

## Wavelets and Fast Numerical Algorithms

GREGORY BEYLKIN

In the usual transform methods, the functions of the basis (e.g. exponentials, Chebyshev polynomials, etc.) are chosen to be eigenfunctions of some differential operator (e.g. solutions of the Sturm-Liouville problem). The choice of the differential operator and, hence, of the basis functions, is dictated by the availability of fast algorithms for expanding an arbitrary function into the basis.

bases are very narrow.

Wavelets, on the other hand, are not solutions of a differential equation. These functions are defined recursively and are generated via an iterative algorithm. They are translations and dilations of a single function.<sup>1</sup> Instead of diagonalizing some differential operator, representations in the wavelet bases reduce a wide class of operators to a sparse form. Using the orthogonality of the wavelet



the non-standard form (which achieves uncoupling among the scales) and the associated fast numerical algorithms. Examples of non-standard forms of several basic operators (e.g. derivatives) will be computed explicitly.

### 1. Multiresolution analysis and wavelets.

We briefly outline here the properties of compactly supported wavelets and refer for details to [16], [17] and [28]. Let us start with the notion of a multiresolution analysis [26], [23] which captures the essential features of all multiresolution approaches developed so far.

DEFINITION 1.1. A multiresolution analysis (MRA) is a sequence of nested

$$(1.1) \quad \dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots$$

such that

- (i)  $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$  and  $\bigcup_{j \in \mathbb{Z}} V_j$  is dense in  $L^2(\mathbb{R}^d)$ .
- (ii) For any  $f \in L^2(\mathbb{R}^d)$  and any  $j \in \mathbb{Z}$ ,  $f(x) \in V_{-j}$  if and only if  $f(2^j x) \in V_0$ .

$$(1.6) \quad V_n \subset \dots \subset V_2 \subset V_1 \subset V_0, \quad V_0 \subset L^2(\mathbb{R}^d)$$

instead of (1.4). In numerical realizations the subspace  $V_0$  is finite dimensional.

(1.9) as

$$(1.10) \quad \hat{\varphi}(\xi) = m_0(\xi/2)\hat{\varphi}(\xi/2),$$

where

$$(1.11) \quad m_0(\xi) = \frac{1}{\sqrt{2}} \sum_{k=0}^{L-1} h_k e^{ik\xi}$$

and the  $2\pi$ -periodic function  $m_0$  is defined as

$$(1.12) \quad m_0(\xi) = 2^{-1/2} \sum_{k=0}^{L-1} h_k e^{ik\xi}$$

Second, the orthogonality of  $\{\varphi(x - k)\}_{k \in \mathbb{Z}}$  implies that

$$(1.13) \quad \delta_{k0} = \int_{-\infty}^{+\infty} \varphi(x - k)\varphi(x) dx = \int_{-\infty}^{+\infty} |\hat{\varphi}(\xi)|^2 e^{-ik\xi} d\xi,$$

and, therefore,

and

$$(1.15) \quad \sum_{l \in \mathbb{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2 = \frac{1}{2\pi}.$$

Using (1.10), we obtain

$$(1.16) \quad \sum_{l \in \mathbb{Z}} |m_0(\xi/2 + \pi l)|^2 |\hat{\varphi}(\xi/2 + \pi l)|^2 = \frac{1}{2\pi},$$

and, by taking the sum in (1.16) separately over odd and even indices, we have

or, equivalently, the Fourier transform of  $\psi$  by

where

$$k=L-1$$

$$k=0$$

it is not difficult to show (see e.g. [28], [16], [17]), that for each fixed scale  $j \in \mathbf{Z}$ , the wavelets  $\{\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)\}_{k \in \mathbf{Z}}$  form an orthonormal basis of  $\mathbf{W}_j$ .

