Potential Integral Equations of the 2D Laplace Operator in Wavelet Basis

Mihai Dorobantu

NADA - KTH
Lindstedsvägen 15
S100 44 Stockholm - Sweden
e-mail: mihai@nada.kth.se

Abstract

In this paper we solve the Laplace equation by solving a boundary integral equation in a wavelet basis. The idea of solving integral equations using a hierarchical approach is used in the Multipole method and Wavelet-Galerkin methods in the non-standard form. Due to the orthogonality of the wavelet transform, the discrete linear systems preserve their good conditioning while their sparse structures speed up iterative solvers. We investigate the link between the two multi-scale methods and the convergence of Wavelet Galerkin discretizations. The behaviour of GCR algorithms under wavelet transformations and in non-standard form is investigated. Finaly we analyze the optimal choice of Daubechies filters in the non-standard form and describe some numerical experiments.

Keywords: Boundary integrals, Laplace Equation, Wavelets, Multipole, Multiresolution

Contents

Acknowledgments

I would like to thank my advisor, Professor Björn Engquist for his guidance and encouragement. I am also in debt to Dr. Anders Szepessy and Dr. Leif Abrahamsson for their fruitful advice. I cannot thank enough the entire staff at NADA for their kind and constant support. Last, but not least, I thank my wife Mariana and my son Andrei for the understanding they showed during many working weekends and white nights.

1 Introduction

Solving the Laplace equation by means of boundary integral equations is no new technique. It is a powerful method in the sense that it reduces the problem's dimension, it agrees with non-rectangular boundaries and can be used to solve exterior problems. In total contrast to Finite Element or Finite Difference methods, the discretization of the boundary integral equation yields well-conditioned but full linear systems. The dense matrices obtained with classical discretization effectively destroy all the savings induced by lowering the dimensions of the problem. A hierarchical approach stands a good chance to lower the computational complexity without loosing the good conditioning of the problem.

The Multipole method [10] illustrates this approach. It is a fast algorithm for applying the dense matrix of a discretize integral equation to a vector. Its speed comes from dropping the exact arithmetics requirements in favor of a hierarchical aggregation of contributions of neighbors to the computation of an integral. This is based on the regularity properties of the integral kernel at long range. In principle one can express the contribution of all far away points as a polynomial and then evaluate rapidly a substantial part of a boundary integral. To be more precise, say we discretize the underlying integral operator so that

$$\int k(x,y)\sigma(y)dy \approx \sum_j a_{i,j}\sigma_j.$$

Multipole will essentially express

$$\Delta x \sum_{|C-x_j|>3\rho} k(x,x_j)\sigma(x_j) \approx \sum_{n=0}^{d} c_n (x-C)^n,$$

valid for any $|x-C| < \rho$. The degree d of the polynomial is low. If we take $x = x_i$, we find that

$$\sum_{|C-x_j|>3\rho} a_{i,j}\sigma_j \approx \int_{|C-y|>3\rho} k(x_i,y)\sigma(y)dy \approx \sum_{n=0}^{d} c_n (x_i-C)^n$$

and thus substitute the computation of a long sum with the evaluation of a polynomial, while the approximation error is bounded by a constant depending only on the polynomial's degree d.

Assume now that we integrate the kernel against a wavelet with sufficient vanishing moments, supported in $|C-y| > 3\rho$. Then we have that

$$\int\int k(x,y)\sigma(y)\psi_{j,p}(x)dydx \approx \sum_{n=0}^{d} c_n \int (x-C)^n \psi_{j,p}(x)dx = 0.$$

We thus find a parallel between Multipole's compression of computations and the compression of an integral operator in a wavelet basis. This is the effect described by Beylkin,

Coifman and Rokhlin in [1] in a more general setting, compression of operators with wavelet tools.

Let us adopt the framework of Zygmund-Calderon operators studied extensively in [1]. We can project the integral operator in a wavelet basis. Assume that away from the diagonal, the kernel is smooth. Since the basis functions are localized, away from the diagonal they pick up only the regular part of the kernel and thus the "energy" of the projection is small. Neglecting these small entries yields a banded linear system.

One cannot neglect the interactions of different scales. These will be important even at long range and spoil the band structure described above. Thus it is important to separate the different scales, which can be done effectively with the non-standard form of the operator introduced in [1].

Note that as with most hierarchical methods, such an approach will never find the exact solution. But for any given $\varepsilon > 0$, the algorithm commits an error bounded by ε.

The paper is organized as follows: Section 2 swiftly defines the basic concepts used. In Section 3 a standard argument is given for the convergence of the full wavelet-Galerkin discretization of the boundary integral equations. Section 4 deals with the structure of the non-standard form of some integral operators. Numerical schemes and their properties are analyzed in Section 5. Finally, numerical examples are given in Section 6.

2 Definitions and classical results

2.1 Boundary integral equations

Let Ω be a bounded domain in the plane and Γ its boundary. We consider the boundary value problems for the Laplace equation:

Interior Dirichlet problem:

$$\Delta u(x) = 0 \quad x \in \Omega, \qquad u(x) = f(x) \quad x \in \Gamma \tag{1}$$

Exterior Dirichlet problem:

$$\Delta u(x) = 0 \quad x \in \mathcal{R}^2 \setminus \overline{\Omega}, \qquad u(x) = f(x) \quad x \in \Gamma \tag{2}$$

The two Dirichlet problems have solutions for any integrable f and they are unique (in the class of bounded functions for the exterior problem).

Let L be the length of Γ and $\gamma : [0\ L]$ be the arc-length parameterization of Γ. For any integrable function f on Γ, let $\int_\Gamma f(x)dx$ denotes the line integral along the curve Γ. It is of course a short-hand notation for $\int_0^L f(\gamma(t))\,|\gamma'(t)|\,dt$. We can always rescale the parameterization by the curve length and thus identify the functions defined on Γ with periodic functions on the unit interval.

Potential Theory gives the solutions of the Dirichlet problems in terms of the *double-layer potentials*:

$$k(x,y) = \frac{N_y^T(x-y)}{||x-y||^2}, \qquad \forall x \in \mathcal{R}^2, \quad y \in \Gamma.$$

where N_y is the normal to Γ at y. Note the possible singularities at $x = y$. The solution of the Dirichlet problem is represented by

$$u(x) = \int_\Gamma k(x,y)\sigma(y)dy. \tag{3}$$

where the weights σ are found by solving integral equations:

Interior Dirichlet problem:

$$\pi\sigma(x) + \int_\Gamma k(x,y)\sigma(y)dy = f(x), \qquad \forall x \in \Gamma \tag{4}$$

Exterior Dirichlet problem:

$$-\pi\sigma(x) + \int_\Gamma k(x,y)\sigma(y)dy = f(x), \qquad \forall x \in \Gamma \tag{5}$$

Solving equations (4)-(5) is no easy numerical task. Even though the singularities of the kernels are not dangerous, their discretizations will have slow decay away from the diagonal giving rise to large, non-sparse linear systems. It is known that the condition number of the linear system, when discretizing equation (4) with a quadrature rule on an equidistant grid, is bounded (as the step-size tends to zero) and the orders of accuracy is that given by the quadrature rule. Thus iterative solvers for the discrete linear systems will converge in a few iterations, apart from the high computational cost of applying a dense matrix to a vector.

2.2 Wavelet-type basis on closed curves

The problem of constructing a periodic wavelet basis on an interval is not trivial. If we are only interested in representing scales much finer the the interval's length, then the "brutal" periodization of any wavelet basis on the real line is efficient. Following [3] and [5], we build an orthonormal basis on the closed curve $\Gamma \subset \mathcal{R}^2$ from the Daubechies wavelets. We sketch

the construction of a periodic, wavelet-type basis on a smooth closed curve in $\mathcal{R}^2$ following the "brutal" construction explained in e.g. [3] and [5].

Let φ and ψ be the *shape function* and the *mother wavelet* of any Daubechies wavelets basis. We can assume that their supports are not longer then the length of Γ, otherwise we just substitute φ and ψ with a scaled version $2^{j_0/2}\varphi(2^{j_0}x)$ and $2^{j_0/2}\psi(2^{j_0}x)$.

