Fast wavelet transforms for matrices arising from boundary element methods *

Dave M. Bond † Stephen A. Vavasis ‡

March 26, 1994

Abstract

For many boundary element methods applied to Laplace's equation in two dimensions, the resulting integral equation has both an integral with a logarithmic kernel and an integral with a discontinuous kernel. If standard collocation methods are used to discretize the integral equation we are left with two dense matrices. We consider expressing these matrices in terms of wavelet bases with compact support via a fast wavelet transform as in Beylkin, Coifman and Rokhlin. Upper bounds on the size of the wavelet transform elements are obtained. These bounds are then used to show that if the original matrices are of size $N \times N$, the resulting transformed matrices are sparse, having only $O(N \log N)$ significant entries. Some numerical results will also be presented.

Unlike Beylkin, Coifman and Rokhlin who use the fast wavelet transform as a numerical approximation to a continuous operator already expressed in a full wavelet basis of $\mathcal{L}_2(\mathbb{R})$, we think of the fast wavelet transform as a change of basis matrix for a finite dimension, and apply it to a discretized function or matrix. As a result, we can use this fast wavelet transform as a "black box" transformation in existing boundary element codes.

*This work supported by an NSF Presidential Young Investigator award with additional support from Xerox Corp.

†Center for Applied Mathematics, Engineering and Theory Center, Cornell University, Ithaca, NY 14853.

‡Department of Computer Science, Upson Hall, Cornell University, Ithaca, NY 14853.

1 Introduction

1.1 The boundary integral element method

The *boundary integral equation method* (BIEM) is a technique for solving certain kinds of boundary value problems, including Laplace's equation. The essential feature of this method is the transformation of the governing differential equation into a set of integral equations defined over the boundary of the domain in question. The boundary of the domain is discretized and the resulting linear equations are then solved. Note that in this case only the boundary of the domain and not the entire domain itself is discretized. It follows that the set of linear equations generated by the BIEM is much smaller than the corresponding set of equations generated by an equivalent finite-difference or finite-element method. However, unlike the latter two methods, the BIEM yields a dense and usually unsymmetric system. The BIEM is usually credited to Hess and Smith [14], Jaswon [16] and Symm [21]. For a more detailed examination of this method, the reader is referred to the series edited by Brebbia [7].

1.2 Wavelets

Wavelets are usually thought of as functions which give rise to a basis of $\mathcal{L}_2(\mathbb{R})$ (see Daubechies [9], [10] and Meyer [17]). However, numerically one can use the coefficients that are used to define wavelets in order to create a change of basis matrix from $\mathbb{R}^N$ to $\mathbb{R}^N$. Primarily, we will view wavelets from the latter perspective, and utilize a "fast wavelet transform", such as is employed by Beylkin, Coifman and Rokhlin [5] (henceforth referred to as (BCR)), Alpert [2] and Strang [20], to numerically transform vectors from one basis to another. Let us denote the change of basis matrix arising from the fast wavelet transform by W. For the BIEM matrices arising from Laplace's equation in two dimensions (denoted by A), we can show that the fast wavelet transform of the columns of these matrices contain many small, albeit nonzero, entries. In order to obtain a sparse matrix we zero out these elements of WA that are smaller than a specified tolerance rule. So, in practice, instead of computing the matrix/vector product $A\mathbf{x}$, we would compute the matrix $A' = WA$, and then zero out the entries in A' that are less than the specified tolerance rule in absolute value, creating a sparse matrix A''

($O(N \log N)$ elements). We would then compute the vector $W^{-1}A''\mathbf{x}$, exploiting the sparsity of A'', giving a result very close to that of $A\mathbf{x}$ (see §6.1 for numerical results). The cost of a fast wavelet transform (and its inverse) of a vector of length N is $O(N)$. Other people have also looked at making integral operators sparse, such as Canning [8], Greenbaum, Greengard and Mayo [13], and Rokhlin [18]. Brandt and Lubrecht [6] have also proposed a multigrid method to make these types of operators sparse. Alpert [2], and [3], Alpert, Beylkin, Coifman and Rokhlin [4], and Jaffard [15] among others have also looked at making integral operators sparse using fast wavelet transforms and matrices, but while Alpert and Alpert et. al use the fast wavelet transform as a linear operator to make general discretized integral operators sparse, we look specifically at boundary integral operators and take into account the effect of the geometry upon the resulting sparse matrices. Jaffard uses a Galerkin method (rather than collocation, as we do) for discretizing the integral operators and he requires certain smoothness properties on the wavelets that he uses. Jaffard also takes the view of using wavelets as a basis of $\mathcal{L}_2(\mathbb{R})$. This seems to also be the usual viewpoint taken among those others who have also look into making operators sparse using a wavelet basis.

