

Wavelets, Multiresolution Analysis and Fast Numerical Algorithms

G. Beylkin ¹

I Introduction

These lectures are devoted to fast numerical algorithms and their relation to a number of important results of Harmonic Analysis. The representation of wide classes of operators in wavelet bases, for example, Calderón-Zygmund or pseudo-differential operators, may be viewed as a method for their “compression”, i.e., conversion to a sparse form. The sparsity of these representations is a consequence of the localization of wavelets in both, space and wave number domains. In addition, the multiresolution structure of the wavelet expansions brings about an efficient organization of transformations on a given scale and interactions between different neighbouring scales. Such an organization of both linear and non-linear transformations has been a powerful tool in Harmonic Analysis (usually referred to as Littlewood-Paley and Calderón-Zygmund theories, see e.g. [33]) and now appears to be an equally powerful tool in Numerical Analysis.

In applications, for example in image processing [12] and in seismics [22], the multiresolution methods were developed in a search for a substitute for the signal processing algorithms based on the Fourier transform. A technique of subband coding with the exact quadrature mirror filters (QMF) was introduced in [37]. It is clear that the stumbling block on the road to both, the simpler analysis and the fast algorithms, was a limited variety of the orthonormal bases of functional spaces. In fact, there were (with some qualifications) only two major choices, the Fourier basis and the Haar basis. These two bases are almost the antipodes in terms of their space-wave number (or time-frequency) localization. Therefore, it is a remarkable discovery that besides the Fourier and Haar bases, there is an infinite number of various orthonormal bases with a controllable localization in the space-wave number domain. The efforts in mathematics and various applied fields culminated in the development of orthonormal bases of wavelets [39], [30], and the notion of Multiresolution Analysis [31], [27]. There are many new constructions of orthonormal bases with a controllable localization in the space-wave number domain, notably [19], [16], [18], [17].

In Numerical Analysis many ingredients of Calderón-Zygmund theory were used in the Fast Multipole algorithm for computing potential interactions [35], [24], [13]. The

¹Program in Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526; Yale University, P.O.Box 2155 Yale Station, New Haven, CT 06520

Fast Multipole algorithm requires order N operation to compute all the sums

$$p_j = \sum_{i \neq j} \frac{q_i q_j}{|x_i - x_j|}, \quad \text{where } x_i \in \mathbf{R}^3 \quad i, j = 1, \dots, N. \quad (1.1)$$

This algorithm is order N independently of the configuration of the charges, therefore it is providing a substantial improvement over *FFT*, even though the problem may be reduced to an evaluation of a convolution. The reduction of the complexity of computing the sums in (1.1) from order N^2 to $(-\log \epsilon)N$, where ϵ is the desired accuracy, 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 organized in a dyadic hierarchy enabling an efficient $O(N)$ algorithm.

The fast algorithms of [7] provide a systematic generalization of the Fast Multipole method and its descendents (e.g. [34], [3], [23]) to all Calderón-Zygmund and pseudo-differential operators. Both, the subdivision of the space into “boxes” and the approximation of the “far field effect”, are provided adaptively (and automatically) by a wavelet basis. 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 wavelet bases is the so-called non-standard form [7]. The remarkable feature of the non-standard form is the uncoupling achieved among the scales. A straightforward realization, or the standard form, by contrast, contains matrix entries reflecting “interactions” between all pairs of scales. The non-standard form leads to an order N algorithm for evaluating operators on functions, whereas the standard form yields, in general, only an order $N \log(N)$ algorithm. 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.

The direct links between practical algorithms and results of Harmonic Analysis established in [7], are extended in these lectures. By generalizing the theorems of J.M. Bony [10], [11], [9], [15] on propagation of singularities for linear and non-linear equations in Section IX, we obtain a novel method for computing $F(u)$ (directly in the wavelet basis), where F is a smooth function and u is represented in a wavelet basis. The complexity of this algorithm is automatically adaptable to the complexity of the wavelet representation of u . This algorithm opens a number of new possibilities for solving linear and non-linear equations and capturing singularities.

The non-standard forms of many basic operators, such as fractional derivatives, the Hilbert and Riesz transforms, etc., may be computed explicitly [5]. The derivative operators and, in general, the operators with homogeneous symbols have an explicit diagonal preconditioner in the wavelet bases. Since the condition number κ_p controls the rate of convergence of a number of iterative algorithms (for example, the number of iterations of the conjugate gradient method is proportional to $\sqrt{\kappa_p}$), this simple remark implies a completely new outlook on finite difference- and finite element- type methods.

These methods may be viewed as devices for reducing a partial differential equation to a sparse linear system for the cost of an inherently high condition number of the resulting matrices. If instead of finite difference or finite element representations, we use the representation of the derivatives in wavelet bases, then a simple modification by a diagonal preconditioner produces a well-conditioned system. This, in turn, “automatically” yields fast solvers for *PDE*’s and removes one of the original reasons for constructing such methods as multigrid.

Another novel and important element of computing in the wavelet bases is that the “compressed” operators in standard and non-standard forms may be multiplied rapidly. The product of two operators in the standard form requires $C(-\log \epsilon)N$ (or maybe $C(-\log \epsilon)N \log^2 N$) operations, where ϵ is the desired accuracy. In Section VII, we describe a new algorithm for the multiplication of operators in the non-standard form in order $(-\log \epsilon)N$ operations. 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. Among the algorithms requiring multiplication of matrices is an iterative algorithm for constructing the generalized inverse [36], [4], [38], the scaling and squaring method for computing the exponential of an operator (see, for example, [40]), 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, an order N operations. For example, if both, the operator and its generalized inverse, admit sparse representations in the wavelet basis, then the iterative algorithm [36] 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 [6], [2] and [8].

As a final remark, we note, that computing in the wavelet bases has the following three distinct attributes:

1. The operators and functions are represented in an *orthonormal* basis
2. The basis functions have *vanishing moments* leading to the sparsity of representations
3. The algorithms are *recursive* due to the multiresolution properties of the basis

One of the goals of these notes is to explain why (1-3) are of importance for fast numerical algorithms. The algorithms of these notes are useful practical tools and at the same time, directly related to several important results of Harmonic Analysis. I hope, that the explicit interrelation between pure and numerical analysis in the algorithms of these notes will be perceived as an accomplishment of the program of developing fast, wavelet based algorithms for Numerical Analysis. This program was initiated jointly with R. R. Coifman and V. Rokhlin at Yale University, and consequently, this work would not have been possible without the collaboration with them.

II Preliminary Remarks

II.1 Multiresolution analysis.

We start with the definition of the multiresolution analysis. This notion, introduced by Meyer [31], and Mallat [27], captures the essential features of all multiresolution approaches developed so far.

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

$$\dots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \dots \quad (2.1)$$

such that

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

Since all constructions in these lecture notes only use orthonormal bases, we will require the basis of Condition 4 to be an orthonormal rather than just a Riesz basis.

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

By defining \mathbf{W}_j as an orthogonal complement of \mathbf{V}_j in \mathbf{V}_{j-1} ,

$$\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{W}_j, \quad (2.2)$$

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

$$\mathbf{L}^2(\mathbf{R}^d) = \bigoplus_{j \in \mathbf{Z}} \mathbf{W}_j. \quad (2.3)$$

If there is the coarsest scale n , then the chain of the subspaces (2.1) is replaced by

$$\mathbf{V}_n \subset \dots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \dots, \quad (2.4)$$

and

$$\mathbf{L}^2(\mathbf{R}^d) = \mathbf{V}_n \bigoplus_{j \leq n} \mathbf{W}_j. \quad (2.5)$$

If there is a finite number of scales, then without loss of generality, we set $j = 0$ to be the finest scale. Instead of (2.4), we then have

$$\mathbf{V}_n \subset \dots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0, \quad \mathbf{V}_0 \subset \mathbf{L}^2(\mathbf{R}^d). \quad (2.6)$$

In numerical realizations the subspace \mathbf{V}_0 is finite dimensional.

II.2 The Haar basis

For a long time the only example of the multiresolution analysis satisfying Definition II.1 with Condition 4' was the Haar basis [25]. Computing in the Haar basis offers a glimpse of the algorithms that we will describe in these lectures and provides a useful prototype for numerical experimentation. If $d = 1$, then the Haar basis $h_{j,k}(x) = 2^{-j/2}h(2^{-j}x - k)$ $j, k \in \mathbf{Z}$, is formed by the dilation and translation of a single function

$$h(x) = \begin{cases} 1 & \text{for } 0 < x < 1/2 \\ -1 & \text{for } 1/2 \leq x < 1 \\ 0 & \text{elsewhere.} \end{cases} \quad (2.7)$$

In this case $\varphi(x) = \chi(x)$, where $\chi(x)$ is the characteristic function of the interval $(0, 1)$. For each j , $\chi_{j,k}(x) = 2^{-j/2}\chi(2^{-j}x - k)$, $k \in \mathbf{Z}$, is the basis of \mathbf{V}_j and $h_{j,k}(x) = 2^{-j/2}h(2^{-j}x - k)$, $k \in \mathbf{Z}$, is the basis of \mathbf{W}_j .

The decomposition of a function into the Haar basis is an order N procedure. Given $N = 2^n$ ‘‘samples’’ of a function, which may for simplicity be thought of as values of scaled averages of f on intervals of length 2^{-n} ,

$$s_k^0 = 2^{n/2} \int_{2^{-n}k}^{2^{-n}(k+1)} f(x)dx, \quad (2.8)$$

we obtain the Haar coefficients

$$d_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j - s_{2k}^j) \quad (2.9)$$

and averages

$$s_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j + s_{2k}^j) \quad (2.10)$$

for $j = 0, \dots, n-1$ and $k = 0, \dots, 2^{n-j-1} - 1$. It is easy to see that evaluating the whole set of coefficients d_k^j, s_k^j in (2.9), (2.10) requires $2(N-1)$ additions and $2N$ multiplications.

In two dimensions, there are two natural ways to construct Haar systems. The first is simply the tensor product $h_{j,j',k,k'}(x, y) = h_{j,k}(x)h_{j',k'}(y)$, so that each basis function $h_{j,j',k,k'}(x, y)$ is supported on a rectangle. Representing an operator in this basis leads

to the standard form. The second basis is defined by the set of three kinds of basis functions supported on squares: $h_{j,k}(x)h_{j,k'}(y)$, $h_{j,k}(x)\chi_{j,k'}(y)$, and $\chi_{j,k}(x)h_{j,k'}(y)$, where $\chi(x)$ is the characteristic function of the interval $(0, 1)$ and $\chi_{j,k}(x) = 2^{-j/2}\chi(2^{-j}x - k)$. Representing an operator in this basis leads to the non-standard form (the terminology will become clear later).

By considering an integral operator

$$T(f)(x) = \int K(x, y)f(y)dy, \quad (2.11)$$

and expanding its kernel in a two-dimensional Haar basis, we find that for Calderón-Zygmund and pseudo-differential operators the decay of entries as a function of the distance from the diagonal is faster in these representations than that in the original kernel. These classes of operators are given by integral or distributional kernels that are smooth away from the diagonal. For example, kernels $K(x, y)$ of Calderón-Zygmund operators satisfy the estimates

$$\begin{aligned} |K(x, y)| &\leq \frac{1}{|x - y|}, \\ |\partial_x^M K(x, y)| + |\partial_y^M K(x, y)| &\leq \frac{C_M}{|x - y|^{1+M}}, \end{aligned} \quad (2.12)$$

for some $M \geq 1$. Let $M = 1$ in (2.12) and consider

$$\beta_{kk'}^j = \int \int K(x, y) h_{j,k}(x) \chi_{j,k'}(y) dx dy, \quad (2.13)$$

where we assume that the distance between $|k - k'| \geq 1$. Since

$$\int h_{j,k}(x) dx = 0, \quad (2.14)$$

we have (2.12)

$$\begin{aligned} |\beta_{kk'}^j| &\leq \left| \int \int [K(x, y) - K(x_{j,k}^0, y)] h_{j,k}(x) \chi_{j,k'}(y) dx dy \right| \\ &\leq \frac{C}{|k - k'|^2}, \end{aligned} \quad (2.15)$$

where $x_{j,k}^0 = 2^j(k + \frac{1}{2})$ denotes the center of the support of $h_{j,k}$. Thus, entries $\beta_{kk'}^j$ decay quadratically as functions of the distance between k and k' . Given a finite precision of calculations, entries $\beta_{kk'}^j$ may be discarded as the distance between k and k' becomes large, leaving only a band around the diagonal.

The rate of decay depends on the number of vanishing moments of the functions of the basis. The Haar functions have only one vanishing moment, $\int h(x)dx = 0$, and for this

reason the gain in the decay is insufficient to make computing in the Haar basis practical. To have a faster decay, it is necessary to use basis functions with several vanishing moments. The vanishing moments are responsible for attaining practical algorithms, i.e. controlling the constants in the complexity estimates of the fast algorithms.

II.3 Orthonormal bases of compactly supported wavelets

The question of the existence of a multiresolution analysis with a smooth function φ of Condition 4' was not resolved until the construction of the orthonormal wavelet bases generalizing the Haar functions by Stromberg [39] and Meyer [30]². We will consider only compactly supported wavelets with vanishing moments constructed by I. Daubechies [19], following the work of Y. Meyer [32] and S. Mallat [28], though most of the results will remain valid for other choices of wavelet bases as well. We will further restrict our choice of wavelet bases by only considering, for dimensions $d \geq 2$, the bases constructed from those for $d = 1$.

Let us consider the multiresolution analysis for $\mathbf{L}^2(\mathbf{R})$ and let $\{\psi(x-k)\}_{k \in \mathbf{Z}}$ be an orthonormal basis of \mathbf{W}_0 . We will require that the function ψ has M vanishing moments,

$$\int_{-\infty}^{+\infty} \psi(x)x^m dx = 0, \quad m = 0, \dots, M-1. \quad (2.16)$$

We outline here the properties of compactly supported wavelets and refer for details to [19], [20] and [33].

There are two immediate consequences of Definition II.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

$$\varphi(x) = \sqrt{2} \sum_{k=0}^{L-1} h_k \varphi(2x - k), \quad (2.17)$$

which we may rewrite as

$$\hat{\varphi}(\xi) = m_0(\xi/2)\hat{\varphi}(\xi/2), \quad (2.18)$$

where

$$\hat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x) e^{ix\xi} dx, \quad (2.19)$$

and the 2π -periodic function m_0 is defined as

$$m_0(\xi) = 2^{-1/2} \sum_{k=0}^{L-1} h_k e^{ik\xi}. \quad (2.20)$$

²We note that the notion of multiresolution analysis is more recent than the constructions of [39], [30] and, of course, [25].

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

$$\delta_{k0} = \int_{-\infty}^{+\infty} \varphi(x - k)\varphi(x) dx = \int_{-\infty}^{+\infty} |\hat{\varphi}(\xi)|^2 e^{-ik\xi} d\xi, \quad (2.21)$$

and therefore,

$$\delta_{k0} = \int_0^{2\pi} \sum_{l \in \mathbf{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2 e^{-ik\xi} d\xi, \quad (2.22)$$

and

$$\sum_{l \in \mathbf{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2 = 1. \quad (2.23)$$

Using (2.18), we obtain

$$\sum_{l \in \mathbf{Z}} |m_0(\xi/2 + \pi l)|^2 |\hat{\varphi}(\xi/2 + \pi l)|^2 = 1. \quad (2.24)$$

By taking the sum in (2.24) separately over odd and even indices, we have

$$\sum_{l \in \mathbf{Z}} |m_0(\xi/2 + 2\pi l)|^2 |\hat{\varphi}(\xi/2 + 2\pi l)|^2 + \sum_{l \in \mathbf{Z}} |m_0(\xi/2 + 2\pi l + \pi)|^2 |\hat{\varphi}(\xi/2 + 2\pi l + \pi)|^2 = 1. \quad (2.25)$$

Using the 2π -periodicity of the function m_0 and (2.23), we obtain, after replacing $\xi/2$ by ξ , a necessary condition

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1. \quad (2.26)$$

Thus, the coefficients h_k in (2.20) are such that the 2π -periodic function m_0 satisfies equation (2.26).