Equation (1.18) can also be viewed as the condition for exact reconstruction for a pair of the quadrature mirror filters (QMFs)  $H$  and  $G$ , where  $H = \{h_k\}_{k=0}^{k=L-1}$  and  $G = \{g_k\}_{k=0}^{k=L-1}$ . Such exact QMF filters were first introduced by Smith and Barnwell [32] for subband coding.

We will not go into a full discussion of the necessary and sufficient conditions for the quadrature mirror filters  $H$  and  $G$  to generate a wavelet basis and refer to [17] for the details. The coefficients of the quadrature mirror filters  $H$  and  $G$  are computed by solving a set of algebraic equations (see e.g. [17]). The

(1.26)

with  $\mathcal{M}_0 = 1$ .

Alternatively, using

$$(1.27) \quad \hat{\varphi}(\xi) = (2\pi)^{-1/2} \prod_{j=1}^{\infty} m_0(2^{-j}\xi),$$

the moments  $\mathcal{M}_m$  may be obtained within the desired accuracy as a limit of a



with kernel  $K(x_{ij})$ . The orthogonal projection operators on the subspace  $V_j$ .

$j \in \mathbb{Z}$ ,

$$(2.3) \quad P_j : L^2(\mathbb{R}) \rightarrow V_j,$$

are given by

$$(2.4) \quad (P_j f)(x) = \sum_k \langle f, \varphi_{j,k} \rangle \varphi_{j,k}(x).$$

Expanding  $T$  in a "telescopic" series, we obtain

$$(2.5) \quad T = \sum_{j \in \mathbb{Z}} (Q_j T Q_j + Q_j T P_j + P_j T Q_j),$$

where

$$(2.6) \quad Q_j = P_{j-1} - P_j$$

is the projection operator on the subspace  $W_j$ . If there is a coarsest scale  $n$ , then instead of (2.5) we have

$$(2.7) \quad T = \sum_{j=-\infty}^n (Q_j T Q_j + Q_j T P_j + P_j T Q_j) + P_n T P_n,$$

and if the scale  $j = 0$  is the finest scale, then

$$(2.8) \quad T = \sum_{j=0}^n (Q_j T Q_j + Q_j T P_j + P_j T Q_j) + P_0 T P_0,$$

where  $T \sim T_0 = P_0 T P_0$  is a discretization of the operator  $T$  on the finest scale.

The non-standard form is a representation (see [7]) of the operator  $T$  as a

where the operators  $T_j$ ,

$$(2.14) \quad T_j : \mathbf{V}_j \rightarrow \mathbf{V}_j,$$

are defined by  $T_j = P_j T P_j$ .

If there is a coarsest scale  $n$  then

$$(2.15) \quad T = \{\{A_j, B_j, \Gamma_j\}_{j \in \mathbb{Z}: j \leq n}, T_n\},$$

where  $T_n = P_n T P_n$ . If the number of scales is finite, then  $j = 1, 2, \dots, n$  in (2.15) and the operators are organized as blocks of a matrix (see Figures 1 and 2).

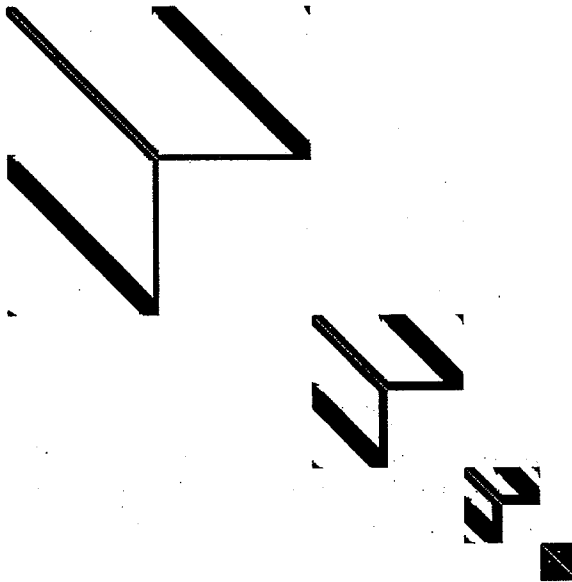


FIGURE 2. An example of a matrix in the non-standard form (see Example 4.2).