1.3 Solving the BIEM system of equations

The traditional way to solve a discretized boundary element method is by Gaussian elimination. Since the system of equations is dense, this is an $O(N^3)$ process (for N boundary points). In any iterative solution, including multigrid, the majority of the work involves applying the coefficient matrix to a vector, generally an $O(N^2)$ operation, at each iteration. For a more detailed examination of iterative methods pertaining to Boundary Integral Equations, the reader is referred to Atkinson and Graham [1], Schippers [19], Yan [23], and Vavasis [22]. If, in an iterative algorithm, we replace the original matrix/vector multiplications by applying the fast wavelet transform as mentioned in §1.2 above, we obtain a "black box" transformation which is easy to implement in an existing numerical iterative algorithm. While the total cost of transforming the matrix A is $O(N^2)$, for an iterative method this need only be computed once as a precursor to the iterations, thus yielding an iterative solution with an $O(N \log N)$ cost per iteration and asymptotically reducing the running time of the iterative solution. Furthermore, if the fast wavelet transform is utilized as mentioned above, all previous theory regarding the approximation of the BIEM to the original problem can still be applied, because the fast wavelet transform is applied to the already discretized matrices. In BCR [5], the integral operator is expressed in terms of a full wavelet basis of $\mathcal{L}_2(\mathbb{R})$, and this expansion is approximated using a finite number of terms. Because of this difference in the usage of the fast wavelet transform, we cannot apply the theory of BCR and have thus developed some theorems to provide upper bounds on the wavelet representations of the transformed vectors and hence on the sparseness that these fast wavelet transforms achieve.

1.4 Outline of paper

The columns of a coefficient matrix of a discretized boundary element method generally represent discretizations of functions with jump discontinuities in their derivatives. If one applies the fast wavelet transform to any column of the coefficient matrix, then depending upon the nature of these discontinuities, the resulting transformed column may have a sparsity pattern that one can then exploit in solving the discretized boundary element method system of equations. In §2 we examine the BIEM as applied to Laplace's equation in two dimensions and give explicit formulae for the functions that give rise to the coefficient matrix of the discretized boundary element method. In §3 we provide upper bounds on these functions and their derivatives and lower bounds on their corresponding vector discretizations. In §4 we define both the discrete wavelet transform and the fast wavelet transform. We also show that the fast wavelet transform (and its inverse) of a vector of length N is an $O(N)$ operation. In §5 we state and prove some theorems regarding the size of the elements of the transformed columns of the coefficient matrix. We also show that the transformed columns have $N - O(\log N)$ "small" entries. These theorems can be applied more generally to any matrix whose columns are interpolants of a piecewise smooth function. In §6 we present some numerical results supporting the theorems, and our conclusions and a discussion of ongoing work are presented in §7.

2 Boundary Integral Equations

2.1 Outline of method

In this section we examine the BIEM as applied to Laplace's equation in two dimensions. The specific problem under consideration is as follows: given a compact connected region $\Omega \in \mathsf{IR}^2$ with boundary Γ, find a solution ϕ to $\nabla^2\phi = 0$ on Ω given Neumann boundary data $(\partial\phi/\partial n)$ on Γ_1 and Dirichlet boundary data (ϕ) on Γ_2, where $\Gamma = \mathrm{cl}(\Gamma_1 \cup \Gamma_2)$, $\Gamma_1 \cap \Gamma_2 = \emptyset$. We will allow Γ to have a finite number of corners, but Γ must not contain any cusps. Note that while the present work focuses on the two-dimensional problem only, it includes all possible compatible boundary data. The BIEM is now outlined for the problem stated above. Let ϕ satisfy Laplace's equation, and let z be a point on Γ. Let $K(x,z)$ be the two-dimensional *Green's function*

$$K(x,z) = \frac{1}{2\pi} \ln \|x - z\|.$$

We assume that for any point $x \in \Gamma$, there is a unit normal direction $n(x)$ with respect to Γ. In case there is a corner at x, we take $n(x)$ to be the average of the limits of $n(y)$ as y approaches x from either side of x. Let $K_n(x,z)$ be the partial derivative of $K(x,z)$ with respect to x in the normal direction. Then

$$K_n(x,z) = \frac{1}{2\pi} \frac{\langle x - z, n(x)\rangle}{\|z - x\|^2}$$

where $\langle \cdot, \cdot \rangle$ is the IR^2 inner product.

