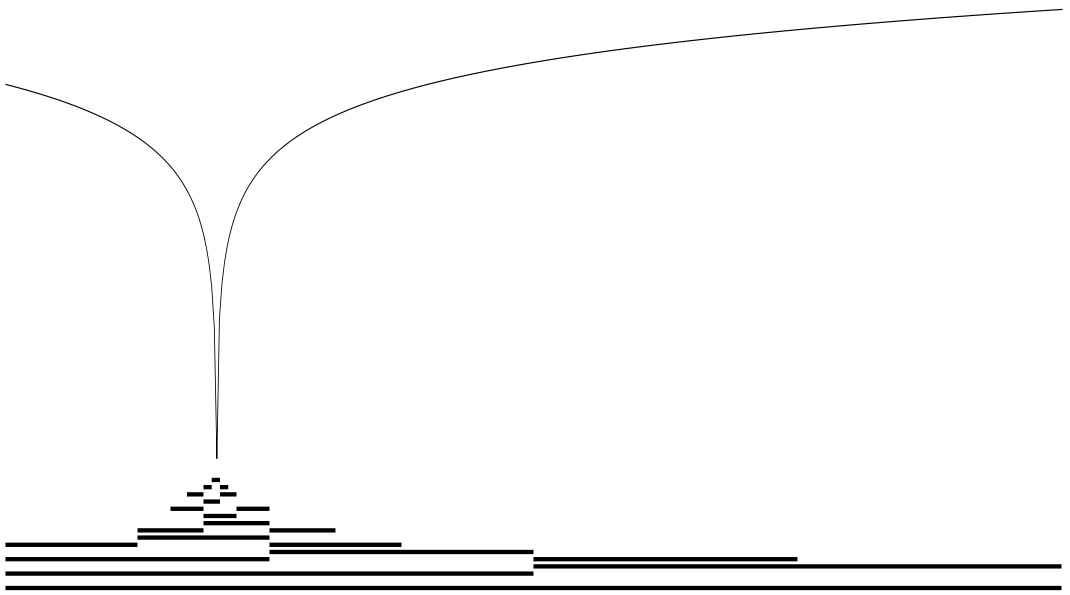

Figure 5. The function $f(x) = \log|x - .2|$, defined on the interval $[0,1]$, is expanded in the multi-wavelet basis of order 8 to six-digit accuracy. The solid line segments indicate the intervals of support of the basis functions in the resulting representation.

where $\alpha_j = \int_I p_j(x)\, f_a(x)\, dx$, and $p_0, p_1, \ldots$ are the orthonormal polynomials for the interval I. If we approximate f_a on I by keeping just the first n terms of the expansion, the truncation error decays exponentially in n (this is easy to show). For instance, keeping 8 terms yields six-digit accuracy; keeping 18 gives fifteen digits.

How does this connect with wavelets? In constructing the multi-wavelet bases, we built spaces V_n consisting of functions which are locally low-order polynomials. The multi-wavelet basis functions are orthogonal to these low-order polynomials, and are themselves locally supported. When f_a is expanded in a multi-wavelet basis, all basis functions lying on intervals separated from a have small coefficients, and can be neglected, up to a precision which depends on the order N. This property is illustrated in Figure 5. The representation of the function consists only of those basis functions supported near the singularity. It is this representational parsimony which leads to the usefulness of wavelets in a variety of applications.

4.1 Integral Equations

A linear integral equation on the interval $[a,b]$ is an equation of the form

$$f(x) \cdot \gamma + \int_a^b K(x,y)\, f(y)\, dy = g(x), \tag{26}$$

where $f : [a,b] \to \mathbb{R}$ is the unknown, $K : [a,b] \times [a,b] \to \mathbb{R}$ is the kernel, and $g : [a,b] \to \mathbb{R}$ is the right hand side. The equation is an equation of the first kind if $\gamma = 0$, otherwise of the second kind. There is a rather complete theory

where U_j' is given by the formula

$$U_j' = \begin{pmatrix} {U_{j,1}}^L & & & \\ & {U_{j,2}}^L & & \\ & & \ddots & \\ & & & {U_{j,n_j}}^L \\ {U_{j,1}}^U & & & \\ & {U_{j,2}}^U & & \\ & & \ddots & \\ & & & {U_{j,n_j}}^U \end{pmatrix},$$

where $n_j = n/(2^j N)$, $U_{j,i}$ is given by

$$ {U_{j,i}}^T = \mathrm{Orth}(M_{j,i}), \tag{24}$$

and $M_{j,i}$ is given by

$$M_{j,i} = \begin{pmatrix} {U_{j-1,2i-1}}^U M_{j-1,2i-1} \\ {U_{j-1,2i}}^U M_{j-1,2i} \end{pmatrix}, \tag{25}$$

for $i = 1, \ldots, n/(2^j N)$. The final basis matrix $U = U_m \cdots U_1$ represents the wavelet-like basis of parameter N on $x_1, \ldots, x_n$. Note that the matrices U and U_j are of dimension $n \times n$, U_j' is $n/2^{j-1} \times n/2^{j-1}$, $U_{j,i}$ and $M_{j,i}$ are $2N \times 2N$, and $U_{j,i}^L$ and $U_{j,i}^U$ are $N \times 2N$. In the computation of the basis, an adjustment must be made to these formulae to ensure numerical stability (see [6]).

§4. Linear Algebraic Operations

Very little indication has been presented, so far, of the value of wavelets for numerical linear algebra. Before embarking on the body of this section, we present an example showing the use of wavelets.

Example. For problems in electromagnetics, it is often necessary to determine the potential field due to a given distribution of charges. This is done by convolving the potential due to a point charge with the actual charge density. For a charge at the origin, this potential has the form $1/r$ in three dimensions and $\log r$ in two dimensions. We consider a simplified example, in which the function $f_a : \mathbb{R} \to \mathbb{R}$, given by $f_a(x) = \log|x - a|$, is to be represented. On any interval *separated* from a, the expansion of f_a in orthogonal polynomials converges rapidly. An interval I is separated from the point a if its distance to a is at least as great as its length. Expanding f_a on such an interval I, we obtain

$$f_a(x) = \sum_{j=0}^{\infty} \alpha_j p_j(x),$$

where $u_i = (i-1)2N$. The first basis matrix U_1 is the $n \times n$-matrix given by the formula

$$U_1 = \begin{pmatrix} {U_{1,1}}^L & & & & \\ & {U_{1,2}}^L & & & \\ & & \ddots & & \\ & & & {U_{1,n_1}}^L & \\ {U_{1,1}}^U & & & & \\ & {U_{1,2}}^U & & & \\ & & \ddots & & \\ & & & {U_{1,n_1}}^U & \end{pmatrix},$$

where ${U_{1,i}}^T = \mathrm{Orth}(M_{1,i})$ and $n_1 = n/(2N)$. The second basis matrix is $U_2 U_1$, with U_2 defined by the formula

$$U_2 = \begin{pmatrix} I_{n/2} & \\ & U_2' \end{pmatrix},$$

where I_j is the $j \times j$ identity matrix and the $n/2 \times n/2$-matrix U_2' is given by the formula

$$U_2' = \begin{pmatrix} {U_{2,1}}^L & & & & \\ & {U_{2,2}}^L & & & \\ & & \ddots & & \\ & & & {U_{2,n_2}}^L & \\ {U_{2,1}}^U & & & & \\ & {U_{2,2}}^U & & & \\ & & \ddots & & \\ & & & {U_{2,n_2}}^U & \end{pmatrix},$$

where $n_2 = n/(4N)$, ${U_{2,i}}^T = \mathrm{Orth}(M_{2,i})$, and the $2N \times 2N$-matrix $M_{2,i}$ is given by

$$M_{2,i} = \begin{pmatrix} {U_{1,2i-1}}^U M_{1,2i-1} \\ {U_{1,2i}}^U M_{1,2i} \end{pmatrix},$$

for $i = 1, \ldots, n/(4N)$. In general, the jth basis matrix, for $j = 2, \ldots, m$, is $U_j \cdots U_1$ with U_j defined by the formula

