Lecture Notes for short course, AMS-93

Wavelets and Fast Numerical Algorithms

G. Beylkin
Program in Applied Mathematics
University of Colorado at Boulder
Boulder, CO 80309-0526

Numerical algorithms using wavelet bases are similar to other transform methods in that vectors and operators are expanded into a basis and the computations take place in the new system of coordinates. As in all transform methods, such approach seeks advantages in that the computation is faster in the new system of coordinates than in the original domain. However, due to the recursive definition of wavelets, their controllable localization in both space and wave number (time and frequency) domains, and the vanishing moments property, wavelet based algorithms exhibit a number of new and important properties.

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. Unfortunately, classes of operators which have a sparse representation in such 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 [1]. Instead of diagonalizing some differential operator, representations in the wavelet bases reduce a wide class of operators to a sparse form. Here the orthogonality of wavelets to the low degree polynomials (the vanishing moments property) plays a crucial role in producing sparse systems [2].

Historically, the orthonormal bases of wavelets were first constructed by Stromberg [33] and then by Meyer [25]. Later, the notion of the Multiresolution Analysis was introduced by Meyer [26] and Mallat [23]. Orthonormal bases of compactly supported wavelets were constructed by Daubechies [16]. There are many new constructions of orthonormal bases with a controllable localization in the time–frequency domain, notably "wavelet-packet" bases in [13] and [15], local trigonometric bases in [14] and [24], and wavelet bases on the interval in [11], [12] and [22]. There exists a very important connection between wavelets and the technique of subband coding in signal processing. In fact,

[1]It is also possible to construct bases with translations and dilations of several functions, see e.g. [1].

[2]This property and the fact that the basis is orthonormal, distinguish the wavelet bases from the hierarchical bases.

the discrete wavelet transform is accomplished by a pair of so-called quadrature mirror filters. Quadrature mirror filters (QMFs) with the "exact reproduction property" were introduced by Smith and Barnwell [32].

Wavelets have some of their historical roots in Littlewood-Paley and Calderón-Zygmund theories (see e.g. [28]) which have been powerful tools in the analysis of linear and non-linear operators. In numerical analysis some of the ingredients of Calderón-Zygmund theory appear in the Fast Multipole Method (FMM) for computing potential interactions [30], [19], [10]. FMM was designed for computing potential interactions between N particles in $O(-N \log \epsilon)$ operations (instead of $O(N^2)$ operations). The reduction of the complexity in FMM is achieved by approximating the far field effect of a cloud of charges located in a box by the effect of a single multipole at the center of the box. All boxes are then organized in a dyadic hierarchy enabling an efficient $O(N)$ algorithm.

The fast wavelet-based algorithms of [8] provide a systematic generalization of the FMM and its descendents (e.g. [29], [2], [18]) to all Calderón-Zygmund and pseudo-differential operators. The subdivision of the space and its organization in a dyadic hierarchy are a consequence of the multiresolution properties of the wavelet bases, while the vanishing moments of the basis functions make them useful tools for approximation.

A novel aspect of representing operators in the wavelet bases is the so-called non-standard form [8]. The remarkable feature of the non-standard form is the uncoupling of the interactions between the scales. The non-standard form leads to an order N algorithm for evaluating operators on functions. It is also quite remarkable that the error estimates for the non-standard form lead to a proof of the celebrated "T(1)" theorem of David and Journé (see [8]). The non-standard forms of many basic operators, such as derivatives, fractional derivatives, the Hilbert and Riesz transforms, may be computed explicitly [4]. A straightforward realization, or the standard form, by contrast, contains matrix entries reflecting "interactions" between all pairs of scales. The standard form yields, in general, only an order $N \log(N)$ algorithm for evaluating operators on functions.

The representation of wide classes of operators in wavelet bases may be viewed as a method for their "compression", i.e., conversion to a sparse form. For these operators sparse representations lead to fast algorithms for matrix multiplications. Since the performance of many algorithms requiring multiplication of dense matrices has been limited by $O(N^3)$ operations, these fast algorithms address a critical numerical issue.

Examples of such algorithms requiring multiplication of matrices are, for instance, an iterative algorithm for constructing the generalized inverse [31], the scaling and squaring method for computing the exponential of an operator, and similar algorithms for sine and cosine of an operator, to mention a few. By replacing the ordinary matrix multiplication in these algorithms by the fast multiplication in the wavelet bases, the number of operations is reduced to, essentially, $O(N)$ operations. For example, if both the operator and its generalized inverse admit sparse representations in the wavelet basis, then the iterative algorithm [31] for computing the generalized inverse requires only $O(N \log \kappa)$

operations, where κ is the condition number of the matrix. Various numerical examples and applications may be found in [7], [1] and [9].

Solving the two-point boundary value problem for the elliptic differential operators in the wavelet "system of coordinates" allows us to construct the Green's function (the inverse operator) in $O(N)$ operation. We note that the ordinary matrix representation of the Green's function requires $O(N^2)$ significant entries but the representation of the Green's function in the wavelet bases requires (for a given accuracy) only $O(N)$ entries. The main tool in constructing the Green's function numerically is the diagonal preconditioner available for periodized differential operators in the wavelet bases [4], [5] (see also [21]).

Unfortunately, the format of one lecture does not allow us to cover all the developments or mention all the results available today. Instead, we will review several features of the new numerical methodology based on the wavelet representations. Starting from the notion of multiresolution analysis, we will consider 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.

I 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 I.1 *A multiresolution analysis is a decomposition of the Hilbert space* $\mathbf{L}^2(\mathbf{R}^\mathbf{d})$, $\mathbf{d} \geq 1$, *into a chain of closed subspaces*

$$\ldots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \ldots \tag{1.1}$$

such that

1. $\bigcap_{j \in \mathbf{Z}} \mathbf{V}_j = \{0\}$ *and* $\bigcup_{j \in \mathbf{Z}} \mathbf{V}_j$ *is dense in* $\mathbf{L}^2(\mathbf{R}^\mathbf{d})$
2. *For any* $f \in \mathbf{L}^2(\mathbf{R}^\mathbf{d})$ *and any* $j \in \mathbf{Z}$, $f(x) \in \mathbf{V}_j$ *if and only if* $f(2x) \in \mathbf{V}_{j-1}$
3. *For any* $f \in \mathbf{L}^2(\mathbf{R}^\mathbf{d})$ *and any* $k \in \mathbf{Z}^\mathbf{d}$, $f(x) \in \mathbf{V}_0$ *if and only if* $f(x-k) \in \mathbf{V}_0$
4. *There exists a function* $\varphi \in \mathbf{V}_0$ *such that* $\{\varphi(x-k)\}_{k \in \mathbf{Z}^\mathbf{d}}$ *is a Riesz basis of* $\mathbf{V}_0$.

In this lecture we use only orthonormal bases, so that we replace Condition 4 by

4'. There exists a function $\varphi \in \mathbf{V}_0$ *such that* $\{\varphi(x-k)\}_{k \in \mathbf{Z}^\mathbf{d}}$ *is an orthonormal basis of* $\mathbf{V}_0$.

Let us define the subspaces $\mathbf{W}_j$ as an orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j-1}$,

$$\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{W}_j, \tag{1.2}$$

and represent the space $\mathbf{L}^2(\mathbf{R}^\mathbf{d})$ as a direct sum

$$\mathbf{L}^2(\mathbf{R}^\mathbf{d}) = \bigoplus_{j \in \mathbf{Z}} \mathbf{W}_j. \tag{1.3}$$

Selecting the coarsest scale n, we may replace the chain of the subspaces (1.1) by

$$\mathbf{V}_n \subset \ldots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \ldots, \tag{1.4}$$

and obtain

$$\mathbf{L}^2(\mathbf{R}^\mathbf{d}) = \mathbf{V}_n \bigoplus_{j \leq n} \mathbf{W}_j. \tag{1.5}$$

If there is a finite number of scales then, without loss of generality, we set $j = 0$ to be the finest scale and consider

$$\mathbf{V}_n \subset \ldots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0, \quad \mathbf{V}_0 \subset \mathbf{L}^2(\mathbf{R}^\mathbf{d}) \tag{1.6}$$

instead of (1.4). In numerical realizations the subspace $\mathbf{V}_0$ is finite dimensional.