Unless otherwise specified the following conventions will be used: $\|\cdot\|$ will denote the 2-norm (l_2 norm for vectors, $\mathcal{L}_2(\mathsf{IR})$ for functions), lowercase letters will be used for functions, scalars and for vectors in IR^2, capital letters will be used to denote matrices and functions of more than one variable, and boldface letters will be used for high-dimensional vectors. With the above definition of $K(x,z)$, the following identity always holds for a Laplace solution ϕ:

$$c(z)\phi(z) = \int_\Gamma K_n(x,z)\phi(x)\,dx - \int_\Gamma K(x,z)\frac{\partial\phi}{\partial n}(x)\,dx. \tag{1}$$

Here $c(z)$ measures the angle at the point z and would be 1 if the boundary were smooth. This is a consequence of Green's Second Integral Identity. Let us now discretize the boundary Γ in the following way: define

$$\gamma : [0,1] \mapsto \Gamma,$$

and define the nodal points t_i via

$$t_i = \frac{i-1}{N}, \qquad i = 1, \ldots, N.$$

We will assume that γ is continuous and piecewise $C^M[0,1]$ (for sufficiently large M; see below) with a finite number of pieces, and that the parameterization of Γ by γ is constant speed, that is

$$\|\gamma'(t)\| = C,$$

for some constant $C > 0$, for all but a finite number of $t \in [0,1]$ (the corners of Γ). The nodal points on the boundary Γ then occur at the equally spaced points $\gamma(t_i)$, $i = 1, \ldots, N$. Naturally we must have $\gamma(0) = \gamma(1)$, since Γ is a simple closed curve.

Now let $\{\psi_i\}_{i=1}^N$ be a sequence of basis functions defined periodically on the interval $[0,1]$ (e.g., polynomial splines). These basis functions should have the property that $\psi_i(t_k) = 0$ if $i \neq k$, and $\psi_i(t_i) = 1$. We also assume these functions are bounded and are translations of a fixed basis function ψ, that is, $\psi_j(s) \equiv \psi(s - t_j)$, and that the support of ψ_j lies in a small interval around t_j, say $[t_j - \frac{c_\psi}{N}, t_j + \frac{c_\psi}{N}]$, for some small integer constant c_ψ. We further assume that ψ is normalized so that

$$\int_0^1 \psi(s)\, ds = \frac{1}{N}.$$

Some of the assumptions in the last paragraph can be relaxed at the expense of a more complicated analysis. We note, though, that many common basis functions for boundary element methods appearing in the literature, such as piecewise constant function and piecewise linear hat functions, satisfy these assumptions.

For $s \in [0,1]$, we can approximate $\phi(\gamma(s))$ and $\frac{\partial\phi}{\partial n}(\gamma(s))$ by $\bar{\phi}(\gamma(s))$ and $\overline{\frac{\partial\phi}{\partial n}}(\gamma(s))$ respectively, via

$$\begin{aligned} \bar{\phi}(\gamma(s)) &= \sum_{i=1}^N p_i \psi_i(s), \\ \overline{\frac{\partial\phi}{\partial n}}(\gamma(s)) &= \sum_{i=1}^N q_i \psi_i(s), \end{aligned}$$

where the p_i's and q_i's are unknown coefficients. Note that by the assumptions above, $\bar{\phi}(\gamma(t_i)) = p_i$ and $\overline{\frac{\partial \phi}{\partial n}}(\gamma(t_i)) = q_i$ for all $1 \leq i \leq N$. Then for each nodal value t_k on Γ, $k = 1, \ldots, N$ we approximate (1) with:

$$\begin{aligned} c(\gamma(t_k))p_k &= \sum_{i=1}^{N} p_i \|\gamma'\| \int_{s=0}^{1} K_n(\gamma(s), \gamma(t_k))\psi_i(s)\, ds \\ &\qquad - \sum_{i=1}^{N} q_i \|\gamma'\| \int_{s=0}^{1} K(\gamma(s), \gamma(t_k))\psi_i(s)\, ds. \end{aligned} \tag{2}$$

This approach to discretizing an integral equation is known as *collocation*, which is the most common means of discretizing a BEM integral equation. Note that the second integral in (2) may be improper because $K(t,t)$ is infinite. The integral itself is has a well-defined value, but these singularities make the numerical methods more complicated. Also note that the kernel of the first integral has a removable singularity at the point $t = s$, unless $\gamma(t)$ coincides with a corner.