Equation (2.26) defines a pair of the quadrature mirror filters (QMF). On denoting

$$m_1(\xi) = e^{-i\xi} \overline{m_0(\xi + \pi)}, \quad (2.27)$$

and defining the function ψ ,

$$\psi(x) = \sqrt{2} \sum_k g_k \varphi(2x - k), \quad (2.28)$$

where

$$g_k = (-1)^k h_{L-k-1}, \quad k = 0, \dots, L-1, \quad (2.29)$$

or, the Fourier transform of ψ ,

$$\hat{\psi}(\xi) = m_1(\xi/2) \hat{\varphi}(\xi/2), \quad (2.30)$$

where

$$m_1(\xi) = 2^{-1/2} \sum_{k=0}^{L-1} g_k e^{ik\xi}, \quad (2.31)$$

it is not difficult to show (see e.g., [33], [19], [20]), that on 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 .

The following lemma (I. Daubechies [19]) characterizes trigonometric polynomial solutions of (2.26) which correspond to the orthonormal bases of compactly supported wavelets with vanishing moments.

Lemma II.1 *Any trigonometric polynomial solution $m_0(\xi)$ of (2.26) is of the form*

$$m_0(\xi) = \left[\frac{1}{2}(1 + e^{i\xi}) \right]^M Q(e^{i\xi}) \quad (2.32)$$

where $M \geq 1$ is the number of vanishing moments, and where Q is a polynomial, such that

$$|Q(e^{i\xi})|^2 = P(\sin^2 \frac{1}{2}\xi) + \sin^{2M}(\frac{1}{2}\xi) R(\frac{1}{2}\cos \xi), \quad (2.33)$$

where

$$P(y) = \sum_{k=0}^{M-1} \binom{M-1+k}{k} y^k, \quad (2.34)$$

and R is an odd polynomial, such that

$$0 \leq P(y) + y^M R(\frac{1}{2} - y) \quad \text{for } 0 \leq y \leq 1, \quad (2.35)$$

and

$$\sup_{0 \leq y \leq 1} [P(y) + y^M R(\frac{1}{2} - y)] < 2^{2(M-1)}. \quad (2.36)$$

The proof of this Lemma contains an algorithm for generating the coefficients of the quadrature mirror filters H and G . The number L of the filter coefficients in (2.20) and (2.31) is related to the number of vanishing moments M . For the wavelets constructed in [19], $L = 2M$. If additional conditions are imposed (see [7] for an example), then the relation might be different, but L is always even.

The decomposition of a function into the wavelet basis is an order N procedure. Given the coefficients s_k^0 , $k = 0, 1, \dots, N$ as ‘‘samples’’ of the function f , the coefficients s_k^j and d_k^j on scales $j \geq 1$ are computed at a cost proportional to N via

$$s_k^j = \sum_{n=0}^{n=L-1} h_n s_{n+2k+1}^{j-1}, \quad (2.37)$$

and

$$d_k^j = \sum_{n=0}^{n=L-1} g_n s_{n+2k+1}^{j-1}, \quad (2.38)$$

where s_k^j and d_k^j may be viewed as periodic sequences with the period 2^{n-j} . Computing via (2.37) and (2.38) is illustrated by the pyramid scheme

$$\begin{array}{ccccccc}
\{s_k^0\} & \longrightarrow & \{s_k^1\} & \longrightarrow & \{s_k^2\} & \longrightarrow & \{s_k^3\} \cdots \\
& & \searrow & & \searrow & & \searrow \\
& & & & \{d_k^1\} & & \{d_k^2\} & & \{d_k^3\} \cdots
\end{array} \tag{2.39}$$

II.4 Multi-wavelet bases with vanishing moments

We note that the Haar system is a degenerate case of Daubechies's wavelets with $M = 1$. There is, however, a different construction of orthonormal bases for $\mathbf{L}^2([0, 1])$ or $\mathbf{L}^2(\mathbf{R})$, which generalizes the Haar system and yields basis functions with several vanishing moments [2], [1].

Let us construct M functions, $f_1, \dots, f_M : \mathbf{R} \rightarrow \mathbf{R}$, supported on the interval $[-1, 1]$ and such that (i) on the interval $(0, 1)$ the function f_i is a polynomial of the degree $M - 1$. Let us (ii) extend f_i to the interval $(-1, 0)$ as an even or odd function according to the parity of $i + M - 1$. We require these functions to be orthogonal (iii)

$$\int_{-1}^1 f_i(x) f_l(x) dx = \delta_{il}, \quad i, l = 1, \dots, M. \tag{2.40}$$

and have vanishing moments (iv)

$$\int_{-1}^1 f_i(x) x^m dx = 0, \quad m = 0, 1, \dots, i + M - 2. \tag{2.41}$$

The properties (i) and (ii) imply that there are M^2 polynomial coefficients that determine the functions f_1, \dots, f_M , while the properties (iii) and (iv) provide M^2 constraints. It turns out that the equations uncouple to give M nonsingular linear systems which may be solved to obtain the coefficients and, thus, yielding the functions f_1, \dots, f_M up to a sign.

The functions f_1, \dots, f_M may be obtained constructively as follows. Let us start with

$$f_m^1(x) = \begin{cases} x^{m-1}, & x \in (0, 1), \\ -x^{m-1}, & x \in (-1, 0), \\ 0, & \text{otherwise,} \end{cases} \tag{2.42}$$

and note that the $2M$ functions $1, x, \dots, x^{M-1}, f_1^1, f_2^1, \dots, f_M^1$ are linearly independent. Then, by the Gram-Schmidt process, we orthogonalize f_m^1 with respect to $1, x, \dots, x^{M-1}$, to obtain f_m^2 , for $m = 1, \dots, M$. This orthogonality is preserved by the remaining orthogonalizations, which only produce linear combinations of the f_m^2 . First, if at least one of f_m^2 is not orthogonal to x^M , we reorder the functions, so that $\langle f_1^2, x^M \rangle \neq 0$.

We then define $f_m^3 = f_m^2 - a_m \cdot f_0^2$, where a_m is chosen so that $\langle f_m^3, x^M \rangle = 0$ for $m = 2, \dots, M$, achieving the desired orthogonality to x^M . Similarly, we continue to orthogonalize with respect to x^{M+1}, \dots, x^{2M-2} to obtain $f_1^2, f_2^3, f_3^4, \dots, f_M^{M+1}$, such that $\langle f_m^{m+1}, x^i \rangle = 0$ for $i \leq m + M - 2$. Finally, we perform the Gram-Schmidt orthogonalization on $f_M^{M+1}, f_{M-1}^M, \dots, f_1^2$, in that order, and normalize to obtain f_M, f_{M-1}, \dots, f_1 . By construction, the functions $\{f_m\}_{m=1}^M$ satisfy properties (i)-(iv). On denoting

$$h_m(x) = 2^{1/2} f_m(2x - 1), \quad m = 1, \dots, M, \quad (2.43)$$

we define the space \mathbf{W}_j^M , $j = 0, -1, -2, \dots$, as a linear span of functions

$$h_{m,j}^n(x) = 2^{-j/2} h_m(2^{-j}x - n), \quad m = 1, \dots, M; \quad n = 0, \dots, 2^{-j} - 1. \quad (2.44)$$

We also define \mathbf{V}_0^M to be the space of polynomials of degree less than M . By defining recursively

$$\mathbf{V}_{j-1}^M = \mathbf{V}_j^M \oplus \mathbf{W}_j^M, \quad (2.45)$$

starting with $j = 0$, we obtain multiresolution analysis

$$\mathbf{V}_0^M \subset \mathbf{V}_{-1}^M \subset \dots \subset \mathbf{V}_j^M \subset \dots. \quad (2.46)$$

It is not difficult to check that

$$\mathbf{L}^2([0, 1]) = \mathbf{V}_0^M \oplus_{j \leq 0} \mathbf{W}_j^M. \quad (2.47)$$

Also, it is easy to see that the orthonormal set

$$h_{m,j}^n(x) = 2^{-j/2} h_m(2^{-j}x - n), \quad m = 1, \dots, M; \quad j, n \in \mathbf{Z}. \quad (2.48)$$

is an orthonormal basis of $\mathbf{L}^2(\mathbf{R})$.

We refer to bases constructed in this manner as *multi-wavelet* bases of order M . Algorithms, various numerical examples, and applications utilizing these bases are described in [2] and [1]. It is interesting to note that the multi-wavelet bases have been constructed only after ‘‘ordinary’’ wavelets had been used in algorithms of [7], though their construction addresses the problem of vanishing moments directly and does not require any prior knowledge of the wavelet theory. If the number of vanishing moments $M \geq 2$, then, on a given scale j , the multi-wavelet basis functions with different labels n do not have an overlapping support unlike the ‘‘ordinary’’ compactly supported wavelets. On the other hand, the non-standard form of operators in the multi-wavelet bases is more complicated than that in the bases of the ‘‘ordinary’’ wavelets. This becomes clear once we outline the construction of bases for $\mathbf{L}^2[0, 1]^d$ and $\mathbf{L}^2(\mathbf{R}^d)$, for any dimension d . We describe this extension by giving the basis for $\mathbf{L}^2([0, 1]^2)$, which is illustrative of the construction for any finite-dimensional space. Let us define the space $\mathbf{V}_j^{M,2}$ as

$$\mathbf{V}_j^{M,2} = \mathbf{V}_j^M \times \mathbf{V}_j^M, \quad j = 0, -1, -2, \dots, \quad (2.49)$$

where \mathbf{V}_j^M is given in (2.45), and the space $\mathbf{W}_j^{M,2}$ as the orthogonal complement of $\mathbf{V}_j^{M,2}$ in $\mathbf{V}_{j-1}^{M,2}$,

$$\mathbf{V}_{j-1}^{M,2} = \mathbf{V}_j^{M,2} \oplus \mathbf{W}_j^{M,2}. \quad (2.50)$$

The space $\mathbf{W}_0^{M,2}$ is spanned by the orthonormal basis

$$\{u_i(x)h_l(y), h_i(x)u_l(y), h_i(x)h_l(y) : i, l = 1, \dots, M\}, \quad (2.51)$$

where $\{u_m\}_{m=1}^M$ is an orthonormal basis for \mathbf{V}_0^M . Each element among these $3M^2$ basis elements has vanishing moments, i.e., it is orthogonal to the polynomials $x^i y^l$, $i, l = 0, 1, \dots, M-1$.

The space $\mathbf{W}_j^{M,2}$ is spanned by dilations and translations of the basis functions of $\mathbf{W}_0^{M,2}$ and the basis of $\mathbf{L}^2([0, 1]^2)$ consists of these functions and the low-order polynomials $x^i y^l$, $i, l = 0, 1, \dots, M-1$.

We note that the two-dimensional multi-wavelet bases require $3M^2$ different combinations of one-dimensional basis functions, where M is the number of vanishing moments. On the other hand, the two-dimensional bases, obtained by using compactly supported wavelets [19], require only three such combinations which simplifies the construction of the non-standard form (see Section III).

II.5 A remark on computing in the wavelet bases

Finally, we note that once the filter H has been chosen, it completely determines the functions φ and ψ and therefore, the multiresolution analysis. It is an interesting observation, that in properly constructed algorithms, the functions φ and ψ are 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,

$$\mathcal{M}_\infty^m = \int x^m \varphi(x) dx, \quad m = 0, \dots, M-1, \quad (2.52)$$

in terms of the filter coefficients $\{h_k\}_{k=1}^{k=L}$, may be found using a formula for $\hat{\varphi}$,

$$\hat{\varphi}(\xi) = (2\pi)^{-1/2} \prod_{j=1}^{\infty} m_0(2^{-j}\xi), \quad (2.53)$$

where

$$m_0(\xi) = 2^{-1/2} \sum_{k=0}^{k=L-1} h_k e^{ik\xi}. \quad (2.54)$$

The moments \mathcal{M}_∞^m are obtained numerically (within the desired accuracy) by recursively generating a sequence of vectors, $\{\mathcal{M}_r^m\}_{m=0}^{m=M-1}$ for $r = 1, 2, \dots$,

$$\mathcal{M}_{r+1}^m = \sum_{j=0}^{j=m} \binom{m}{j} 2^{-jr} \mathcal{M}_r^{m-j} \mathcal{M}_1^j, \quad (2.55)$$

starting with

$$\mathcal{M}_1^m = 2^{-m-\frac{1}{2}} \sum_{k=0}^{k=L-1} h_k k^m, \quad m = 0, \dots, M-1. \quad (2.56)$$

Each vector $\{\mathcal{M}_r^m\}_{m=0}^{m=M-1}$ represents M moments of the product in (2.53) with r terms, and this iteration converges rapidly. Notice, that we never computed the function φ itself.

III The non-standard and standard forms

III.1 The Non-Standard Form

Let T be an operator

$$T : \mathbf{L}^2(\mathbf{R}) \rightarrow \mathbf{L}^2(\mathbf{R}), \quad (3.1)$$

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

$$P_j : \mathbf{L}^2(\mathbf{R}) \rightarrow \mathbf{V}_j, \quad (3.2)$$

as

$$(P_j f)(x) = \sum_k \langle f, \varphi_{j,k} \rangle \varphi_{j,k}(x), \quad (3.3)$$

and expanding T in a “telescopic” series, we obtain

$$T = \sum_{j \in \mathbf{Z}} (Q_j T Q_j + Q_j T P_j + P_j T Q_j), \quad (3.4)$$

where

$$Q_j = P_{j-1} - P_j \quad (3.5)$$

is the projection operator on the subspace \mathbf{W}_j . If there is the coarsest scale n , then instead of (3.4) we have

$$T = \sum_{j=-\infty}^n (Q_j T Q_j + Q_j T P_j + P_j T Q_j) + P_n T P_n, \quad (3.6)$$

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

$$T_0 = \sum_{j=1}^n (Q_j T Q_j + Q_j T P_j + P_j T Q_j) + P_n T P_n, \quad (3.7)$$

where $T \sim T_0 = P_0 T P_0$ is a discretization of the operator T on the finest scale. Expansions (3.4), (3.6) and (3.7) decompose the operator T into a sum of contributions from different scales.

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

$$T = \{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}} \quad (3.8)$$

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

$$A_j : \mathbf{W}_j \rightarrow \mathbf{W}_j, \quad (3.9)$$

$$B_j : \mathbf{V}_j \rightarrow \mathbf{W}_j, \quad (3.10)$$

$$\Gamma_j : \mathbf{W}_j \rightarrow \mathbf{V}_j, \quad (3.11)$$

where 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 $\Gamma_j = P_j T Q_j$.

The operators $\{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}}$ admit a recursive definition via the relation

$$T_j = \begin{pmatrix} A_{j+1} & B_{j+1} \\ \Gamma_{j+1} & T_{j+1} \end{pmatrix}, \quad (3.12)$$

where operators $T_j = P_j T P_j$,

$$T_j : \mathbf{V}_j \rightarrow \mathbf{V}_j, \quad (3.13)$$

and the operator represented by the 2×2 matrix in (3.12) is a mapping

$$\begin{pmatrix} A_{j+1} & B_{j+1} \\ \Gamma_{j+1} & T_{j+1} \end{pmatrix} : \mathbf{W}_{j+1} \oplus \mathbf{V}_{j+1} \rightarrow \mathbf{W}_{j+1} \oplus \mathbf{V}_{j+1} \quad (3.14)$$

If there is a coarsest scale n , then

$$T = \{\{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}: j \leq n}, T_n\}, \quad (3.15)$$

where $T_n = P_n T P_n$. If the number of scales is finite, then $j = 1, 2, \dots, n$ in (3.15) and the operators are organized as blocks of the matrix (see Figures 1 and 2).