First, let us consider bases in $\mathbf{L}^2(\mathbf{R})$, $d = 1$. The function φ is the so-called *scaling function* and, with its help, we may define the function ψ, the *wavelet*, such that the set of functions $\{\psi(x-k)\}_{k\in\mathbf{Z}}$ is an orthonormal basis of $\mathbf{W}_0$,

An example of a multiresolution analysis satisfying Definition I.1 with Condition 4' is the chain of subspaces generated by the Haar basis [20]. The scaling function in this case is the characteristic function of the interval $(0,1)$. The Haar function is defined as

$$h(x) = \begin{cases} 1 & \text{for } \ 0 < x < 1/2 \\ -1 & \text{for } \ 1/2 \leq x < 1 \\ 0 & \text{elsewhere.} \end{cases}, \tag{1.7}$$

and the Haar basis is formed by functions $h_{j,k}(x) = 2^{-j/2}h(2^{-j}x - k)$, $j,k \in \mathbf{Z}$.

Wavelet bases (with a smooth scaling function φ in Condition 4') generalizing the Haar functions were first constructed by Stromberg [33] and then Meyer [25]. The notion of the Multiresolution Analysis was introduced by Meyer [26] and Mallat [23] and is more recent than the constructions of [33], [25] and, of course, of [20]. Compactly supported wavelets with vanishing moments were constructed by I. Daubechies [16]; those are the ones we will use in this lecture. However, most of the results that we discuss do not depend on this particular choice of the wavelet basis.

The vanishing moments property simply means that the basis functions are chosen to be orthogonal to the low degree polynomials, namely, if the set of functions $\{\psi(x-k)\}_{k\in\mathbf{Z}}$ is an orthonormal basis of $\mathbf{W}_0$, then

$$\int_{-\infty}^{+\infty} \psi(x)x^m dx = 0, \qquad m = 0, \ldots, M-1. \tag{1.8}$$

For the Haar function in (1.7) $M = 1$ and it is indeed trivially orthogonal to constants.

There are two immediate consequences of Definition I.1 with Condition 4'. First, the function φ may be expressed as a linear combination of the basis functions of $\mathbf{V}_{-1}$. Since the functions $\{\varphi_{j,k}(x) = 2^{-j/2}\varphi(2^{-j}x-k)\}_{k\in\mathbf{Z}}$ form an orthonormal basis of $\mathbf{V}_j$, we have

$$\varphi(x) = \sqrt{2}\sum_{k=0}^{L-1} h_k \varphi(2x-k). \tag{1.9}$$

In general, the sum in (1.9) does not have to be finite and, by choosing a finite sum in (1.9), we are selecting compactly supported wavelets. We may rewrite (1.9) as

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

where

$$\hat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x)\, \mathrm{e}^{\mathrm{i}x\xi}\, dx, \tag{1.11}$$

and the 2π-periodic function m_0 is defined as

$$m_0(\xi) = 2^{-1/2} \sum_{k=0}^{L-1} h_k \mathrm{e}^{\mathrm{i}k\xi}. \tag{1.12}$$

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

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

and, therefore,

$$\delta_{k0} = \int_0^{2\pi} \sum_{l\in\mathbf{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2\, \mathrm{e}^{-\mathrm{i}k\xi}\, d\xi, \tag{1.14}$$

and

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

Using (1.10), we obtain

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

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

$$\sum_{l\in\mathbf{Z}} |m_0(\xi/2 + 2\pi l)|^2 |\hat{\varphi}(\xi/2 + 2\pi l)|^2 + \sum_{l\in\mathbf{Z}} |m_0(\xi/2 + 2\pi l + \pi)|^2 |\hat{\varphi}(\xi/2 + 2\pi l + \pi)|^2 = \frac{1}{2\pi}. \tag{1.17}$$

Using the 2π-periodicity of the function m_0 and (1.15), we obtain (after replacing $\xi/2$ by ξ) a necessary condition

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1, \tag{1.18}$$

for the coefficients h_k in (1.12). On defining the function ψ by

$$\psi(x) = \sqrt{2} \sum_k g_k \varphi(2x - k), \tag{1.19}$$

where

$$g_k = (-1)^k h_{L-k-1}, \qquad k = 0, \ldots, L-1, \tag{1.20}$$

or, equivalently, the Fourier transform of ψ by

$$\hat{\psi}(\xi) = m_1(\xi/2)\hat{\varphi}(\xi/2), \tag{1.21}$$

where

$$m_1(\xi) = 2^{-1/2} \sum_{k=0}^{k=L-1} g_k \mathrm{e}^{\mathrm{i}k\xi} = \mathrm{e}^{-\mathrm{i}\xi} \overline{m}_0(\xi + \pi), \tag{1.22}$$

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 the 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 number L of the filter coefficients in (1.12) and (1.22) is related to the number of vanishing moments M, and $L = 2M$ for the wavelets constructed in [16]. If additional conditions are imposed (see [8] for an example), then the relation might be different, but L is always even.

We observe that once the filter H has been chosen, it completely determines the functions φ and ψ and therefore, the multiresolution analysis. Moreover, in properly constructed algorithms, the values of the functions φ and ψ are usually never computed. Due to the recursive definition of the wavelet bases, all the manipulations are performed with the quadrature mirror filters H and G, even if they involve quantities associated with φ and ψ.

As an example, let us compute the moments of the scaling function ϕ. The expressions for the moments,

$$\mathcal{M}_m = \int x^m \, \varphi(x) \; dx, \quad m = 0, \ldots, M-1, \tag{1.23}$$

may be found in terms of the filter coefficients $\{h_k\}_{k=0}^{k=L-1}$. Applying operator $(\frac{1}{\mathrm{i}} d/d\xi)^m$ to both sides of (1.10) and setting $\xi = 0$, we obtain

$$\mathcal{M}_m = 2^{-m} \sum_{j=0}^{j=m} \begin{pmatrix} m \\ j \end{pmatrix} \mathcal{M}_j \, \mathcal{M}^h_{m-j}, \tag{1.24}$$

where

$$\mathcal{M}^h_l = 2^{-\frac{1}{2}} \sum_{k=0}^{k=L-1} h_k k^l, \quad l = 0, \ldots, M-1. \tag{1.25}$$

Thus, we have from (1.24)

$$\mathcal{M}_m = \frac{1}{2^m - 1} \sum_{j=0}^{j=m-1} \begin{pmatrix} m \\ j \end{pmatrix} \mathcal{M}_j \, \mathcal{M}^h_{m-j}, \tag{1.26}$$

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

Alternatively, using

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

the moments $\mathcal{M}_m$ may be obtained within the desired accuracy as a limit of recursively generated sequence of vectors, $\{\mathcal{M}_m^{(r)}\}_{m=0}^{m=M-1}$ for $r = 1, 2, \ldots,$

$$\mathcal{M}_m^{(r+1)} = \sum_{j=0}^{j=m} \left(\begin{array}{c} m \\ j \end{array} \right) 2^{-j(r+1)} \mathcal{M}_{m-j}^{(r)} \mathcal{M}_j^h, \tag{1.28}$$

starting with

$$\mathcal{M}_m^{(1)} = 2^{-m} \mathcal{M}_m^h, \quad m = 0, \ldots, M-1. \tag{1.29}$$

Each vector $\{\mathcal{M}_m^{(r)}\}_{m=0}^{m=M-1}$ represents M moments of the product in (1.27) with r terms, and the iteration converges rapidly. Notice that in both algorithms we never computed the values of the function φ itself.

II The non-standard form

A wavelet basis in $\mathbf{L}^2(\mathbf{R}^\mathbf{d})$, $d \geq 2$, may be constructed as a tensor product of one-dimensional bases. Considering $d = 2$ and using the Haar basis as an example, we note that the supports of the basis functions are rectangles of various dyadic sizes. Representing operators in such bases leads to the standard form which we will discuss in the next Section.