$$U_j = \begin{pmatrix} I_{n-n/2^{j-1}} & \\ & U_j' \end{pmatrix},$$

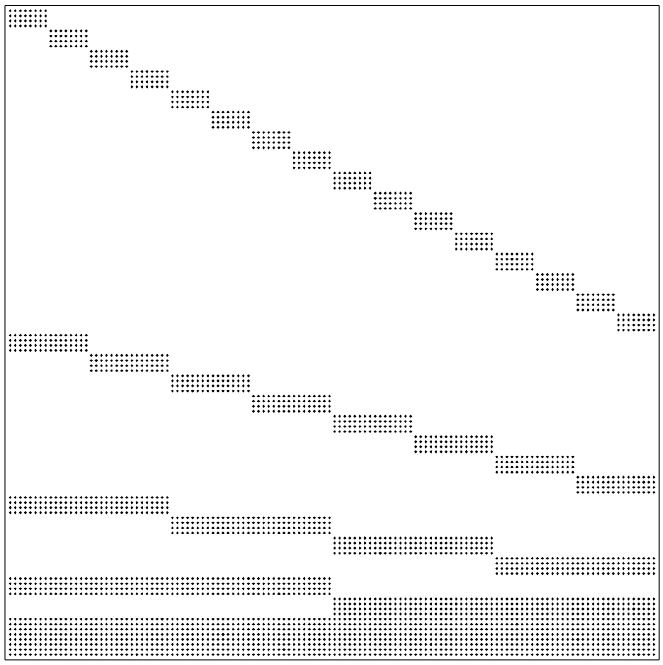

Figure 4. The matrix represents a wavelet-like basis for a discretization with $n = 128$ points, for $N = 4$. Each row denotes one basis vector, with the dots depicting non-zero elements. The first $n/2$ rows form a basis for W_{1-m}, the next $n/4$ form a basis for W_{2-m}, and so forth. All but the final N rows are orthogonal to polynomials of degree less than N, of which V_0 is composed.

Suppose S is a matrix whose columns $s_1, \ldots, s_{2N}$ are linearly independent. We define $T = \mathrm{Orth}(S)$ to be the matrix which results from the column-by-column Gram-Schmidt orthogonalization of S. Namely, denoting the columns of T by $t_1, \ldots, t_{2N}$, we have

$$\begin{aligned} \text{linear span}\{t_1, \ldots, t_i\} &= \text{linear span}\{s_1, \ldots, s_i\} \\ {t_i}^T t_j &= \delta_{ij} \end{aligned} \qquad i, j = 1, \ldots, 2N.$$

For a $2N \times 2N$-matrix S we let S^U and S^L denote two $N \times 2N$-matrices, S^U consisting of the upper N rows and S^L the lower N rows of S,

$$S = \begin{pmatrix} S^U \\ S^L \end{pmatrix}.$$

Now we proceed to the definition of the basis matrices. Given the set of points $\{x_1, \ldots, x_n\} \in \mathbb{R}$ with $x_1 < \cdots < x_n$, where $n = 2^m N$, we define the $2N \times 2N$ moments matrices $M_{1,i}$ for $i = 1, \ldots, n/(2N)$ by the formula

$$M_{1,i} = \begin{pmatrix} 1 & x_{u_i+1} & \cdots & {x_{u_i+1}}^{2N-1} \\ 1 & x_{u_i+2} & \cdots & {x_{u_i+2}}^{2N-1} \\ \vdots & & & \vdots \\ 1 & x_{u_i+2N} & \cdots & {x_{u_i+2N}}^{2N-1} \end{pmatrix}, \tag{23}$$

on a discrete set of points $\{x_1, \ldots, x_n\} \subset \mathbb{R}$. The structure of this analogue is essentially similar to that of the multi-wavelet bases, but the discrete construction is more convenient when the representation of a function (and its related operators) is based on its values at a finite set of points. Such representations arise in finite-difference computations and in Nyström discretizations for integral equations; these are discussed in this connection in §5.

The primary difference in the present construction from the multi-wavelet bases is the lack of complete scale invariance. In other words, the spaces V_n in the discrete construction are not dilates of a single space V_0, rather only nearly so.

Construction of Bases. This construction follows that in [6]. The discrete set of points $\{x_1, \ldots, x_n\}$ is ordered so that $x_1 < \cdots < x_n$. We assume that the number n of points satisfies $n = 2^m N$, where N is the order of approximation required, and m is a positive integer. We define V_0 to be the N-dimensional vector space of polynomials of degree less than N on $\{x_1, \ldots, x_n\}$,

$$V_0 = \text{span}\left\{ \langle {x_1}^j, \ldots, {x_n}^j \rangle \;\middle|\; j = 0, \ldots, N-1 \right\}.$$

We define V_{-1} to be the $2N$-dimensional space of vectors which are polynomial of degree less than N on $\{x_1, \ldots, x_{n/2}\}$ and on $\{x_{n/2+1}, \ldots, x_n\}$. In general, V_{-j} is the $2^j N$-dimensional space consisting of vectors which are polynomial of degree less than N on $\{x_1, \ldots, x_{n/2^j}\}$, on $\{x_{n/2^j+1}, \ldots, x_{2n/2^j}\}$, and so forth, up to $\{x_{n-n/2^j+1}, \ldots, x_n\}$, for $j = 0, \ldots, m$. Thus V_{-m} is the entire n-dimensional vector space.

The difference space W_{-j}, for $j = 0, 1, \ldots, m-1$, is again defined by (8). The procedure to construct a basis for V_{-m}, which exploits the decomposition

$$V_{-m} = W_{1-m} \oplus W_{2-m} \oplus \cdots \oplus W_0 \;\oplus\; V_0,$$

consists of an orthogonalization procedure to construct a basis for each of the W_{-j}. The result is an orthogonal matrix with rows which are the basis vectors, as shown in Figure 4. We construct the basis by the construction of a sequence of bases, for the decompositions

$$\begin{aligned}
&W_{1-m} \oplus V_{1-m} \\
&W_{1-m} \oplus W_{2-m} \oplus V_{2-m} \\
&\qquad \vdots \\
&W_{1-m} \oplus W_{2-m} \oplus \cdots \oplus W_{-1} \oplus V_{-1} \\
&W_{1-m} \oplus W_{2-m} \oplus \cdots \oplus W_{-1} \oplus W_0 \oplus V_0.
\end{aligned}$$

The bases are given by the finite sequence of matrix products U_1, $U_2 U_1$, ..., $U_m \cdots U_1$. Before we can specify the matrices $U_1, \ldots, U_m$, we require some additional notation.

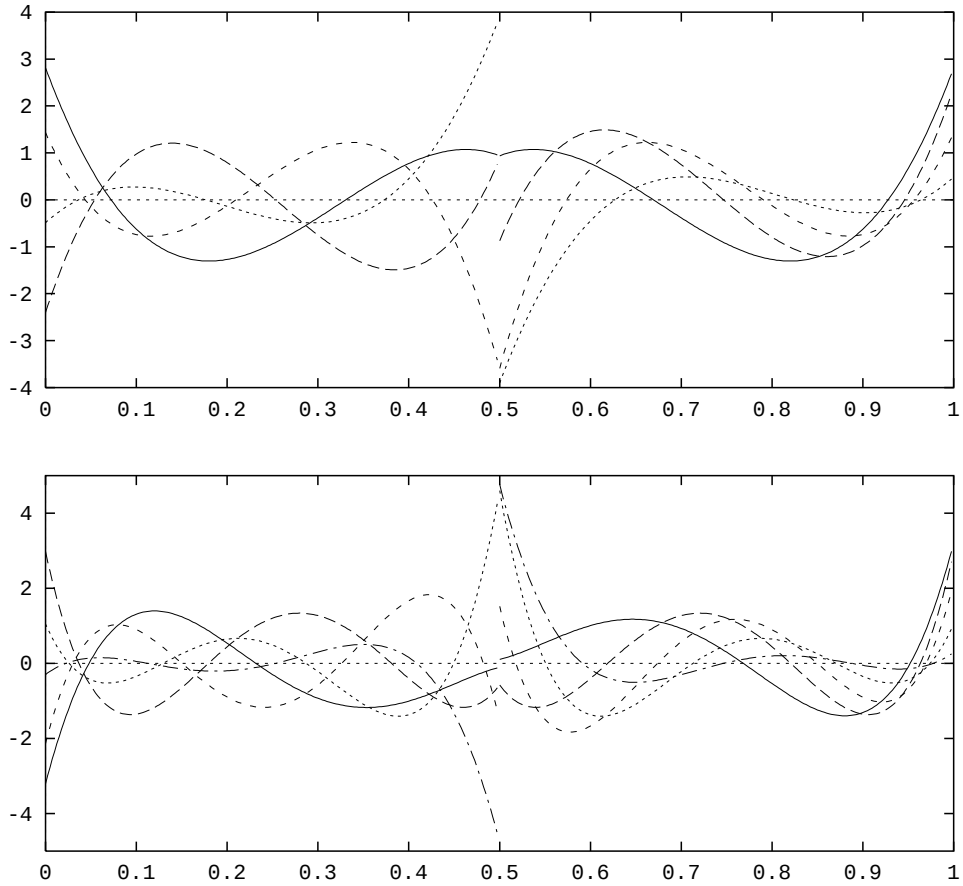

Figure 3. Wavelet functions for the piecewise polynomial multiwavelet bases for $N = 4$ (top) and $N = 5$. Note that each function is even or odd about the interval midpoint.