Let us make the following observations:

1. The operator A_j describes the interaction on the scale j only, since the subspace \mathbf{W}_j in (3.9) is an element of the direct sum in (2.3).
2. The operators B_j, Γ_j in (3.10) and (3.11) describe the interaction between the scale j and all coarser scales. Indeed, the subspace \mathbf{V}_j contains all the subspaces $\mathbf{V}_{j'}$ with $j' > j$ (see (2.1)).
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 ,

$$\alpha_{k,k'}^j = \int \int K(x, y) \psi_{j,k}(x) \psi_{j,k'}(y) dx dy, \quad (3.16)$$

$$\beta_{k,k'}^j = \int \int K(x, y) \psi_{j,k}(x) \varphi_{j,k'}(y) dx dy, \quad (3.17)$$

and

$$\gamma_{k,k'}^j = \int \int K(x, y) \varphi_{j,k}(x) \psi_{j,k'}(y) dx dy. \quad (3.18)$$

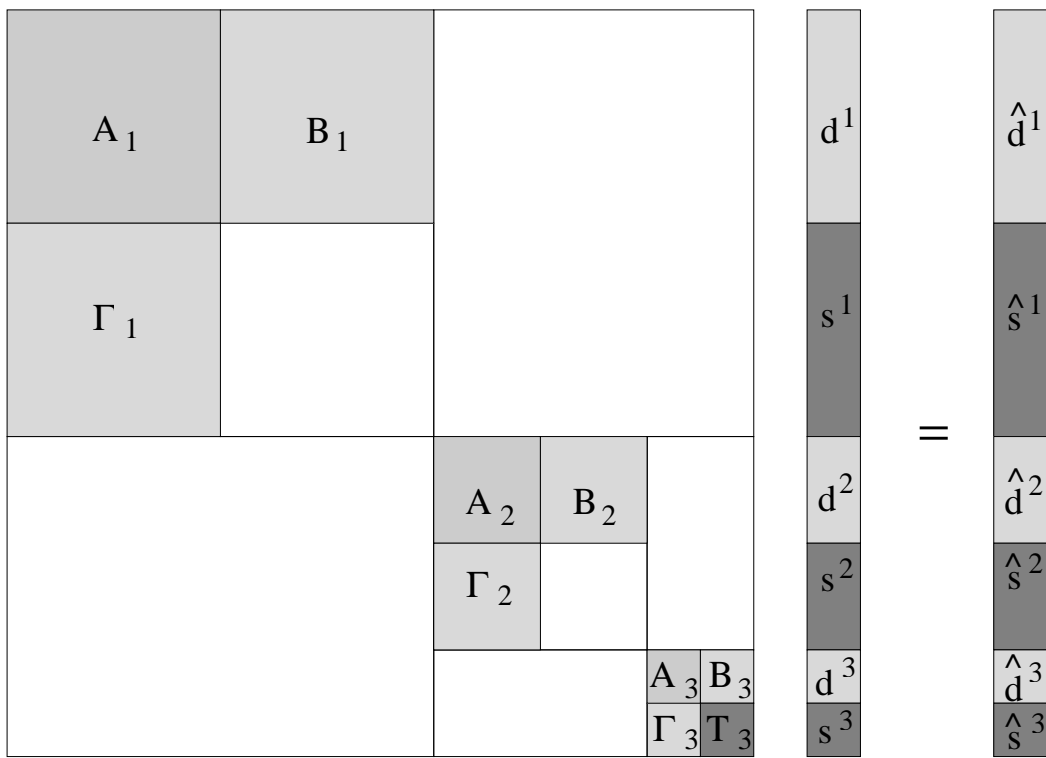


Figure 1: Organization of the non-standard form of a matrix. The submatrices A_j , B_j , and Γ_j , $j = 1, 2, 3$, and T_3 are the only non-zero submatrices.

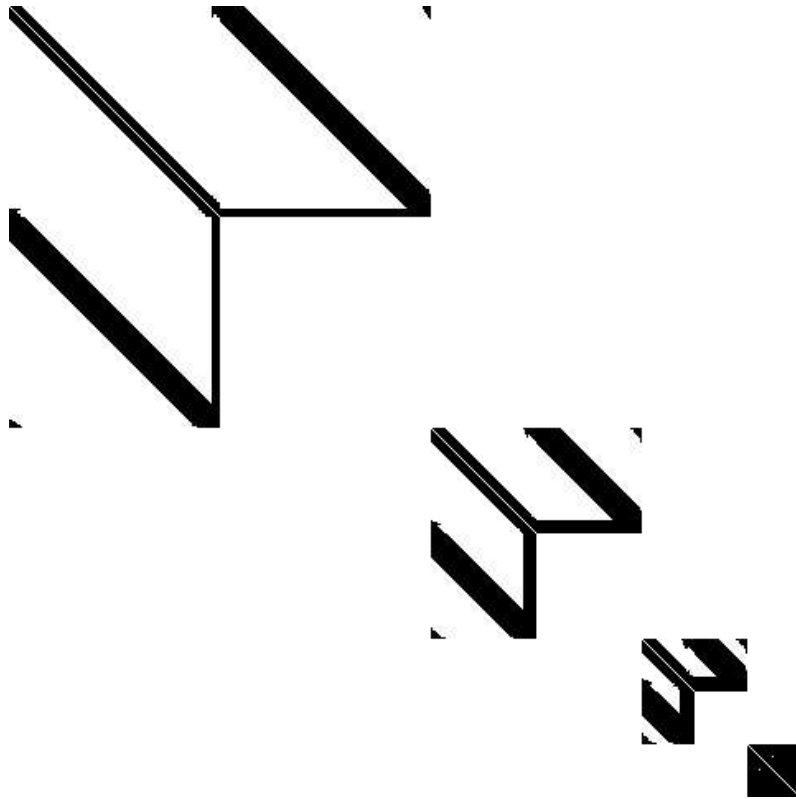


Figure 2: An example of a matrix in the non-standard form (see Example 1)

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

$$s_{k,k'}^j = \int \int K(x, y) \varphi_{j,k}(x) \varphi_{j,k'}(y) dx dy. \quad (3.19)$$

Given a set of coefficients $s_{k,k'}^0$ with $k, k' = 0, 1, \dots, N-1$, a repeated application of the formulas (2.37), (2.38) produces

$$\alpha_{i,l}^j = \sum_{k,m=0}^{L-1} g_k g_m s_{k+2i+1, m+2l+1}^{j-1}, \quad (3.20)$$

$$\beta_{i,l}^j = \sum_{k,m=0}^{L-1} g_k h_m s_{k+2i+1, m+2l+1}^{j-1}, \quad (3.21)$$

$$\gamma_{i,l}^j = \sum_{k,m=0}^{L-1} h_k g_m s_{k+2i+1, m+2l+1}^{j-1}, \quad (3.22)$$

$$s_{i,l}^j = \sum_{k,m=0}^{L-1} h_k h_m s_{k+2i+1, m+2l+1}^{j-1}, \quad (3.23)$$

with $i, l = 0, 1, \dots, 2^{n-j} - 1$, $j = 1, 2, \dots, n$. Clearly, formulas (3.20) - (3.23) provide an order N^2 scheme for the evaluation of the elements of all matrices $\alpha^j, \beta^j, \gamma^j$ with $j = 1, 2, \dots, n$.

A fast scheme for the evaluation of the elements of matrices $\alpha^j, \beta^j, \gamma^j$ with $j = 1, 2, \dots, n$, requires that the locations of the singularities were known in advance. Given the locations of the singularities, only the significant coefficients above a threshold of accuracy are computed, and the only issue is how to compute averages $s_{i,l}^j$ in a neighborhood of the singularities without (3.23). We refer to [7] for the details of such computations using quadrature formulas.

Remark. It follows from (3.7) that after applying the non-standard form to a vector, we arrive at the representation

$$(T_0 u_0)(x) = \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} \hat{d}_k^j \psi_k^j(x) + \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} \hat{s}_k^j \phi_k^j(x), \quad (3.24)$$

where $u_0 = P_0 u$ (see 1). We note, that in order to reconstruct a vector from the first and the second sums in (3.24), we may apply the same algorithm. The reconstruction from the first sum is just a reconstruction from the wavelet expansion and is accomplished by using the quadrature mirror filters H and G . The reconstruction from the second sum may be accomplished by using the same algorithm but with the pair of filters H and H . Both results are then added. The number of operations required for this computation is proportional to N .

III.2 The Standard Form

The standard form is obtained by representing

$$\mathbf{V}_j = \bigoplus_{j' > j} \mathbf{W}_{j'}, \quad (3.25)$$

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

$$B_j^{j'} : \mathbf{W}_{j'} \rightarrow \mathbf{W}_j, \quad (3.26)$$

$$\Gamma_j^{j'} : \mathbf{W}_j \rightarrow \mathbf{W}_{j'}. \quad (3.27)$$

If there is the coarsest scale n , then instead of (3.25), we have

$$\mathbf{V}_j = \mathbf{V}_n \bigoplus_{j'=j+1}^{j'=n} \mathbf{W}_{j'}. \quad (3.28)$$

In this case, the operators $\{B_j^{j'}, \Gamma_j^{j'}\}$ for $j' = j + 1, \dots, n$ are the same as in (3.26) and (3.27) and, in addition, for each scale j , there are operators $\{B_j^{n+1}\}$ and $\{\Gamma_j^{n+1}\}$,

$$B_j^{n+1} : \mathbf{V}_n \rightarrow \mathbf{W}_j, \quad (3.29)$$

$$\Gamma_j^{n+1} : \mathbf{W}_j \rightarrow \mathbf{V}_n. \quad (3.30)$$

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

$$T_0 = \{A_j, \{B_j^{j'}\}_{j'=j+1}^{j'=n}, \{\Gamma_j^{j'}\}_{j'=j+1}^{j'=n}, B_j^{n+1}, \Gamma_j^{n+1}, T_n\}_{j=1, \dots, n}. \quad (3.31)$$

The operators (3.31) are organized as blocks of the matrix (see Figure 3 and Figure 4).

If the operator T is a Calderón-Zygmund or a pseudo-differential operator then, for a fixed accuracy, all the operators in (3.31), 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. First consists in applying the one-dimensional transform (see (2.37) and (2.38)) 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 [7] for details.

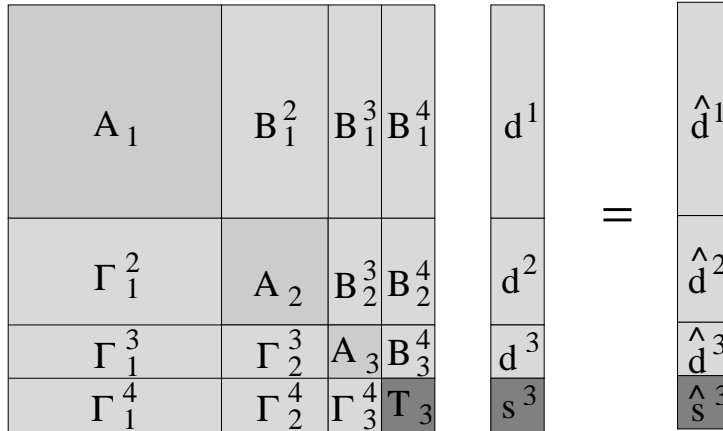


Figure 3: Organization of a matrix in the standard form

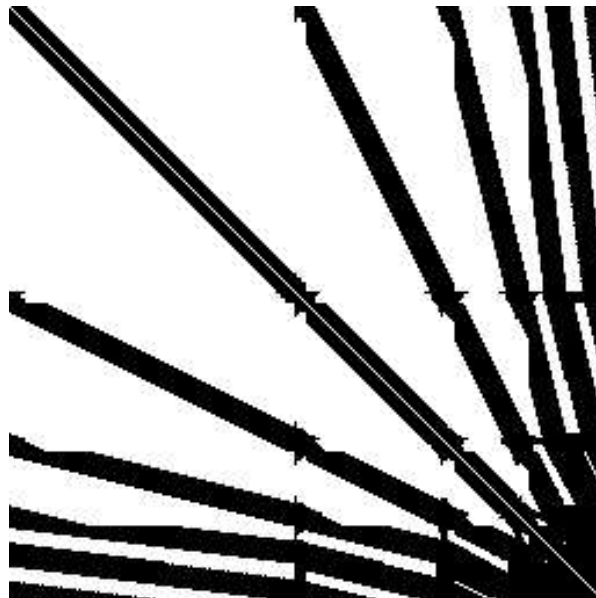


Figure 4: An example of a matrix in the standard form (see Example 1)

IV The compression of operators

The compression of operators or, in other words, the construction of their sparse representations in orthonormal bases, directly affects the speed of computational algorithms. While the compression of data (of images, for example) achieved by methods other than finding a sparse representation in some basis may be adequate for some applications, the compression of operators calls for a representation in a basis in order to effectively compute in the sparse form. The standard and non-standard forms of operators in the wavelet bases may be viewed as compression schemes for a wide class of operators. We restrict our attention to two specific classes frequently encountered in analysis and applications, Calderón-Zygmund and pseudo-differential operators.

The non-standard form of an operator T with a kernel $K(x, y)$ is obtained by evaluating the expressions

$$\alpha_{II'} = \int \int K(x, y) \psi_I(x) \psi_{I'}(y) dx dy, \quad (4.1)$$

$$\beta_{II'} = \int \int K(x, y) \psi_I(x) \varphi_{I'}(y) dx dy, \quad (4.2)$$

and

$$\gamma_{II'} = \int \int K(x, y) \varphi_I(x) \psi_{I'}(y) dx dy, \quad (4.3)$$

where the coefficients are labeled by the intervals $I = I_k^j$ and $I' = I_{k'}^j$ denoting the supports of φ_k^j and $\varphi_{k'}^j$.

If K is smooth on the square $I \times I'$, then expanding K into a Taylor series around the center of the square and combining (2.16) with (4.1) - (4.3) and remembering that the functions ψ_I and $\psi_{I'}$ are supported on the intervals I and I' , we obtain the estimate

$$|\alpha_{II'}| + |\beta_{II'}| + |\gamma_{II'}| \leq C|I|^{M+1} \sup_{(x,y) \in I \times I'} (|\partial_x^M K(x, y)| + |\partial_y^M K(x, y)|). \quad (4.4)$$

Obviously, the right-hand side of (4.4) is small whenever either $|I|$ or the derivatives involved are small, and we use this fact to “compress” matrices of integral operators by converting them to the non-standard form, and discarding the coefficients that are smaller than a chosen threshold.

For most numerical applications, the following Proposition IV.1 is quite adequate, as long as the singularity of K is integrable across each row and each column.

Proposition IV.1 *If the wavelet basis has M vanishing moments, then for any kernel K satisfying the conditions*

$$|K(x, y)| \leq \frac{1}{|x - y|}, \quad (4.5)$$

$$|\partial_x^M K(x, y)| + |\partial_y^M K(x, y)| \leq \frac{C_0}{|x - y|^{1+M}} \quad (4.6)$$

the matrices $\alpha^j, \beta^j, \gamma^j$ (3.16) - (3.18) of the non-standard form satisfy the estimate

$$|\alpha_{i,l}^j| + |\beta_{i,l}^j| + |\gamma_{i,l}^j| \leq \frac{C_M}{1 + |i - l|^{M+1}}, \quad (4.7)$$

for all $|i - l| \geq 2M$.

Similar considerations apply in the case of pseudo-differential operators. Let T be a pseudo-differential operator with symbol $\sigma(x, \xi)$ defined by the formula

$$T(f)(x) = \int e^{ix\xi} \sigma(x, \xi) \hat{f}(\xi) d\xi = \int K(x, y) f(y) dy, \quad (4.8)$$

where K is the distributional kernel of T .

Proposition IV.2 *If the wavelet basis has M vanishing moments, then for any pseudo-differential operator with symbol σ of T and σ^* of T^* satisfying the standard conditions*

$$|\partial_\xi^\alpha \partial_x^\beta \sigma(x, \xi)| \leq C_{\alpha,\beta} (1 + |\xi|)^{\lambda - \alpha + \beta} \quad (4.9)$$

$$|\partial_\xi^\alpha \partial_x^\beta \sigma^*(x, \xi)| \leq C_{\alpha,\beta} (1 + |\xi|)^{\lambda - \alpha + \beta}, \quad (4.10)$$

the matrices $\alpha^j, \beta^j, \gamma^j$ (3.16) - (3.18) of the non-standard form satisfy the estimate

$$|\alpha_{i,l}^j| + |\beta_{i,l}^j| + |\gamma_{i,l}^j| \leq \frac{2^{\lambda j} C_M}{(1 + |i - l|)^{M+1}}, \quad (4.11)$$

for all integer i, l .