Alternatively, wavelet bases in $\mathbf{L}^2(\mathbf{R}^\mathbf{d})$, $d \geq 2$ may be constructed using the scaling function in addition to the wavelets. Such construction is specific to wavelet bases. Considering $d = 2$ as an example, we note that the triplet of functions

$$\{\psi_{j,k}(x)\,\psi_{j,k'}(y),\;\; \psi_{j,k}(x)\,\varphi_{j,k'}(y),\;\; \varphi_{j,k}(x)\,\psi_{j,k'}(y)\}, \tag{2.1}$$

where $j, k, k' \in \mathbf{Z}$, forms a basis of $\mathbf{L}^2(\mathbf{R}^\mathbf{2})$. We note that the basis functions have square supports. Representing operators in these bases leads to the non-standard form [8].

Let us introduce the non-standard form in the context of Multiresolution Analysis, independently of the specific choice of the wavelet basis. Let T be an operator

$$T : \mathbf{L}^2(\mathbf{R}) \to \mathbf{L}^2(\mathbf{R}), \tag{2.2}$$

with kernel $K(x, y)$. We define projection operators on the subspace $\mathbf{V}_j$, $j \in \mathbf{Z}$,

$$P_j : \mathbf{L}^2(\mathbf{R}) \to \mathbf{V}_j, \tag{2.3}$$

as follows

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

Expanding T in a "telescopic" series, we obtain

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

where

$$Q_j = P_{j-1} - P_j \tag{2.6}$$

is the projection operator on the subspace $\mathbf{W}_j$. If there is a coarsest scale n, then instead of (2.5) we have

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

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

$$T_0 = \sum_{j=1}^{n} (Q_j T Q_j + Q_j T P_j + P_j T Q_j) + P_n T P_n, \tag{2.8}$$

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 [8]) of the operator T as a chain of triplets

$$T = \{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}} \tag{2.9}$$

acting on the subspaces $\mathbf{V}_j$ and $\mathbf{W}_j$,

$$A_j : \mathbf{W}_j \to \mathbf{W}_j, \tag{2.10}$$

$$B_j : \mathbf{V}_j \to \mathbf{W}_j, \tag{2.11}$$

$$\Gamma_j : \mathbf{W}_j \to \mathbf{V}_j. \tag{2.12}$$

The operators $\{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}}$ are defined as $A_j = Q_j T Q_j$, $B_j = Q_j T P_j$ and $\Gamma_j = P_j T Q_j$. These operators admit a recursive definition via the relation

$$T_j = \begin{pmatrix} A_{j+1} & B_{j+1} \\ \Gamma_{j+1} & T_{j+1} \end{pmatrix}, \tag{2.13}$$

where the operators T_j,

$$T_j : \mathbf{V}_j \to \mathbf{V}_j, \tag{2.14}$$

are defined by $T_j = P_j T P_j$.

If there is a coarsest scale n, then

$$T = \{\{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}: j \leq n}, T_n\}, \tag{2.15}$$

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

1). The map (2.10) implies that the operator A_j describes the interaction on the scale j only, since the subspace $\mathbf{W}_j$ is an element of the direct sum in (1.5).

2). The operators B_j, Γ_j in (2.11) and (2.12) describe the interaction between scale j and all coarser scales. Indeed, the subspace $\mathbf{V}_j$ contains all the subspaces $\mathbf{V}_{j'}$ with $j' > j$ (see (1.1)).

3). The operator T_j is an "averaged" version of the operator T_{j-1}.

The operators A_j, B_j and Γ_j are represented by the matrices α^j, β^j and γ^j,

$$\alpha^j_{k,k'} = \int\int K(x,y)\, \psi_{j,k}(x)\, \psi_{j,k'}(y)\, dxdy, \tag{2.16}$$

$$\beta^j_{k,k'} = \int\int K(x,y)\, \psi_{j,k}(x)\, \varphi_{j,k'}(y)\, dxdy, \tag{2.17}$$

and

$$\gamma^j_{k,k'} = \int\int K(x,y)\, \varphi_{j,k}(x)\, \psi_{j,k'}(y)\, dxdy. \tag{2.18}$$

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

$$s^j_{k,k'} = \int\int K(x,y)\, \varphi_{j,k}(x)\, \varphi_{j,k'}(y)\, dxdy. \tag{2.19}$$

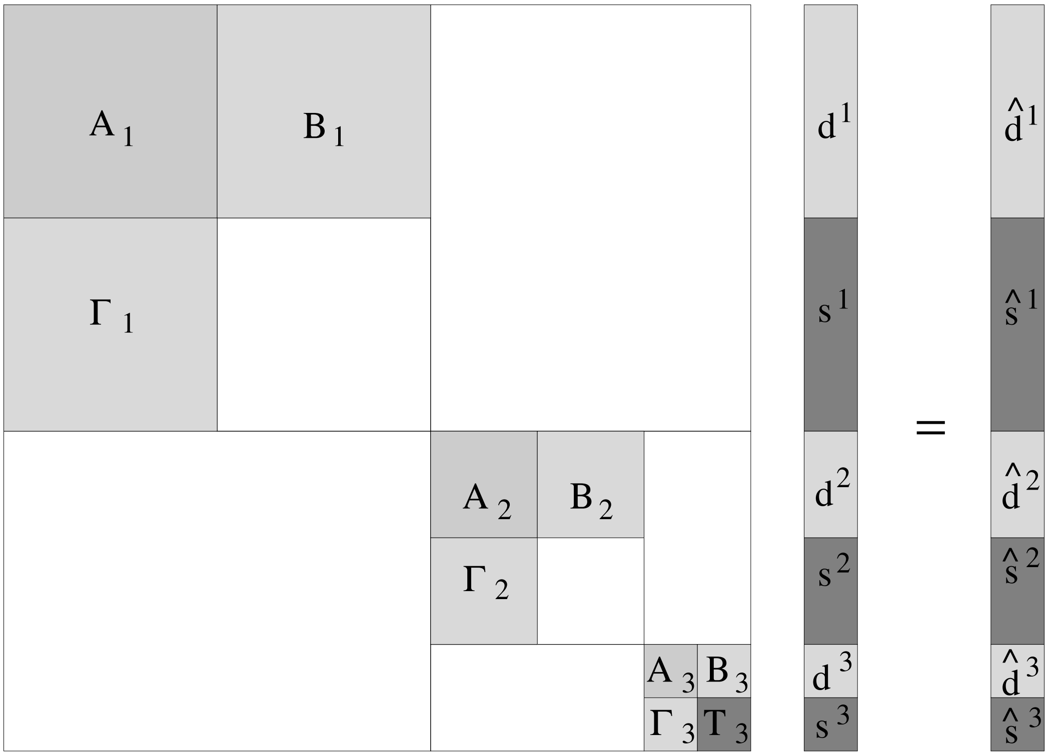

Figure 1: Organization of the non-standard form of a matrix. The submatrices A_j, B_j, and Γ_j, $j = 1, 2, 3$, and T_3 are the only non-zero submatrices.