If we define the $N \times N$ matrices U and V as follows:

$$\begin{aligned} U(i,j) &= \int_{s=0}^{1} K_n(\gamma(s), \gamma(t_i))\psi_j(s)\, ds - \delta_{i,j} c(\gamma(t_j)) \\ V(i,j) &= \int_{s=0}^{1} K(\gamma(s), \gamma(t_i))\psi_j(s)\, ds. \end{aligned} \tag{3}$$

then we can express (2) as the matrix equation

$$A\mathbf{x} = \mathbf{b}. \tag{4}$$

The coefficient matrix A in (4) is an $N \times N$ matrix composed of some columns of U and some columns of V. The vector $\mathbf{x}$ in (4) is composed of the unknowns from the $2N$ coefficients $p_1, \ldots, p_N$ and $q_1, \ldots, q_N$ in (2) (N of which are given as boundary data) and is of length N, and the vector $\mathbf{b}$ is a known vector also of length N. The matrices U and V can be approximated by suitable numerical quadrature formulae. For this paper we assume that the quadrature is exact, but our results hold for approximate quadrature as well.

2.2 Definition of column functions

Consider now the columns of A coming from U. We see that each column from this part of A can be regarded as the evaluation of the function

$$u_j(t) = \int_{s=0}^{1} K_n(\gamma(s), \gamma(t))\psi_j(s)\, ds - \Delta(t - t_j)c(\gamma(t_j)) \qquad (5)$$

for t discretized at the equally spaced points $t_1, \ldots, t_N$. Here,

$$\Delta(s) = \begin{cases} 1 & \text{if } s = 0 \\ 0 & \text{if } s \neq 0 \end{cases} .$$

The columns of A arising from V are evaluations of the function

$$v_j(t) = \int_{s=0}^{1} K(\gamma(s), \gamma(t))\psi_j(s)\, ds \qquad (6)$$

at the same equally-spaced points for the variable t over the interval $[0, 1]$.

3 Bounds on the derivatives and discretizations of the column functions

In this section we will obtain bounds on the derivatives of the functions arising from the BEM equation (1), namely bounds on the functions

$$u_j(t) = \int_{s=0}^{1} K_n(\gamma(s), \gamma(t))\psi_j(s)\, ds - \Delta(t - t_j)c(\gamma(t_j))$$

and

$$v_j(t) = \int_{s=0}^{1} K(\gamma(s), \gamma(t))\psi_j(s)\, ds,$$

defined by (5) and (6) in §2.2.

3.1 Bounds on $u_j(t)$ and its derivatives

Let us begin with the integral in the equation defining $u_j(t)$. We are interested in obtaining bounds on the function

$$\left|\frac{d^p \tilde{u}_j(t)}{dt^p}\right| = 2\pi \left|\frac{d^p \left(\int_{s=0}^{1} K_n(\gamma(s), \gamma(t))\psi_j(s)\, ds\right)}{dt^p}\right|$$

$$
\begin{aligned}
&= 2\pi \left| \int_{s=0}^{1} \frac{d^p \left(K_n(\gamma(s),\gamma(t))\right)}{dt^p} \psi_j(s)\, ds \right| \\
&= \left| \int_{s=0}^{1} \frac{d^p}{dt^p} \left(\frac{\langle \gamma(t)-\gamma(s), n(\gamma(s)) \rangle}{\langle \gamma(t)-\gamma(s), \gamma(t)-\gamma(s) \rangle} \right) \psi_j(s)\, ds \right| \\
&\le \int_{s=0}^{1} \left| \frac{d^p}{dt^p} \left(\frac{\langle \gamma(t)-\gamma(s), n(\gamma(s)) \rangle}{\langle \gamma(t)-\gamma(s), \gamma(t)-\gamma(s) \rangle} \right) \right| |\psi_j(s)|\, ds,
\end{aligned}
$$

for integers $p \ge 0$. To make things more compact, let us introduce some notation that we will use to study the above term. Let us define

$$f \equiv f_s(t) \equiv \langle \gamma(t)-\gamma(s), n(\gamma(s)) \rangle,$$

$$g \equiv g_s(t) \equiv \langle \gamma(t)-\gamma(s), \gamma(t)-\gamma(s) \rangle,$$