If we approximate the operator T_0^N by the operator $T_0^{N,B}$ obtained from T_0^N by setting to zero all coefficients of matrices $\alpha_{i,l}^j, \beta_{i,l}^j$, and $\gamma_{i,l}^j$ outside of bands of width $B \geq 2M$ around their diagonals, then it is easy to see that

$$\|T_0^{N,B} - T_0^N\| \leq \frac{C}{B^M} \log_2 N, \quad (4.12)$$

where C is a constant determined by the kernel K . 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 B has to be such that

$$\|T_0^{N,B} - T_0^N\| \leq \frac{C}{B^M} \log_2 N \leq \epsilon, \quad (4.13)$$

or, equivalently,

$$B \geq \left(\frac{C}{\varepsilon} \log_2 N \right)^{1/M}. \quad (4.14)$$

The estimate (4.12) is sufficient for practical purposes. It is possible, however, to obtain

$$\|T_0^{N,B} - T_0^N\| \leq \frac{C}{B^M} \quad (4.15)$$

instead of (4.12). In order to obtain this tighter estimate and avoid the factor $\log_2 N$, we use some of the ideas arising in the proof of the ‘‘T(1)’’ theorem of David and Journé.

The estimates (4.5), (4.6) are not sufficient to conclude that $\alpha_{i,l}^j, \beta_{i,l}^j, \gamma_{i,l}^j$ are bounded for $|i-l| \leq 2M$ (for example, consider $K(x,y) = \frac{1}{|x-y|}$). Therefore, we need to assume that T defines a bounded operator on L^2 or, a substantially weaker condition,

$$\left| \int_{I \times I} K(x,y) dx dy \right| \leq C|I|, \quad (4.16)$$

for all dyadic intervals I . This is the so-called ‘‘weak cancellation condition’’ (see [32]). Under this condition and the conditions (4.5), (4.6) Proposition IV.1 may be extended so that the estimate (4.7) holds for all integer i, l (see [32]).

Let us evaluate an operator T in the non-standard form on a function f ,

$$T(f)(x) = \sum_{I,I'} \psi_I(x) \alpha_{II'} d_{I'} + \sum_{I,I'} \psi_I(x) \beta_{II'} s_{I'} + \sum_{I,I'} \varphi_I(x) \gamma_{II'} d_{I'} \quad (4.17)$$

and write (4.17) as a sum of three terms,

$$T = \mathcal{A} + L_\beta + L_\gamma^*, \quad (4.18)$$

$$\mathcal{A}(f)(x) = \sum_{I,I'} \psi_I(x) \alpha_{I,I'} d_{I'} + \sum_{I,I'} \psi_I(x) \beta_{I,I'} (s_{I'} - s_I) + \sum_{I,I'} (\varphi_I(x) - \varphi_{I'}(x)) \gamma_{I,I'} d_{I'} \quad (4.19)$$

$$L_\beta(f)(x) = \sum_I \psi_I(x) s_I \beta_I, \quad (4.20)$$

$$L_\gamma^*(f)(x) = \sum_{I'} \varphi_{I'}(x) d_{I'} \gamma_{I'} \quad (4.21)$$

where

$$\beta_I = \sum_{I'} \beta_{II'} = \frac{1}{|I|^{1/2}} \int \int \psi_I(x) K(x,y) dx dy = \langle \psi_I, \beta(x) \rangle \frac{1}{|I|^{1/2}} \quad (4.22)$$

$$\gamma_{I'} = \sum_I \gamma_{II'} = \frac{1}{|I'|^{1/2}} \int \int \psi_{I'}(y) K(x,y) dx dy = \langle \psi_{I'}, \gamma(y) \rangle \frac{1}{|I'|^{1/2}} \quad (4.23)$$

and

$$\beta(x) = T(1)(x), \quad (4.24)$$

$$\gamma(y) = T^*(1)(y) \tag{4.25}$$

It is a remarkable fact that by analysing the functions (4.24) and (4.25) (and, therefore, the operators L_β and L_γ^*), it is possible to decide if a Calderón-Zygmund operator is bounded.

Theorem IV.1 (G. David, J.L. Journé) *Suppose that the operator (3.1) satisfies the conditions (4.5), (4.6), and (4.16). Then a necessary and sufficient condition for T to be bounded on L^2 is that $\beta(x)$ in (4.24) and $\gamma(y)$ in (4.25) belong to dyadic B.M.O., i.e. satisfy condition*

$$\sup_J \frac{1}{|J|} \int_J |\beta(x) - m_J(\beta)|^2 dx \leq C, \tag{4.26}$$

where J is a dyadic interval and

$$m_J(\beta) = \frac{1}{|J|} \int_J \beta(x) dx. \tag{4.27}$$

Splitting the operator T into the sum of three terms (4.18) and estimating them separately leads to the estimate (4.15). We note that the functions $T(1)$ and $T^*(1)$ are easily computed in the process of constructing the non-standard form and may be used to provide a useful estimate of the norm of the operator.

V The differential operators in wavelet bases.

V.1 The operator d/dx in wavelet bases

The non-standard forms of several commonly used operators may be computed explicitly. First we construct the non-standard form of the operator d/dx . The matrix elements α_{il}^j , β_{il}^j , γ_{il}^j of A_j , B_j , Γ_j , and r_{il}^j of $T_j = P_j T P_j$, $i, l, j \in \mathbf{Z}$, for the operator d/dx are easily computed as

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

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

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

and

$$r_{il}^j = 2^{-j} \int_{-\infty}^{\infty} \varphi(2^{-j}x - i) \varphi'(2^{-j}x - l) 2^{-j} dx = 2^{-j} r_{i-l}, \quad (5.4)$$

where

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

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

$$\gamma_l = \int_{-\infty}^{+\infty} \varphi(x - l) \frac{d}{dx} \psi(x) dx. \quad (5.7)$$

and

$$r_l = \int_{-\infty}^{+\infty} \varphi(x - l) \frac{d}{dx} \varphi(x) dx. \quad (5.8)$$

Moreover, using (2.17) and (2.28) we have

$$\alpha_i = 2 \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k g_{k'} r_{2i+k-k'}, \quad (5.9)$$

$$\beta_i = 2 \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k h_{k'} r_{2i+k-k'}, \quad (5.10)$$

and

$$\gamma_i = 2 \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} h_k g_{k'} r_{2i+k-k'}, \quad (5.11)$$

and, therefore, the representation of d/dx is completely determined by r_l in (5.8) or in other words, by the representation of d/dx on the subspace \mathbf{V}_0 .

Rewriting (5.8) in terms of $\hat{\varphi}(\xi)$, where

$$\hat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x) e^{ix\xi} dx, \quad (5.12)$$

we obtain

$$r_l = \int_{-\infty}^{+\infty} (-i\xi) |\hat{\varphi}(\xi)|^2 e^{-il\xi} d\xi. \quad (5.13)$$

We prove the following

Proposition V.1 *1. If the integrals in (5.8) or (5.13) exist, then the coefficients r_l , $l \in \mathbf{Z}$ in (5.8) satisfy the following system of linear algebraic equations*

$$r_l = 2 \left[r_{2l} + \frac{1}{2} \sum_{k=1}^{L/2} a_{2k-1} (r_{2l-2k+1} + r_{2l+2k-1}) \right], \quad (5.14)$$

and

$$\sum_l l r_l = -1, \quad (5.15)$$

where

$$a_{2k-1} = 2 \sum_{i=0}^{L-2k} h_i h_{i+2k-1}, \quad k = 1, \dots, L/2. \quad (5.16)$$

2. If $M \geq 2$, then equations (5.14) and (5.15) 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

$$r_l = -r_{-l}, \quad (5.17)$$

Remark 1. If $M = 1$, then equations (5.14) and (5.15) have a unique solution but the integrals in (5.8) or (5.13) may not be absolutely convergent. For the Haar basis ($h_1 = h_2 = 2^{-1/2}$) $a_1 = 1$ and $r_1 = -1/2$ and we obtain the simplest finite difference operator $(1/2, 0, -1/2)$. In this case the function φ is not continuous and

$$\hat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \frac{\sin \frac{1}{2}\xi}{\frac{1}{2}\xi} e^{i\frac{1}{2}\xi}.$$

Any trigonometric polynomial $m_0(\xi)$ satisfying (2.26) may be written as

$$|m_0(\xi)|^2 = \frac{1}{2} + \frac{1}{2} \sum_{k=1}^{L/2} a_{2k-1} \cos(2k-1)\xi, \quad (5.18)$$

where a_n are the autocorrelation coefficients of the filter $H = \{h_k\}_{k=0}^{L-1}$,

$$a_n = 2 \sum_{i=0}^{L-1-n} h_i h_{i+n}, \quad n = 1, \dots, L-1. \quad (5.19)$$

It is easy to see that the autocorrelation coefficients a_n with even indices are zero,

$$a_{2k} = 0, \quad k = 1, \dots, L/2 - 1. \quad (5.20)$$

This may be verified by using (2.20) to compute $|m_0(\xi)|^2$ and $|m_0(\xi + \pi)|^2$,

$$|m_0(\xi)|^2 = \frac{1}{2} + \frac{1}{2} \sum_{n=1}^{L-1} a_n \cos n\xi, \quad (5.21)$$

$$|m_0(\xi + \pi)|^2 = \frac{1}{2} - \frac{1}{2} \sum_{k=1}^{L/2} a_{2k-1} \cos(2k-1)\xi + \frac{1}{2} \sum_{k=1}^{L/2-1} a_{2k} \cos 2k\xi, \quad (5.22)$$

where a_n are given in (5.19). Combining (5.21) and (5.22) to satisfy (2.26), we obtain

$$\sum_{k=1}^{L/2-1} a_{2k} \cos 2k\xi = 0, \quad (5.23)$$

and hence, (5.20) and (5.18).

The even moments of the coefficients a_{2k-1} from (5.19) vanish, namely

$$\sum_{k=1}^{k=L/2} a_{2k-1} (2k-1)^{2m} = 0 \quad \text{for } 1 \leq m \leq M-1, \quad (5.24)$$

since

$$\left[\left(\frac{1}{i} \partial_\xi \right)^m |m_0(\xi)|^2 \right]_{\xi=0} = 0, \quad \text{for } 1 \leq m \leq 2M-1, \quad (5.25)$$

which follows from the explicit representation in (2.32).

Proof.

Using (2.17) for both $\varphi(x-l)$ and $\frac{d}{dx}\varphi(x)$ in (5.8) we obtain

$$r_i = 2 \sum_{k=0}^{L-1} \sum_{l=0}^{L-1} h_k h_l \int_{-\infty}^{+\infty} \varphi(2x-2i-k) \varphi'(2x-l) 2 dx \quad (5.26)$$

and hence,

$$r_i = 2 \sum_{k=0}^{L-1} \sum_{l=0}^{L-1} h_k h_l r_{2i+k-l}. \quad (5.27)$$

Substituting $l = k - m$, we rewrite (5.27) as

$$r_i = 2 \sum_{k=0}^{L-1} \sum_{m=k}^{k-L+1} h_k h_{k-m} r_{2i+m}. \quad (5.28)$$

Changing the order of summation in (5.28) and using the fact that $\sum_{k=0}^{L-1} h_k^2 = 1$, we arrive at

$$r_l = 2r_{2l} + \sum_{n=1}^{L-1} a_n (r_{2l-n} + r_{2l+n}), \quad l \in \mathbf{Z}, \quad (5.29)$$

where a_n are given in (5.19). Using (5.20) we obtain (5.14) from (5.29).

In order to obtain (5.15) we use the following relation

$$\sum_{l=-\infty}^{l=+\infty} l^m \varphi(x-l) = x^m + \sum_{l=1}^{l=m} (-1)^l \binom{m}{l} M_l^\varphi x^{m-l}, \quad (5.30)$$

where

$$M_l^\varphi = \int_{-\infty}^{+\infty} \varphi(x) x^l dx, \quad l = 1, \dots, m, \quad (5.31)$$

are the moments of the function $\varphi(x)$. Relation (5.31) follows simply on taking Fourier transforms and using Leibniz' rule. Using (5.8) and (5.30) with $m = 1$ we obtain (5.15).

If $M \geq 2$, then

$$|\hat{\varphi}(\xi)|^2 |\xi| \leq C(1 + |\xi|)^{-1-\epsilon}, \quad (5.32)$$

where $\epsilon > 0$, and hence, the integral in (5.13) is absolutely convergent. This assertion follows from Lemma 3.2 of [19], where it is shown that

$$|\hat{\varphi}(\xi)| \leq C(1 + |\xi|)^{-M+\log_2 B}, \quad (5.33)$$

where

$$B = \sup_{\xi \in \mathbf{R}} |Q(e^{i\xi})|.$$

Due to the condition (2.36), we have $\log_2 B = M - 1 - \epsilon$ with some $\epsilon > 0$.

The existence of a solution of the system of equations (5.14) and (5.15) follows from the existence of the integral in (5.13). Since the scaling function φ has a compact support there are only finite number of non-zero coefficients r_l . The specific interval $-L + 2 \leq l \leq L - 2$ is obtained by the direct examination of (5.14).

Let us show now that

$$\sum_l r_l = 0. \quad (5.34)$$

Multiplying (5.14) by $e^{il\xi}$ and summing over l , we obtain

$$\hat{r}(\xi) = 2 \left[\hat{r}_{even}(\xi/2) + \frac{1}{2} \hat{r}_{odd}(\xi/2) \sum_{k=1}^{L/2} a_{2k-1} \left(e^{-i(2k-1)\xi/2} + e^{i(2k-1)\xi/2} \right) \right], \quad (5.35)$$

where

$$\hat{r}(\xi) = \sum_l r_l e^{il\xi}, \quad (5.36)$$

$$\hat{r}_{even}(\xi/2) = \sum_l r_{2l} e^{il\xi}, \quad (5.37)$$

and

$$\hat{r}_{odd}(\xi/2) = \sum_l r_{2l+1} e^{i(2l+1)\xi/2}. \quad (5.38)$$

Noticing that

$$2\hat{r}_{even}(\xi/2) = \hat{r}(\xi/2) + \hat{r}(\xi/2 + \pi) \quad (5.39)$$

and

$$2\hat{r}_{odd}(\xi/2) = \hat{r}(\xi/2) - \hat{r}(\xi/2 + \pi), \quad (5.40)$$

and using (5.18), we obtain from (5.35)

$$\hat{r}(\xi) = \left[\hat{r}(\xi/2) + \hat{r}(\xi/2 + \pi) + \left(2|m_0(\xi/2)|^2 - 1 \right) (\hat{r}(\xi/2) - \hat{r}(\xi/2 + \pi)) \right]. \quad (5.41)$$

Finally, using (2.26) we arrive at

$$\hat{r}(\xi) = 2 \left(|m_0(\xi/2)|^2 \hat{r}(\xi/2) + |m_0(\xi/2 + \pi)|^2 \hat{r}(\xi/2 + \pi) \right). \quad (5.42)$$

Setting $\xi = 0$ in (5.42) we obtain $\hat{r}(0) = 2\hat{r}(0)$ and, thus, (5.34).

Uniqueness of the solution of (5.14) and (5.15) follows from the uniqueness of the representation of d/dx . Given the solution r_l of (5.14) and (5.15) we consider the operator T_j defined by these coefficients on the subspace \mathbf{V}_j and apply it to a sufficiently smooth function f . Since $r_l^j = 2^{-j} r_l$ (5.4), we have

$$(T_j f)(x) = \sum_{k \in \mathbf{Z}} \left(2^{-j} \sum_l r_l f_{j,k-l} \right) \varphi_{j,k}(x), \quad (5.43)$$

where

$$f_{j,k-l} = 2^{-j/2} \int_{-\infty}^{+\infty} f(x) \varphi(2^{-j}x - k + l) dx. \quad (5.44)$$

Rewriting (5.44) as

$$f_{j,k-l} = 2^{-j/2} \int_{-\infty}^{+\infty} f(x - 2^j l) \varphi(2^{-j}x - k) dx, \quad (5.45)$$

and expanding $f(x - 2^j l)$ in the Taylor series at the point x , we have

$$f_{j,k-l} = \int_{-\infty}^{+\infty} f(x) \varphi_{j,k}(x) dx - 2^j l \int_{-\infty}^{+\infty} f'(x) \varphi_{j,k}(x) dx + 2^{2j} \frac{l^2}{2} \int_{-\infty}^{+\infty} f''(\tilde{x}) \varphi_{j,k}(x) dx, \quad (5.46)$$

where $\tilde{x} = \tilde{x}(x, x - 2^j l)$ and $|\tilde{x} - x| \leq 2^j l$. Substituting (5.46) in (5.43) and using (5.34) and (5.15), we obtain

$$(T_j f)(x) = \sum_{k \in \mathbf{Z}} \left(\int_{-\infty}^{+\infty} f'(x) \varphi_{j,k}(x) dx \right) \varphi_{j,k}(x) + 2^j \sum_{k \in \mathbf{Z}} \left(\frac{1}{2} \sum_l r_l l^2 \int_{-\infty}^{+\infty} f''(\tilde{x}) \varphi_{j,k}(x) dx \right) \varphi_{j,k}(x). \quad (5.47)$$

It is clear that as $j \rightarrow -\infty$, operators T_j and d/dx coincide on smooth functions. Using standard arguments it is easy to prove that $T_{-\infty} = d/dx$ and hence, the solution to (5.14) and (5.15) is unique. The relation (5.17) follows now from (5.13).