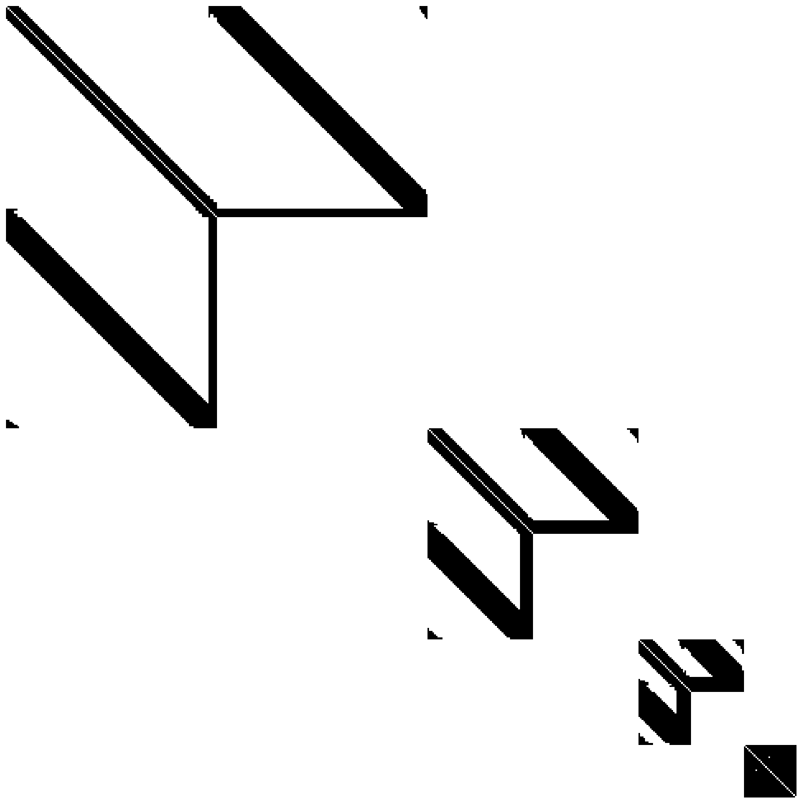

Figure 2: An example of a matrix in the non-standard form (see Example 1)

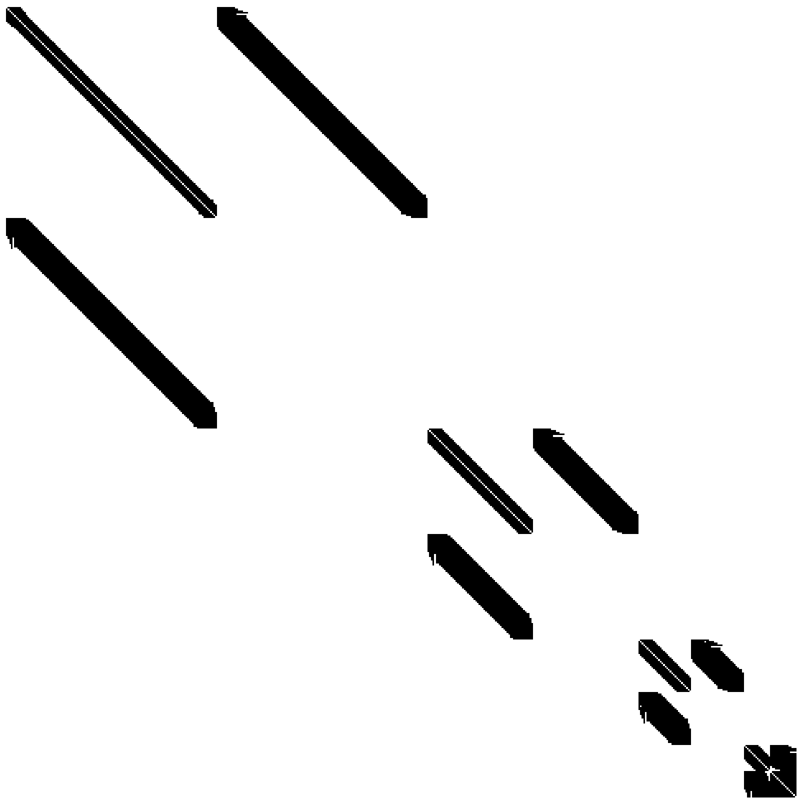

Figure 3: The non-standard form of the same matrix as in Figure 2 using a basis of wavelets on the interval [6]. The vertical and horizontal bands (which are present in Figure 2 due to periodization) do not appear in this representation

III 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

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

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

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

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

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

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

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 $\{\Gamma_j^{n+1}\}$,

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

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

(In this notation, $\Gamma_n^{n+1} = \Gamma_n$ and $B_n^{n+1} = B_n$). If there are finitely many scales and $\mathbf{V}_0$ is finite dimensional, then the standard form is a representation of $T_0 = P_0 T P_0$ as

$$T_0 = \{A_j, \{B_j^{j'}\}_{j'=j+1}^{j'=n}, \{\Gamma_j^{j'}\}_{j'=j+1}^{j'=n}, B_j^{n+1}, \Gamma_j^{n+1}, T_n\}_{j=1,\dots,n}. \tag{3.7}$$

The operators (3.7) are organized as blocks of a matrix (see Figures 3 and 4).

If the operator T is a Calderón-Zygmund or a pseudo-differential operator then, for a fixed accuracy, all the operators in (3.7) (except T_n) are banded. As a result, the standard form has several "finger" bands which correspond to the interaction between different scales. For a large class of operators (pseudo-differential, for example), the interaction between different scales characterized by the size of the coefficients of "finger" bands, decays as the distance $j' - j$ between the scales increases. Therefore, if the scales j and j' are well separated, then for a given accuracy, the operators $B_j^{j'}, \Gamma_j^{j'}$ can be neglected.

There are two ways of computing the standard form of a matrix. The first consists in applying the one-dimensional transform to each column (row) of the matrix and, then, to each row (column) of the result. Alternatively, one can compute the non-standard form and then apply the one-dimensional transform to each row of all operators B^j and each column of all operators Γ_j. We refer to [8] for details.

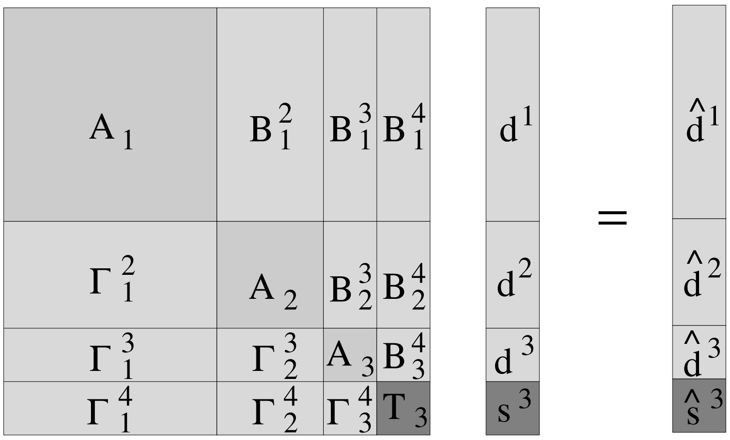

Figure 4: Organization of a matrix in the standard form

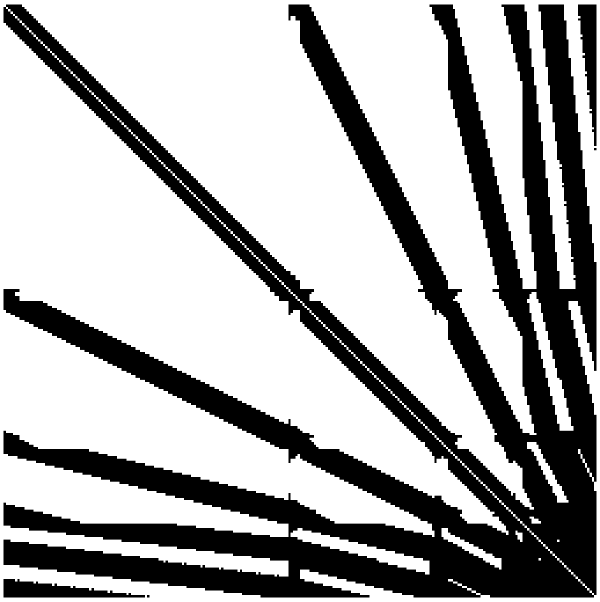

Figure 5: An example of a matrix in the standard form (see Example 1)

IV Compression of operators

If the operator T is a Calderón-Zygmund or a pseudo-differential operator then, by using the wavelet basis with M vanishing moments, we force the operators $\{A_j, B_j, \Gamma_j\}_{j\in\mathbf{Z}}$ to decay roughly as $1/d^{M+1}$, where d is the distance from the diagonal. For example, let the kernel satisfy the conditions

$$
\begin{aligned}
|K(x,y)| &\leq \frac{1}{|x-y|}, && (4.1)\\
|\partial_x^M K(x,y)| + |\partial_y^M K(x,y)| &\leq \frac{C_0}{|x-y|^{1+M}} && (4.2)
\end{aligned}
$$

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

$$
|\alpha^j_{i,l}| + |\beta^j_{i,l}| + |\gamma^j_{i,l}| \leq \frac{C_M}{1+|i-l|^{M+1}}, \tag{4.3}
$$

for all

$$
|\ i-l\ | \geq 2M. \tag{4.4}
$$

If, in addition to (4.1), (4.2),

$$
|\int_{I\times I} K(x,y)\ dxdy\ | \leq C|I| \tag{4.5}
$$

for all dyadic intervals I (this is the "weak cancellation condition", see [27]), then (4.3) holds for all i, l.