Denote the space generated by the integer translations of φ with V_0. Let V_j be the space generated by

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$$

and let W_j be the orthogonal complement of V_j in V_{j+1}. It is known that $\{\varphi_{j,k}\}_k$ and $\{\psi_{j,k}\}_k$ are orthonormal basis in V_j and W_j respectively. Since

$$V_{J+1} = W_J \oplus W_{J-1} \oplus \cdots \oplus W_0 \oplus V_0,$$

we also have that

$$\{\psi_{j,k}\}_{j,k} \cup \{\varphi_{0,k}\}_k$$

is also an orthonormal basis in V_{j+1}. Let h_n and g_n be the wavelet filters, i.e.

$$\varphi = \sum_{k\in\mathrm{Z}} h_k\varphi_{1,k}, \qquad \psi = \sum_{k\in\mathrm{Z}} g_k\varphi_{1,k}.$$

This expansion holds on any two adjacent scales:

$$\varphi_{j,k} = \sum_{n=0}^{R-1} h_m\varphi_{j+1,2k+m} \quad \text{and} \quad \psi_{j,k} = \sum_{n=0}^{R-1} g_m\varphi_{j+1,2k+m}. \tag{6}$$

and is the key to the fast transform between the two basis of V_{J+1}. Let us mention that the wavelets $\psi_{j,k}$ have a certain number of vanishing moments. The Daubechies shape function is supported in $[0\ R-1]$ and the maximal number of vanishing moments is $M = R/2$. We can ask that the shape function should also have M shifted vanishing, but then $R = 3M$.

The brutal Periodization is done in the following way. We look at the interval $I_0 = [0\ 1]$ and retain all the basis functions that are fully supported there. If the support of a basis function, say $\psi_{j,k}$, exists I_0 at the left end, the missing part is matched by $\psi_{j,k+2^j}$ at the right end. Adding the two functions and restricting their sum to I_0, creates a new function which is orthonormal on all the other basis functions.

Let $\tilde{\varphi}_{j,k}$ and $\tilde{\psi}_{j,k}$ be the periodized versions of $\varphi_{j,k}$ and $\psi_{j,k}$. In [5] we prove the following results:

Theorem 2.1 *$\tilde{\varphi}_{j,k}$ and $\tilde{\psi}_{j,k}$ defined above form orthonormal basis in $L^2(I_0)$. For each $j \geq 0$ they generate a sequence of subspaces V_j and W_j such that*

$$V_{j+1} = V_j + W_j, \qquad V_j \perp W_j, \qquad \overline{\bigcup_{j=-\infty}^{\infty} V_j} = L^2(I_0)$$

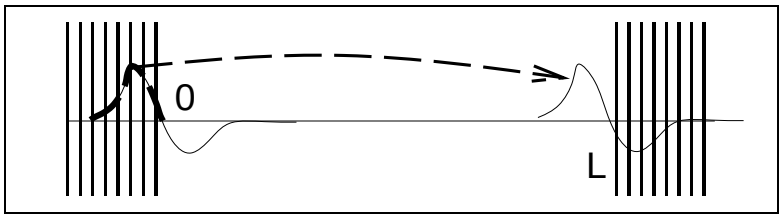

Figure 1: How to periodize without loss of orthogonality: If the support of $\psi_{j,k}$ goes over L, add $\psi_{j,k+L2^j}$. The contribution to inner products "lost" at the left of $\psi_{j,k}$ is matched by the left part of $\psi_{j,k+L2^j}$.

Theorem 2.2 *Let f be the periodic extension to whole real axis of $\tilde{f} \in L^2(I_0)$. Then*

$$< f, \varphi_{j,k} >=< \tilde{f}, \tilde{\varphi}_{j,k} >$$

and

$$< f, \psi_{j,k} >=< \tilde{f}, \tilde{\psi}_{j,k} > .$$

The results above show that the decomposition of $\tilde{f}$ can be thought of as the standard decomposition of f restricted to those wavelets whose support intersect I_0. The projection and reconstruction between scales can therefore use the pyramid scheme, which is the key to the *fast* transformation between the standard and wavelet-type basis.

Note that $\tilde{\psi}_{j,k}$ has the same number of vanishing moments as ψ, at least as long as its support does not contain the origin. Since translation of the other basis functions moves the origin out of their support, for all purposes we can still talk about vanishing moments for periodic wavelets.

From now on we drop the $\tilde{}$ notation and call the periodized basis a wavelet basis. The fast wavelet transform given by (6) is preserved; it is nothing else then an orthonormal coordinate change between two basis:

$$\mathcal{B}' = (\varphi_{J+1,k})_k \longleftrightarrow (\varphi_{0,k})_k \cup \bigcup_{0 \le j \le J} (\psi_{j,k})_k = \mathcal{B}$$

2.3 Non-standard form of operators

The non-standard form of an operator is an elegant instrument for decoupling the different built-in scales. This concept was introduced by Beylkin, Coifman and Rokhlin [1] together with the general analysis of the compression of Calderon-Zygmund kernels.

Let T be a linear operator on $L^2(I_0)$. We have two representations of the projection of T on the subspace V_{J+1} corresponding to the two basis $\mathcal{B}$ and $\mathcal{B}'$. The representation on the

finest scale (i.e. in $\mathcal{B}'$) is roughly equivalent to a standard uniform grid representation:

$$\int \varphi_{J+1,l'}(x) \left(T\varphi_{J+1,l}\right)(x)dx \qquad \forall l, l'.$$

The representation in the wavelet basis $\mathcal{B}$ has elements of type

$$\int \psi_{j',l'}(x) \left(T\psi_{j,l}\right)(x)dx, \qquad \forall j, j', l, l',$$

This representation is efficient if T has localized interactions within each scale $0 \leq j \leq J$. If the operator T has interactions across different scales, the representation of T becomes complicated.

Let P_j and Q_j denote the orthogonal projections of $L^2(I_0)$ onto V_j and W_j. We have that

$$P_{j+1} = Q_j + P_j$$

and if T is a linear operator on $L^2(I_0)$, and $T_j = P_j T P_j$ its projection V_j, then

$$T_{j+1} = Q_j T Q_j + Q_j T P_j + P_j T Q_j + T_j.$$

If we denote the projection of T onto W_j by A_j and $B_j = Q_j T P_j$ and $C_j = P_j T Q_j$, we can write

$$T_{j+1} v_{j+1} = \begin{bmatrix} A_j & B_j \\ C_j & T_j \end{bmatrix} \begin{pmatrix} w_j \\ v_j \end{pmatrix} \tag{7}$$

for any $v_{j+1} \in V_{j+1}$ which is decomposed into $w_j + v_j$ that lie in W_j and V_j respectively. Note that $T_j v_j$ can be treated in a similar way down to the coarsest scale.

It is clear that the projection T_{J+1} of T on the finest scale subspace V_{J+1} can be represented by the operators

$$A_J, B_J, C_J, \cdots, A_0, B_0, C_0 \qquad \text{and} \qquad T_0$$

and the coupling conditions $v_{j+1} = w_j + v_j$. This is the non-standard form of the operator T as introduced in [1]. What is the advantage of such a representation against the representation in the wavelet basis $\mathcal{B}$? In the case of operators describing localized interactions, In the non-standard form all the different scales are decoupled. Due to the localized interactions on each scale, the non-standard form matrices have a simpler structure. The price we may pay is additional storage space and more complex computations. In many applications however, the opposite is true. Due to simpler structures, storage becomes more efficient while applying the operator to a vector can become a $O(N)$ procedure.

If $k(x, y)$ is I_0-periodic in both arguments, then its projection on the space $V_{J+1} \times V_{J+1}$ can be represented in the non-standard form by representing the projection of

$$(Tu)(x) = \int_{I_0} k(x, y)u(y)dy.$$

Thus the projection of $k(x,y)$ is represented by the matrices $\{A_j, B_j, C_j\}_{0 \le j \le J}$ and T_0 with elements

$$\alpha_{p,q}^j = \int\int k(x,y)\psi_{j,q}(x)\psi_{j,p}(y)dxdy$$

$$\beta_{p,q}^j = \int\int k(x,y)\varphi_{j,q}(x)\psi_{j,p}(y)dxdy$$

$$\gamma_{p,q}^j = \int\int k(x,y)\psi_{j,q}(x)\varphi_{j,p}(y)dxdy$$

$$t_{p,q}^0 = \int\int k(x,y)\varphi_{0,q}(x)\varphi_{0,p}(y)dxdy$$

This decomposition will be called the non-standard form of the wavelet discretization of the kernel $k(x,y)$.

Note also that any smooth closed curve $\Gamma \subset \mathcal{R}^2 = \mathcal{C}$ can be identified with the interval $I_0 = [0\ L]$ via the unique arc-length parameterization (L is the length of Γ). Functions that live on Γ can be identified (isometrically) with periodic functions on I_0. Thus we can use all the periodic wavelet tools on such a curve.