Remark 2. We note that expressions (5.9) and (5.10) for α_l and β_l ($\gamma_l = -\beta_{-l}$) may be simplified by changing the order of summation in (5.9) and (5.10) and introducing the correlation coefficients $2 \sum_{i=0}^{L-1-n} g_i h_{i+n}$, $2 \sum_{i=0}^{L-1-n} h_i g_{i+n}$ and $2 \sum_{i=0}^{L-1-n} g_i g_{i+n}$. The expression for α_l is especially simple, $\alpha_l = 4r_{2l} - r_l$.

Examples. For the examples we will use Daubechies' wavelets constructed in [19]. First, let us compute the coefficients a_{2k-1} , $k = 1, \dots, M$, where M is the number of vanishing moments and $L = 2M$. Using relation (4.22) of [19],

$$|m_0(\xi)|^2 = 1 - \frac{(2M-1)!}{[(M-1)!]^2 2^{2M-1}} \int_0^\xi \sin^{2M-1} \xi d\xi, \quad (5.48)$$

we find, by computing $\int_0^\xi \sin^{2M-1} \xi d\xi$, that

$$|m_0(\xi)|^2 = \frac{1}{2} + \frac{1}{2} C_M \sum_{m=1}^M \frac{(-1)^{m-1} \cos(2m-1)\xi}{(M-m)!(M+m-1)!(2m-1)}, \quad (5.49)$$

where

$$C_M = \left[\frac{(2M-1)!}{(M-1)! 4^{M-1}} \right]^2. \quad (5.50)$$

Thus, by comparing (5.49) and (5.18), we have

$$a_{2m-1} = \frac{(-1)^{m-1} C_M}{(M-m)!(M+m-1)!(2m-1)}, \quad \text{where } m = 1, \dots, M. \quad (5.51)$$

We note that by virtue of being solutions of a linear system with rational coefficients (a_{2m-1} in (5.51) are rational by construction), the coefficients r_l are rational numbers. The coefficients r_l are the same for all Daubechies' wavelets with a fixed number of vanishing moments M , while there are several wavelet bases for a given M depending on the choice of the roots of polynomials in the construction described in [19].

Solving equations of Proposition 1, we present the results for Daubechies' wavelets with $M = 2, 3, 4, 5, 6$.

1. $M = 2$

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

and

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

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

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

3. $M = 4$

$$a_1 = \frac{1225}{1024}, \quad a_3 = -\frac{245}{1024}, \quad a_5 = \frac{49}{1024}, \quad a_7 = -\frac{5}{1024},$$

and

$$\begin{aligned} r_1 &= -\frac{39296}{49553}, & r_2 &= \frac{76113}{396424}, & r_3 &= -\frac{1664}{49553}, \\ r_4 &= \frac{2645}{1189272}, & r_5 &= \frac{128}{743295}, & r_6 &= -\frac{1}{1189272}. \end{aligned}$$

4. $M = 5$

$$a_1 = \frac{19845}{16384}, \quad a_3 = -\frac{2205}{8192}, \quad a_5 = \frac{567}{8192}, \quad a_7 = -\frac{405}{32768}, \quad a_9 = \frac{35}{32768},$$

and

$$\begin{aligned} r_1 &= -\frac{957310976}{1159104017}, & r_2 &= \frac{265226398}{1159104017}, & r_3 &= -\frac{735232}{13780629}, \\ r_4 &= \frac{17297069}{2318208034}, & r_5 &= -\frac{1386496}{5795520085}, & r_6 &= -\frac{563818}{10431936153}, \\ r_7 &= -\frac{2048}{8113728119}, & r_8 &= -\frac{5}{18545664272}. \end{aligned}$$

5. $M = 6$

$$a_1 = \frac{160083}{131072}, \quad a_3 = -\frac{38115}{131072}, \quad a_5 = \frac{22869}{262144},$$

$$a_7 = -\frac{5445}{262144}, \quad a_9 = \frac{847}{262144}, \quad a_{11} = -\frac{63}{262144},$$

and

$$r_1 = \frac{3986930636128256}{4689752620280145}, \quad r_2 = \frac{4850197389074509}{18759010481120580}, \quad r_3 = \frac{1019185340268544}{14069257860840435},$$

$$r_4 = \frac{136429697045009}{9379505240560290}, \quad r_5 = \frac{7449960660992}{4689752620280145}, \quad r_6 = \frac{483632604097}{112554062886723480},$$

$$r_7 = \frac{78962327552}{6565653668392203}, \quad r_8 = \frac{31567002859}{75036041924482320}, \quad r_9 = \frac{2719744}{937950524056029},$$

$$r_{10} = \frac{1743}{2501201397482744}.$$

Coefficients for $M = 5$ and $M = 6$ can be compared with the corresponding output from the following iterative algorithm.

Iterative algorithm for computing the coefficients r_l .

As a way of solving equations (5.14) and (5.15) we may also use an iterative algorithm. We start with $r_{-1} = 0.5$ and $r_1 = -0.5$ and iterate using (5.14) to recompute r_l . It is easy to verify using (5.42) that (5.15) and (5.17) are satisfied due to the choice of initialization. The following 1 for Daubechies' wavelets with $M = 5, 6, 7, 8, 9$ was computed using this algorithm. It displays the coefficients $\{r_l\}_{l=1}^{L-2}$. We note, that $r_{-l} = -r_l$ and $r_0 = 0$.

V.2 The operators d^n/dx^n in the wavelet bases

Similar to the operator d/dx , the non-standard form of the operator d^n/dx^n is completely determined by its representation on the subspace \mathbf{V}_0 , i.e., by the coefficients

$$r_l^{(n)} = \int_{-\infty}^{+\infty} \varphi(x-l) \frac{d^n}{dx^n} \varphi(x) dx, \quad l \in \mathbf{Z}, \quad (5.52)$$

or, alternatively,

$$r_l^{(n)} = \int_{-\infty}^{+\infty} (-i\xi)^n |\hat{\varphi}(\xi)|^2 e^{-il\xi} d\xi. \quad (5.53)$$

if the integrals in (5.52) or (5.53) exist (see also Remark 3 below).

		Coefficients				Coefficients	
		l	r_l	l		l	r_l
$M = 5$	1	1	-0.82590601185015	$M = 8$	1	1	-0.88344604609097
		2	0.22882018706694		2	2	0.30325935147672
		3	-5.3352571932672E-02		3	3	-0.10636406828947
		4	7.4613963657755E-03		4	4	3.1290147839488E-02
		5	-2.3923582002393E-04		5	5	-6.9583791164537E-03
		6	-5.4047301644748E-05		6	6	1.0315302133757E-03
		7	-2.5241171135682E-07		7	7	-7.6677069083796E-05
		8	-2.6960479423517E-10		8	8	-2.4519921109537E-07
$M = 6$	1	1	-0.85013666155592	9	9	-3.9938104563894E-08	
		2	0.25855294414146	10	10	7.2079482385949E-08	
		3	-7.2440589997659E-02	11	11	9.6971849256415E-10	
		4	1.4545511041994E-02	12	12	7.2522069166503E-13	
		5	-1.5885615434757E-03	13	13	-1.2400785360984E-14	
		6	4.2968915709948E-06	14	14	1.5854647516841E-19	
		7	1.2026575195723E-05	$M = 9$	1	1	-0.89531640583699
		8	4.2069120451167E-07		2	2	0.32031206224855
		9	-2.8996668057051E-09		3	3	-0.12095364936000
		10	6.9686511520083E-13		4	4	3.9952721886694E-02
			5		5	-1.0616930669821E-02	
			6		6	2.1034028106558E-03	
$M = 7$	1	1	-0.86874391452377	7	7	-2.7812077649932E-04	
		2	0.28296509452594	8	8	1.9620437763642E-05	
		3	-9.0189066217795E-02	9	9	-4.8782468879634E-07	
		4	2.2687411014648E-02	10	10	1.0361220591478E-07	
		5	-3.8814546576295E-03	11	11	-1.5966864798639E-08	
		6	3.3734404776409E-04	12	12	-8.1374108294110E-10	
		7	4.2363946800701E-06	13	13	-5.4025197533630E-13	
		8	-1.6501679210868E-06	14	14	-4.7814005916812E-14	
		9	-2.1871130331900E-07	15	15	-1.6187880013009E-18	
		10	4.1830548203747E-10	16	16	-4.8507474310747E-24	
		11	-1.2035273999989E-11				
		12	-6.6283900594600E-16				

Table 1: The coefficients $\{r_l\}_{l=1}^{l=L-2}$ for Daubechies' wavelets, where $L = 2M$ and $M = 5, \dots, 9$.

Proposition V.2 1. If the integrals in (5.52) or (5.53) exist, then the coefficients $r_l^{(n)}$, $l \in \mathbf{Z}$ satisfy the following system of linear algebraic equations

$$r_l^{(n)} = 2^n \left[r_{2l} + \frac{1}{2} \sum_{k=1}^{L/2} a_{2k-1} (r_{2l-2k+1}^{(n)} + r_{2l+2k-1}^{(n)}) \right], \quad (5.54)$$

and

$$\sum_l l^n r_l^{(n)} = (-1)^n n!, \quad (5.55)$$

where a_{2k-1} are given in (5.19).

2. Let $M \geq (n+1)/2$, where M is the number of vanishing moments in (2.16). If the integrals in (5.52) or (5.53) exist, then the equations (5.54) and (5.55) have a unique solution with a finite number of non-zero coefficients $r_l^{(n)}$, namely, $r_l^{(n)} \neq 0$ for $-L+2 \leq l \leq L-2$. Also, for even n

$$r_l^{(n)} = r_{-l}^{(n)}, \quad (5.56)$$

$$\sum_l l^{2\tilde{n}} r_l^{(n)} = 0, \quad \tilde{n} = 1, \dots, n/2 - 1, \quad (5.57)$$

and

$$\sum_l r_l^{(n)} = 0, \quad (5.58)$$

and for odd n

$$r_l^{(n)} = -r_{-l}^{(n)}, \quad (5.59)$$

$$\sum_l l^{2\tilde{n}-1} r_l^{(n)} = 0, \quad \tilde{n} = 1, \dots, (n-1)/2. \quad (5.60)$$

The proof of Proposition 2 is analogous to that of Proposition 1.

Remark 3. If $M \geq (n+1)/2$, then the solution of the linear system in Proposition 2 may exist while integrals (5.52) and (5.53) are not absolutely convergent.

A case in point is the Daubechies' wavelet with $M = 2$. The representation of the first derivative in this basis is described in the previous Section. The equations (5.54) and (5.55) do not have a solution for the second derivative $n = 2$. However, the system of equations (5.54) and (5.55) has a solution for the third derivative $n = 3$. We have

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

and

$$r_{-2} = -\frac{1}{2}, \quad r_{-1} = 1, \quad r_0 = 0, \quad r_1 = -1, \quad r_2 = \frac{1}{2}.$$

We note that among the wavelets with $L = 4$, the wavelets with two vanishing moments $M = 2$ do not have the best Hölder exponent (see [21]), but the representation of the third derivative exists only if the number of vanishing moments $M = 2$.

The equations for computing the coefficients $r_l^{(n)}$ may be viewed as an eigenvalue problem. Let us derive the equation corresponding to (5.42) for d^n/dx^n directly from (5.53). We rewrite (5.53) as

$$r_l^{(n)} = \int_0^{2\pi} \sum_{k \in \mathbf{Z}} |\hat{\varphi}(\xi + 2\pi k)|^2 i^n (\xi + 2\pi k)^n e^{-il\xi} d\xi. \quad (5.61)$$

Therefore,

$$\hat{r}(\xi) = \sum_{k \in \mathbf{Z}} |\hat{\varphi}(\xi + 2\pi k)|^2 i^n (\xi + 2\pi k)^n, \quad (5.62)$$

where

$$\hat{r}(\xi) = \sum_l r_l^{(n)} e^{il\xi}. \quad (5.63)$$

Substituting the relation

$$\hat{\varphi}(\xi) = m_0(\xi/2)\hat{\varphi}(\xi/2), \quad (5.64)$$

into the right hand side of (5.62), and summing over even and odd indices in (5.62) separately, we arrive at

$$\hat{r}(\xi) = 2^n \left(|m_0(\xi/2)|^2 \hat{r}(\xi/2) + |m_0(\xi/2 + \pi)|^2 \hat{r}(\xi/2 + \pi) \right). \quad (5.65)$$

Let us consider the operator M_0 on 2π -periodic functions

$$(M_0 f)(\xi) = |m_0(\xi/2)|^2 f(\xi/2) + |m_0(\xi/2 + \pi)|^2 f(\xi/2 + \pi). \quad (5.66)$$

We rewrite (5.65) as

$$M_0 \hat{r} = 2^{-n} \hat{r}, \quad (5.67)$$

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 (5.67) and vice versa. (The operator M_0 is also introduced in [14] and [26], where the problem (5.67) with the eigenvalue 1 is considered).

V.3 The explicit preconditioning of the derivative operators

While theoretically it is well understood that the derivative operators (or, more generally, operators with homogeneous symbols) have an explicit diagonal preconditioner

in wavelet bases, the numerical evidence illustrating this fact is of interest, since it represents one of the advantages of computing in the wavelet bases.

For operators with a homogeneous symbol the bound on the condition number depends on the particular choice of the wavelet basis (by the condition number we understand the ratio of the largest singular value to the smallest singular value above the threshold of accuracy). After applying such a preconditioner, the condition number κ_p of the operator is uniformly bounded with respect to the size of the matrix. We recall that the condition number controls the rate of convergence of a number of iterative algorithms; for example the number of iterations of the conjugate gradient method is $O(\sqrt{\kappa_p})$.

We present here two tables, 2 and 3, illustrating such preconditioning applied to the standard form of the second derivative. In the following examples the standard form of the periodized second derivative D_2 of size $N \times N$, where $N = 2^n$, is preconditioned by the diagonal matrix P ,

$$D_2^p = PD_2P.$$

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 $P_{NN} = 2^n$.

In the tables we compare the original condition number κ of D_2 and κ_p of D_2^p .

A remark the boundary value problems.

The wavelet bases are not well adapted for the boundary value problems (unless the problem is reduced to solving equations on the boundary). In one dimension, there is a simple construction of the wavelet basis for an interval due to Y. Meyer. Unfortunately, it does not generalize to higher dimensions.

There is, however, a simple method for setting the boundary conditions. This method is based on an observation, that the value of a function at a given point may be obtained by considering a linear combination of the basis functions which contain this point inside their supports. For a fixed point, the number of such functions is proportional to the number of scales, $\log_2 N$. In two-dimensional problems, the number of basis functions involved in the determination of the boundary conditions is proportional to the length of the boundary times the number of scales. Thus, the boundary conditions may be imposed as *sparse* linear constraints.

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

VI The convolution operators in wavelet bases

In this Section we consider the computation of the non-standard form of convolution operators. For convolution operators the quadrature formulas for representing the kernel on \mathbf{V}_0 are of the simplest form due to the fact that the moments of the autocorrelation of the scaling function φ vanish. Moreover, by combining the asymptotics of the wavelet coefficients of the operator with the system of linear algebraic equations (similar to those in Section V) we arrive at an effective method for computing these coefficients [5]. This method is especially simple if the symbol of the operator is homogeneous of some degree.