and

$$h_s(t) = \frac{f}{g}.$$

We will denote the pth derivative of $f_s(t)$ with respect to t, $\frac{d^p f}{dt^p}$, by $f^{(p)}$, which is not to be confused with the pth power of f, which will be denoted by $(f)^p$. We will use the convention that the zeroth derivative of a function is the function itself, i.e. $f^{(0)} \equiv f$. Let us define the *chord-arc* number of Γ to be

$$\chi_\Gamma = \sup_{\substack{0 \le s,t \le 1 \\ s \ne t}} \frac{|||s-t|||}{\|\gamma(s)-\gamma(t)\|},$$

where $|||s-t|||$ measures the distance from s to t in the metric on $[0,1]$ induced by identifying 0 and 1. We will continue to use the above definition of $|||\cdot|||$ throughout the paper. Because Γ is a piecewise C^M curve which has no cusps, a compactness argument (which we omit) can be used to show that this supremum is finite. The chord-arc number arises often in the analysis of boundary operators; see, e.g., David [11]. We will also need to define the *diameter* of Γ to be

$$\partial\Gamma \equiv \sup_{0 \le s,t \le 1} \|\gamma(t)-\gamma(s)\|.$$

Henceforth C will denote a constant which may change in value from line to line within a proof. With these new definitions now set, let us first concentrate on obtaining bounds on the part of the integrand of $\frac{d^p \tilde{u}_j(t)}{dt^p}$ that contains the derivatives, specifically

$$\left|\frac{d^p h_s(t)}{dt^p}\right| = \left|\frac{d^p}{dt^p}\left(\frac{\langle\gamma(t)-\gamma(s), n(\gamma(s))\rangle}{\langle\gamma(t)-\gamma(s), \gamma(t)-\gamma(s)\rangle}\right)\right|.$$

The first lemma, whose proof is omitted, provides a formula for $\frac{d^p h_s(t)}{dt^p}$ in terms of derivatives of f and g.

Lemma 1 *Let p be an integer, $p \geq 1$. Then using the definitions of f, g and $h_s(t)$ above, we have for $0 \leq s, t \leq 1$,*

$$\frac{d^p h_s(t)}{dt^p} = \sum_{\substack{k,l_1,l_2,\ldots,l_p=0 \\ k+l_1+l_2+\cdots+l_p=p}}^{p} C_{k,l_1,l_2,\ldots,l_p} \frac{f^{(k)} g^{(l_1)} g^{(l_2)} \cdots g^{(l_p)}}{(g)^{p+1}}.$$

Let us further concentrate our efforts by examining the generic term in the formula for $\frac{d^p h_s(t)}{dt^p}$,

$$\frac{f^{(k)} g^{(l_1)} g^{(l_2)} \cdots g^{(l_p)}}{(g)^{p+1}}. \tag{7}$$

Lemma 2 *Let p be an integer, $p \geq 1$, and let $0 \leq s, t \leq 1$, $s \neq t$. Assume that $\gamma(t)$ is not a corner of Γ. Then the generic term of the form*

$$\frac{f^{(k)} g^{(l_1)} g^{(l_2)} \cdots g^{(l_p)}}{(g)^{p+1}}$$

in the formula for $\frac{d^p h_s(t)}{dt^p}$ has the following bound:

$$\left|\frac{f^{(k)} g^{(l_1)} g^{(l_2)} \cdots g^{(l_p)}}{(g)^{p+1}}\right| \leq \frac{C(p) R_\Gamma(p)}{|||t-s|||^{p+1}},$$

where $R_\Gamma(p)$ is defined by (8) below and depends on the geometry of Γ, and $C(p)$ depends only on p.

Proof: We begin by first noting some facts which will be used later on in the proof. First,

$$g^{(l)} = \sum_{m=0}^{l} \binom{l}{m} \langle(\gamma(t)-\gamma(s))^{(m)}, (\gamma(t)-\gamma(s))^{(l-m)}\rangle.$$