3 Convergence of wavelet discretizations

Let us now look at the convergence properties of the solution of the discretized version of equation (4). The continuous equation can be written as

$$T\sigma = (I + K)\sigma = f. \tag{8}$$

Projection onto V_J yields the discrete version

$$T_J v = (I + K_J)v = P_J f, \tag{9}$$

which is in fact a Galerkin discretization with V_J as the test functions' space. Here we denote $K_J = P_J K$ whose restriction to V_J is nothing else then the projection $P_J K P_J$ of K. We can thus adapt standard Boundary Element Method techniques (e.g. [9]) to prove convergence. The only "new" feature is the evaluation of the wavelet approximation error in a suitable norm. Following [8], we have that

$$||u - P_J u|| \le C 2^{-J} ||u||_{H^1}. \tag{10}$$

On the other hand the operator K is compact, i.e.

$$||Ku||_{H^1} \le C ||u||.$$

The discretization error follows:

$$||(K - K_J)u|| = ||(I - P_J)Ku|| \leq C2^{-J} ||Ku||_{H^1} \leq C2^{-J} ||u||$$

which yields

$$||K - K_J|| \leq C2^{-J}. \tag{11}$$

We have that

$$||u|| \leq C ||(I + K)u||$$

by choosing $f = (I + K)u$ and using the classical estimates for the solution of the equation (8). Using the discretization error (11), we have

$$\begin{aligned} ||u|| &\leq C ||(I + K)u|| \\ &\leq C (||(I + K_J)u|| + ||(K - K_J)u||) \\ &\leq C ||(I + K_J)u|| + C_1 2^{-J} ||u||. \end{aligned}$$

Thus, for large enough J, we find the classical stability estimate:

$$||u|| \leq \frac{C}{1 - C_1 2^{-J}} ||(I + K_J)u|| \leq C ||(I + K_J)u||.$$

Since $I + K_J$ is bounded, convergence now follows from the approximation and discretization errors (10),(11). From the continuous and discrete equations (8), (9), we have

$$(I + K_J)(\sigma - v) = (K_J - K)\sigma + (I - P_J)f.$$

Due to stability, we find

$$||\sigma - v|| \leq C ||(I + K_J)(\sigma - v)|| \leq C (||(K_J - K)\sigma|| + ||(I - P_J)f||)$$

and finally

$$||\sigma - v|| \leq C2^{-J} (||\sigma|| + ||f||_{H^1}).$$

If more is known about the regularity of the exact solution, these estimates can be improved. It is easily seen that the convergence order is governed by the approximation error (10). In particular, if σ and f have more then M bounded derivatives (M is the number of vanishing moments of the wavelets orthogonal to V_J), then the approximation error becomes $O(2^{-JM})$ and the wavelet-Galerkin discretization converges like an order M method.

4 Compression of potentials in non-standard form

In this section we show that for kernels with integrable diagonal singularities that are otherwise smooth, the non-standard form matrices are sparse. This is the case of the double-layer potential on a regular curve Γ. The result is by no means new, it is known for Calderon-Zygmund kernels.

We try to establish a connection between the Multipole method and the compression of the potential equations in non-standard form. Let us first pick a scale j and look at the wavelet $\psi_{j,p}$. Denote its support $I_{j,p}$ centered at C_p. Then pick another wavelet $\psi_{j,q}$ localized "far away" and denote its support $I_{j,q}$ and its center C_q. Multipole will express an approximation to

$$y \mapsto \int_{I_{j,q}} \frac{N(x)}{y-x} dx$$

on $I_{j,p}$ as a polynomial in the unknown $y - C_q$. Multiplication against a wavelet with sufficiently many vanishing moments "codes" such a contribution to zero or a very small entry. This is explained by the following:

Lemma 4.1 *Assume the curve Γ is smooth and the wavelets have M vanishing moments. If the distance between the supports $I_{j,p}$ and $I_{j,q}$ of the wavelets $\psi_{j,p}$ and $\psi_{j,q}$ is not less then $2^{-j} R\,|p-q|$, we have*

$$\left| \int \int \frac{N(x)}{y-x} \psi_{j,p}(x) \psi_{j,q}(y) dx dy \right| \leq \frac{C}{(2\,|p-q|)^{M+1}} \tag{12}$$

Proof: Let us first note that if $|x - C_q| < |y - C_q|$, we have

$$\frac{N(x)}{y-x} = \frac{N(x)}{y-C_q} \sum_{l \geq 0} \left(\frac{x-C_q}{y-C_q} \right)^l = \sum_{l \geq 1} \frac{1}{(y-C_q)^l} N(x) (x-C_q)^{l-1}.$$

Expand $N(x)$ in a Taylor series around C_q and then we have that

$$\frac{N(x)}{y-x} = \sum_{l \geq 1} \frac{r_l}{(y-C_q)^l} (x-C_q)^{l-1}.$$

The sum is convergent when $y \in I_{j,p}$ and $x \in I_{j,q}$, so we can multiply with $\psi_{j,q}(x)$ and integrate on $I_{j,q}$ to find

$$\int_{I_{j,q}} \psi_{j,q}(x) \frac{N(x)}{y-x} dx = \sum_{l \geq 1} \frac{r_l}{(y-C_q)^l} \int_{I_{j,q}} \psi_{j,q}(x) (x-C_q)^{l-1} dx.$$

Now we use the first M vanishing moments of $\psi_{j,q}$ to find that

$$\left| \int_{I_{j,q}} \psi_{j,q}(x) \frac{N(x)}{y-x} dx \right| \leq \frac{C}{(y-C_q)^{M+1}} |I_{j,q}| \left(\frac{|I_{j,q}|}{2} \right)^M .$$

We have that all intervals $I_{j,q}$ have the same length at scale j:

$$|I_{j,q}| = Q_M 2^{-j}.$$

The constant Q_M is the length of the support of the shape function φ and it depends on the choice of the filters but it is a modest multiple of M.

Finally we multiply with $\psi_{j,p}(y)$ and integrate on its support $I_{j,p}$ to find:

$$\left| \int_{I_{j,p}} \int_{I_{j,q}} \frac{N(x)}{y-x} \psi_{j,q}(x) \psi_{j,p}(y) dx dy \right| \leq \frac{C}{(2^{-j} R \, |p-q|)^{M+1}} (2^{-j} Q_M)^{M+1} 2^{-M}.$$

Picking $R = Q_M/2$ yields the bound we are looking for and also makes precise the loose notion of two intervals being "far away". □

The lemma tells us that entries of the non-standard form matrix A_j that lie away from the diagonal can be thrown away without much loss. Note that Multipole and the wavelet discretization make use of the regularity of

$$y \mapsto \int_{I_{j,q}} \frac{N(x)}{y-x} dx$$

on $I_{j,p}$. The former encodes this behavior with a few polynomial coefficients, while the latter does the same job using a band matrix. Similar computations as in the proposition, or using directly the results of [1], show that the entries away from the diagonals of B_j and C_j are also very small:

Lemma 4.2 *The matrices A_j, B_j and C_j of the non-standard form decomposition of the double layer potential are banded and the bandwidth is independent of the scale j:*

$$\left| \alpha^j_{p,q} \right|, \left| \beta^j_{p,q} \right|, \left| \gamma^j_{p,q} \right| \leq \frac{C}{(2 \, |p-q|)^{M+1}} \tag{13}$$

Note that the actual band-width [1] is much thinner then that given by (12), since the "true" support of highly regular mother wavelets ψ is much less then their exact support.

[1] Throughout this paper we view the non-standard form matrices as "periodic" and thus a diagonal band also contains the lower-left and upper-right corners.

An extreme example is that of the unit-circle. The integral equation of the interior Dirichlet problem becomes:

$$\pi\sigma(x) + \int_\Gamma \left(\frac{-x^T(y-x)}{||y-x||^2} \right) \sigma(y)dy = f(x), \qquad \forall x \in \Gamma, x \neq y \tag{14}$$

with the constant kernel

$$k(x,y) = \frac{1}{2}.$$

The kernel has no decay but maximum regularity and projecting equation (14) even in Haar wavelet coordinates yields the linear system's matrix is just a multiple of the identity.

The estimate (13) together with the "far away" condition give a pessimistic image of the true decay of the non-standard form matrices which contrasts sharply with the case of the circle. Assume that we have a domain with a smooth boundary such that the kernel is smooth away from the diagonal. This implies that the domain is not of the type illustrated in Figure 2 [2] Further assume that the normal field is twice differentiable. Around any fixed point $y \in \Gamma$, the boundary can be approximated by a circle of radius R_y given by curvature at y. After a translation of the origin, we have

$$k(x,y) = \frac{N_y^T(x-y)}{||x-y||^2} \approx \frac{-\frac{1}{R_y}y^T(x-y)}{||x-y||^2} = \frac{1}{2R_y}$$

valid locally near y. As we saw in the case of the circle, if R_y is smooth, the decay away from the diagonal of non-standard form matrices is much faster then (13).