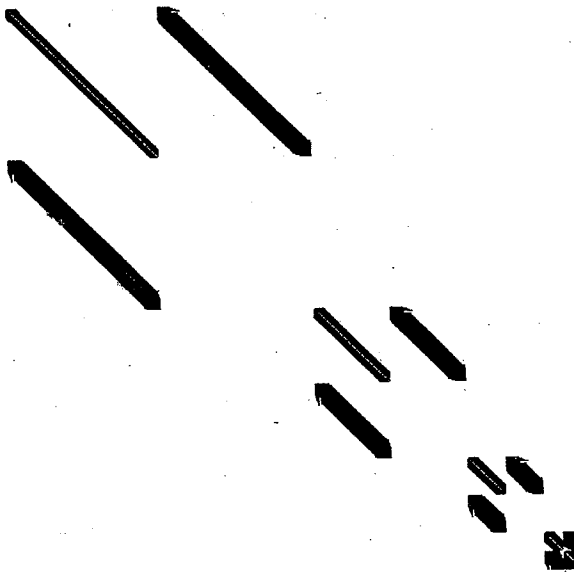


FIGURE 3. The non-standard form of the same matrix as in Figure 2. The blocks are arranged in a staircase pattern.

$$(2.17) \quad \beta_{k,k'}^j = \int \int K(x,y) \psi_{j,k}(x) \varphi_{j,k'}(y) dx dy,$$

and

$$(2.18) \quad \gamma_{k,k'}^j = \int \int K(x,y) \varphi_{j,k}(x) \psi_{j,k'}(y) dx dy.$$

The operator  $T_j$  is represented by the matrix  $s^j$ ,

$$(2.19) \quad s_{k,k'}^j = \int \int K(x,y) \varphi_{j,k}(x) \varphi_{j,k'}(y) dx dy.$$

### 3. The standard form.

The standard form is the representation of an operator in the tensor product basis. Instead of introducing the standard form in this manner, we emphasize the connection with the non-standard form. The standard form is obtained by representing

$$(3.1) \quad \mathbf{V}_j = \bigoplus_{j' > j} \mathbf{W}_{j'},$$

and considering for each scale  $j$  the operators  $\{B_j^{j'}, \Gamma_j^{j'}\}_{j' > j}$ ,

$$(3.2) \quad B_j^{j'} : \mathbf{W}_{j'} \rightarrow \mathbf{W}_j,$$

$$(3.3) \quad \Gamma_j^{j'} : \mathbf{W}_j \rightarrow \mathbf{W}_{j'}.$$

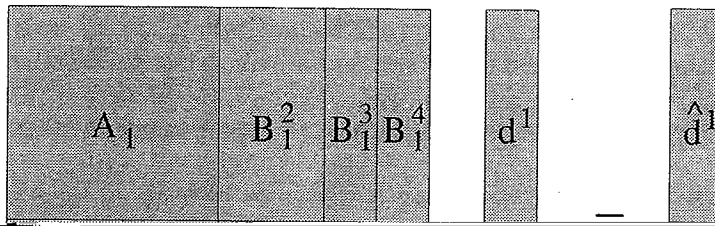
If there is a coarsest scale  $n$ , then instead of (3.1) we have

$$(3.4) \quad \mathbf{V}_j = \mathbf{V}_n \bigoplus_{j'=j+1}^{j'=n} \mathbf{W}_{j'}.$$