For $N = 2$, the wavelet functions are given by

$$ {}_2w_0(x) = \begin{cases} \sqrt{3}\,(1-4x), & \text{if } x \in [0,1/2); \\ \sqrt{3}\,(4x-3), & \text{if } x \in [1/2,1); \\ 0, & \text{otherwise;} \end{cases} $$

$$ {}_2w_1(x) = \begin{cases} 6x-1, & \text{if } x \in [0,1/2); \\ 6x-5, & \text{if } x \in [1/2,1); \\ 0, & \text{otherwise.} \end{cases} $$

For larger N, the wavelet functions can also be derived explicitly, but the expressions naturally get rather long. Figure 3 shows the graphs of the wavelet functions for $N = 4$ and $N = 5$.

3.5 Discrete Wavelet-Like Bases

The Daubechies wavelet bases of §3.3 and the multi-wavelet bases of §3.4 lie in the space of square-integrable functions $L^2(\mathbb{R})$. We consider now an analogue of the multi-wavelet bases which lies in the space of functions defined

where

$$\alpha_j^{n,k} = \int_{-\infty}^{\infty} \phi_j^{n,k}(x)\, f(x)\, dx.$$

The approximation error $\|f_n - f\|$ is bounded, according to the following lemma.

Lemma 1. *Suppose that the function $f : [0,1] \to \mathbb{R}$ is N times continuously differentiable, $f \in C^N[0,1]$. Then f_n (given by equation (21)) approximates f with mean error bounded as follows:*

$$\|f_n - f\| \le 2^{-nN} \frac{2}{4^N N!} \sup_{x\in[0,1]} |f^{(N)}(x)|. \tag{22}$$

Proof: We divide the interval $[0,1)$ into subintervals on which f_n is a polynomial; the restriction of f_n to one such subinterval $I_{n,k}$ is the polynomial of degree less than N that approximates f with minimum mean error. We then use the maximum error estimate for the polynomial which interpolates f at Chebyshev nodes of order N on $I_{n,k}$.

We define $I_{n,k} = [2^n k, 2^n(k+1))$ for $n = 0, -1, \ldots$; $k = 0, \ldots, 2^n - 1$, and obtain

$$\begin{aligned}
\|f_n - f\|^2 &= \int_0^1 \left[f_n(x) - f(x)\right]^2 dx \\
&= \sum_k \int_{I_{n,k}} \left[f_n(x) - f(x)\right]^2 dx \\
&\le \sum_k \int_{I_{n,k}} \left[C_{n,k}^N f(x) - f(x)\right]^2 dx \\
&\le \sum_k \int_{I_{n,k}} \left(\frac{2^{1-nN}}{4^N N!} \sup_{x\in I_{n,k}} \left|f^{(N)}(x)\right|\right)^2 dx \\
&\le \left(\frac{2^{1-nN}}{4^N N!} \sup_{x\in[0,1]} \left|f^{(N)}(x)\right|\right)^2,
\end{aligned}$$

and by taking square roots we have bound (22). Here $C_{n,k}^N f$ denotes the polynomial of degree $N-1$ which agrees with f at the Chebyshev nodes of order N on $I_{n,k}$, and we have used the well-known maximum error bound for Chebyshev interpolation (see, *e.g.*, [1]). ■

The error of the approximation f_n of f therefore decays like 2^{-nN} and, since it requires $2^n N$ basis elements, we have convergence of order N. Despite the lack of regularity of the scaling functions ϕ_j (and the projections f_n), the convergence is similar to that obtained with the Daubechies wavelets.

Examples. The multi-wavelets case $N = 1$ coincides with the Haar basis (again).

and on $[\frac{1}{2},1)$, then orthogonalize N of them, first to the functions $\phi_0,\ldots,\phi_{N-1}$, then to the functions $\phi_N,\ldots,\phi_{2N-1}$, and finally among themselves. We define $f_0^1, f_1^1, \ldots, f_{N-1}^1$ by the formula

$$f_j^1(x) = \begin{cases} -\phi_j(2x), & \text{if } x \in [0,1/2); \\ \phi_j(2x-1), & \text{if } x \in [1/2,1); \\ 0, & \text{otherwise;} \end{cases}$$

note that the $2N$ functions $\phi_0,\ldots,\phi_{N-1}, f_0^1,\ldots,f_{N-1}^1$ are linearly independent, hence span the space of functions that are polynomials of degree less than N on $[0,\frac{1}{2})$ and on $[\frac{1}{2},0)$.

1. By the Gram-Schmidt process we orthogonalize f_j^1 with respect to $\phi_0,\ldots,\phi_{N-1}$, to obtain f_j^2, for $j=0,\ldots,N-1$. This orthogonality is preserved by the remaining orthogonalizations, which only produce linear combinations of the f_j^2.
2. The next sequence of steps yields $N-1$ functions orthogonal to ϕ_N, of which $N-2$ functions are orthogonal to ϕ_{N+1}, and so forth, down to 1 function which is orthogonal to ϕ_{2N-2}. First, if at least one of f_j^2 is not orthogonal to ϕ_N, we reorder the functions so that it appears first, $\langle f_0^2,\phi_N\rangle \neq 0$. We then define $f_j^3 = f_j^2 - a_j \cdot f_0^2$ where a_j is chosen so $\langle f_j^3,\phi_N\rangle = 0$ for $j=1,\ldots,N-1$, achieving the desired orthogonality to ϕ_N. Similarly, we orthogonalize to $\phi_{N+1},\ldots,\phi_{2N-2}$, each in turn, to obtain $f_0^2, f_1^3, f_2^4, \ldots, f_N^{N+1}$ such that $\langle f_j^{j+2},\phi_i\rangle = 0$ for $i=0,\ldots,N-1+j$.
3. Finally, we do Gram-Schmidt orthogonalization on $f_{N-1}^{N+1}, f_{N-2}^N \ldots, f_0^2$, in that order, and normalize to obtain ${}_Nw_{N-1}$, ${}_Nw_{N-2},\ldots,{}_Nw_0$.

The resulting basis functions ${}_Nw_0,\ldots,{}_Nw_{N-1}$ are polynomials of degree $N-1$ on the intervals $[0,\frac{1}{2})$ and $[\frac{1}{2},1)$. The basis for $L^2(\mathbb{R})$ consists of the translates and dilates of ${}_Nw_0,\ldots,{}_Nw_{N-1}$. In addition, we can construct in this fashion a basis for a finite interval. For instance, $L^2[0,1]$ has an orthonormal basis consisting of $\phi_0,\ldots,\phi_{N-1}$ plus those translates and dilates of ${}_Nw_0,\ldots,{}_Nw_{N-1}$ that are supported in the interval $[0,1]$.

Approximation. The space V_0 consists of functions which are piecewise polynomial, of degree less than N. The expansion of an arbitrary function $f \in C^N(\mathbb{R})$ in a multi-wavelet basis of order N is Nth-order convergent. The argument for this case is simpler than the general argument of Strang and Fix [20], and we give it here. We define translates and dilates $\phi_j^{n,k}$ by the formula

$$\phi_j^{n,k}(x) = 2^{-n/2}\phi_j(2^{-n}x-k), \qquad n,k \in \mathbb{Z},$$

such that the set $\{\,\phi_j^{n,k} \mid j=0,\ldots,N-1;\ k\in\mathbb{Z}\,\}$ is a basis for V_n. The projection of f on V_n is defined (as before) by the formula

$$f_n(x) = \sum_{k\in\mathbb{Z}} \sum_{j=0}^{N-1} \alpha_j^{n,k}\,\phi_j^{n,k}(x), \tag{21}$$

integer nodes are obtained by solving a linear system and are used to obtain the values at half-integer nodes, then quarter-integer nodes, and so forth, by the dilation equation. The linear equations

$$\phi(j) = \sum_{k=0}^{2N-1} a_k \, \phi(2j-k), \qquad j = 1, \ldots, 2N-2$$

are deficient in rank by one (they determine ϕ only within a constant factor). Discarding one of them and adding the equation

$$1 = \sum_{j=1}^{2N-2} \phi(j),$$

we obtain a non-singular system; it is solved to get $\phi(1), \ldots, \phi(2N-2)$.