Let us assume that the matrix $t_{i-l}^{(j-1)}$ ($i, l \in \mathbf{Z}$) represents the operator $P_{j-1}TP_{j-1}$ on the subspace \mathbf{V}_{j-1} . To compute the representation of P_jTP_j , we have the following formula (3.26) of [7]

$$t_l^{(j)} = \sum_{k=0}^{L-1} \sum_{m=0}^{L-1} h_k h_m t_{2l+k-m}^{(j-1)}. \quad (6.1)$$

It easily reduces to

$$t_l^{(j)} = t_{2l}^{(j-1)} + \frac{1}{2} \sum_{k=0}^{L/2} a_{2k-1} (t_{2l-2k+1}^{(j-1)} + t_{2l+2k-1}^{(j-1)}). \quad (6.2)$$

where the coefficients a_{2k-1} are given in (5.19).

We also have

$$t_l^{(j)} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x-y) \varphi_{j,0}(y) \varphi_{j,l}(x) dx dy, \quad (6.3)$$

and by changing the order of integration obtain

$$t_l^{(j)} = 2^j \int_{-\infty}^{+\infty} K(2^j(l-y)) \Phi(y) dy, \quad (6.4)$$

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

$$\Phi(y) = \int_{-\infty}^{+\infty} \varphi(x) \varphi(x-y) dx. \quad (6.5)$$

Let us verify that

$$\int_{-\infty}^{+\infty} \Phi(y) dy = 1 \quad (6.6)$$

and

$$\mathcal{M}_{\Phi}^m = \int_{-\infty}^{+\infty} y^m \Phi(y) dy = 0, \quad \text{for } 1 \leq m \leq 2M-1. \quad (6.7)$$

Clearly, we have

$$\mathcal{M}_{\Phi}^m = \left[\left(\frac{1}{i} \partial_{\xi} \right)^m |\hat{\varphi}(\xi)|^2 \right]_{\xi=0}. \quad (6.8)$$

Using (6.8) and the identity $\hat{\varphi}(\xi) = \hat{\varphi}(\xi/2)m_0(\xi/2)$ (see [19]), it follows from (5.25) that (6.7) holds.

Since the moments of the function Φ vanish, (6.7), equation (6.4) leads to a one-point quadrature formula for computing the representation of convolution operators on the finest scale for all compactly supported wavelets. This formula is obtained in exactly the same manner as for the special choice of the wavelet basis described in [7] (eqns. 3.8-3.12), where the shifted moments of the function φ vanish; we refer to this paper for the details.

Here we introduce a different approach, which consists in solving the system of linear algebraic equations (6.2) subject to asymptotic conditions. This method is especially simple if the symbol of the operator is homogeneous of some degree since in this case the operator is completely defined by its representation on \mathbf{V}_0 .

Let us consider two examples of such operators, the Hilbert transform and the operator of fractional differentiation (or anti-differentiation).

VI.1 The Hilbert Transform

We apply our method to the computation of the non-standard form of the Hilbert transform

$$g(x) = (\mathcal{H}f)(y) = \frac{1}{\pi} \text{p.v.} \int_{-\infty}^{\infty} \frac{f(s)}{s-x} ds, \quad (6.9)$$

where p.v. denotes a principal value at $s = x$.

The representation of \mathcal{H} on \mathbf{V}_0 is defined by the coefficients

$$r_l = \int_{-\infty}^{\infty} \varphi(x-l) (\mathcal{H}\varphi)(x) dx, \quad l \in \mathbf{Z}, \quad (6.10)$$

which, in turn, completely define all other coefficients of the non-standard form. Namely, $\mathcal{H} = \{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}}$, $A_j = A_0$, $B_j = B_0$, and $\Gamma_j = \Gamma_0$, where matrix elements α_{i-l} , β_{i-l} , and γ_{i-l} of A_0 , B_0 , and Γ_0 are computed from the coefficients r_l ,

$$\alpha_i = \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k g_{k'} r_{2i+k-k'}, \quad (6.11)$$

$$\beta_i = \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k h_{k'} r_{2i+k-k'}, \quad (6.12)$$

and

$$\gamma_i = \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} h_k g_{k'} r_{2i+k-k'}. \quad (6.13)$$

	Coefficients		Coefficients	
	l	r_l	l	r_l
$M = 6$	1	-0.588303698	9	-0.035367761
	2	-0.077576414	10	-0.031830988
	3	-0.128743695	11	-0.028937262
	4	-0.075063628	12	-0.026525823
	5	-0.064168018	13	-0.024485376
	6	-0.053041366	14	-0.022736420
	7	-0.045470650	15	-0.021220659
	8	-0.039788641	16	-0.019894368

Table 4: The coefficients r_l , $l = 1, \dots, 16$ of the Hilbert transform for Daubechies' wavelet with six vanishing moments.

The coefficients r_l , $l \in \mathbf{Z}$ in (6.10) satisfy the following system of linear algebraic equations

$$r_l = r_{2l} + \frac{1}{2} \sum_{k=1}^{L/2} a_{2k-1} (r_{2l-2k+1} + r_{2l+2k-1}), \quad (6.14)$$

where the coefficients a_{2k-1} are given in (5.19). Using (6.4), (6.6) and (6.7) we obtain the asymptotics of r_l for large l ,

$$r_l = -\frac{1}{\pi l} + O\left(\frac{1}{l^{2M}}\right). \quad (6.15)$$

By rewriting (6.10) in terms of $\hat{\varphi}(\xi)$,

$$r_l = -2 \int_0^\infty |\hat{\varphi}(\xi)|^2 \sin(l\xi) d\xi. \quad (6.16)$$

we obtain $r_l = -r_{-l}$ and set $r_0 = 0$. We note that the coefficient r_0 cannot be determined from equations (6.14) and (6.15).

Solving (6.14) with the asymptotic condition (6.15), we compute the coefficients r_l , $l \neq 0$ with any prescribed accuracy.

Example.

We compute the coefficients r_l of the Hilbert transform for Daubechies' wavelets with six vanishing moments with accuracy 10^{-7} . The coefficients for $l > 16$ are obtained using asymptotics (6.15). (We note that $r_{-l} = -r_l$ and $r_0 = 0$).

VI.2 The fractional derivatives

We use the following definition of fractional derivatives

$$(\partial_x^\alpha f)(x) = \int_{-\infty}^{+\infty} \frac{(x-y)_+^{-\alpha-1}}{\Gamma(-\alpha)} f(y) dy, \quad (6.17)$$

where we consider $\alpha \neq 1, 2, \dots$. If $\alpha < 0$, then (6.17) defines fractional anti-derivatives.

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

$$r_l = \int_{-\infty}^{+\infty} \varphi(x-l) (\partial_x^\alpha \varphi)(x) dx, \quad l \in \mathbf{Z}, \quad (6.18)$$

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 ,

$$\alpha_i = 2^\alpha \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k g_{k'} r_{2i+k-k'}, \quad (6.19)$$

$$\beta_i = 2^\alpha \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} g_k h_{k'} r_{2i+k-k'}, \quad (6.20)$$

and

$$\gamma_i = 2^\alpha \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} h_k g_{k'} r_{2i+k-k'}. \quad (6.21)$$

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

$$r_l = 2^\alpha \left[r_{2l} + \frac{1}{2} \sum_{k=1}^{L/2} a_{2k-1} (r_{2l-2k+1} + r_{2l+2k-1}) \right], \quad (6.22)$$

where the coefficients a_{2k-1} are given in (5.19). Using (6.4), (6.6) and (6.7) we obtain the asymptotics of r_l for large l ,

$$r_l = \frac{1}{\Gamma(-\alpha)} \frac{1}{l^{1+\alpha}} + O\left(\frac{1}{l^{1+\alpha+2M}}\right) \quad \text{for } l > 0, \quad (6.23)$$

$$r_l = 0 \quad \text{for } l < 0. \quad (6.24)$$

Example.

	l	Coefficients r_l	l	Coefficients r_l
$M = 6$	-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 5: The coefficients $\{r_l\}_l$, $l = -7, \dots, 14$ of the fractional derivative $\alpha = 0.5$ for Daubechies' wavelet with six vanishing moments.

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

$$r_l = -\frac{1}{2\sqrt{\pi}} \frac{1}{l^{1+\frac{1}{2}}} + O\left(\frac{1}{l^{3+\frac{1}{2}}}\right) \quad \text{for } l > 0, \quad (6.25)$$

$$r_l = 0 \quad \text{for } l < 0. \quad (6.26)$$

VII Multiplication of operators in wavelet bases

VII.1 Multiplication of matrices in the standard form

The multiplication of matrices of Calderón-Zygmund and pseudo-differential operators in the standard form requires at most $O(N \log^2 N)$ operations. In addition, it is possible to control the width of the “finger” bands by setting to zero all the entries in the product below a threshold of accuracy.

Let us compute $T = \tilde{T}\hat{T}$, where

$$\tilde{T} = \{\tilde{A}_j, \{\tilde{B}_j^{j'}\}_{j'=j+1}^{j'=n}, \{\tilde{\Gamma}_j^{j'}\}_{j'=j+1}^{j'=n}, \tilde{B}_j^{n+1}, \tilde{\Gamma}_j^{n+1}, \tilde{T}_n\}_{j=1, \dots, n} \quad (7.1)$$

and

$$\hat{T} = \{\hat{A}_j, \{\hat{B}_j^{j'}\}_{j'=j+1}^{j'=n}, \{\hat{\Gamma}_j^{j'}\}_{j'=j+1}^{j'=n}, \hat{B}_j^{n+1}, \hat{\Gamma}_j^{n+1}, \hat{T}_n\}_{j=1, \dots, n}. \quad (7.2)$$

Since the standard form is a representation in an orthonormal basis (which is the tensor product of one-dimensional bases), the result of the multiplication of two standard forms is also a standard form in the same basis. Thus, the product T must have a representation

$$T = \{A_j, \{B_j^{j'}\}_{j'=j+1}^{j'=n}, \{\Gamma_j^{j'}\}_{j'=j+1}^{j'=n}, B_j^{n+1}, \Gamma_j^{n+1}, T_n\}_{j=1, \dots, n}. \quad (7.3)$$

Due to the block structure of the corresponding matrix, each element of (7.3) is obtained as a sum of products of the corresponding blocks of \tilde{T} and \hat{T} . For example,

$$\Gamma_1^2 = \tilde{\Gamma}_1^2 \hat{A}_1 + \tilde{A}_2 \hat{\Gamma}_1^2 + \sum_{j'=3}^{j'=n+1} \tilde{B}_2^{j'} \hat{\Gamma}_1^{j'}. \quad (7.4)$$

If the operators \tilde{T} and \hat{T} are Calderón-Zygmund or pseudo-differential operators, then all the blocks of (7.1) and (7.2) (except for \tilde{T}_n and \hat{T}_n) are banded and it is clear that Γ_1^2 is banded. This example is generic for all operators in (7.3) except for $B_j^{n+1}, \Gamma_j^{n+1}$, ($j = 1, \dots, n$) and T_n . The latter are dense due to the terms involving \tilde{T}_n and \hat{T}_n . It is easy now to estimate the number of operations necessary to compute T . It takes no more than $O(N \log^2 N)$ operations to obtain T , where $N = 2^n$.

If, in addition, when the scales j and j' are well separated, the operators $B_j^{j'}, \Gamma_j^{j'}$ can be neglected for a given accuracy (as in the case of pseudo-differential operators), then the number of operations reduces asymptotically to $O(N)$.

We note, that we may set to zero all the entries of T below the threshold of accuracy and, thus, prevent the widening of the bands in the product. On denoting \tilde{T}_ϵ and \hat{T}_ϵ the approximations to \tilde{T} and \hat{T} obtained by setting all entries that are less than ϵ to zero, and assuming (without a loss of generality) $\|\tilde{T}\| = \|\hat{T}\| = 1$, we obtain using the result of [7]

$$\|\tilde{T} - \tilde{T}_\epsilon\| \leq \epsilon, \quad \|\hat{T} - \hat{T}_\epsilon\| \leq \epsilon, \quad (7.5)$$

and, therefore,

$$\|\tilde{T}\hat{T} - (\tilde{T}_\epsilon\hat{T}_\epsilon)_\epsilon\| \leq \epsilon + \epsilon(1 + \epsilon) + \epsilon(1 + \epsilon)^2. \quad (7.6)$$

The right hand side of (7.6) is dominated by 3ϵ . For example, if we compute T^4 then we might lose one significant digit.

VII.2 Multiplication of matrices in the non-standard form

We will now outline an algorithm for the multiplication of the operators in the non-standard form. This new algorithm is remarkable in the way it decouples the scales in the process of multiplication. Let \hat{T} and \tilde{T} be two operators

$$\hat{T}, \tilde{T} : \mathbf{L}^2(\mathbf{R}) \rightarrow \mathbf{L}^2(\mathbf{R}). \quad (7.7)$$

Given the non-standard forms of \hat{T} and \tilde{T} , $\{\hat{A}_j, \hat{B}_j, \hat{\Gamma}_j\}_{j \in \mathbf{Z}}$ and $\{\tilde{A}_j, \tilde{B}_j, \tilde{\Gamma}_j\}_{j \in \mathbf{Z}}$, we compute the non-standard form $\{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}}$ of $T = \hat{T}\tilde{T}$.

We recall that the operators of the non-standard form are defined using the projection operators P_j (on the subspace \mathbf{V}_j), and $Q_j = P_{j-1} - P_j$ (on the subspace \mathbf{W}_j), $j \in \mathbf{Z}$, as follows $\hat{A}_j = Q_j\hat{T}Q_j$, $\hat{B}_j = Q_j\hat{T}P_j$, $\hat{\Gamma}_j = P_j\hat{T}Q_j$, $\tilde{A}_j = Q_j\tilde{T}Q_j$, $\tilde{B}_j = Q_j\tilde{T}P_j$ and $\tilde{\Gamma}_j = P_j\tilde{T}Q_j$.

We will also need the operators $\hat{T}_j = P_j\hat{T}P_j$ and $\tilde{T}_j = P_j\tilde{T}P_j$, $j \in \mathbf{Z}$. These operators may not be sparse but the algorithm requires only a band around the diagonal of their matrices.

In order to decouple the scales, we write a “telescopic” series,

$$\hat{T}_0\tilde{T}_0 - \hat{T}_n\tilde{T}_n = \sum_{j=1}^{j=n} \left[(P_{j-1}\hat{T}P_{j-1})(P_{j-1}\tilde{T}P_{j-1}) - (P_j\hat{T}P_j)(P_j\tilde{T}P_j) \right] \quad (7.8)$$

Since $P_{j-1} = P_j + Q_j$, we obtain from (7.8)

$$\begin{aligned} \hat{T}_0\tilde{T}_0 - \hat{T}_n\tilde{T}_n = \sum_{j=1}^{j=n} [& (P_j\hat{T}P_j)(P_j\tilde{T}Q_j) + (Q_j\hat{T}P_j)(P_j\tilde{T}P_j) + \\ & (Q_j\hat{T}P_j)(P_j\tilde{T}Q_j) + (P_j\hat{T}Q_j)(Q_j\tilde{T}P_j) + \\ & (P_j\hat{T}Q_j)(Q_j\tilde{T}Q_j) + (Q_j\hat{T}Q_j)(Q_j\tilde{T}P_j) + \\ & (Q_j\hat{T}Q_j)(Q_j\tilde{T}Q_j)], \end{aligned} \quad (7.9)$$

or

$$\begin{aligned} \hat{T}_0\tilde{T}_0 - \hat{T}_n\tilde{T}_n = \sum_{j=1}^{j=n} [& (\hat{A}_j\tilde{A}_j + \hat{B}_j\tilde{\Gamma}_j) + (\hat{B}_j\tilde{T}_j + \hat{A}_j\tilde{B}_j) + \\ & (\hat{T}_j\tilde{\Gamma}_j + \hat{\Gamma}_j\tilde{A}_j) + \hat{\Gamma}_j\tilde{B}_j] \end{aligned} \quad (7.10)$$

Finally, we rewrite (7.10) as a sum of two terms,

$$\hat{T}_0 \tilde{T}_0 = F + R, \quad (7.11)$$

where

$$F = \sum_{j=1}^{j=n} \left[(\hat{A}_j \tilde{A}_j + \hat{B}_j \tilde{\Gamma}_j) + (\hat{B}_j \tilde{T}_j + \hat{A}_j \tilde{B}_j) + (\hat{T}_j \tilde{\Gamma}_j + \hat{\Gamma}_j \tilde{A}_j) \right] \quad (7.12)$$

and

$$R = \hat{T}_n \tilde{T}_n + \sum_{j=1}^{j=n} P_j \hat{\Gamma}_j \tilde{B}_j P_j \quad (7.13)$$

The operators in the sum (7.12) are acting on following the subspaces,

$$\hat{A}_j \tilde{A}_j + \hat{B}_j \tilde{\Gamma}_j : \mathbf{W}_j \rightarrow \mathbf{W}_j, \quad (7.14)$$

$$\hat{B}_j \tilde{T}_j + \hat{A}_j \tilde{B}_j : \mathbf{V}_j \rightarrow \mathbf{W}_j, \quad (7.15)$$

$$\hat{T}_j \tilde{\Gamma}_j + \hat{\Gamma}_j \tilde{A}_j : \mathbf{W}_j \rightarrow \mathbf{V}_j, \quad (7.16)$$

and the operators in the sum (7.13),

$$\hat{\Gamma}_j \tilde{B}_j : \mathbf{V}_j \rightarrow \mathbf{V}_j. \quad (7.17)$$

where $j = 1, \dots, n$, and

$$\hat{T}_n \tilde{T}_n : \mathbf{V}_n \rightarrow \mathbf{V}_n. \quad (7.18)$$

We now describe a fast $O(N)$ algorithm for the multiplication of the non-standard forms of Calderón-Zygmund and pseudo-differential operators.