In this case, the operators  $\{B_j^{j'}, \Gamma_j^{j'}\}$  for  $j' = j+1, \dots, n$  are as in (3.2) and (3.3) and, in addition, for each scale  $j$  there are operators  $\{B_j^{n+1}\}$  and  $\{\Gamma_j^{n+1}\}$ ,

$$(3.5) \quad B_j^{n+1} : \mathbf{V}_n \rightarrow \mathbf{W}_j,$$

$$(3.6) \quad \Gamma_j^{n+1} : \mathbf{W}_j \rightarrow \mathbf{V}_n.$$



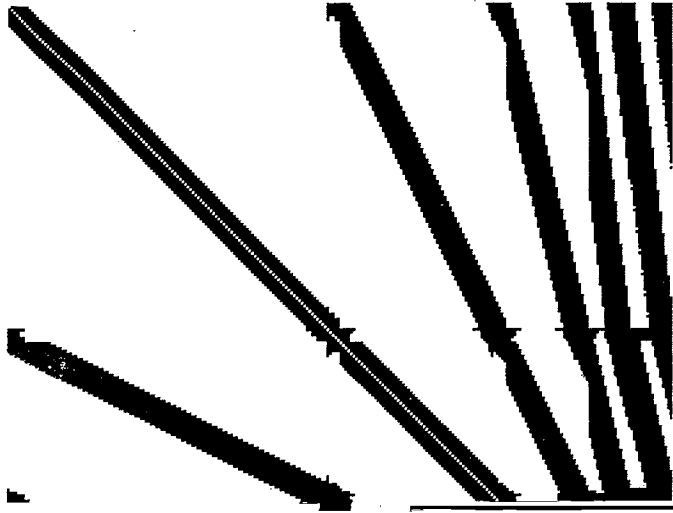


FIGURE 5. An example of a matrix in the standard form (see Example 4.2).

for some  $M \geq 1$ . Then by choosing the wavelet basis with  $M$  vanishing moments, the coefficients  $\alpha_{i,l}^j, \beta_{i,l}^j, \gamma_{i,l}^j$  of the non-standard form (see (2.16) - (2.18)) satisfy the estimate

$$(4.3) \quad |\alpha_{i,l}^j| + |\beta_{i,l}^j| + |\gamma_{i,l}^j| \leq \frac{C_M}{1 + |i-l|^{M+1}},$$

for all

where  $K$  is the distributional kernel of  $T$ , then assuming that the symbols  $\sigma$  of

$$(4.7) \quad |\partial_{\xi}^{\alpha} \partial_x^{\beta} \sigma(x, \xi)| \leq C_{\alpha, \beta} (1 + |\xi|)^{\lambda - \alpha + \beta}$$

$$(4.8) \quad |\partial_{\xi}^{\alpha} \partial_x^{\beta} \sigma^*(x, \xi)| \leq C_{\alpha, \beta} (1 + |\xi|)^{\lambda - \alpha + \beta},$$

we have the inequality

$$(4.9) \quad |\alpha_{i,l}^j| + |\beta_{i,l}^j| + |\gamma_{i,l}^j| \leq \frac{2^{\lambda j} C_M}{(1 + |i - l|)^{M+1}},$$

for all integer  $i, l$ .

Suppose now that we approximate the operator  $T_0$  by the operator  $T_0^B$  obtained from  $T_0$  by setting to zero all coefficients of matrices  $\alpha^j$ ,  $\beta^j$  and  $\gamma^j$  outside

bands of width  $B \geq 2M$  around their diagonals. We obtain

$$(4.10) \quad \|T_0^B - T_0\| \leq \frac{C}{B^M} \log_2 N,$$

where  $C$  is a constant determined by the kernel  $K$  and  $\log_2 N$  is the number of scales in the representation. In most numerical applications, the accuracy  $\varepsilon$  of calculations is fixed, and the parameters of the algorithm (in our case, the band

$$(4.15) \quad \gamma(y) = T^*(1)(y)$$

satisfy the dyadic bounded mean oscillation (B.M.O.) condition,

$$(4.16) \quad \sup_J \frac{1}{|J|} \int_J |\beta(x) - m_J(\beta)|^2 dx \leq C,$$

where  $J$  is a dyadic interval and

$$(4.17) \quad m_J(\beta) = \frac{1}{|J|} \int_J \beta(x) dx.$$

Again we refer to [7] for details.