If T is a pseudo-differential operator with symbol $\sigma(x,\xi)$ of order λ defined by the formula

$$
T(f)(x) = \sigma(x,D)f = \int e^{ix\xi}\ \sigma(x,\xi)\hat{f}(\xi)\ d\xi = \int K(x,y)f(y)\ dy, \tag{4.6}
$$

where K is the distributional kernel of T, then assuming that the symbols σ of T and σ^* of T^* satisfy the standard conditions

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

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

we have the inequality

$$
|\alpha^j_{i,l}| + |\beta^j_{i,l}| + |\gamma^j_{i,l}| \leq \frac{2^{\lambda j}\ C_M}{(1+|i-l|)^{M+1}}, \tag{4.9}
$$

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 α^j, β^j and γ^j outside bands of width $B \geq 2M$ around their diagonals. We obtain

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

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 ε of calculations is fixed, and the parameters of the algorithm (in our case, the band width B and order M) have to be chosen in such a manner that the desired precision of calculations is achieved. If M is fixed, then we choose B so that

$$B \geq \left(\frac{C}{\varepsilon} \log_2 N \right)^{1/M}. \tag{4.11}$$

In other words, T_0 has been approximated to precision ε with its truncated version, which can be applied to arbitrary vectors for a cost proportional to $N \left((C/\varepsilon) \log_2 N \right)^{1/M}$, which for all practical purposes does not differ from N.

A more detailed investigation [8] permits the estimate (4.10) to be replaced with the estimate

$$\| T_0^B - T_0 \| \leq \frac{C}{B^M}, \tag{4.12}$$

making the application of the operator T_0 to an arbitrary vector with arbitrary fixed accuracy into a procedure of order N. Obtaining this uniform estimate leads to a proof of

Theorem (G. David, J.L. Journé) Suppose that the operator

$$T(f) = \int K(x, y)\ f(y)\ dy \tag{4.13}$$

satisfies the conditions (4.1), (4.2), (4.5). Then a necessary and sufficient condition for T to be bounded on L^2 is that

$$\beta(x) = T(1)(x), \tag{4.14}$$

$$\gamma(y) = T^*(1)(y) \tag{4.15}$$

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

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

where J is a dyadic interval and

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

Again we refer to [8] for details.

The compression of operators results in fast algorithms for evaluation of operators on functions. We present here one example and refer to [8] for additional examples.

Example 1.

In this example, we consider the matrix

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

and convert it to the non-standard form using wavelets with six vanishing moments. 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 experiments in applying this sparse matrix to a vector are tabulated in Table 1. The standard form of the operator A with $N = 256$ is depicted in Figure 5.

Column 1 of Table 1 contains the number N indicating the size of $N \times N$ matrix A_{ij}, columns 2, 3 contain CPU times T_s, T_w required by the standard order $O(N^2)$ and the fast $O(N)$ schemes to multiply a vector by the matrix, and column 4 contains the CPU T_d time used to produce the non-standard form of the operator. Columns 5, 6 contain the L_2 and L_∞ errors of the direct calculation, and columns 7, 8 contain the same information for the result obtained 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 of the matrix.

Input Size	Time			Error of Single Precision Multiplication		Error of FWT Multiplication		Compression Coefficient
N	T_s	T_w	T_d	L_2 - norm	L_∞ - norm	L_2 - norm	L_∞ - norm	C_{comp}
64	0.12	0.16	7.76	$1.26 \cdot 10^{-7}$	$3.65 \cdot 10^{-7}$	$8.89 \cdot 10^{-8}$	$1.72 \cdot 10^{-7}$	1.39
128	0.48	0.38	32.62	$2.17 \cdot 10^{-7}$	$8.64 \cdot 10^{-7}$	$1.12 \cdot 10^{-7}$	$9.94 \cdot 10^{-7}$	2.22
256	1.92	0.80	96.44	$2.81 \cdot 10^{-7}$	$1.12 \cdot 10^{-6}$	$1.25 \cdot 10^{-7}$	$5.30 \cdot 10^{-7}$	3.93
512	7.68	1.80	252.72	$4.21 \cdot 10^{-7}$	$1.75 \cdot 10^{-6}$	$1.23 \cdot 10^{-7}$	$5.16 \cdot 10^{-7}$	7.33
1024	30.72	3.72	605.74	$6.64 \cdot 10^{-7}$	$3.90 \cdot 10^{-6}$	$1.36 \cdot 10^{-7}$	$5.04 \cdot 10^{-7}$	14.09

Table 1: Numerical results for Example 1

V The operator d/dx in wavelet bases.

For a number of operators (e.g., differential operators, fractional derivatives, 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 α^j_{il}, β^j_{il}, and γ^j_{il} of A_j, B_j, and Γ_j, where $i,l,j \in \mathbf{Z}$ for the operator d/dx are easily computed as

$$\alpha^j_{il} = 2^{-j}\int_{-\infty}^{\infty} \psi(2^{-j}x-i)\,\psi'(2^{-j}x-l)\,2^{-j}dx = 2^{-j}\alpha_{i-l}, \tag{5.1}$$

$$\beta^j_{il} = 2^{-j}\int_{-\infty}^{\infty} \psi(2^{-j}x-i)\,\varphi'(2^{-j}x-l)\,2^{-j}dx = 2^{-j}\beta_{i-l}, \tag{5.2}$$

and

$$\gamma^j_{il} = 2^{-j}\int_{-\infty}^{\infty} \varphi(2^{-j}x-i)\,\psi'(2^{-j}x-l)\,2^{-j}dx = 2^{-j}\gamma_{i-l}, \tag{5.3}$$

where

$$\alpha_l = \int_{-\infty}^{+\infty} \psi(x-l)\,\frac{d}{dx}\psi(x)\,dx, \tag{5.4}$$

$$\beta_l = \int_{-\infty}^{+\infty} \psi(x-l)\,\frac{d}{dx}\varphi(x)\,dx, \tag{5.5}$$

and

$$\gamma_l = \int_{-\infty}^{+\infty} \varphi(x-l)\,\frac{d}{dx}\psi(x)\,dx. \tag{5.6}$$

Moreover, using (1.9) and (1.19) we have

$$\alpha_i = 2\sum_{k=0}^{L-1}\sum_{k'=0}^{L-1} g_k\, g_{k'}\, r_{2i+k-k'}, \tag{5.7}$$

$$\beta_i = 2\sum_{k=0}^{L-1}\sum_{k'=0}^{L-1} g_k\, h_{k'}\, r_{2i+k-k'}, \tag{5.8}$$

and

$$\gamma_i = 2\sum_{k=0}^{L-1}\sum_{k'=0}^{L-1} h_k\, g_{k'}\, r_{2i+k-k'}, \tag{5.9}$$

where

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

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 $\mathbf{V}_0$. Rewriting (5.10) in terms of $\hat{\varphi}(\xi)$ (see (1.11)), we obtain

$$r_l = \int_{-\infty}^{+\infty} |\hat{\varphi}(\xi)|^2 (\mathrm{i}\xi) \mathrm{e}^{-\mathrm{i}l\xi}\,d\xi. \tag{5.11}$$

Thus, the coefficients r_l depend only on the autocorrelation function of the scaling function φ, rather that the scaling function itself since the integral in (5.11) depends just on $|\hat{\varphi}(\xi)|^2$. The same holds, in fact, for all convolution operators [4].

Remark. The autocorrelation function of the scaling function (see (5.24)) has $2M-1$ vanishing moments and its "zero moment" is equal to one (see (5.25) and (5.26)). This implies that if we consider the representation of the derivative operator on the subspace $\mathbf{V}_0$ as a finite-difference scheme, such scheme has order $2M$. For integral convolution operators, it implies that the asymptotics is accurate to order $2M$ (see [4] and below).