This technique is fast, easy, and accurate; it was used to generate the graphs in Figures 1 and 2.

3.4 Multi-Wavelets

We introduce now another class of bases for $L^2(\mathbb{R})$ which differ from wavelets in that instead of a single scaling function ϕ, there are several functions $\phi_0, \ldots, \phi_{N-1}$ whose translates span the space V_0. This difference enables high-order approximation with basis functions supported on non-overlapping intervals; the price is the lack of regularity of the functions.

In the framework of multiresolution analysis, this class is very simple. On the interval $[0,1)$, each scaling function ϕ_i is a dilated, translated, and normalized Legendre polynomial:

$$\phi_i(x) = \begin{cases} \sqrt{2i+1}\, P_i(2x-1), & \text{if } x \in [0,1); \\ 0, & \text{otherwise;} \end{cases} \qquad i = 0, \ldots, N-1,$$

where $P_0, P_1, \ldots$ are the Legendre polynomials. The rest of the basis construction (almost) follows, just by turning the "crank" of multiresolution analysis. The space V_0 consists of the span of the integer translates of the scaling functions $\phi_0, \ldots, \phi_{N-1}$. The spaces V_n for $n \in \mathbb{Z}$ are dilates of V_0, and the difference spaces W_n are defined, as before, by (8).

Construction of Basis. The basis functions for W_n are not unique, of course, until we add additional constraints. Following [2], we construct orthonormal basis functions ${}_N w_0, \ldots, {}_N w_{N-1}$, which vanish outside the interval $[0,1)$, whose integer translates span W_0, and which are orthogonal to polynomials of maximum degree,

$$\int_{-\infty}^{\infty} {}_N w_j(x)\, x^i \, dx, \qquad i = 0, \ldots, N-1+j.$$

We start with $2N$ dilates and translates of $\phi_0, \ldots, \phi_{N-1}$, which span the space of functions that are polynomials of degree less than N on the interval $[0, \frac{1}{2})$

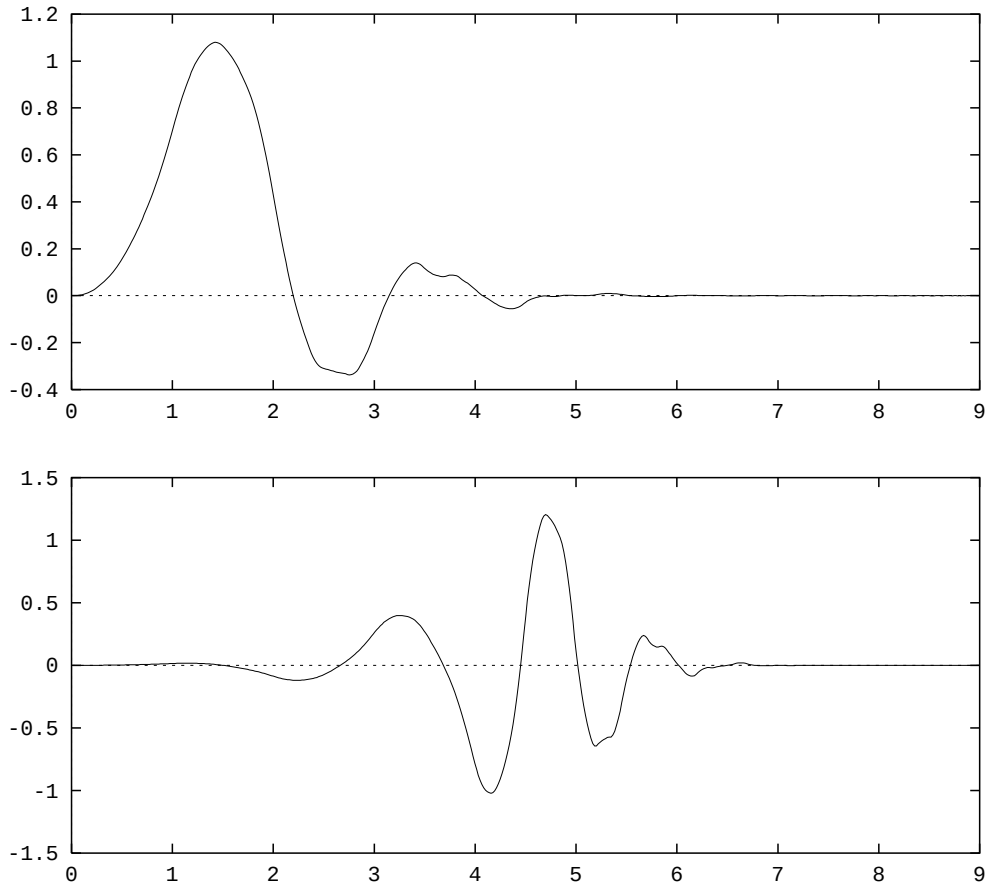

Figure 2. Scaling function ${}_5\phi$ (top) and corresponding wavelet function ${}_5w$, translated. In addition to their fifth-order approximation, they have greater smoothness than the second-order functions (Figure 1).

linear combinations of the scaling functions

$$1 = \sum_{k\in\mathbb{Z}} {}_2\phi(x-k),$$
$$x = \sum_{k\in\mathbb{Z}} \left(k + \frac{3-\sqrt{3}}{2}\right) {}_2\phi(x-k), \qquad x \in \mathbb{R}.$$

For increasing N, the scaling functions ${}_N\phi$ are increasingly regular. Figure 2 shows the graphs of ${}_5\phi$ and ${}_5w$, each of which has a continuous first derivative.

Computation of ϕ. We have produced no explicit representation for ϕ; how it is computed? There are at least three approaches. The most obvious is to begin with some initial estimate ϕ_0 and calculate $\phi_1, \phi_2, \ldots$ by iteration of the dilation equation (12). Another method [9] calculates the Fourier transform $\hat{\phi}$ from an explicit formula involving the dilation coefficients. In the third, very elegant approach, suggested by Strang [19], the values $\phi(1), \ldots, \phi(2N-2)$ at