The compression of operators results in fast algorithms for evaluation of operators on functions. We present here one example and refer to [7] for additional examples.

EXAMPLE 4.2. In this example, we consider the matrix

$$A_{ij} = \begin{cases} \frac{1}{i-j} & i \neq j, \\ 0 & i = j, \end{cases}$$

ments. Setting to zero all entries whose absolute values are smaller than  $10^{-7}$ , we obtain the non-standard form where the non-zero elements are shown in black in Figure 2. The results of numerical experiments are shown in Figure 3.



by computing in the wavelet system of coordinates. Finally, the last column contains the compression coefficients  $C_{comp}$ , defined by the ratio of  $N^2$  to the number of non-zero elements in the non-standard form of the matrix.

### 5. The operator $d/dx$ in wavelet bases.

Hilbert and Riesz transforms) we may compute the non-standard form in the wavelet bases by solving a small system of linear algebraic equations [4]. As an example, we construct the non-standard form of the operator  $d/dx$ . The matrix elements  $\alpha_{il}^j$ ,  $\beta_{il}^j$ , and  $\gamma_{il}^j$  of  $A_j$ ,  $B_j$ , and  $\Gamma_j$ , where  $i, l, j \in \mathbf{Z}$  for the operator  $d/dx$  are easily computed as

$r^\infty$

where

$$(5.10) \quad r_l = \int_{-\infty}^{+\infty} \varphi(x-l) \frac{d}{dx} \varphi(x) dx, \quad l \in \mathbf{Z}.$$

Therefore, the representation of  $d/dx$  is completely determined by the coefficients  $r_l$  in (5.10) or in other words, by the representation of  $d/dx$  on the subspace  $V_0$ .

Rewriting (5.10) in terms of  $\hat{\varphi}(\xi)$  (see (1.11)), we obtain

$$(5.11) \quad r_l = \int_{-\infty}^{+\infty} |\hat{\varphi}(\xi)|^2 (i\xi) e^{-il\xi} d\xi.$$

Thus, the coefficients  $r_l$  depend only on the autocorrelation function of the scaling function  $\varphi$ , rather than the scaling function itself since the integral in (5.11)

Solving equations (5.12), (5.13), we present the results for Daubechies' wavelets with  $M = 2, 3$ . For further examples we refer to [4].

1.  $M = 2$

$$a_1 = \frac{9}{8}, \quad a_3 = -\frac{1}{8},$$

and

$$2 \quad 1$$

We note that the coefficients  $(-1/12, 2/3, 0, -2/3, 1/12)$  of this example can be found in many books on numerical analysis as a choice of coefficients for numerical differentiation.

2.  $M = 3$

$$a_1 = \frac{75}{64}, \quad a_3 = -\frac{25}{128}, \quad a_5 = \frac{3}{128},$$

and

$$r_1 = -\frac{272}{365}, \quad r_2 = \frac{53}{365}, \quad r_3 = -\frac{16}{1095}, \quad r_4 = -\frac{1}{2920}.$$

The structure of non-standard and standard forms of derivative operators is illustrated in Figures 6 and 7.



FIGURE 6. Sparse structure of the non-standard form of derivative operators. The width of the bands depends only on the choice of the basis and is equal to  $2L - 3$ .