The following proposition [4] reduces the computation of the coefficients r_l to solving a system of linear algebraic equations.

1. If the integrals in (5.10) or (5.11) exist, then the coefficients r_l, $l \in \mathbf{Z}$ in (5.10) satisfy the following system of linear algebraic equations

$$r_l = 2\left[r_{2l} + \tfrac{1}{2}\sum_{k=1}^{L/2} a_{2k-1}(r_{2l-2k+1} + r_{2l+2k-1})\right], \tag{5.12}$$

and

$$\sum_l l\, r_l = -1, \tag{5.13}$$

where

$$a_{2k-1} = 2\sum_{i=0}^{L-2k} h_i\, h_{i+2k-1}, \quad k = 1, \dots, L/2 \tag{5.14}$$

are the autocorrelation coefficients of the filter H.

2. If $M \geq 2$, then equations (5.12) and (5.13) have a unique solution with a finite number of non-zero r_l, namely, $r_l \neq 0$ for $-L+2 \leq l \leq L-2$ and

$$r_l = -r_{-l}, \tag{5.15}$$

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. $M = 2$

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

and

$$r_1 = -\frac{2}{3}, \qquad r_2 = \frac{1}{12},$$

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.

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

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.

For the coefficients $r_l^{(n)}$ of d^n/dx^n, $n>1$, the system of linear algebraic equations is similar to that for the coefficients of d/dx. This system (and (5.12)) may be written in terms of

$$\hat{r}(\xi) = \sum_l r_l^{(n)} \mathrm{e}^{\mathrm{i}l\xi}, \tag{5.16}$$

as

$$\hat{r}(\xi) = 2^n \left(|m_0(\xi/2)|^2 \, \hat{r}(\xi/2) + |m_0(\xi/2+\pi)|^2 \, \hat{r}(\xi/2+\pi) \right), \tag{5.17}$$

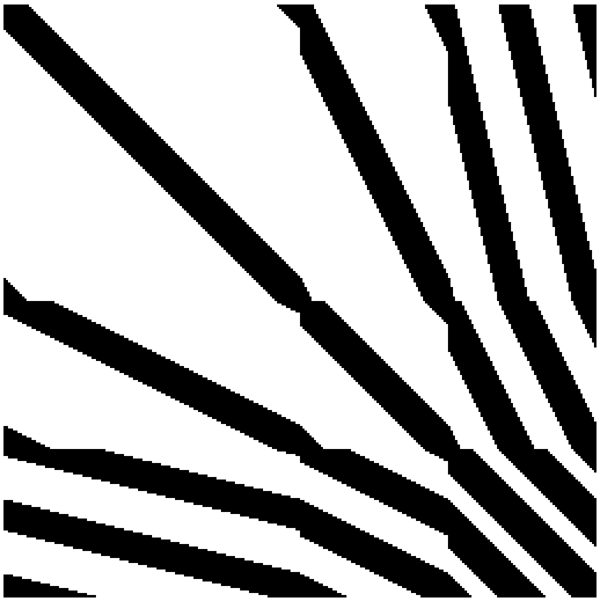

Figure 7: Sparse structure of the standard form of derivative operators.

where m_0 is the 2π-periodic function in (1.12). Considering the operator M_0 on 2π-periodic functions

$$(M_0 f)(\xi) = |m_0(\xi/2)|^2 f(\xi/2) + |m_0(\xi/2+\pi)|^2 f(\xi/2+\pi), \tag{5.18}$$

we rewrite (5.17) as

$$M_0 \hat{r} = 2^{-n}\hat{r}, \tag{5.19}$$

so that $\hat{r}$ is an eigenvector of the operator M_0 corresponding to the eigenvalue 2^{-n}. Thus, finding the representation of the derivatives in the wavelet basis is equivalent to finding trigonometric polynomial solutions of (5.19) and vice versa [4].

An important property of the wavelet representation of the (periodized) derivative operators (and, in general, pseudodifferential operators with homogeneous symbols) is that these operators have an explicit diagonal preconditioner in wavelet bases.

We present here two tables illustrating such preconditioning applied to the standard form of the second derivative. In the following examples the standard form of periodized second derivative $\mathbf{D}_2$ of size $N \times N$, where $N = 2^n$, is preconditioned in the wavelet basis by the diagonal matrix $\mathbf{P}$,

$$\mathbf{D}_2^p = \mathbf{P}\mathbf{D}_2^w\mathbf{P},$$

where $P_{il} = \delta_{il}2^j$, $1 \leq j \leq n$, and where j is chosen depending on i,l so that $N - N/2^{j-1} + 1 \leq i,l \leq N - N/2^j$, and $P_{NN} = 2^n$. The matrix $\mathbf{P}$ is illustrated in Figure 8.

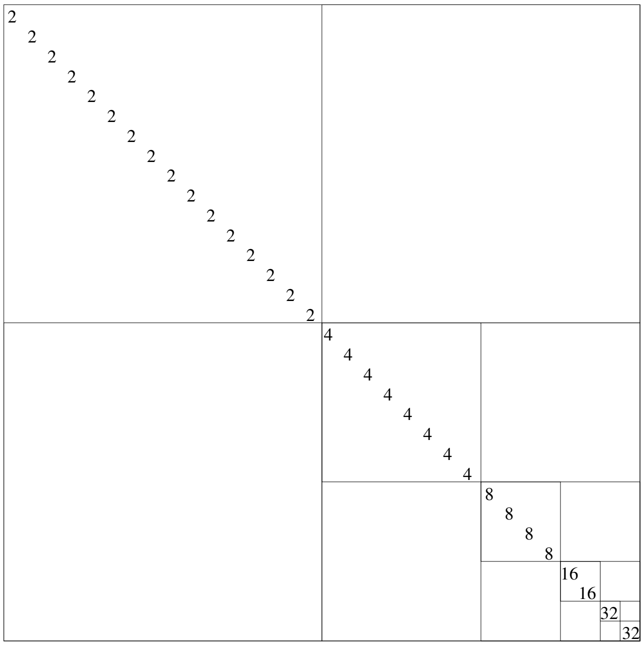

Figure 8: An example ($n = 5$) of the diagonal martix $\mathbf{P}$ used to rescale the matrix of the periodized second derivative $\mathbf{D}_2^w$ in the wavelet system of coordinates.

The following Tables 2 and 3 compare the original condition number κ of $\mathbf{D}_2^w$ and κ_p of $\mathbf{D}_2^p$.

N	κ	κ_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 numbers of the matrix of the periodized second derivative (with and without preconditioning) in the basis of Daubechies' wavelets with three vanishing moments $M = 3$.

N	κ	κ_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	0.43633E+01
1024	0.26807E+06	0.43640E+01

Table 3.
Condition numbers of the matrix of the periodized second derivative (with and without preconditioning) in the basis of Daubechies' wavelets with six vanishing moments $M = 6$.

Fractional derivatives

First, let us consider a convolution operator T and the infinite matrix $t_{i-l}^{(j-1)}$, $i, l \in \mathbf{Z}$, representing $P_{j-1}TP_{j-1}$ on the subspace $\mathbf{V}_{j-1}$. To compute the representation of P_jTP_j, we have (see e.g., formula (3.26) of [8])

$$t_l^{(j)} = \sum_{k=0}^{L-1} \sum_{m=0}^{L-1} h_k \, h_m \, t_{2l+k-m}^{(j-1)}. \tag{5.20}$$

It easily reduces to

$$t_l^{(j)} = t_{2l}^{(j-1)} + \tfrac{1}{2} \sum_{k=0}^{L/2} a_{2k-1} \, (t_{2l-2k+1}^{(j-1)} + t_{2l+2k-1}^{(j-1)}). \tag{5.21}$$

where the coefficients a_{2k-1} are given in (5.14).