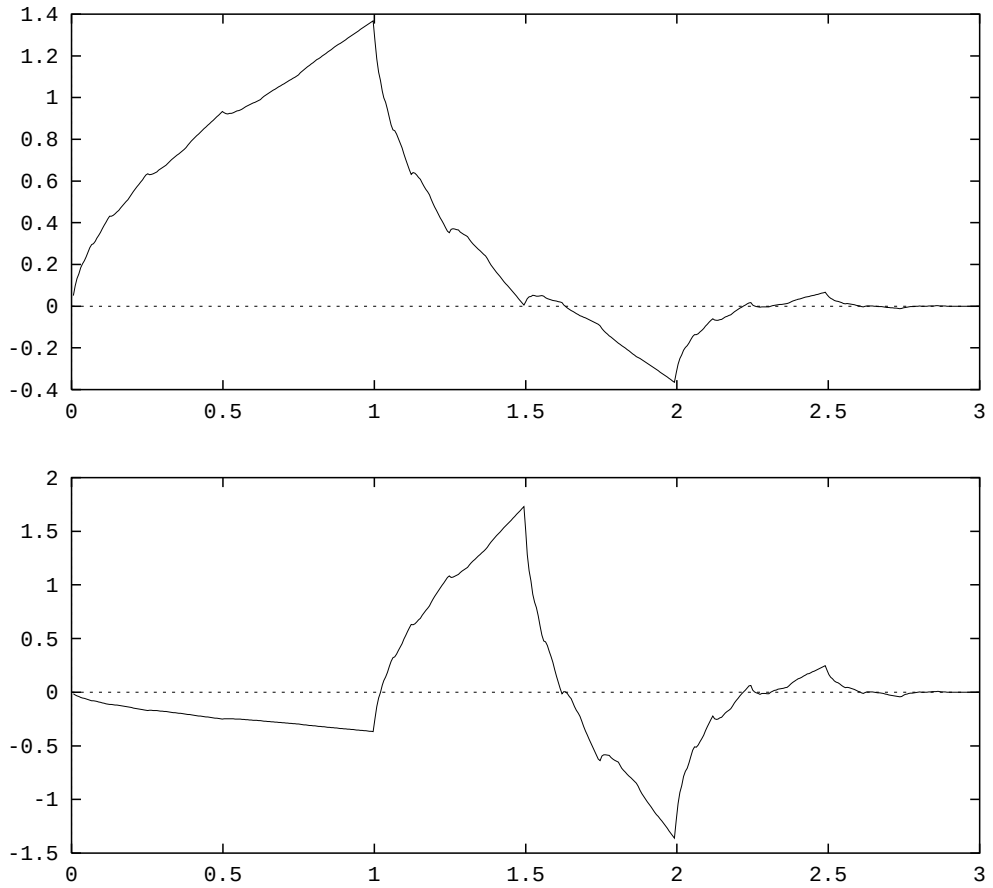

Figure 1. Daubechies scaling function ${}_2\phi$ (top) and corresponding wavelet function ${}_2w$. Here ${}_2w$ has been translated to the interval $[0,3]$ by using coefficients $b_k = (-1)^k a_{2N-1-k}$.

Daubechies [9] constructs the scaling function ${}_N\phi$ by the choice of nonzero coefficients $a_0, \ldots, a_{2N-1}$ to satisfy equations (11) and (17), as well as a regularity condition on ${}_N\phi$. It is remarkable that the regularity requirement leads to good approximation properties, and is equivalent to (20).

Examples. The scaling function ${}_1\phi$ coincides with the scaling function $\chi_{[0,1)}$ of the Haar basis.

The function ${}_2\phi$ is given by the formula

$$
\begin{aligned}
{}_2\phi(x) = \frac{1}{4}\Bigg[&\left(1+\sqrt{3}\right)\ {}_2\phi(2x) + \left(3+\sqrt{3}\right)\ {}_2\phi(2x-1) \\
&+ \left(3-\sqrt{3}\right)\ {}_2\phi(2x-2) + \left(1-\sqrt{3}\right)\ {}_2\phi(2x-3)\Bigg].
\end{aligned}
$$

Figure 1 shows the graphs of ${}_2\phi$ and the corresponding wavelet function ${}_2w$. Note that, unlike ${}_1\phi$, they are continuous. They do not, however, possess continuous derivatives. What is not obvious from the figure is their second-order approximation property: the functions $f(x) = 1$ and $f(x) = x$ are given by

$$f_n(x) = \sum_{k \in \mathbb{Z}} \alpha_{nk}\, \phi_{nk}(x),$$

where the (orthonormal) basis functions for V_n are given by

$$\phi_{nk}(x) = 2^{-n/2} \phi(2^{-n}x - k)$$

and the coefficients are

$$\alpha_{nk} = \int_{-\infty}^{\infty} \phi_{nk}(x)\, f(x)\, dx.$$

The rate of convergence of the sequence $\ldots, f_1, f_0, f_{-1}, \ldots \to f$ is of interest, for it dictates the computational cost of representing f to a prescribed accuracy using some f_n. A theory of approximation by translates was developed by G. Strang and G. Fix [20]. Although their result was derived in the context of finite element analysis, it is directly applicable here. They proved that for an arbitrary $f \in C^N(\mathbb{R})$ the sequence $f_0, f_{-1}, f_{-2}, \ldots$ has order of convergence N if and only if polynomials of degree less than N are contained in V_0. This is the property given in equation (14), for $V_0 = \text{span}\{ {}_N\phi(x-k) \mid k \in \mathbb{Z} \}$. Thus the Daubechies scaling function ${}_N\phi$ gives rise to wavelets whose expansions are Nth-order convergent.

The low-order polynomials are contained in V_0, so they are orthogonal to W_0, hence to w. In terms of the dilation equation coefficients, we may write for $j = 0, \ldots, N-1$

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty} x^j\, w(x)\, dx \\
&= \int_{-\infty}^{\infty} \sum_k \left(\frac{2x - k + k}{2} \right)^j b_k\, \phi(2x - k)\, dx \\
&= 2^{-j-1} \sum_{r=0}^{j} \binom{j}{r} \sum_k k^{j-r}\, b_k \int_{-\infty}^{\infty} (2x-k)^r\, \phi(2x-k)\, 2dx \\
&= 2^{-j-1} \sum_{r=0}^{j} \binom{j}{r} \sum_k k^{j-r}\, b_k \int_{-\infty}^{\infty} x^r\, \phi(x)\, dx
\end{aligned} \tag{19}$$

.

The integrals $\int x^r\, \phi(x)\, dx$ are all non-zero (since x^r can be written as a linear combination of translates of ϕ for $r = 0, \ldots, N-1$). Consecutively substituting $j = 0, 1, \ldots, N-1$ into equation (19) yields $\sum_k b_k\, k^j = 0$, $j = 0, \ldots, N-1$, from which we obtain

$$\sum_k (-1)^k a_k\, k^j = 0, \qquad j = 0, \ldots, N-1. \tag{20}$$

hence

$$\frac{1}{2}\sum_{m\in\mathbb{Z}} a_{2k+m}\, a_{2l+m} = \delta_{kl}, \qquad k,l \in \mathbb{Z}. \tag{17}$$

Equation (17) is the fundamental relation which ensures the orthogonality of the translates of the scaling function ϕ. To establish the orthogonality between the translates of the wavelet function

$$w(x) = \sum_{k\in\mathbb{Z}} b_k\, \phi(2x-k), \tag{18}$$

we use the coefficients b_k given by equation (11) and apply equation (17) to obtain

$$\begin{aligned}
&\int_{-\infty}^{\infty} w(x-k)\, w(x-l)\, dx \\
&\qquad = \int_{-\infty}^{\infty} \sum_m b_m \phi(2(x-k)-m) \sum_n b_n\, \phi(2(x-l)-n)\, dx \\
&\qquad = \frac{1}{2}\sum_m b_{2k+m}\, b_{2l+m} \\
&\qquad = \frac{1}{2}\sum_m (-1)^{2k+m} a_{1-2k-m}\, (-1)^{2l+m} a_{1-2l-m} \\
&\qquad = \frac{1}{2}\sum_m a_{1-2k-m}\, a_{1-2l-m} \\
&\qquad = \delta_{kl}.
\end{aligned}$$

Finally, the orthogonality between the ϕ and w translates is established (without equation (17)!) by

$$\begin{aligned}
\int_{-\infty}^{\infty} \phi(x-k)\, w(x-l)\, dx &= \frac{1}{2}\sum_m a_{2l+m}\, b_{2k+m} \\
&= \frac{1}{2}\sum_m a_{2l+m}\, (-1)^{2k+m} a_{1-2k-m} \\
&= 0.
\end{aligned}$$

The last statement follows from the observation that in the summation, each product $a_i\, a_j$ occurs twice, with opposite signs.

To summarize, equation (17), combined with the coefficient definition $b_k = (-1)^k a_{1-k}$, ensures the orthogonality relations $\int \phi(x-k)\, \phi(x-l)\, dx = \delta_{kl}$, $\int \phi(x-k)\, w(x-l)\, dx = 0$, and $\int w(x-k)\, w(x-l)\, dx = \delta_{kl}$.

Approximation. The projection of a function $f \in L^2(\mathbb{R})$ on the space V_n is defined by the formula

other than $a_m, a_{m+1}, \ldots, a_n$ vanish, then ϕ vanishes outside the interval $[m, n]$. To see this, observe that the iteration of

$$\phi_{i+1}(x) = \sum_{k=m}^{n} a_k \, \phi_i(2x - k)$$

maps a function ϕ_i supported on $[a, b]$ to the function ϕ_{i+1} supported on [(a+m)/2,(b+n)/2].