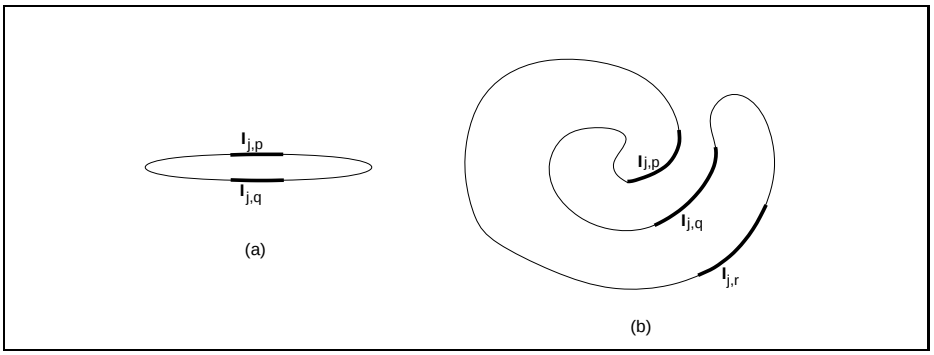

Figure 2: "Distant" supports of strongly interacting neighbors

In order to speed up computations, we are interested in ignoring the small entries of these matrices by setting to zero all entries outside a given band of width B. According to [1],

[2] Geometricly this restriction means that for any scale $j \geq j_0$, p is close to q whenever the distance between $I_{j,p}$ and $I_{j,q}$ is small.

should the error of this procedure (in operator norm) be bounded by a given ε, we need

$$B \geq \left(\frac{C}{\varepsilon}\right)^{1/M}. \tag{15}$$

An important aspect that we read off from the circle is that the operators T_j are full. Indeed, there are no cancellations in the coefficients

$$t^j_{p,q} = \frac{1}{2}\int\int \varphi_{j,q}(x)\varphi_{j,p}(y)dxdy = \frac{1}{2^{j+1}}.$$

This property is not dangerous as long as the full matrix T_0 is set up at a scale coarse enough to ensure modest dimensions.

The second important step of Multipole is to aggregate the expansions of finer scales to get the expansions of the coarser scale. In the wavelet approach this is built in the reconstruction algorithm of the wavelet transform.

5 Iterative solvers for the potential equations

5.1 Equivalence of GCR iterations

We show that the linear systems produced by classical and wavelet discretizations of potential equations are equivalent via an orthogonal transformation. It then follows that they have the same conditioning. In particular, the iterative solver GCR (Generalized Conjugate Gradient) which is known to converge fast for some classical discretization of (4), will converge equally fast in wavelet coordinates.

Let us start with an integral equation

$$\sigma(x) + \int_\Gamma k(x,y)\sigma(y)dy = f(x), \qquad \forall x \in \Gamma. \tag{16}$$

Take a uniform grid $(x_p)_{p=1:N}$ on Γ with respect to the arc-length parameterization. The integral equation can be discretized by e.g. the trapeze rule

$$s_p + \sum_{q\neq p} a_{p,q}s_q = f_p, \tag{17}$$

where

$$f_p = f(x_p), \quad a_{p,q} = \frac{1}{N}k(x_q, x_p).$$

We assume that the kernel is smooth away from the diagonal $x = y$ and that the possible singularity at $x = y$ is integrable. Solving the system (17) yields a second order accurate approximation to $s_p = \sigma(x_p)$ and the procedure is called the Nyström method.

Now we want a wavelet discretization of the integral equation (16) Assume that we have a wavelet basis $\mathcal{B}$ in a space V_{J+1} and $\mathcal{B}' = (\varphi_{J+1,p})_p$ is the standard basis of V_{J+1}. Let T be the integral operator

$$(T\sigma)(x) = \sigma(x) + \int k(x,y)\sigma(y)dy.$$

Then we can write (16) as

$$T\sigma = f.$$

This equation we project onto V_{J+1} to find

$$T_{J+1}P_{J+1}\sigma = P_{J+1}TP_{J+1}P_{J+1}\sigma = P_{J+1}f$$

or

$$\sigma_p + \sum_q k_{p,q}\sigma_q = f_p,$$

where

$$s_p = \int \sigma(x)\varphi_{J+1,p}(x)dx, \quad f_p = \int f(x)\varphi_{J+1,p}(x)dx,$$

$$k_{p,q} = \int\int k(x,y)\varphi_{J+1,p}(y)\varphi_{J+1,q}(x)dxdy.$$

Note that

$$k_{p,q} \approx 2^{-(J+1)}k(C_p, C_q) \approx \frac{1}{N}k(x_p, x_q) = a_{p,q}$$

if $N = 2^{J+1}$ is sufficiently large. This is no surprise since the basis $\mathcal{B}'$ is very "near" the uniform grid basis on fine enough scales, and thus $A_N \approx T_{J+1}$.

Denote by $\mathcal{W}$ the discrete wavelet transform, which is nothing else but an orthonormal basis change from $\mathcal{B}'$ to $\mathcal{B}$. Let S_{J+1} be the matrix of T_{J+1} in the basis $\mathcal{B}$. Then the wavelet discretization of the equation (16) is

$$S_{J+1}w = g, \qquad S_{J+1} = \mathcal{W}T_{J+1}\mathcal{W}^{-1} \approx \mathcal{W}A_N\mathcal{W}^{-1} \tag{18}$$

where g is the wavelet transform of $P_{J+1}f$. Since $\mathcal{W}$ is an orthogonal transformation, we have that

$$\kappa(A_N) \approx \kappa(T_{J+1}) = \kappa(S_{J+1}).$$

It is known that the Nyström method for (4) yields a well-conditioned matrix A_N. This is why GCR solves the linear system (17) in a few iterations. Due to the good conditioning of the wavelet-based linear system (18), we may hope that the GCR iteration will converge just as fast. More is in fact true since GCR is invariant to orthogonal transformations (see Appendix A): If $s^{(k)}$ and $w^{(k)}$ are the sequences of approximations produced by GCR applied to the discrete equations (17) and (18) with the initial guesses $s^{(0)} = w^{(0)} = 0$, then $S_{J+1} \approx \mathcal{W}A_N\mathcal{W}^{-1}$ implies that

$$\mathcal{W}s^{(k)} \approx w^{(k)}, \qquad \forall k \geq 0$$

while the residuals are invariant

$$\left|\left|A_N s^{(k)} - f\right|\right| \approx \left|\left|S_J w^{(k)} - g\right|\right|, \qquad \forall k \geq 0.$$

All the approximations are due to $T_{J+1} \approx A_N$. In practice we never can hope for $T_{J+1} = A_N$ since we want to replace T_{J+1} with an approximation. This approximation comes from either computing only diagonal bands in the non-standard form of T_{J+1}, or from ignoring the small entries of S_{J+1}. The truncated matrix S_{J+1} has a more complicated sparse structure then the non-standard form matrices and its entries are more expensive to produce. But when the kernel has singularities away from the diagonal (due to "corners" or "thin" geometries), applying S_{J+1} may well prove better then the non-standard form and T_{J+1}, while the condition number is unchanged.

5.2 Description of the Algorithm

We now have the necessary ingredients to develop a fast algorithm for solving the boundary integral equations (4) and (5). In a wavelet basis $\mathcal{B}$, GCR will converge in a few iterations, truncation outside a diagonal band of the non-standard form matrices makes each iteration fast.

For $v_{J+1} \in V_{J+1}$, let $\underline{v_{J+1}}$ be the sequence of projections

$$\{\ w_j, v_j\ :\ v_j = P_j v_{j+1}, w_j = Q_j v_{j+1}\ , j = J, \cdots, 0\ \}.$$

In Appendix B we describe a procedure that computes $\underline{\tilde{v}_{J+1}}$ from the non-standard form of T_{J+1}, $\underline{v_{J+1}}$ and $\underline{z_{J+1}}$ where $\tilde{v}_{J+1} = z_{J+1} + T_{J+1} v_{J+1}$.

Each GCR iteration computes a residual r_{n+1} and a search direction p_{n+1} using linear combinations of r_n, all the previous search directions $p_0, \cdots, p_n$ and $T_{J+1} p_n$.

We have the following procedure for solving the linear system

$$T_{J+1} v_{J+1} = f_{J+1}:$$

Step 1 Using the intermediate steps of the Fast Wavelet Transform, compute

$$\underline{p_0} := \underline{r_0} := \underline{f_{J+1}},$$

the non-standard form of T_{J+1} and put $\underline{x_0} = 0$.

Step 2 Until convergence, iterate

$$\underline{q_n} = \underline{Ap_n}$$

$$\alpha_n = r_n^T q / \left|\left|q\right|\right|^2$$

$$\underline{x_{n+1}} = \underline{x_n} + \alpha_n \underline{p_n}$$

$$\underline{r_{n+1}} = \underline{r_n} + \alpha_n \underline{q_n}$$

$$\underline{p_{n+1}} = \underline{r_{n+1}} + \sum_{i=1}^{n} \beta_{n,i} \underline{p_i}.$$

The coefficients $\beta_{n,i}$ are chosen such that $q_n \perp \psi_{n+1}$

Step 3 From $\underline{x_n}$ recover the approximate solution v_{J+1}.

Note that in the computations of α_n we do not need to recover r_n and q_n. Their orthonormal decompositions in $V_j \oplus W_j$ are contained in $\underline{r_n}$ and $\underline{q_n}$ and can be used as such to compute inner-products in V_{J+1}.