1. We compute the operators (7.14)-(7.18). This involves the multiplication of banded matrices. There is no problem in obtaining (7.14) or (7.17), since all matrices in the products are banded. *We are yet to explain that in computing (7.15) and (7.16) we need to use only the band around the diagonal of \hat{T}_j and \tilde{T}_j , $j = 1, \dots, n$ (see below).*

Since the number of operations to compute (7.14)-(7.18) on a given scale j is half of that on the previous finer scale $j - 1$, the total number of operations is twice that on the finest scale $j = 1$. Thus, the total number of operations at this step is proportional to N .

2. We compute the non-standard form of all the operators in (7.17). First, we observe that $\hat{\Gamma}_j \tilde{B}_j$, $j = 1, \dots, n$ are banded. Starting from the finest scale $j = 1$, we expand $\hat{\Gamma}_j \tilde{B}_j$ to obtain \bar{A}_{j+1} , \bar{B}_{j+1} , $\bar{\Gamma}_{j+1}$ and \bar{T}_{j+1} . We then add \bar{T}_{j+1} to $\hat{\Gamma}_{j+1} \tilde{B}_{j+1}$ and expand the sum of the two, etc. As a result, we obtain \bar{A}_j , \bar{B}_j , $\bar{\Gamma}_j$, $j = 2, \dots, n$ and \bar{T}_n . Since at all scales we are expanding a banded matrix, and the number

of operations is halved each time we go to the sparser scale, the total number of operations at this step is proportional to N .

The resulting operators $\bar{A}_j, \bar{B}_j, \bar{\Gamma}_j, j = 2, \dots, n$, and \bar{T}_n are acting on the subspaces,

$$\bar{A}_j : \mathbf{W}_j \rightarrow \mathbf{W}_j, \quad (7.19)$$

$$\bar{B}_j : \mathbf{V}_j \rightarrow \mathbf{W}_j, \quad (7.20)$$

$$\bar{\Gamma}_j : \mathbf{W}_j \rightarrow \mathbf{V}_j, \quad (7.21)$$

and

$$\bar{T}_n : \mathbf{V}_n \rightarrow \mathbf{V}_n, \quad (7.22)$$

3. At this step we add the corresponding operators computed at step 1 and step 2, to obtain the non-standard form of the operator $T = \hat{T}\tilde{T}$,

$$A_j = \bar{A}_j + \hat{A}_j\tilde{A}_j + \hat{B}_j\tilde{\Gamma}_j, \quad (7.23)$$

$$B_j = \bar{B}_j + \hat{B}_j\tilde{T}_j + \hat{A}_j\tilde{B}_j, \quad (7.24)$$

$$\Gamma_j = \bar{\Gamma}_j + \hat{T}_j\tilde{\Gamma}_j + \hat{\Gamma}_j\tilde{A}_j, \quad (7.25)$$

and

$$T_n = \bar{T}_n + \hat{T}_n\tilde{T}_n. \quad (7.26)$$

This step obviously requires $O(N)$ operations.

If the product of two operators satisfies the estimates of Section IV, then the operators $\{A_j, B_j, \Gamma_j\}_{j \in \mathbf{Z}}$ of the non-standard form of $T = \hat{T}\tilde{T}$ are banded (for a given precision). Examining (7.23)-(7.25), we find that in (7.24) and (7.25) there is only one term that may potentially be dense. However, if the multiplicands and the product are banded then all terms are banded, and we conclude that $\hat{B}_j\tilde{T}_j$ and $\hat{T}_j\tilde{\Gamma}_j$ must be banded. Thus, we need only a banded version of the operators \hat{T}_j and \tilde{T}_j . The banded versions of \hat{T}_j and \tilde{T}_j are computed in the process of constructing the non-standard form and, therefore, we only need to store the results of these computations.

These is an alternative (direct) argument to show that $\hat{B}_j\tilde{T}_j$ and $\hat{T}_j\tilde{\Gamma}_j$ are banded. It requires a proof that the first several moments of rows of \hat{B}_j and columns of $\tilde{\Gamma}_j$ are negligible, which may be found in [29] for pseudo-differential operators.

VIII Fast iterative algorithms in wavelet bases

The fast multiplication algorithms of Section VII give a second life to a number of iterative algorithms.

VIII.1 An iterative algorithm for computing the generalized inverse

In order to construct the generalized inverse A^\dagger of the matrix A , we use the following result [36]:

Proposition VIII.1 *Consider the sequence of matrices X_k*

$$X_{k+1} = 2X_k - X_k A X_k \tag{8.1}$$

with

$$X_0 = \alpha A^*, \tag{8.2}$$

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

When this result is combined with the fast multiplication algorithms of Section VII, we obtain an algorithm for constructing the generalized inverse in the standard form in at most $O(N \log^2 N \log R)$ operations and in the non-standard form in $O(N \log R)$, 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 appear in [6]. Throughout the iteration (8.1)-(8.2), it is necessary to maintain the “finger” band structure of the standard form or the banded structure of the submatrices of the non-standard form of the matrices X_k . Hence, the standard and non-standard forms of both the operator and its generalized inverse must admit such structures. Since pseudo-differential operators of order zero form an algebra, the operators of this class satisfy this condition. Also, since any pseudo-differential operator may be multiplied by a compressible operator so that the product is a pseudo-differential operator of order zero, it is easy to see that the iteration (8.1)-(8.2) is applicable to all pseudo-differential operators.

Example.

The following table contains timings and accuracy comparison of the construction of the generalized inverse via the singular value decomposition (SVD), which is $O(N^3)$

procedure, and via the iteration (8.1)-(8.2) in the wavelet basis using Fast Wavelet Transform (FWT). The computations were performed on 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 $i, j = 1, \dots, N$. The accuracy threshold 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.)	

Let us describe several iterative algorithms indicating that numerical functional calculus with operators can be implemented efficiently (at least for pseudo-differential operators). Numerical results and relative performance of these algorithms will be reported separately.

VIII.2 An iterative algorithm for computing the projection operator on the null space.

Let us consider the following iteration

$$X_{k+1} = 2X_k - X_k^2 \tag{8.3}$$

with

$$X_0 = \alpha A^* A, \tag{8.4}$$

where A^* is the adjoint matrix and α is chosen so that the largest eigenvalue of $\alpha A^* A$ is less than two.

Then $I - X_k$ converges to P_{null} . This can be shown either directly or by combining an invariant representation for $P_{null} = I - A^*(AA^*)^\dagger A$ with the iteration (8.1)-(8.2) to compute the generalized inverse $(AA^*)^\dagger$. The fast multiplication algorithm makes the iteration (8.3)-(8.4) fast for a wide class of operators (with the same complexity as the algorithm for the generalized inverse). The important difference is, however, that (8.3)-(8.4) does not require compressibility of the inverse operator but only of the powers of the operator.

VIII.3 An iterative algorithm for computing a square root of an operator.

Let us describe an iteration to construct both $A^{1/2}$ and $A^{-1/2}$, where A is, for simplicity, a self-adjoint and non-negative definite operator. We consider the following iteration

$$Y_{l+1} = 2Y_l - Y_l X_l Y_l, \quad (8.5)$$

$$X_{l+1} = \frac{1}{2}(X_l + Y_l A), \quad (8.6)$$

with

$$\begin{aligned} Y_0 &= \frac{\alpha}{2}(A + I), \\ X_0 &= \frac{\alpha}{2}(A + I), \end{aligned} \quad (8.7)$$

where α is chosen so that the largest eigenvalue of $\frac{\alpha}{2}(A + I)$ is less than $\sqrt{2}$.

The sequence X_l converges to $A^{1/2}$ and Y_l to $A^{-1/2}$. By writing $A = V^* D V$, where D is a diagonal and V is a unitary, it is easy to verify that both X_l and Y_l can be written as $X_l = V^* P_l V$ and $Y_l = V^* Q_l V$, where P_l and Q_l are diagonal and

$$\begin{aligned} Q_{l+1} &= 2Q_l - Q_l P_l Q_l, \\ P_{l+1} &= \frac{1}{2}(P_l + Q_l D), \end{aligned} \quad (8.8)$$

with

$$\begin{aligned} Q_0 &= \frac{\alpha}{2}(D + I), \\ P_0 &= \frac{\alpha}{2}(D + I). \end{aligned} \quad (8.9)$$

Thus, the convergence need to be checked only for the scalar case, which we leave to the reader. If the operator A is a pseudo-differential operator, then the iteration (8.5) - (8.6) leads to a fast algorithm due to the same considerations as in the case of the generalized inverse in Section VIII.

VIII.4 Fast algorithms for computing the exponential, sine and cosine of a matrix

The exponential of a matrix (or an operator), as well as sine and cosine functions are among the first to be considered in any calculus of operators. As in the case of the generalized inverse, we use previously known algorithms (see e.g. [40]), which obtain a completely different complexity estimates when we use them in conjunction with the wavelet representations.

An algorithm for computing the exponential of a matrix is based on the identity

$$\exp(A) = \left[\exp(2^{-L}A) \right]^{2^L}. \quad (8.10)$$

First, $\exp(2^{-L}A)$ is computed by, for example, using the Taylor series. 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 $2^{-L}A$ is squared L times to obtain the result.

Similarly, sine and cosine of a matrix may be computed using the elementary double-angle formulas. On denoting

$$Y_l = \cos(2^{l-L}A) \quad (8.11)$$

$$X_l = \sin(2^{l-L}A), \quad (8.12)$$

we have for $l = 0, \dots, L-1$

$$Y_{l+1} = 2Y_l^2 - I \quad (8.13)$$

$$X_{l+1} = 2Y_lX_l, \quad (8.14)$$

where I is the identity. Again, we choose L so that the largest singular value of $2^{-L}A$ is less than one, compute the sine and cosine of $2^{-L}A$ using the Taylor series, and then use (8.13) and (8.14).

Ordinarily such algorithms require at least $O(N^3)$ operations, since a number of multiplications of dense matrices has to be performed [40]. The fast algorithm for the multiplication of matrices in the standard form reduces the complexity to no more than $O(N \log^2 N)$ operations, and the fast algorithm for the multiplication of matrices in the non-standard form to $O(N)$ operations (see Section VII).

To achieve such performance, it is necessary to maintain the “finger” band structure of the standard form or the banded structure of the submatrices of the non-standard form throughout the iteration. Whether it is possible to do depends on the particular operator and, usually, may be verified analytically.

Unlike the algorithm for the generalized inverse, the algorithms of this Remark are not self-correcting. Thus, it is necessary to maintain sufficient accuracy initially so as to obtain the desired accuracy after all the multiplications have been performed.

IX Computing $F(u)$ in the wavelet bases

In this Section we describe a fast, adaptive algorithm for computing $F(u)$, where F is an infinitely differentiable function and u is represented in a wavelet basis. An important example is $F(u) = u^2$. Our analytic results generalize theorems of J. M. Bony [10], [11], [9], [15] on the propagation of singularities of solutions of non-linear equations. Our numerical approach, however, is novel. We expect a wide range of applications of this algorithm.

IX.1 The algorithm for evaluating u^2

We start with an algorithm to compute $F(u) = u^2$. Let us project $u \in \mathbf{L}^2(\mathbf{R})$ on subspaces \mathbf{V}_j , $j \in \mathbf{Z}$, so that

$$u_j = P_j u, \quad u_j \in \mathbf{V}_j. \quad (9.1)$$

In order to decouple the scales, we write a “telescopic” series,

$$u_0^2 - u_n^2 = \sum_{j=1}^{j=n} [(P_{j-1}u)^2 - (P_ju)^2] = \sum_{j=1}^{j=n} (P_{j-1}u + P_ju)(P_{j-1}u - P_ju) \quad (9.2)$$

Using $P_{j-1} = P_j + Q_j$, we obtain

$$u_0^2 - u_n^2 = \sum_{j=1}^{j=n} (2P_ju + Q_ju)(Q_ju), \quad (9.3)$$

or

$$u_0^2 = 2 \sum_{j=1}^{j=n} (P_ju)(Q_ju) + \sum_{j=1}^{j=n} (Q_ju)(Q_ju) + u_n^2. \quad (9.4)$$

In (9.4) there is no interaction between different scales j and j' , $j \neq j'$.

For the numerical purposes, we need formulas (9.3) or (9.4) with a finite number of scales, though it is clear that by taking limits $j \rightarrow \infty$ and $j \rightarrow -\infty$ we have

$$u^2 = \sum_{j \in \mathbf{Z}} (2P_ju + Q_ju)(Q_ju), \quad (9.5)$$

which is essentially the para-product of J.M. Bony. In what follows, we will consider each term of (9.5) as a bilinear mapping

$$M_{VW}^j : \mathbf{V}_{j-1} \times \mathbf{W}_j \rightarrow \mathbf{L}^2(\mathbf{R}) = \mathbf{V}_j \bigoplus_{j' \leq j} \mathbf{W}_{j'}, \quad (9.6)$$

and show that, for a given precision, very few coefficients are needed to describe mapping (9.6) for all j .

Before proceeding further, let us consider an example of (9.4) in the Haar basis. We have the following explicit relations,

$$\begin{aligned}(\chi_k^j(x))^2 &= 2^{-j/2} \chi_k^j(x), \\(h_k^j(x))^2 &= 2^{-j/2} \chi_k^j(x), \\ \chi_k^j(x) h_k^j(x) &= 2^{-j/2} h_k^j(x).\end{aligned}\tag{9.7}$$

All other products on the same scale are zero.

Expanding u_0 explicitly into the Haar basis,

$$u_0(x) = \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} d_k^j h_k^j(x) + \sum_{k \in \mathbf{Z}} s_k^n \chi_k^n(x),\tag{9.8}$$

and using (9.7), we obtain from (9.4)

$$u_0^2(x) = 2 \sum_{j=1}^{j=n} 2^{-j/2} \sum_{k \in \mathbf{Z}} d_k^j s_k^j h_k^j(x) + \sum_{j=1}^{j=n} 2^{-j/2} \sum_{k \in \mathbf{Z}} (d_k^j)^2 \chi_k^j(x) + 2^{-n/2} \sum_{k \in \mathbf{Z}} (s_k^n)^2 \chi_k^n(x).\tag{9.9}$$

On denoting

$$\begin{aligned}\hat{d}_k^j &= 2^{-j/2+1} d_k^j s_k^j, \\ \hat{s}_k^j &= 2^{-j/2} (d_k^j)^2, \\ \hat{s}_k^n &= 2^{-n/2} (s_k^n)^2,\end{aligned}\tag{9.10}$$

we rewrite (9.9) as

$$u_0^2(x) = \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} \hat{d}_k^j h_k^j(x) + \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} \hat{s}_k^j \chi_k^j(x) + \sum_{k \in \mathbf{Z}} \hat{s}_k^n \chi_k^n(x).\tag{9.11}$$

We note that *if the coefficient d_k^j is zero then there is no need to keep the corresponding average s_k^j* . In other words, we need to keep averages only near the singularities, i.e., where the wavelet coefficients d_k^j (or products $s_k^j d_k^j$) are significant for a given accuracy.