Consistency. Integration of the dilation equation (12) determines $\sum a_k$:

$$\begin{aligned}\int_{-\infty}^{\infty} \phi(x)\, dx &= \int_{-\infty}^{\infty} \sum_{k \in \mathbb{Z}} a_k \, \phi(2x - k)\, dx \\ &= \frac{1}{2} \sum_k a_k \int_{-\infty}^{\infty} \phi(2x - k)\, 2dx \\ &= \frac{1}{2} \sum_k a_k \int_{-\infty}^{\infty} \phi(x)\, dx.\end{aligned}$$

But $\int \phi(x)\, dx \neq 0$, since the sum of translates of ϕ is a nonzero constant (equation (14) with $j = 0$). Dividing through by $\int \phi(x)\, dx$ yields

$$\sum_k a_k = 2. \tag{15}$$

The dilation equation is homogeneous, so it determines ϕ only up to a constant factor. It is convenient to choose the factor so that

$$\int_{-\infty}^{\infty} \phi(x)\, dx = 1, \tag{16}$$

which will be assumed in the following discussion.

Orthogonality. We examine the consequences of equation (13) by combining it with the dilation equation (12):

$$\begin{aligned}\delta_{kl} &= \int_{-\infty}^{\infty} \phi(x - k)\, \phi(x - l)\, dx \\ &= \int_{-\infty}^{\infty} \sum_{m \in \mathbb{Z}} a_m \, \phi(2(x - k) - m) \sum_{n \in \mathbb{Z}} a_n \, \phi(2(x - l) - n)\, dx \\ &= \frac{1}{2} \sum_m \sum_n a_m \, a_n \int_{-\infty}^{\infty} \phi(2(x - k) - m)\, \phi(2(x - l) - n)\, 2dx \\ &= \frac{1}{2} \sum_m \sum_n a_m \, a_n \, \delta_{2k+m, 2l+n} \\ &= \frac{1}{2} \sum_m a_m \, a_{2(k-l)+m},\end{aligned}$$

for coefficients $\ldots a_{-1}, a_0, a_1, \ldots$. In §3.3 we will see that the choice for b_k given by

$$b_k = (-1)^k a_{1-k} \tag{11}$$

yields the desired orthogonality $W_0 \perp V_0$. The task of choosing a_k so that the translates of ϕ are mutually orthogonal, in addition to possessing other properties, is also addressed in §3.3.

Example. For the Haar basis, the scaling function satisfies the recurrence equation

$$\phi(x) = \phi(2x) + \phi(2x-1)$$

and the wavelet function is given by the formula

$$w(x) = \phi(2x) - \phi(2x-1),$$

following equation (11). The inner products $\int \phi(x-k)\, w(x)\, dx$, as well as $\int \phi(x-k)\, \phi(x)\, dx$ and $\int w(x-k)\, w(x)\, dx$, vanish for $k \neq 0$.

3.3 Daubechies Wavelets

Constraints can be placed on the coefficients a_k in the expansion for the scaling function ϕ,

$$\phi(x) = \sum_{k \in \mathbb{Z}} a_k\, \phi(2x-k), \tag{12}$$

so that ϕ satisfies various properties. I. Daubechies [9] constructed a class $\{ {}_N\phi \mid N \in \mathbb{N} \}$ of scaling functions, such that each ${}_N\phi$ vanishes outside a finite interval, satisfies the orthonormality requirement

$$\int_{-\infty}^{\infty} {}_N\phi(x-k)\, {}_N\phi(x-l)\, dx = \delta_{kl}, \qquad k, l \in \mathbb{Z}, \tag{13}$$

and has regularity ${}_N\phi \in C^{\mu N}(\mathbb{R})$, where $\mu \approx .35$. An important additional property of her scaling function ${}_N\phi$ is that low-order polynomials can be expressed as a linear combination of its translates:

$$x^j = \sum_{k \in \mathbb{Z}} \alpha_{jk}^N\, {}_N\phi(x-k), \qquad j = 0, \ldots, N-1, \tag{14}$$

where $\alpha_{jk}^N = \int {}_N\phi(x-k)\, x^j\, dx$. This latter property leads to good approximation properties; functions with several continuous derivatives have rapidly convergent expansions in bases of Daubechies wavelets.

We now explore these properties of ${}_N\phi$ and the corresponding constraints on the dilation coefficients a_k. We also present a method for computing the values of ${}_N\phi$. (We drop the prefixed subscript for ϕ where convenient.)

Compact Support. If only a finite number of the coefficients a_k for ϕ are nonzero, then ϕ vanishes outside a finite interval. In particular, if all coefficients

they have intersection that is trivial and union that is dense in $L^2(\mathbb{R})$,

$$\bigcap_n V_n = \{\mathbf{0}\}, \qquad \overline{\bigcup_n V_n} = L^2(\mathbb{R}), \tag{5}$$

they are dilates of one another,

$$f(x) \in V_n \quad \Leftrightarrow \quad f(2^n x) \in V_0, \tag{6}$$

and there exists a *scaling* function $\phi \in V_0$ whose integer translates span V_0,

$$V_0 = \left\{ f \in L^2(\mathbb{R}) \;\middle|\; f(x) = \sum \alpha_k \, \phi(x-k) \right\}. \tag{7}$$

The spaces V_n are all alike when the scale is ignored; nevertheless, a journey up the containment hierarchy can be thought of (roughly) as adding ever higher frequencies while retaining low frequencies.

The multiresolution analysis leads directly to a scalewise, orthogonal decomposition of $L^2(\mathbb{R})$. The orthogonal complement of V_n in V_{n-1}, denoted by W_n, is the building block:

$$W_n \oplus V_n = V_{n-1}, \qquad W_n \perp V_n. \tag{8}$$

The spaces W_n, $n \in \mathbb{Z}$, are dilates of W_0 and their direct sum is L^2,

$$\bigoplus_n W_n = \overline{\bigcup_n V_n} = L^2(\mathbb{R}). \tag{9}$$

The space W_0 is spanned by integer translates of a function w. Intuitively, the space V_{-1}, spanned by integer translates of two functions ($\phi(2x)$ and $\phi(2x-1)$), is twice the size of V_0, spanned by integer translates of ϕ; W_0 is the difference between these two spaces. The argument is made rigorous using group representations (omitted here). Using this characterization of W_0, and equation (9), we write L^2 as the space spanned by dilates and translates of w,

$$L^2(\mathbb{R}) = \left\{ f \;\middle|\; f(x) = \sum \alpha_{nk} \, w_{nk}(x) \text{ with } \sum {\alpha_{nk}}^2 < \infty \right\}, \tag{10}$$

where $w_{nk}(x) = 2^{-n/2} w(2^{-n}x - k)$.

What can be said about the wavelet function w? We have $w \in W_0 \subset V_{-1}$, so

$$w(x) = \sum_{k \in \mathbb{Z}} b_k \, \phi(2x-k),$$

for some coefficients $\ldots b_{-1}, b_0, b_1, \ldots$. Furthermore, a similar expansion holds for the scaling function ϕ (since $\phi \in V_0 \subset V_{-1}$):

$$\phi(x) = \sum_{k \in \mathbb{Z}} a_k \, \phi(2x-k),$$

Here the translated function $w : \mathbb{R} \to \mathbb{R}$ is

$$w(x) = \begin{cases} 1, & \text{if } x \in [0,1/2); \\ -1, & \text{if } x \in [1/2,1); \\ 0, & \text{otherwise.} \end{cases}$$

Through (3) L^2 can be decomposed into a direct sum of the W_n,

$$\bigoplus_n W_n = \overline{\bigcup_n V_n} = L^2(\mathbb{R}),$$

so L^2 is spanned by dilates and translates of w,

$$L^2(\mathbb{R}) = \left\{ f \;\middle|\; f(x) = \sum \alpha_{nk}\, w_{nk}(x) \text{ with } \sum {\alpha_{nk}}^2 < \infty \right\}.$$

The normalized dilates and translates $w_{nk}(x) = 2^{-n/2} w(2^{-n}x - k)$, $n, k \in \mathbb{Z}$ form an orthonormal basis for $L^2(\mathbb{R})$. They are wavelets, according to the definition (1), with coefficients a, b taking the discrete values $a = 2^n$ and $b = 2^n k$ for $n, k \in \mathbb{Z}$.

One last point about the Haar basis. Toward the goal of accurate, practical representation of functions the Haar basis provides little help. Given a function f with several continuous derivatives, the expansion of f in the Haar basis,

$$f(x) = \sum_{n,k \in \mathbb{Z}} \alpha_{nk}\, w_{nk}(x),$$

converges only slowly. In general, as the number of terms in a truncated expansion for f doubles, the error from neglecting the discarded terms is cut in half, so the expansion is first-order convergent. For typical functions, high accuracy is achieved only with a very large number of terms. This slow convergence limits the practical value of the Haar basis for numerical applications.