The underlying condition in all the above statements was that we can find R such that the distance between two supports $I_{j,p}$ and $I_{j,q}$ on the same scale can be approximated by

$$2^{-j} R \left|p - q\right|.$$

This condition must hold on all the scales j involved. The condition can be expressed in terms of bounding the curvature of Γ. But under these circumstances, we have seen that the kernel is smooth at the diagonal too, since it locally resembles the kernel of a circular domain. For such domains, we may ask what is the optimal number of vanishing moments M of the wavelets we use.

5.3 Optimal wavelet filters

Note that broadly speaking there are four error sources which influence the choice of the wavelets' vanishing moments and other parameters. These are:

1. The discretization error, i.e. the error of the discretization viewed as a Boundary Element Method. If nothing is known about the right-hand side f, then we have an error proportional to 2^{-J}. As soon as we have more information about the regularity of f and the exact solution, the error estimate can be drastically improved (in principle, one power for each bounded derivative). The finest scale $J+1$ is chosen so that this error is under the given tolerance. This choice is equivalent to choosing the number of gridpoints in the Nyström method.

2. The quadrature error, i.e. the error committed in computing the coefficients of the non-standard form due to the quadrature rule employed. In the case of the one-point quadrature previously described, this error is proportional to at most $2^{-j_0(M+1)}$, where j_0 is the coarsest scale used.

3. The banding error, i.e. the error in the discrete operator due to ignoring all the non-standard form entries outside a given band of width B. It is proportional to $B^{-M} \log N$.

4. The iterative solver's error. The threshold on the residual will affect the number of iterations needed by GCR to produce an approximate solution of the linear system. Note that for the interior problem, the condition number $\kappa(T_{J+1})$ is independent of the size of the problem and small, thus a "constant" number of GCR iterations is needed for a fixed threshold. Furthermore, this constant grows very slowly as the threshold is lowered.

Let us first assume that we have no error in the coefficients of the non-standard form and N has been chosen large enough so that the discretization error is well under a given threshold ε. The work W needed to solve the linear system is

$$W \sim NMB, \tag{19}$$

where B is the band-width of the non-standard form matrices and N is the size of the problem. Assume that are interested in committing an error not larger then a given ε and N has been fixed such that the non-standard form coefficients can be computed within the given tolerance. We estimate B from (15) to find

$$W \sim NM \left(\frac{C}{\varepsilon}\right)^{1/M}.$$

For a given ε, the optimal M is found by differentiating W. We find

$$M_\varepsilon \sim \log \frac{C}{\varepsilon}$$

and then we have the optimal band-width

$$B_\varepsilon \sim \left(\frac{C}{\varepsilon}\right)^{1/\log \frac{C}{\varepsilon}} = \left(e^{\log \frac{C}{\varepsilon}}\right)^{1/\log \frac{C}{\varepsilon}} = e.$$

Note that M_ε increases slowly in terms of the tolerance ε, while B_ε is independent of ε!

Note in Figure (3) that the band-width B and the work's proportionality constant $C = BM$ increase violently for less vanishing moments then the optimum.

Now consider the more complex situation of the one-point quadrature evaluation of the non-standard form. Let j_0 be the coarsest scale and δ the error in the coefficients. Note that

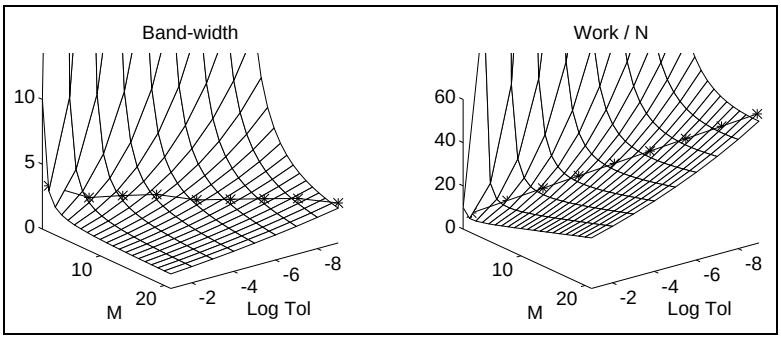

Figure 3: Variation of band-width (left) and work's proportionality constant (right) as functions of the error threshold $TOL = \varepsilon$ and number of vanishing moments M. The stars represent the optimal values B_ε and $B_\varepsilon M_\varepsilon$ for given ε.

applying T_{j_0} to a vector produces an error bounded by $2^{j_0}\delta$ while all the other matrices produce errors bounded by $B\delta$. Setting $B\delta \sim \varepsilon$ gives the choice of j_0:

$$\left(\frac{C}{\varepsilon}\right)^{1/M} 2^{-j_0(M+1)} \sim \varepsilon$$

which implies

$$j_0 \sim -\frac{1}{M}\log\varepsilon + \frac{1}{M(M+1)}\log C.$$

This shows that when many vanishing moments are at hand, the wavelet decomposition can be carried up to very coarse scales! On the other hand, on such very coarse scales the wrap-around due to the periodicity of the basis blurs the significance of vanishing moments. On the coarsest scale we apply a full matrix T_{j_0} and thus the work becomes

$$W \sim NMB + 2^{2j_0}$$

and the optimization problem is to solve

$$\min_M NM\left(\frac{C}{\varepsilon}\right)^{1/M} + \varepsilon^{2/M} C^{1/(M^2+M)}.$$

However, the second term is small, so that for all practical purposes we may still consider $W \sim NMB$.

If we are interested in minimizing the work needed for setting up the non-standard form matrices. In the one-point quadrature case, the work is

$$W \sim NM^2B$$

and the optimal number of vanishing moments and bandwidth are

$$M_\varepsilon \sim \frac{1}{2}\log\frac{C}{\varepsilon}, \qquad B_\varepsilon \sim \sqrt{\epsilon}.$$

Let us also look at the case where exact solution and the righthand-side f are smooth. Then we have an order M interpolation error and we can choose $N \sim \varepsilon^{-1/M}$. The work is

$$W \sim NMB \sim M \left(\frac{1}{\varepsilon}\right)^{2/M}$$

which yields

$$M_\varepsilon \sim 2 \log \frac{1}{\varepsilon}, \qquad B_\varepsilon \sim e^2 \qquad \text{and} \qquad W_\varepsilon \sim e^2 \log \frac{1}{\varepsilon},$$

which shows that the banded-wavelet based discretizations behave like spectral methods when allowing for large values of M. Even if extremely high accuracy is demanded, the band-width does not grow provided we can find wavelets with increasingly high number of vanishing moments. For such problems, the wavelet approach is strongly superior to the Nyström method complemented with a fast solver of the linear system (e.g. Multipole). In this case the work is at best proportional to the number of grid-points $N = \varepsilon^{-1/p}$, where p is the accuracy order of the quadrature involved.

We can then ask to minimize various combinations of work depending on the exact nature of the problem, but it is important to realize that the optimal values of M and B are small and independent or increase slowly as $\varepsilon \to 0$, so that W is essentially

$$\alpha \frac{1}{\varepsilon^\beta} \log \frac{C}{\varepsilon}, \qquad \beta \leq 1.$$

6 Numerical results

The most simple numerical exercise is the Laplace problem on a circle. The integral kernel is constant, so any uniform grid discretization has constant coefficients. In the operator's matrix T_{J+1} all entries are equal to $1/2$. Direct wavelet transform of the discrete system yields

$$\pi I + \frac{\pi}{N} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

where the only nonzero entry corresponds to the shape function itself. The non-standard form matrices A_j, B_j and C_j are reduced to multiples of the identity, while the coarsest scale operator T_0 is a scaled version of T_{J+1} regardless of the wavelets used.

In the case of an ellipse with aspect-ratio 4:1 we computed the non-standard form matrices, truncated outside different band-width B and compared the original and reassembled operators. We tried different Daubechies wavelets with $M = 1, 2 \ldots, 10$ vanishing moments. The work of applying the operator to a vector is

$$W \sim CN, \qquad C = BM.$$

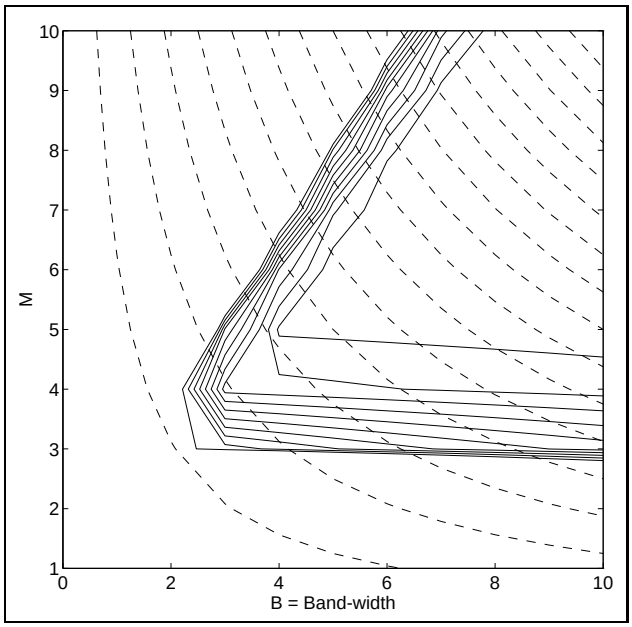

Figure 4: Error levels (full lines) and computation constants C=BM (dotted lines) for an ellipse.