We also have

$$t_l^{(j)} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x-y) \, \varphi_{j,0}(y) \, \varphi_{j,l}(x) \, dx dy, \tag{5.22}$$

and, by changing the order of integration, we obtain

$$t_l^{(j)} = 2^j \int_{-\infty}^{+\infty} K(2^j(l-y)) \, \Phi(y) \, dy, \tag{5.23}$$

where Φ is the autocorrelation function of the scaling function φ,

$$\Phi(y) = \int_{-\infty}^{+\infty} \varphi(x) \, \varphi(x-y) \, dx. \tag{5.24}$$

It is easy to verify (see [4]) that

$$\int_{-\infty}^{+\infty} \Phi(y)dy = 1, \tag{5.25}$$

and

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

The vanishing moments of the autocorrelation function Φ allow us to compute the elements of the matrix $t_l^{(j)}$ for large l and sufficiently fine scales $j \leq 0$. Expanding the kernel K in the Taylor series, we obtain from (5.23)

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

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 choice of M) and (5.27) implies a one-point quadrature formula $t_l^{(j)} \approx 2^j K(2^j l)$ for large l and sufficiently fine scales $j \leq 0$.

Computing representations of convolution operators simplifies further if the symbol of the operator is homogeneous of some degree. Let us illustrate this using example of fractional derivatives. We define fractional derivatives as

$$(\partial_x^\alpha f)(x) = \int_{-\infty}^{+\infty} \frac{(x-y)_+^{-\alpha-1}}{\Gamma(-\alpha)} f(y)dy, \tag{5.28}$$

where we consider $\alpha \neq 1, 2 \ldots$. If $\alpha < 0$, then (5.28) defines fractional anti-derivatives.

The representation of ∂_x^α on $\mathbf{V}_0$ is determined by the coefficients

$$r_l = \int_{-\infty}^{+\infty} \varphi(x-l)\; (\partial_x^\alpha \varphi)(x)\, dx, \quad l \in \mathbf{Z}, \tag{5.29}$$

provided that this integral exists.

The non-standard form $\partial_x^\alpha = \{A_j, B_j, \Gamma_j\}_{j\in\mathbf{Z}}$ is computed via $A_j = 2^{-\alpha j} A_0$, $B_j = 2^{-\alpha j} B_0$, and $\Gamma_j = 2^{-\alpha j}\Gamma_0$, where matrix elements α_{i-l}, β_{i-l}, and γ_{i-l} of A_0, B_0, and Γ_0 are obtained from the coefficients r_l,

$$\alpha_i = 2^\alpha \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k\, g_{k'}\, r_{2i+k-k'}, \tag{5.30}$$

$$\beta_i = 2^\alpha \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k\, h_{k'}\, r_{2i+k-k'}, \tag{5.31}$$

and

$$\gamma_i = 2^{\alpha} \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} h_k \, g_{k'} \, r_{2i+k-k'}. \tag{5.32}$$

It easy to verify that the coefficients r_l satisfy the following system of linear algebraic equations

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

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

$$r_l = \frac{1}{\Gamma(-\alpha)} \frac{1}{l^{1+\alpha}} + O(\frac{1}{l^{1+\alpha+2M}}) \quad \text{for} \quad l > 0, \tag{5.34}$$

$$r_l = 0 \qquad \text{for} \quad l < 0. \tag{5.35}$$

Example.

We compute the coefficients r_l of the fractional derivative with $\alpha = 0.5$ for Daubechies' wavelets with six vanishing moments with accuracy 10^{-7}. The coefficients for r_l, $l > 14$ or $l < -7$ are obtained using asymptotics

$$r_l = -\frac{1}{2\sqrt{\pi}} \frac{1}{l^{1+\frac{1}{2}}} + O(\frac{1}{l^{13+\frac{1}{2}}}) \quad \text{for} \quad l > 0, \tag{5.36}$$

$$r_l = 0 \qquad \text{for} \quad l < 0. \tag{5.37}$$

l	Coefficients r_l	l	Coefficients r_l
-7	-2.82831017E-06	4	-2.77955293E-02
-6	-1.68623867E-06	5	-2.61324170E-02
-5	4.45847796E-04	6	-1.91718816E-02
-4	-4.34633415E-03	7	-1.52272841E-02
-3	2.28821728E-02	8	-1.24667403E-02
-2	-8.49883759E-02	9	-1.04479500E-02
-1	0.27799963	10	-8.92061945E-03
0	0.84681966	11	-7.73225246E-03
1	-0.69847577	12	-6.78614593E-03
2	2.36400139E-02	13	-6.01838599E-03
3	-8.97463780E-02	14	-5.38521459E-03

Table 4. The coefficients $\{r_l\}_l$, $l = -7, \ldots, 14$ of the fractional derivative $\alpha = 0.5$ for Daubechies' wavelets with six vanishing moments.

VI Multiplication of matrices and fast iterative construction of the generalized inverse

The standard and non-standard forms may be multiplied in fast manner if the matrices represent Calderón-Zygmund or pseudo-differential operators. Multiplication of matrices in the standard form is a straightforward algorithm [9], [1] and requires at most $O(N \log^2 N)$ operations. The algorithm for the multiplication of matrices in the non-standard form has been outlined in [3] and requires $O(N)$ operations. This is a significant improvement over $O(N^3)$ operations for dense matrices which arise in the ordinary discretization of the operators from these classes.

Fast multiplication algorithms give a second life to a great number of iterative algorithms. Indeed, powers of matrices may be computed as well as other functions of matrices. Let us consider an iterative construction of the generalized inverse. In order to construct the generalized inverse $A^{\dagger}$ of the matrix A, we use the following result [31]:

Let σ_1 be the largest singular value of the $m \times n$ matrix A. Consider the sequence of matrices X_k

$$X_{k+1} = 2X_k - X_k A X_k \tag{6.1}$$

with

$$X_0 = \alpha A^*, \tag{6.2}$$

where A^* is the adjoint matrix and α 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 number we understand the ratio of the largest singular value to the smallest singular value above the threshold of accuracy).

The details of this algorithm (in the context of computing in wavelet bases) will be described in [7]. We note that throughout the iteration (6.1), it is necessary to maintain the "finger" band structure of the standard form of matrices X_k. Hence, the standard form of both the operator and its generalized inverse must admit such structure. We note that the pseudo-differential operators satisfy this condition.

The following table contains timings and accuracy comparisons for the construction of the generalized inverse via the singular value decomposition (SVD), which is an $O(N^3)$ procedure, and via the iteration (6.1)-(6.2) in the wavelet basis using the Fast Wavelet Transform (FWT). The computations were performed on a Sun Sparc workstation and we used a routine from LINPACK for computing the singular value decomposition. For tests we used the following full rank matrix

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

where $i, j = 1, \ldots, N$. The accuracy theshold was set to 10^{-4}, i.e., entries of X_k below 10^{-4} were systematically removed after each iteration.

Size $N \times N$	SVD	FWT Generalized Inverse	L_2-Error
128×128	20.27 sec.	25.89 sec.	$3.1 \cdot 10^{-4}$
256×256	144.43 sec.	77.98 sec.	$3.42 \cdot 10^{-4}$
512×512	1,155 sec. (est.)	242.84 sec.	$6.0 \cdot 10^{-4}$
1024×1024	9,244 sec. (est.)	657.09 sec.	$7.7 \cdot 10^{-4}$
...	...	...	...
$2^{15} \times 2^{15}$	9.6 years (est.)	1 day (est.)	

We note that the iteration in (6.1) also allows us to compute the projector on the null space (see [9] for this and several other examples).

The algorithm for the exponential is based on the identity

$$\exp(A) = \left[\exp(2^{-L}A)\right]^{2^L}. \tag{6.3}$$

First, $\exp(2^{-L}A)$ is computed by means of the Taylor series, for instance. The number L is chosen so that the largest singular value of $2^{-L}A$ is less than one. At the second stage of the algorithm the matrix $\exp(2^{-L}A)$ is squared L times to obtain the result. Similarly, sine and cosine of a matrix can be computed using the elementary double-angle formulas. Unlike the algorithm for the generalized inverse, this algorithm is not self-correcting. Thus, it is necessary to maintain sufficient accuracy initially so as to obtain the desired accuracy after all the mutiplications have been performed.