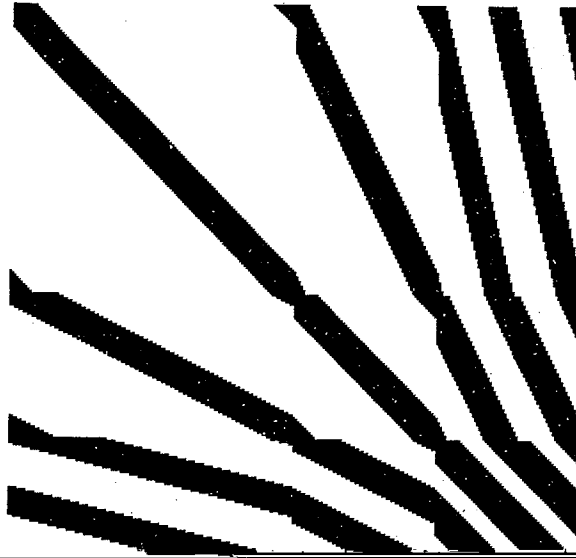


FIGURE 7. Sparse structure of the standard form of derivative operators.

For the coefficients  $r_i^{(n)}$  of  $d^n/dx^n$ ,  $n > 1$ , the system of linear algebraic equations is similar to that for the coefficients of  $d/dx$ . This part of (1.5.10)

We present here two tables illustrating such preconditioning applied to the standard form of the second derivative. In the following we shall let  $\lambda = 1$ .

N	$\kappa$	$\kappa_p$
64	0.14545E+04	0.10792E+02
128	0.58181E+04	0.11511E+02
256	0.23272E+05	0.12091E+02
512	0.93089E+05	0.12604E+02
1024	0.37236E+06	0.13045E+02

TABLE 2. Condition number and spectral ratio of the modified

N	$\kappa$	$\kappa_p$
64	0.10472E+04	0.43542E+01
128	0.41886E+04	0.43595E+01
256	0.16754E+05	0.43620E+01
512	0.67018E+05	0.43628E+01

and

$$(5.26) \quad \mathcal{M}_{\Phi}^m = \int_{-\infty}^{+\infty} y^m \Phi(y) dy = 0, \quad \text{for } 1 \leq m \leq 2M - 1.$$

The vanishing moments of the autocorrelation function  $\Phi$  allow us to compute the elements of the matrix  $t_i^{(j)}$  for large  $l$  and sufficiently fine scales  $i < 0$ .

Expanding the kernel  $K$  in its Taylor series, we obtain from (5.23)

$$(5.27) \quad t_i^{(j)} = 2^j K(2^j l) + \frac{(-1)^{2M} 2^{(2M+1)j}}{(2M)!} \int_{-\infty}^{+\infty} K^{(2M)}(2^j(l - \tilde{y})) \Phi(y) dy,$$

where  $\tilde{y} = \tilde{y}(y, l)$  and  $K^{(2M)}$  denotes the  $(2M)$ th derivative of  $K$ . The decay of  $K^{(2M)}(2^j(l - \tilde{y}))$  for large  $l$  is faster than that of the original kernel (see (4.1) and (4.2) with an appropriate change of variables).



algebraic equations

$$(5.33) \quad r_l = 2^\alpha \left[ r_{2l} + \frac{1}{2} \sum_{k=1}^{L/2} a_{2k-1} (r_{2l-2k+1} + r_{2l+2k-1}) \right],$$

where the coefficients  $a_{2k-1}$  are given in (5.14). Using (5.27), we obtain the asymptotics of  $r_l$  for large  $l$ .

where  $A^*$  is the adjoint matrix and  $\alpha$  is chosen so that the largest eigenvalue of  $\alpha A^*A$  is less than one. Then the sequence  $X_k$  converges to the generalized inverse  $A^\dagger$ .

Combining this iteration with fast multiplication algorithms, we obtain an algorithm for constructing the generalized inverse in at most  $O(N \log^2 N \log R)$  operations, where  $R$  is the condition number of the matrix. (By the condition





18. L. Greengard, *Potential flow in channels*, SIAM J. Sci. Stat. Comput. **11** (1990), no. 4, 603-620.
19. L. Greengard and V. Rokhlin, *A fast algorithm for particle simulations*, J. Comp. Phys.