The dotted lines of Figure 4 represent constant C, increasing from left to right, while the full lines represent the operator norm of the error committed by truncating outside a band of width B in the non-standard form matrices. The error levels decrease from left to right. Note that the optimal band-width is found at around $B_0 = 2 \cdot 4$ while the optimal number of vanishing moments M_0 grows slowly, as the theory predicted.

Consider the case of a athletics track shaped domain bounded by the curve Γ in Figure 5.

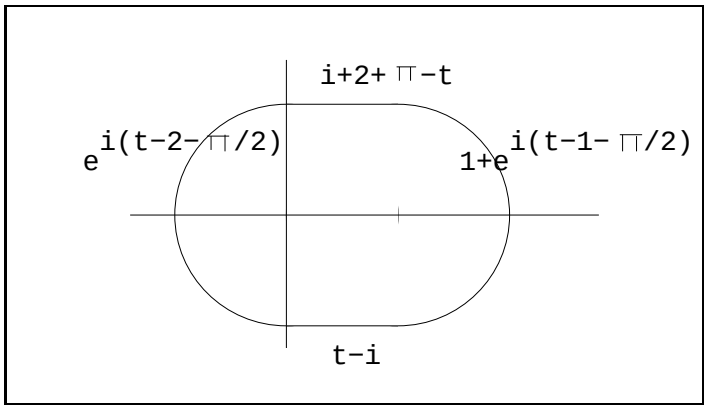

Figure 5: Athletics track arc-length parameterization.

Let us first discretize the integral kernel of the Dirichlet problems (4) and (5) in both classical

and wavelet coordinates. In classical coordinates, we use the Nyström method with stepsize L/N and the trapezoidal rule, setting up the linear system (17). For modest values of N we can set up the matrix A_N explicitly, find $\kappa(\pm\pi I + A_N)$ and solve the equation (17) with GCR.

Then we set up the "exact" (i.e. without truncations) linear system (18) using different wavelet discretizations by simply rewriting the matrix of (17) in the new basis.

Runs were made for $N = 32$, 64, 128 and 256 with the Haar basis and Daubechies wavelets with filters of length 4,8, and 20. As expected, the condition numbers of the systems had very little variation across the different basis (order of 10^{-8}). There is a slight variation for different N. For the initial projection $P_{J+1}f$, we assumed f to be constant on the supports of the basis functions. Results are given in Table 1; GCR was stopped when the residual's norm was smaller then 10^{-5}.

N	Nyström Algorithm	Haar	D_4	D_8	D_{20}	κ
32	8	8	8	8	8	2.5471
64	8	8	8	8	8	2.5424
128	9	9	9	9	8	2.5425
256	9	9	9	9	8	2.5432
512	9	9	9	9	9	2.5435

Table 1: Interior problem: Iterations of GCR with an error of 10^{-6}

Then we truncate all the entries of the matrices that lie below a given threshold ε and solve a problem with smooth right-hand side. It is remarkable that the condition number and number of GCR iterations are almost invariant with respect to the threshold ε, as seen in Table 2.

Truncation threshold ε	κ	Matrix density %	Tolerance of residual in GCR			
			10^{-1}	10^{-2}	10^{-3}	10^{-4}
			l^2-error in solution σ			
0	2.5435	100	0.0097	0.0022	0.0001	0.0000
10^{-4}	2.5435	1.11	0.0097	0.0022	0.0004	0.0004
10^{-3}	2.5436	0.52	0.0107	0.0050	0.0044	0.0044
10^{-2}	2.5150	0.25	0.0231	0.0213	0.0213	0.0213
			Number of GCR iterations			
			4	5	7	8

Table 2: Full wavelet transform of the linear system using Daubechies D_8 wavelets with $M = 4$.

The right-hand side used was

$$f(t) = \sin \frac{t\pi}{1+\pi}.$$

Due to the discontinuity in the 2^{ed} derivative of the normal field at the "corners" of the boundary, the solution of equation (4) is not smooth even if f is so. The exact solution is illustrated in Figure 6(top).

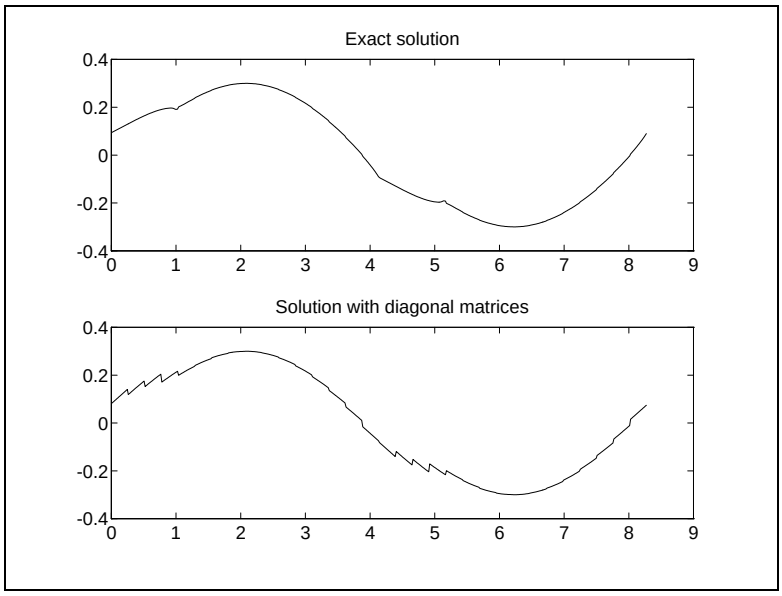

Figure 6: Solutions of using full and diagonal non-standard form matrices.

As in the case of the ellipse, we compute the non-standard form matrices, truncate outside a band of width B and reassemble the operator. Note in Figure 7 the same behavior as for the smooth case of the ellipse. The error-levels are higher in this case due to the mild, corner discontinuity. The dotted lines are $C = BM = constant$ and the value of C grows from left to right. The full lines are the error levels, decreasing from left to right. There is now a growth in the optimal band-width as well as in the optimal filter and these optimal values are not as sharp as in the smoother case. The left figure represents a discretization with $N = 128$ points while the right one uses $N = 512$ points. The error levels on the right figure are about 10% of those on the left.

We also tried an exterior problem (5); the conditioning of the system is proportional to N, but GCR still solves the linear system fast as seen in Table 3.

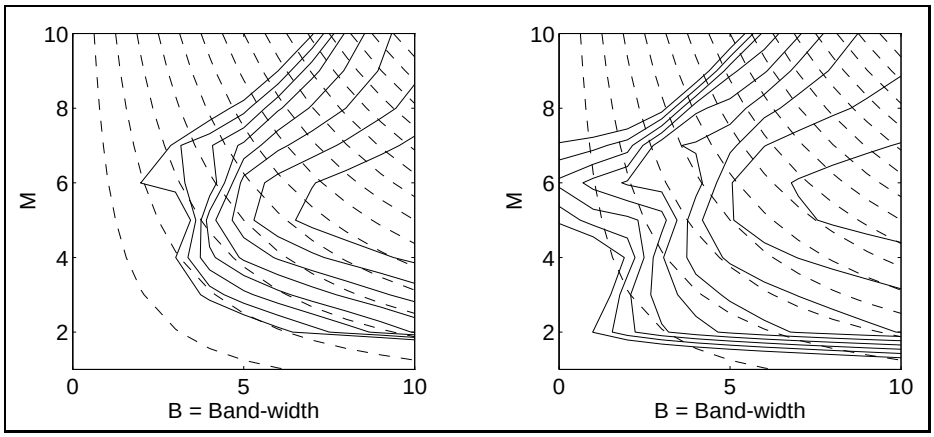

Figure 7: Error levels (full) and cost constants C levels (dotted) for banded non-standard form matrices

N	κ	Iterations
32	36.9447	8
64	73.4331	8
128	146.3410	9

Table 3: Exterior problem

A Invariance of GCR to orthogonal transformations

Lemma A.1 *Assume we solve the linear systems*

$$Ax = f, \qquad By = g$$

using GCR with the initial guesses x_0 and y_0. Let $x_1, x_2, \ldots, x_n, ..$ and $y_1, y_2, \ldots, y_n, ..$ be the sequences of approximations to x and y produced. If there exists a orthogonal matrix U such that

$$B = UAU^T, \qquad g = Uf, \qquad y_0 = Ux_0,$$

then

$$y_n = Ux_n, \qquad \forall n \geq 1$$

and the residuals at each step are equal, i.e.