Finally, as an example, let us consider the matrix

$$\mathbf{D_b} = \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -2 \end{pmatrix}. \tag{6.4}$$

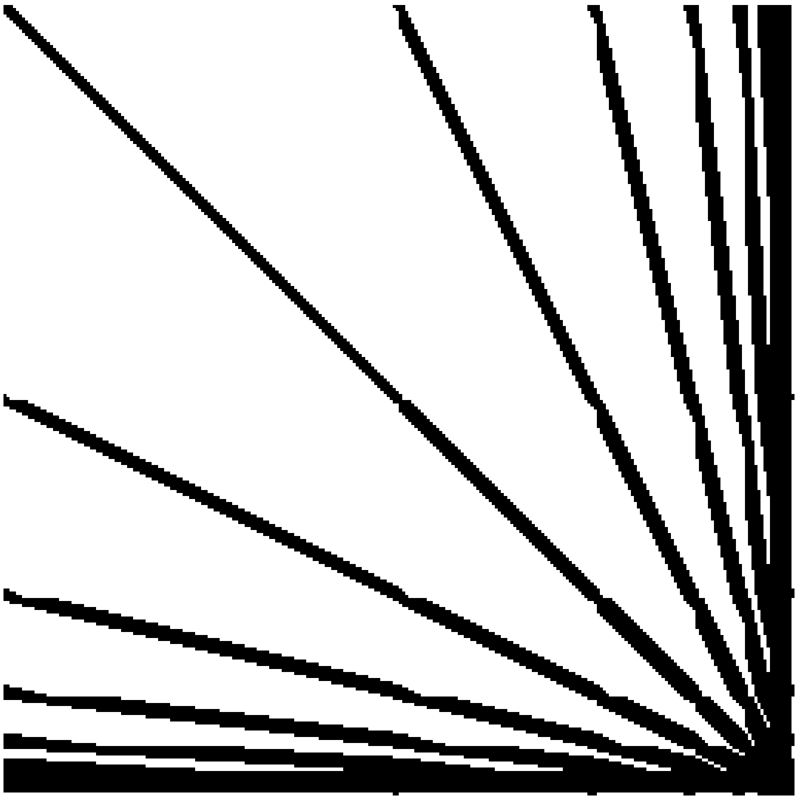

Figure 9: Standard form of the matrix ${\mathbf{D_b}}^{-1}$ computed via the iterative algorithm of this Section with diagonal rescaling. Entries with absolute value greater than 10^{-8} are shown black.

which arises in the finite-difference formulation of the two-point boundary value problem. We note that the inverse of this matrix is sparse in the wavelet basis. As an illustration, we display in Figure 9 the matrix ${\mathbf{D_b}}^{-1}$ obtained via the algorithm sketched above for computing the generalized inverse. Using the diagonal preconditioning (see Figure 8), this computation involves only well-conditioned matrices [5].

Acknowledgments.

The research was partially supported by ONR grant N00014-91-J4037 and a grant from Chevron Oil Field Research Company.

References

[1] B. Alpert, G. Beylkin, R. R. Coifman, and V. Rokhlin. Wavelets for the fast solution of second-kind integral equations. *SIAM Journal of Scientific and Statistical Computing*, 14(1):159–174, 1993. Technical report, Department of Computer Science, Yale University, New Haven, CT, 1990.

[2] B. Alpert and V. Rokhlin. A Fast Algorithm for the Evaluation of Legendre expansions. *SIAM J. on Sci. Stat. Comput.*, 12(1):158–179, 1991. Yale University Technical Report, YALEU/DCS/RR-671 (1989).

[3] G. Beylkin. Wavelets, Multiresolution Analysis and Fast Numerical Algorithms. *A draft of INRIA Lecture Notes*, 1991.

[4] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.*, 29(6):1716–1740, 1992.

[5] G. Beylkin. On wavelet-based algorithms for solving differential equations. *Preprint*, 1992.

[6] G. Beylkin and M. E. Brewster. Fast Numerical Algorithms using Wavelet Bases on the Interval. in progress.

[7] G. Beylkin, R. R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms II. in progress.

[8] G. Beylkin, R. R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. *Comm. Pure and Appl. Math.*, 44:141–183, 1991. Yale University Technical Report YALEU/DCS/RR-696, August 1989.

[9] G. Beylkin, R. R. Coifman, and V. Rokhlin. Wavelets in Numerical Analysis. In *Wavelets and Their Applications*, pages 181–210. Jones and Bartlett, 1992.

[10] J. Carrier, L. Greengard, and V. Rokhlin. A fast adaptive multipole algorithm for particle simulations. *SIAM Journal of Scientific and Statistical Computing*, 9(4), 1988. Yale University Technical Report, YALEU/DCS/RR-496 (1986).

[11] A. Cohen, I. Daubechies, B. Jawerth, and P. Vial. Multiresolution analysis, wavelets and fast algorithms on an interval. *Comptes Rendus Acad. Sc. Paris*, 1992.

[12] A. Cohen, I. Daubechies, and P. Vial. Wavelets on the interval and fast wavelet transforms. *preprint*, 1992.

[13] R. R. Coifman and Y. Meyer. Nouvelles bases orthonormées de $L^2(\mathbf{R})$ ayant la structure du sysème de Walsh. 1989. preprint.

[14] R. R. Coifman and Y. Meyer. Nouvelles bases orthogonales. *C.R. Acad. Sci., Paris*, 1990.

[15] R. R. Coifman and V. Wickerhauser. Best-adapted wave packet bases. 1990.

[16] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure and Appl. Math.*, 41:909–996, 1988.

[17] I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NSF Series in Applied Mathematics. SIAM, 1992.

[18] L. Greengard. Potential flow in channels. *SIAM J. Sci. Stat. Comput.*, 11(4):603–620, 1990.

[19] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comp. Phys.*, 73(1):325–348, 1987.

[20] A. Haar. Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, pages 331–371, 1910.

[21] S. Jaffard. Wavelet methods for fast resolution of elliptic problems. *SIAM Journal on Numerical Analysis*, 29(4):965–986, 1992.

[22] A. Jouini and P.G. Lemarié-Rieusset. Analyse multi-résolution biorthogonale sur l'intervalle et applications. *Annales de l'Institut Poincare, Analyse Non-lineaire*, to appear.

[23] S. Mallat. Multiresolution approximation and wavelets. Technical report, GRASP Lab, Dept. of Computer and Information Science, University of Pennsylvania.

[24] H. S. Malvar. Lapped Transforms for Efficient Transform/Subband Coding. *IEEE Trans. Acoust., Speech, Signal Processing*, 38(6):969–978, 1990.

[25] Y. Meyer. Principe d'incertitude, bases hilbertiennes et algébres d'opérateurs. In *Séminaire Bourbaki,*, page 662. Société Mathématique de France, 1985-86. Astérisque.

[26] Y. Meyer. Ondelettes et fonctions splines. Technical report, séminaire edp, Ecole Polytechnique, Paris, France, 1986.

[27] Y. Meyer. Wavelets and operators. In N.T. Peck E. Berkson and J. Uhl, editors, *Analysis at Urbana*. London Math. Society, Lecture Notes Series 137, 1989. v.1.

[28] Yves Meyer. *Ondelettes et Opérateurs*. Hermann, Paris, 1990.

[29] S.T. O'Donnel and V. Rokhlin. A fast algorithm for the numerical evaluation of conformal mappings. *SIAM J. Sci. Stat. Comput.*, 10(3):475–487, 1989. Yale University Technical Report, YALEU/DCS/RR-554 (1987).

[30] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *J. Comp. Phys.*, 60(2), 1985.

[31] G. Schulz. Iterative Berechnung der reziproken Matrix. *Z. Angew. Math. Mech.*, 13:57–59, 1933.

[32] M. J. Smith and T. P. Barnwell. Exact reconstruction techniques for tree-structured subband coders. *IEEE Transactions on ASSP*, 34:434–441, 1986.

[33] J. O. Stromberg. A Modified Franklin System and Higher-Order Spline Systems on $\mathbf{R}^n$ as Unconditional Bases for Hardy Spaces. In *Conference in harmonic analysis in honor of Antoni Zygmund, Wadworth math. series*, pages 475–493, 1983.