3.2 Multiresolution Analysis

The Haar basis is not new, but until recently there was no known orthogonal basis of wavelets in which the wavelet function w was even continuous, much less differentiable. In 1985 Y. Meyer [16] constructed such a basis with $w \in C^\infty(\mathbb{R})$. This was a surprise which seemed very improbable, and Meyer [15] and S. Mallat [17] developed a framework, the *multiresolution analysis*, in which to understand these bases. The Haar basis was presented in §3.1 within this framework; we now make the framework explicit.

A multiresolution analysis [9] consists of a sequence of closed subspaces V_n, $n \in \mathbb{Z}$, in $L^2(\mathbb{R})$ such that they lie in a containment hierarchy,

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots, \tag{4}$$

3.1 Haar Basis

The simplest example of a basis of wavelets, the Haar basis, consists of piecewise constant functions. The "pieces," or intervals on which the functions are constant, are of arbitrarily small size, and the basis is complete for $L^2(\mathbb{R})$. We start by defining the function $\phi : \mathbb{R} \to \mathbb{R}$ to be the characteristic function of the interval $[0,1)$,

$$\phi(x) = \begin{cases} 1, & \text{if } x \in [0,1); \\ 0, & \text{otherwise.} \end{cases}$$

The integer translates of ϕ span the space V_0 of functions constant on unit intervals,

$$V_0 = \left\{ f \in L^2(\mathbb{R}) \;\middle|\; f(x) = \sum \alpha_k \, \phi(x-k) \right\}.$$

The sum here is taken over all $k \in \mathbb{Z}$; the requirement that $f \in L^2(\mathbb{R})$ is equivalent to $\sum_k {\alpha_k}^2 < \infty$. From the space V_0 we define for each $n \in \mathbb{Z}$ the space V_n by taking dilates of functions in V_0,

$$f(x) \in V_n \quad \Leftrightarrow \quad f(2^n x) \in V_0.$$

V_1 consists of functions constant on intervals of length 2, V_{-1} consists of functions constant on intervals of length $\frac{1}{2}$, and so forth. We therefore have the containment hierarchy

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots \tag{2}$$

(note the decreasing subscripts). The only function in all subspaces is constant on $\mathbb{R}$ (and contained in L^2) so is identically zero; the union of the subspaces contains functions arbitrarily close to any function in L^2 so its closure coincides with L^2,

$$\bigcap_n V_n = \{\mathbf{0}\}, \qquad \overline{\bigcup_n V_n} = L^2(\mathbb{R}).$$

The closure $\overline{\cup V_n}$ is spanned by dilates and translates $\phi(2^n x - k)$ of ϕ, but the dilates and translates are not linearly independent (so do not form a basis). This fact is evident from the containment hierarchy (2). To construct a basis, one can exploit the hierarchy and construct *difference* spaces: for each $n \in \mathbb{Z}$ we define the space W_n to be the orthogonal complement of V_n in V_{n-1},

$$W_n \oplus V_n = V_{n-1}, \qquad W_n \perp V_n. \tag{3}$$

It is easily verified that W_n is a dilate of W_0,

$$f(x) \in W_n \quad \Leftrightarrow \quad f(2^n x) \in W_0,$$

and that, analogous to V_0, the space W_0 is spanned by integer translates of a single function,

$$W_0 = \left\{ f \in L^2(\mathbb{R}) \;\middle|\; f(x) = \sum \alpha_k \, w(x-k) \right\}.$$

sequence forms a basis for the functions defined on the interval I which are square integrable with weight ω. The sequence is uniquely determined by ω up to leading coefficients, and can be computed by the Gram-Schmidt orthogonalization process. For $I = [-1,1]$ and $\omega(x) = 1$, the sequence is the Legendre polynomials; they form an orthogonal basis for $L^2[-1,1]$. The transformation of tabulated functions to expansions of orthogonal polynomials is inexpensive in certain cases (including Legendre [5]). Differentiation and integration is generally simple and fast; the formulae for these operations depend on the weight ω.

2.3 Time-Frequency Localization

One issue which arises commonly in physical problems, and to which classical bases are not well suited, is the representation of very short, high-frequency, signals. An example in music is the attack, or beginning, of a played note, which introduces high-frequency components which decay rapidly as the note is sustained. In image processing, edges (localized high freqencies) are encountered at irregular spacings. The same is true for seismic data. Generally, there is a need for bases for which the elements representing the highest frequencies are most localized in time (or space).

The short-time Fourier transform, in which the exponential e^{inx} is multiplied by a localized window function such as $e^{-(x-a)^2/b}$, for various a, is an attempt to localize the Fourier basis elements. It does not however, localize different frequencies differently. The window width parameter b must be chosen for the degree of localization desired. The short-time Fourier transform also possesses the complication that no choice of window function leads to an orthogonal basis.

Wavelet bases have been developed to handle time-frequency localization cleanly.

§3. Construction of Bases of Wavelets

The term *wavelets* denotes a family of functions of the form

$$w_{a,b}(x) = |a|^{-1/2}\, w\left(\frac{x-b}{a}\right), \qquad a,b \in \mathbb{R},\ a \neq 0, \tag{1}$$

obtained from a single function w by the operations of dilation and translation. Such families, while named rather recently (Grossman and Morlet [13]), were used much earlier in the study of certain integral operators. Recent developments, however, have created widespread interest in wavelets among mathematicians and engineers. Explicit constructions of functions w leading to bases for $L^2(\mathbb{R})$ have propelled wavelets into applications in signal processing and the numerical solution of integral equations and partial differential equations.

In this section we present constructions of wavelets and wavelet-like bases; in §4 we give a sampling of their numerical applications.

2.2 Classical Bases

We now give a few concrete examples of bases in use for a variety of computational tasks.

The fundamental representation for periodic functions is the Fourier basis for $L^2[-\pi,\pi]$,

$$f_n(x) = e^{inx}, \qquad n \in \mathbb{Z},$$

which is an orthogonal basis under the (complex) inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\,\tilde{g}(x)\,dx,$$

where $\tilde{g}$ is the complex conjugate of g. Many properties of the Fourier basis make it well-suited to computation. Differentiation and integration of functions represented in the Fourier basis is simple:

$$\frac{df_n}{dx}(x) = n\,f_n(x), \qquad \int f_n(x)\,dx = \frac{1}{n}\,f_n(x).$$

The transformation of a function tabulated at n equispaced points on the interval $[-\pi,\pi]$ into an n-term Fourier expansion, or its inverse transformation, is accomplished rapidly by the Fast Fourier Transform. In addition, convolutions are diagonal operators in the Fourier basis. Suppose we are given an integral operator

$$(\mathcal{K}f)(x) = \int_{-\infty}^{\infty} K(x,y)\,f(y)\,dy,$$

where the kernel $K(x,y)$ is convolutional, *i.e.*,

$$K(x,y) = k(x-y).$$

The element K_{mn} of the matrix representing the operator $\mathcal{K}$ in the Fourier basis, given by the formula

$$K_{mn} = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} k(x-y)\,e^{imx}\,e^{-iny}\,dx\,dy, \qquad m,n \in \mathbb{Z},$$

satisfies $K_{mn} = 0$ if $m \neq n$, as can be seen by integrating. The ability in the Fourier basis to manipulate convolutions efficiently, as diagonal operators, leads to much of the strength of the basis as a computational tool.

Another class of bases are the orthogonal polynomials. Given an interval $I \subset \mathbb{R}$ and a positive weight function $\omega : I \to \mathbb{R}$, we can define an inner product

$$\langle f, g \rangle_\omega = \int_I f(x)\,g(x)\,\omega(x)\,dx.$$

There is a sequence of polynomials, $p_0, p_1, \ldots$, with p_j of degree j, which is orthogonal with respect to the weight ω, *i.e.*, $\langle p_m, p_n \rangle_\omega = 0$ if $m \neq n$. The

Two functions f and g are *orthogonal* if $\langle f, g\rangle = 0$. An orthogonal basis is a basis in which the functions $f_1, f_2, \ldots$ are pairwise orthogonal: $\langle f_i, f_j\rangle = 0$ for $i \neq j$. The basis is orthonormal if $\langle f_i, f_j\rangle = \delta_{ij}$.