Finally, to compute the coefficients of the wavelet expansion of the function u_0^2 , we need to expand the second sum in (9.11) into the wavelet basis. Starting from the scale $j = 1$, we compute the differences and averages \bar{d}_k^{j+1} and \bar{s}_k^{j+1} from s_k^j . We then add \bar{s}_k^{j+1} to \hat{s}_k^{j+1} before expanding it further. As a result, we compute \bar{d}_k^j , $j = 2, \dots, n$, (we set $\bar{d}_k^1 = 0$) and \hat{s}_k^n and obtain

$$u_0^2(x) = \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} (\hat{d}_k^j + \bar{d}_k^j) h_k^j(x) + \sum_{k \in \mathbf{Z}} (\hat{s}_k^n + \hat{s}_k^n) \chi_k^n(x).\tag{9.12}$$

It is clear, that the number of operations for computing the Haar expansion of u_0^2 is proportional to the number of significant coefficients d_k^j in the wavelet expansion of u_0 . If the original function is represented by a vector of the length N , then, in the worst case, the number of operations is proportional to N . If the original function is smooth with a finite number of singularities, then the number of operations is proportional to $\log_2 N$.

We now return to the general case and derive an algorithm to expand (9.4) into the wavelet basis. For each scale j , and for each term in (9.4), we consider the bilinear mappings,

$$M_{VW}^j : \mathbf{V}_j \times \mathbf{W}_j \rightarrow \mathbf{L}^2(\mathbf{R}) = \mathbf{V}_j \bigoplus_{j' \leq j} \mathbf{W}_{j'}, \quad (9.13)$$

$$M_{WW}^j : \mathbf{W}_j \times \mathbf{W}_j \rightarrow \mathbf{L}^2(\mathbf{R}) = \mathbf{V}_j \bigoplus_{j' \leq j} \mathbf{W}_{j'}, \quad (9.14)$$

and

$$M_{VV}^n : \mathbf{V}_n \times \mathbf{V}_n \rightarrow \mathbf{L}^2(\mathbf{R}) = \mathbf{V}_n \bigoplus_{j' \leq n} \mathbf{W}_{j'}, \quad (9.15)$$

which map a product of two functions into its expansion in the wavelet basis. The choice of the representation of $\mathbf{L}^2(\mathbf{R})$, which depends on the scale j , turns out to be important for the speed of the algorithm.³

The mappings (9.13)-(9.15) are tabulated by computing the coefficients

$$M_{VWW}^{j,j'}(k, k', l) = \int_{-\infty}^{+\infty} \phi_k^j(x) \psi_{k'}^j(x) \psi_l^{j'}(x) dx, \quad (9.16)$$

$$M_{WWW}^{j,j'}(k, k', l) = \int_{-\infty}^{+\infty} \psi_k^j(x) \psi_{k'}^j(x) \psi_l^{j'}(x) dx, \quad (9.17)$$

$$M_{VVW}^{j,j'}(k, k', l) = \int_{-\infty}^{+\infty} \phi_k^j(x) \phi_{k'}^j(x) \psi_l^{j'}(x) dx, \quad (9.18)$$

and

$$M_{VWV}^j(k, k', l) = \int_{-\infty}^{+\infty} \phi_k^j(x) \psi_{k'}^j(x) \phi_l^j(x) dx, \quad (9.19)$$

$$M_{WWV}^j(k, k', l) = \int_{-\infty}^{+\infty} \psi_k^j(x) \psi_{k'}^j(x) \phi_l^j(x) dx, \quad (9.20)$$

$$M_{VVV}^j(k, k', l) = \int_{-\infty}^{+\infty} \phi_k^j(x) \phi_{k'}^j(x) \phi_l^j(x) dx, \quad (9.21)$$

where $j' \leq j$. It is clear, that the coefficients are identically zero for $|k - k'| > k_0$, where k_0 depends on the overlap of the supports of the basis functions. The number

³There might be fewer significant coefficients if we write, for example, $\mathbf{L}^2(\mathbf{R}) = \mathbf{V}_{j-j_0} \bigoplus_{j' \leq j-j_0} \mathbf{W}_{j'}$ in (9.13)-(9.15) with some $j_0 \geq 1$.

of coefficients which need to be stored may be reduced further by observing that, for example,

$$M_{WWW}^{j,j'}(k, k', l) = 2^{-j'/2} \int_{-\infty}^{+\infty} \psi_0^{j-j'}(x) \psi_{k-k'}^{j-j'}(x) \psi_{2^{j-j'}k-l}^0(x) dx, \quad (9.22)$$

so that

$$M_{WWW}^{j,j'}(k, k', l) = 2^{-j'/2} \tilde{M}_{WWW}^{j-j'}(k - k', 2^{j-j'}k - l). \quad (9.23)$$

However, the most significant reduction in the number of coefficients is a consequence of the fact that the coefficients in (9.16)-(9.18) decay as the distance $r = j - j'$ between the scales increases. For example, rewriting (9.23) as

$$\tilde{M}_{WWW}^r(k - k', 2^r k - l) = 2^{-r} \int_{-\infty}^{+\infty} \psi(2^{-r}x) \psi(2^{-r}x - k + k') \psi(x - 2^r k + l) dx, \quad (9.24)$$

and remembering that the regularity of the product $\psi(2^{-r}x) \psi(2^{-r}x - k + k')$ increases linearly with the number of vanishing moments of the function ψ , we obtain

$$|\tilde{M}_{WWW}^r(k - k', 2^r k - l)| \leq C 2^{-r\lambda M} \quad (9.25)$$

with some λ (see [19], [21]). It might be possible to estimate λ in (9.25), but since the coefficients in (9.16)-(9.18) are computed numerically, the rate of decay and the number of significant coefficients may be measured directly. To summarize, for a given precision, only the coefficients with labels $r = j - j'$, such that $0 \leq r \leq j_0$, where j_0 is a small constant, need to be stored.

Evaluating the mappings (9.13)-(9.15), we arrive at the following representation of u_0^2 ,

$$u_0^2(x) = \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} \hat{d}_k^j \psi_k^j(x) + \sum_{j=1}^{j=n} \sum_{k \in \mathbf{Z}} \hat{s}_k^j \phi_k^j(x), \quad (9.26)$$

where \hat{d}_k^j , \hat{s}_k^j are obtained by adding contributions (on the appropriate scales) of the results of evaluating (9.13)-(9.15) for all scales j in (9.4).

The next stage of the algorithm is the expansion of the second term in (9.26) into the wavelet basis and is similar to that for the Haar basis.

Let us evaluate the complexity of computing u_0^2 in the worst case. We assume that the original function u_0 is represented by a vector of length $N = 2^n$ and its expansion into the wavelet basis also requires order N significant entries of the vector d_k^j . In the process of evaluating mappings (9.13)-(9.15), each coefficient d_k^j (j and k are fixed) will combine with the coefficients $d_{k'}^j$ (or $s_{k'}^j$), where $|k - k'| \leq k_0$, to produce a non-zero contribution on scales j' , $0 \leq j - j' \leq j_0$. Since j_0 and k_0 are (small) constants and are independent of the size N , the total number of operations is proportional to the number of significant coefficients d_k^j , which is $O(N)$ in this case.

If the number of significant coefficients d_k^j is proportional to the number of scales, $\log_2 N$, so will be the number of operations required to evaluate the mappings (9.13)-(9.15). It is also clear, that it is necessary to store only those averages $s_{k'}^j$, which combine with the significant coefficients d_k^j to produce a non-zero contribution. Therefore, it is sufficient to store only those $s_{k'}^j$, for which there exist the coefficient d_k^j , such that $|k - k'| \leq k_0$ and the product $s_{k'}^j d_k^j$ is above the threshold of accuracy. It means that we need to store averages only in the neighborhood of singularities.

The number of operation for expanding of the second term in (9.26) into the wavelet basis is proportional to the number of significant entries, and the estimate is completely similar to that for the Haar basis.

Remark. The algorithm for evaluation $F(u) = u^2$ in the wavelet basis allows us to evaluate the product of two functions, since $uv = \frac{1}{4}[(u+v)^2 - (u-v)^2]$.

IX.2 The algorithm for evaluating $F(u)$

Let F be an infinitely differentiable function. In order to decouple the scales, we again use a “telescopic” series,

$$F(u_0) - F(u_n) = \sum_{j=1}^{j=n} [F(P_{j-1}u) - F(P_j u)]. \quad (9.27)$$

Expanding the function F in the Taylor series at the point $(x+y)/2$, we have

$$F(x) - F(y) = \sum_{k=1}^{\infty} \frac{1}{(2k-1)! 2^{2k-2}} F^{(2k-1)}\left(\frac{x+y}{2}\right) (x-y)^{2k-1}. \quad (9.28)$$

Using (9.28) and $P_{j-1} = P_j + Q_j$, we obtain from (9.27)

$$F(u_0) - F(u_n) = \sum_{j=1}^{j=n} \sum_{k=1}^{\infty} \frac{1}{(2k-1)! 2^{2k-2}} F^{(2k-1)}\left(P_j u + \frac{1}{2} Q_j u\right) (Q_j u)^{2k-1}. \quad (9.29)$$

Let us write out several terms of this formula explicitly. If we keep only the first term, we obtain

$$F(u_0) - F(u_n) \approx \sum_{j=1}^{j=n} F^{(1)}\left(P_j u + \frac{1}{2} Q_j u\right) (Q_j u). \quad (9.30)$$

If $F(u) = u^2$, then formula (9.30) is exact and we obtain (9.3). If we keep two terms, we obtain

$$F(u_0) - F(u_n) \approx \sum_{j=1}^{j=n} \left[F^{(1)}\left(P_j u + \frac{1}{2} Q_j u\right) (Q_j u) + \frac{1}{24} F^{(3)}\left(P_j u + \frac{1}{2} Q_j u\right) (Q_j u)^3 \right] \quad (9.31)$$

We note that there is no term with the second derivative of F in (9.31) (there are no even derivatives in (9.29)). Using (9.30) and considering the remainder of the series in (9.29) as an error term, we obtain the results of J.M. Bony. The error term, however, will be slightly smoother than in Bony's results, since the remainder has a factor $(Q_j u)^3$ instead of $(Q_j u)^2$. We may also keep more terms to make the remainder arbitrarily smooth.

It is not clear at this point, if there is an advantage in computing $F(u)$ via (9.29), or a repeated application of the algorithm for $F(u) = u^2$ is sufficient to compute various functions F . However, there are several analytic advantages in considering (9.29). In particular, by using theorems characterizing the functional spaces, for example the Hölder spaces $C^s(\mathbf{R})$, in terms of their wavelet coefficients (see [33]), the series (9.29) may be truncated for given precision.

Finally, we remark that the adaptive numerical algorithms based on the expansion in (9.29) are novel and we expect to develop new methods for solving P.D.E.'s using these algorithms.

References

- [1] B. Alpert. Sparse representation of smooth linear operators. Ph.D. thesis, Yale University, 1990.
- [2] B. Alpert, G. Beylkin, R. R. Coifman, and V. Rokhlin. Wavelets for the fast solution of second-kind integral equations. Technical report, Department of Computer Science, Yale University, New Haven, CT, 1990.
- [3] 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).
- [4] A. Ben-Israel and D. Cohen. On iterative computation of generalized inverses and associate projections. *SIAM J. Numer. Anal.*, 3(3):410–419, 1966.
- [5] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.*, 1991. submitted.
- [6] G. Beylkin, R. R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms II. in progress.
- [7] 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.
- [8] G. Beylkin, R. R. Coifman, and V. Rokhlin. Wavelets in Numerical Analysis. In *Wavelets and Their Applications*, pages 181–210. Jones and Bartlett, 1992.
- [9] J. M. Bony. Interaction des singularités pour les équations aux dérivées partielles non-linéaires. Sémin. e.d.p., 1979/80, 22, 1981/82, 2 et 1983/84, 10, Centre de Mathématique, Ecole Polytechnique, 91128-Palaisau, France.
- [10] J. M. Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non-linéaires. *Ann. Scient. E.N.S.*, 14:209–246, 1981.
- [11] J. M. Bony. Propagation et interaction des singularités pour les solutions des équations aux dérivées partielles non-linéaires. In *Proceedings of the International Congress of Mathematicians, Warszawa*, pages 1133–1147, 1983.
- [12] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. Communications*, 31(4):532–540, Apr. 1983.

- [13] 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).
- [14] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *preprint*, 1990.
- [15] R. R. Coifman and Y. Meyer. Au delà des opérateurs pseudo-différentiels. In *Astérisque*, 57, (seconde édition revue et augmentée). Société Mathématique de France.
- [16] R. R. Coifman and Y. Meyer. Nouvelles bases orthonormée de $l^2(\mathbf{r})$ ayant la structure du système de walsh. 1989. *preprint*.
- [17] R. R. Coifman and Y. Meyer. Nouvelles bases orthogonales. *C.R. Acad. Sci., Paris*, 1990.
- [18] R. R. Coifman and V. Wickerhauser. Best-adapted wave packet bases. 1990.
- [19] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure and Appl. Math.*, 41:909–996, 1988.
- [20] I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NSF Series in Applied Mathematics. SIAM, 1991. in press.
- [21] I. Daubechies and J. Lagarias. Two scale difference equations: I-II. *SIAM J. Math. Anal.*, 1990.
- [22] P. Goupillaud, A. Grossmann, and J. Morlet. Cycle-Octave and related transforms in seismic signal analysis. *Geoplotation*, 23:85–102, 1984.
- [23] L. Greengard. Potential flow in channels. *SIAM J. Sci. Stat. Comput.*, 11(4):603–620, 1990.
- [24] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comp. Phys.*, 73(1):325–348, 1987.
- [25] A. Haar. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, pages 331–371, 1910.
- [26] W. M. Lawton. Necessary and sufficient conditions for constructing orthonormal wavelet bases. *J. Math. Phys.*, 32(1), 1991.
- [27] S. Mallat. Multiresolution approximation and wavelets. Technical report, GRASP Lab, Dept. of Computer and Information Science, University of Pennsylvania.

- [28] S. Mallat. Review of multifrequency channel decomposition of images and wavelet models. Technical Report 412, Courant Institute of Mathematical Sciences, New York University, 1988.
- [29] Y. Meyer. Le calcul scientifique, les ondelettes et filtres miroirs en quadrature. CEREMADE, Université Paris-Dauphine.
- [30] Y. Meyer. Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. In *Séminaire Bourbaki*, page 662. Société Mathématique de France, 1985-86. Astérisque.
- [31] Y. Meyer. Ondelettes et fonctions splines. Technical report, séminaire edp, Ecole Polytechnique, Paris, France, 1986.
- [32] 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.
- [33] Yves Meyer. *Ondelettes et Opérateurs*. Hermann, Paris, 1990.
- [34] S.T. O'Donnell 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).
- [35] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *J. Comp. Phys.*, 60(2), 1985.
- [36] G. Schulz. Iterative berechnung der reziproken matrix. *Z. Angew. Math. Mech.*, 13:57–59, 1933.
- [37] M. J. Smith and T. P. Barnwell. Exact reconstruction techniques for tree-structured subband coders. *IEEE Transactions on ASSP*, 34:434–441, 1986.
- [38] T. Söderström and G.W. Stewart. On the numerical properties of an iterative method for computing the Moore-Penrose generalized inverse. *SIAM J. Numer. Anal.*, 11(1):61–74, 1974.
- [39] 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.
- [40] R. C. Ward. Numerical computation of the matrix exponential with accuracy estimates. *SIAM. J. Numer. Anal.*, 14(4):600–610, 1977.