$$||f - Ax_n|| = ||g - By_n|| .$$

Proof: For $Ax = f$, the GCR algorithm produces the sequences p_n,r_n,α_n and β_n that satisfy

$$\begin{aligned}
\alpha_n &= r_n^T A p_n / \left|\left| A p_n \right|\right|^2 \\
x_{n+1} &= x_n + \alpha_n p_n \\
r_{n+1} &= r_n + \alpha_n A p_n \\
p_{n+1} &= r_{n+1} + \textstyle\sum_{j=1}^{n} \beta_{n,j} p_j
\end{aligned}$$

with $p_0 = r_0 = f - Ax_0$ and $\beta_{n,j}$ such that $p_{n+1}^T A^T A p_j = 0$. For the system $By = g$, denote by s_n and q_n the vectors corresponding to r_n and p_n. Assume that for some n we have

$$y_n = U x_n, \qquad s = U r_n, \qquad q_n = U p_n$$

and apply one step of GCR to find

$$\frac{s_n^T B q_n}{\left|\left| B q_n \right|\right|^2} = \frac{r_n^T U^T U A U^T U p_n}{\left|\left| U A U^T U p_n \right|\right|^2} = \frac{r_n^T A p_n}{\left|\left| A p_n \right|\right|^2} = \alpha_n.$$

Using this fact, it follows that

$$y_{n+1} = y_n + \alpha_n q_n = U(x_n + \alpha_n p_n) = U x_{n+1}$$

and

$$s_{n+1} = s_n + \alpha_n B q_n = U r_n + \alpha_n U A U^T U q_n = U(r_n + \alpha_n A p_n) = U r_{n+1}.$$

Finally we have that

$$q_{n+1} = q_n + \sum \delta_{n,j} q_j = U \left(p_n + \sum \delta_{j,n} p_j \right).$$

The numbers $\delta_{n,j}$ are chosen such that $q_{n+1}^T B^T B q_j = 0$. This implies that δ_n solves the linear system

$$\sum_{i=1}^{n} q_i^T B^T B q_j \delta_{n,i} = -s_{n+1}^T B^T B q_j$$

This system can be written as

$$\sum_{i=1}^{n} p_i^T A^T A p_j \delta_{n,i} = -r_{n+1}^T A^T A p_j$$

which has the solution $\beta_{n,j}$. Thus $\delta_n = \beta_n$ and then we have $q_{n+1} = U p_{n+1}$. This shows that GCR produces the same solutions up to the orthogonal transformation defined by U. For the residuals we have that

$$\left|\left| f - A x_n \right|\right| = \left|\left| r_n \right|\right| = \left|\left| U s_n \right|\right| = \left|\left| g - B y_n \right|\right|.$$

□

Note that GCR is usually started with the zero guess and the usual termination criterion is the size of the residual. Then the lemma tells that the GCR solution of a linear system is invariant under any orthogonal basis change. In particular, the wavelet transform is such an orthonormal basis change, so GCR will work in a wavelet basis exactly as it does in the classical basis.

B Matrix-vector multiplication non-standard form

We show how the multiplication

$$\tilde{v}_{J+1} = T_{J+1} v_{J+1}$$

can be done if we have the non-standard form representation of T_{J+1}.

Assume we have computed the wavelet decomposition of v_{J+1} and we retained the sequence of intermediate projections

$$v_{j+1} = w_j + v_j, \qquad w_j = Q_j v_{j+1}, \quad v_j = P_j v_{j+1}, \qquad j = J, J-1, \cdots, 0$$

Such an operation costs $O(N)$ operations and $2N$ memory locations. At any scale $j+1$, the identity (7) shows us how to compute

$$\tilde{v}_{j+1} = T_{j+1} v_{j+1}$$

should we have w_j, v_j, the non-standard form band matrices A_j, B_j, C_j and $T_j v_j$. Going from the coarser to finer scales, we can compute all $\tilde{v}_{j+1}$ by doing 3 banded matrix - vector multiplications and one step of the inverse fast wavelet transform. At each scale, this procedure costs $C2^j$ operations giving again a $O(N)$ total cost.

To summarize, we have the following $O(N)$ procedure:

Down pass:

For $j := J$ to j_0 (step -1)
begin
$\tilde{w}_j := A_j w_j + B_j v_j$
$\tilde{v}_j := C_j w_j$
end

Coarse scale: $\tilde{v}_{j_0} := \tilde{v}_{j_0} + T_{j_0} v_{j_0}$

Up pass:

For $j := j_0$ to J
begin
$\tilde{v}_{j+1} := \tilde{v}_{j+1} + R_j(\tilde{w}_j, \tilde{v}_j)$
end

R_j is the reconstruction operator, i.e.:

$$R_j^{-1} = \begin{pmatrix} P_j \\ Q_j \end{pmatrix}.$$

Note that if we are interested in computing

$$z_{J+1} + T_{J+1} v_{J+1}$$

and we also have the sequence of projections of z_{J+1}, we only need to add these at the down-pass. Thus we can think of this procedure as a fast method for converting the sequences of projections of v_{J+1} and z_{J+1} into the sequence of projections of $z_{J+1} + T_{J+1} v_{J+1}$.

C Evaluation of the non-standard form

When solving the potential equations, the essential information will be concentrated around the diagonal of the non-standard form matrices, as will be shown later on. Dropping all the other entries will result in a strong data compression that will allow for fast computations. This in it turn implies the need of rapid evaluation of the non-standard form coefficients within a diagonal band.

Following [1], the "shifted" moments of the coiflets' shape function φ can be used to find a fast one-point quadrature for

$$t^{j+1}_{p,q} = \int\int k(x,y)\varphi_{j+1,q}(x)\varphi_{j+1,p}(y)dxdy.$$

Together with the interscale relation (6), we compute the entries

$$\alpha^j_{p,q}, \quad \beta^j_{p,q}, \quad \gamma^j_{p,q}, \quad \text{for} \quad |p-q| < B.$$

First note that if $\int \varphi = 1$ and the next $M-1$ moments of $\varphi(x-\tau_M)$ are zero, then

$$\begin{aligned}
\int f(x)\varphi_{j,n}(x)dx &= 2^{j/2}\int f(x)\varphi(2^j x - n)dx \\
&= 2^{-j/2}\int f(2^{-j}(x+n))\varphi(x)dx \\
&= 2^{-j/2} f(2^{-j}(\tau_M + n)) + O(2^{-j(M+1/2)})
\end{aligned}$$

This one-point quadrature can be used to evaluate any coefficient of $P_{j+1}TP_{j+1}$. Using (6) we find

$$\begin{aligned}
\alpha^j_{p,q} &= \int\int k(x,y)\psi_{j,q}(x)\psi_{j,p}(y)dxdy \\
&= \sum_{n,m=0}^{R-1} g_n g_m \int\int k(x,y)\varphi_{j+1,2q+n}(x)\varphi_{j+1,2p+m}(y)dxdy \\
&\approx 2^{-(j+1)} \sum_{n,m=0}^{R-1} g_n g_m k(2^{-j-1}(\tau_M + 2q + n), 2^{-j-1}(\tau_M + 2p + m))
\end{aligned}$$

Similar formulae express $\beta^j_{p,q}$ and $\gamma^j_{p,q}$ in terms of $k(x,y)$ evaluated point-wise in $BR^2 2^j$ operations at any scale j and bandwidth B.

This technique works as long as the one-point quadrature is accurate. In the case of coiflets with M vanishing moments, it is enough to have kernels with rapidly decaying derivatives of order M. For smooth kernels we have a bound on the error in the non-standard form coefficients valid also near to the diagonal:

$$C 2^{-j(M+1)}. \tag{20}$$

In general, i.e. for Daubechies wavelets where the shape function lacks vanishing moments, it is possible to compute a M point quadrature such that

$$\int \varphi(x) f(x) dx \approx \sum_{l=0}^{M-1} c_l f(l).$$

Based on this quadrature formula, we can approximate coefficients of the projection $P_j f$ by

$$\int \varphi_{j,p}(x) f(x) dx \approx 2^{-j/2} \sum_{l=0}^{M-1} c_l f(2^{-j}(l+p)).$$

Then we find the non-standard form coefficients of the kernel $k(x,y)$

$$\begin{aligned}
\alpha^j_{p,q} &= \int\int k(x,y) \psi_{j,q}(x) \psi_{j,p}(y) dx dy \\
&= \sum_{n,m=0}^{R-1} g_n g_m \int\int k(x,y) \varphi_{j+1,2q+n}(x) \varphi_{j+1,2p+m}(y) dx dy \\
&\approx 2^{-(j+1)} \sum_{n,m=0}^{R-1} \sum_{s,t=0}^{M-1} g_n g_m c_s c_t k(2^{-(j+1)}(s+2q+n), 2^{-(j+1)}(t+2p+m)).
\end{aligned}$$

Note that now $BR^2 M^2 2^j$ operations are needed to evaluate the band. R is the length of the wavelet filters, usually proportional to M.