An orthonormal basis is, in a sense, a natural representation language for functions. Perturbation of the coefficients in the representation of a function produces a commensurate perturbation in the function. Similarly, if the function is perturbed the coefficients change nearly the same amount: the representation is stable. Another important characteristic of orthonormal bases is that it is simple to determine the expansion coefficients. Given an orthonormal basis $\{f_1, f_2, \ldots\}$ for $L^2(\mathbb{R})$ and a function $f \in L^2(\mathbb{R})$, the coefficients in the expansion

$$f(x) = \sum_{n=1}^{\infty} \alpha_n \, f_n(x)$$

are given by the inner product

$$\alpha_n = \langle f_n, f\rangle = \int_{-\infty}^{\infty} f_n(x)\, f(x)\, dx \qquad n \in \mathbb{N}.$$

Without orthogonality, the coefficients must be obtained by the (often expensive) solution of a system of equations. We mention at this point that recent results suggest that *frames* may offer the simplicity of orthogonal bases without their rigidity [10], [11].

Truncated Expansions. Though a function f may be specified by an infinite expansion, actual computations require finite representations. Generally, an infinite basis is abbreviated to a finite basis, which corresponds to truncating the expansion. If the first n terms are retained, we have

$$f(x) = \sum_{j=1}^{n} \alpha_j \, f_j(x) \; + \; E_n(x),$$

where the truncation error $E_n(x)$ is given by

$$E_n(x) = \sum_{j=n+1}^{\infty} \alpha_j \, f_j(x).$$

The computation cost generally increases with n, so it is desirable for the error E_n to decay with increasing n as rapidly as possible. If k is the largest integer such that the quantity

$$n^k \|E_n\|$$

is bounded as $n \to \infty$, then we say the expansion for f is kth-order convergent. If, instead,

$$\sup_n n^k \|E_n\| < \infty$$

for all k, the expansion is super-algebraicly convergent.

The *linear span*, or closure $\overline{F}$, of a set of functions $F = \{f_1, f_2, \ldots\} \subset L^2$ is the set of linear combinations

$$f(x) = \sum_{n=1}^{\infty} \alpha_n f_n(x)$$

that are contained in L^2. The restriction that the sum be square integrable is equivalent to the requirement that the sequence of coefficients $\langle \alpha_1 \|f_1\|, \alpha_2 \|f_2\|, \ldots \rangle$ be square summable, or in l^2. A set of functions F is *linearly independent* if any proper subset F' of F has linear span $\overline{F'}$ which is a proper subset of the linear span $\overline{F}$ of F:

$$F' \subset F \text{ and } F' \neq F \quad \Rightarrow \quad \overline{F'} \neq \overline{F}.$$

A set of functions F is a *basis* for a space S if $\overline{F} = S$ and F is linearly independent. We will concern ourselves with bases for $L^2(\mathbb{R})$.

Stability and Orthogonality. If a function is represented in terms of functions that are not linearly independent, the representation is not unique. For example, if

$$\text{linear span}\{f_1, f_2, f_3, \ldots\} = \text{linear span}\{f_2, f_3, \ldots\},$$

and

$$f(x) = \sum_{n} \alpha_n f_n(x),$$

then the coefficient α_1 can be chosen arbitrarily. On the other hand, linear independence is enough to guarantee representational uniqueness in theory, but is not sufficient when numerical computations are carried out to finite precision. The functions $\langle f_0, f_1, \ldots \rangle$, defined by the formula

$$f_n(x) = \begin{cases} x^n & \text{if } x \in [0,1]; \\ 0 & \text{otherwise;} \end{cases}$$

are linearly independent but f_{11} can be represented as a linear combination of $f_0, \ldots, f_{10}$ to 6 digit accuracy. Perturbation of a function

$$f(x) = \sum_{n=0}^{11} \alpha_n f_n(x)$$

by one part in a million can correspond to variations of $\|\alpha\|_{l^2}$ by several per cent (even if each f_j were normalized so $\|f_j\| = 1$). The monomials are therefore seriously deficient for numerical use as a basis. In computations, they are replaced by the orthogonal polynomials.

The *inner product* of two functions $f, g \in L^2(\mathbb{R})$ is the integral

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)\, g(x)\, dx.$$

structure. Furthermore, the sparse band structure represents an approximation (to arbitrary precision) of the original integral operator.

Despite these caveats, the solution of a wide range of integral equations is transformed using wavelets from a direct procedure requiring order $O(n^3)$ operations to one requiring only order $O(n)$. Here n is the number of points in the discretization of the domain.

Many time-dependent problems formulated as partial differential equations require adaptive representations in carrying out the time evolution. Typically, a small part of the domain has most of the activity, and the representation must have high resolution there. In the rest of the domain, such high resolution is wasted (and costly). Various adaptive mesh techniques have been developed to address this issue, but they often suffer accuracy problems in the application of operators, multiplication of functions, and so forth. Wavelets offer promise in providing a consistent, simple adaptive framework.

In §2 we provide background to the study of wavelets by reviewing the definition and some examples of mathematical bases. In §3 we present several constructions of wavelets and wavelet-like bases and discuss their fundamental properties. In §4 we introduce the application of wavelets to the solution of integral equations and to the representation of differential operators. Finally, we summarize in §5.

§2. Function Representations in Mathematical Physics

2.1 Mathematical Bases

The *norm* of a sequence $\alpha = \langle \alpha_1, \alpha_2, \ldots \rangle$, a measure of its size, will be defined by the formula

$$\|\alpha\|_{l^2} = \Big(\sum_n {\alpha_n}^2 \Big)^{1/2}.$$

The space l^2 consists of the square summable sequences α: $\|\alpha\|_{l^2} < \infty$.

The norm of a function $f : \Omega \to \mathbb{R}$, will be defined as the L^2-norm

$$\|f\|_{L^2} = \left(\int_\Omega f(x)^2 \, dx \right)^{1/2},$$

where Ω is the domain of f. In this chapter we restrict ourselves primarily to $\Omega = \mathbb{R}$ and to functions f with $\|f\|_{L^2} < \infty$ (*i.e.*, $f \in L^2$); the function norm $\|f\|_{L^2}$ will be abbreviated $\|f\|$. A sequence of functions $f_1, f_2, \ldots$ *converges* to a function f if the difference $f_n - f$ becomes arbitrarily small, *i.e.*, if $\|f_n - f\| \to 0$ as $n \to \infty$. A series of functions $\sum_n f_n$ converges to f if the sequence of partial sums $S_1, S_2, \ldots$ converges to f. The partial sum S_n is defined by the formula

$$S_n(x) = \sum_{j=1}^{n} f_j(x).$$

Wavelets and Other Bases for Fast Numerical Linear Algebra

Bradley K. Alpert

Abstract. *The fundamental ideas of wavelets are introduced within the context of mathematical physics. We present essential background notions of mathematical bases, and discuss Fourier, polynomial, and wavelets bases in this light. We construct several types of wavelets and wavelet-like bases and illustrate their use in algorithms for the solution of a variety of integral and differential equations.*

§1. Introduction

The solution of problems of physics, requiring the numerical solution of differential and integral equations, ranks among the most compute-intensive applications currently feasible. The field of scientific computation is concerned with both hardware and algorithmic improvement for the modelling of increasingly complex problems. Recently, the development of new mathematical bases for scientific computation has enabled the construction of algorithms which are dramatically more efficient than earlier algorithms. Wavelets permit the accurate representation of a variety of functions and operators without redundancy. Through the ability to represent local, high-frequency information with localized basis elements, wavelets allow adaption in a straightforward, consistent fashion.

For a variety of applications, sparse matrix representation of differential operators has been the key to efficient algorithms. Integral operators, by contrast, are represented in classical bases as dense matrices. These representations lead to algorithms which, for large-scale problems, are often prohibitively slow. The most notable exception to this rule is for convolutional operators, which are represented in the Fourier basis as diagonal matrices, and which have correspondingly fast algorithms. Wavelets can be viewed as a "diagonalizing" basis for a wider class of integral operators. The quotes are necessary here, because the statement is only approximately true. Wavelet expansions of integral operators are not exactly diagonal; rather, they have a peculiar band

Wavelets–A Tutorial
C. K. Chui (ed.), pp. 1–4.
Copyright © 1991 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN x-yy-zzzzzz-X