D Moments of the shape function

We give a derivation for a quadrature formula for the Daubechies wavelets. We ask for the coefficients $c_0, c_1, \ldots, c_{M_1}$ such that the quadrature

$$\int \varphi(x) f(x) dx \approx \sum_{l=0}^{M-1} c_l f(l)$$

is exact for all polynomials of degree less then M. The coefficients are found by solving the linear system

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 2 & \dots & M-1 \\ 0 & 1^2 & 2^2 & \dots & (M-1)^2 \\ & & \vdots & & \\ 0 & 1^{M-1} & 2^{M-1} & \dots & (M-1)^{M-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{M-1} \end{pmatrix} = \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_{M-1} \end{pmatrix}$$

where μ_l are the moments of the shape function φ:

$$\mu_l = \int \varphi(x) x^l dx.$$

Similar computations are done in [1]. A recursive algorithm is developed for computing μ_l with any given accuracy. We now present a direct derivation of the moments. Note that we know $\mu_0 = 1$. For the higher moments we do by differentiating the Fourier transform of φ

$$\frac{d^l}{d\xi^l} \int \varphi(x) e^{-ix\xi} dx = (-i)^l \int \varphi(x) x^l e^{-ix\xi} dx$$

and evaluating it in $\xi = 0$:

$$\mu_l = \int x^l \varphi(x) dx = i^l \frac{d^l}{d\xi^l} \widehat{\varphi}(0).$$

On the other hand, Fourier transforming the dilation equation

$$\varphi(x) = \sqrt{2} \sum h_n \varphi(2x - n)$$

yields

$$\widehat{\varphi}(\xi) = \frac{\sqrt{2}}{2} \left(\sum_n h_n e^{-i\xi n/2} \right) \widehat{\varphi}(\xi/2) = m_0(\xi/2) \widehat{\varphi}(\xi/2), \tag{21}$$

where m_0 is the standard trigonometric polynomial

$$m_0(\xi) = \frac{\sqrt{2}}{2} \sum_n h_n e^{-i\xi n}.$$

Differentiating equation (21) l times and setting $\xi = 0$ yields:

$$(-i)^l \mu_l = \sum_{j=0}^{l} \begin{pmatrix} l \\ j \end{pmatrix} \frac{1}{2^l} (-i)^j \mu_j \frac{d^{l-j}}{d\xi^{l-j}} m_0(0). \tag{22}$$

The derivatives of m_0 are easy to compute and then we find

$$2^l \mu_l = \frac{\sqrt{2}}{2} \sum_{j=0}^{l} \begin{pmatrix} l \\ j \end{pmatrix} \mu_j \sum_n h_n n^{l-j}.$$

Setting $\xi = 0$ in (21) yields $\sum_n h_n = \sqrt{2}$ and then we have all the moments we need:

$$\begin{cases} \mu_0 = 1 \\ \\ (2^l - 1)\mu_l = \dfrac{\sqrt{2}}{2} \displaystyle\sum_{j=0}^{l-1} \begin{pmatrix} l \\ j \end{pmatrix} \mu_j \sum_n h_n n^{l-j}, \qquad l \geq 1 \end{cases} \tag{23}$$

There is a major disadvantage in using the recurrence formula (23): for long filters where n can have large values, only a few moments can be computed accurately. But we can make use the orthogonality condition

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 \equiv 1 \tag{24}$$

and the vanishing moments of ψ expressed as:

$$\frac{d^l m_0}{d\xi^l}(\pi) = 0, \quad l = 0, 1, ..., M-1. \tag{25}$$

It is obvious that

$$\overline{\frac{dm_0}{d\xi}} = \frac{d\overline{m_0}}{d\xi}$$

and by differentiating l times the orthogonality condition (24), we find

$$\sum_j \begin{pmatrix} l \\ j \end{pmatrix} m_0^{(l)}(\xi)\overline{m_0^{(l-j)}(\xi)} + m_0^{(l)}(\xi + \pi)\overline{m_0^{(l-j)}(\xi + \pi)}, \equiv 0 \qquad \forall l \geq 1.$$

From the condition (25) we have that m_0 has a multiple zero at $x = \pi$, with multiplicity M and by putting $x = \pi$ in the last identity we find:

$$\sum_j \begin{pmatrix} l \\ j \end{pmatrix} m_0^{(j)}(0)\overline{m_0^{(l-j)}(0)} = 0, \qquad \forall 1 \leq l \leq M-1.$$

Note that $m^{(k)}(0)$ is real for even k and imaginary for odd k. Let M_k be $m^{(k)}(0)$ or $im^{(k)}(0)$ for k even and odd respectively. For even l, we have that j and $l - j$ have the same parity and then

$$M_l = -\frac{1}{2} \sum_{j=1}^{l-1} \begin{pmatrix} l \\ j \end{pmatrix} M_j M_{l-j}.$$

For odd l, simple combinatorics cancels out the sums since

$$\sum \binom{2l+1}{j} m_0^{(j)}(0)\overline{m_0^{(2l+1-j)}(0)} =$$

$$= \sum \binom{2l+1}{2j} M_{2j} i M_{2l+1-2j} + \sum \binom{2l+1}{2j+1} (-i) M_{2j+1} M_{2l-2j}$$

$$= i\sum \binom{2l+1}{2j} M_{2j} M_{2l+1-2j} - i\sum \binom{2l+1}{2l-2j+1} M_{2l-2j+1} M_{2j}$$

$$= i\left(\sum \binom{2l+1}{2j} (M_{2j} M_{2l+1-2j} - M_{2l-2j+1} M_{2j})\right) = 0,$$

and thus we obtain no information about M_{2l+1}. This is no surprise, since if $m(\xi) = \sum \alpha_n e^{-in\xi}$ is an arbitrary trigonometric polynomial, then all the odd derivatives of $|m|^2$ cancel in the origin.

Note that (22) becomes:

$$\begin{aligned} 2^l \mu_l &= \sum \binom{l}{j} i^j \mu_{l-j} m_0^{(j)}(0) \\ &= \sum \binom{l}{2j} i^{2j} \mu_{l-2j} M_{2j} + \binom{l}{2j+1} i^{2j+1} \mu_{l-2j-1} (-i) M_{2j+1} \\ &= \sum (-1)^{[j]} \binom{l}{j} \mu_{l-j} M_j. \end{aligned} \tag{26}$$

An improvement over the recursion (23) is to compute in parallel μ_l and M_l using (23) which poses no difficulties. The quantities M_l can be computed in a stable manner for any even index. The unpleasant computation of $\sum h_n n^l$ is done only for the odd powers l, which is an improved, though not complete, solution. Analogue considerations are found in [7].

References

[1] G. Beylkin, R. Coifman, and V. Rokhlin, "Fast Wavelet Transform and Numerical Algorithms", *Yale University Tech. Report* YALE/DCS/RR-696, August 1989.

[2] I. Daubechies, "Orthonormal bases of compactly supported wavelets", *Communications in Pure and Applied Mathematics*, vol. 41, p.909-996, November 1988.

[3] I. Daubechies, "Ten Lectures on Wavelets", *Society for Industrial and Applied Mathematics Philadelphia, Pennsylvania,* 1992, ISBN 0-89871-274 - 2.

[4] G. David, "Wavelets and Singular Integrals on Curves and Surfaces", *Springer-Verlag, Berlin,* 1992, ISBN 3-540-53902-6

[5] M. Dorobantu, "Wavelet Based Algorithms for One-dimensional Parabolic Equations", *NADA Tech. Report* TRITA-NA-9214, September 1992.

[6] B. Engquist, S.Osher and S. Zhong, "Fast wavelet based algorithms for linear evolution operators", preprint, 1991.

[7] R.A. Gopinath, C.S. Burrus, "On the Moments of the Scaling Function $\psi_0(t)$", *Proceedings of ISCAS '92*, 1992(?)

[8] S. Jaffard and Ph. Laurençot, "Orthonormal Wavelets, Analysis of Operators and Applications to Numerical Analysis", *Wavelets - A Tutorial in Theory and Applications, vol.2, Academic Press, Inc.*, 1992, ISBN 0- 12- 174590-2.

[9] C. Johnson "Numerical Solutions of Partial Differential Equations by Finite Element Method",*Studentlitteratur, Lund,* 1989, ISBN 91-44-25241-2.

[10] V. Rokhlin "Rapid Solution of Integral Equations of Classical Potential Theory",*J. of Computational Physics,* vol.60, 1985

[11] W.L. Wedland "Elliptic Systems in the Plane", *Pitman, London,* 1979, ISBN 0-273-01013-1.