(where, as usual, these derivatives are with respect to t) and for $k \geq 1$,

$$f^{(k)} = \langle \gamma^{(k)}(t), n(\gamma(s)) \rangle.$$

Let $r_0 = \|\gamma(t) - \gamma(s)\|$ and $r_k = \|\gamma^{(k)}(t)\|$ for $k \geq 1$. Then we easily obtain the following bounds using the above equations and the Cauchy-Schwartz inequality:

$$\left| f^{(k)} \right| \leq r_k$$

and

$$\left| g^{(l)} \right| \leq C(l) \max_{0 \leq m \leq l} r_m r_{l-m}.$$

In particular, $\left| g^{(0)} \right| = r_0^2$ and $\left| g^{(1)} \right| \leq 2 r_0 r_1$. We can use these inequalities to bound the generic term in the formula for $\frac{d^p h_s(t)}{dt^p}$. Observe that the denominator is exactly r_0^{2p+2}. Thus,

$$\left| \frac{f^{(k)} g^{(l_1)} g^{(l_2)} \cdots g^{(l_p)}}{(g)^{p+1}} \right| \leq C(p) \max \frac{r_k (r_{m_1} r_{l_1 - m_1}) \cdots (r_{m_p} r_{l_p - m_p})}{r_0^{2p+2}}$$

where the max is taken over all choices for all the integers m_i such that $0 \leq m_i \leq l_i$. Since $k + l_1 + \cdots + l_p = p$, this can be further bounded by

$$\left| \frac{f^{(k)} g^{(l_1)} g^{(l_2)} \cdots g^{(l_p)}}{(g)^{p+1}} \right| \leq C(p) \max \left\{ \frac{r_{\lambda_1} \cdots r_{\lambda_{2p+1}}}{r_0^{2p+2}} : \lambda_1 + \cdots + \lambda_{2p+1} = p \right\}$$

where now the max is taken over all nonnegative integer choices of the λ_i's adding up to p. Since there are $2p+1$ nonnegative integral λ_i's adding up to p, at least $p+1$ of them must be zero. Thus, if we delete exactly $p+1$ λ_i's that are zeros, we can rewrite this bound as

$$\begin{aligned} \left| \frac{f^{(k)} g^{(l_1)} g^{(l_2)} \cdots g^{(l_p)}}{(g)^{p+1}} \right| &\leq C(p) \max \left\{ \frac{r_0^{p+1} r_{\lambda_1} \cdots r_{\lambda_p}}{r_0^{2p+2}} : \lambda_1 + \cdots + \lambda_p = p \right\} \\ &= C(p) \max \left\{ \frac{r_{\lambda_1} \cdots r_{\lambda_p}}{r_0^{p+1}} : \lambda_1 + \cdots + \lambda_p = p \right\}. \end{aligned}$$

By the definition of the chord-arc number, χ_Γ, we can deduce that for $0 \leq s, t \leq 1$, $s \neq t$,

$$\frac{1}{\|\gamma(t) - \gamma(s)\|^{p+1}} \leq \frac{\chi_\Gamma^{p+1}}{|||s - t|||^{p+1}}.$$

Thus, the left-hand side of the above equation may be bounded by

$$C(p)\max\left\{\frac{r_{\lambda_1}\cdots r_{\lambda_p}\chi_\Gamma^{p+1}}{|||s-t|||^{p+1}} : \lambda_1+\cdots+\lambda_p = p\right\}.$$

To get an overall bound, let us define

$$R_0 = \partial\Gamma$$

and for $k \geq 1$,

$$R_k = \sup_{t\in[0,1]} \|\gamma^{(k)}(t)\|$$

where the sup is taken over values of t such that $\gamma(t)$ is not a corner. Clearly $r_k \leq R_k$ for all k. Let us define

$$R_\Gamma(p) = (\chi_\Gamma)^{p+1}\max\{R_{\lambda_1}\cdots R_{\lambda_p} : \lambda_1+\cdots+\lambda_p = p\} \tag{8}$$

where each λ_i is a nonnegative integer. Then finally we can conclude that

$$\left|\frac{f^{(k)}g^{(l_1)}g^{(l_2)}\cdots g^{(l_p)}}{(g)^{p+1}}\right| \leq \frac{C(p)R_\Gamma(p)}{|||s-t|||^{p-1}}.$$

□

A different choice for $C(p)$ then yields a bound on $\left|\frac{d^p h_s(t)}{dt^p}\right|$.

Lemma 3 *Let p be an integer, $p \geq 1$, and let $0 \leq s,t \leq 1$, $s \neq t$ and $\gamma(t)$ is not a corner of Γ. Then we have the following bounds on the pth derivative of $K_n(\gamma(s),\gamma(t))$:*

$$\left|\frac{d^p K_n(\gamma(s),\gamma(t))}{dt^p}\right| \leq \frac{C(p)\,R_\Gamma(p)}{|||s-t|||^{p+1}},$$

where $R_\Gamma(p)$ is defined by (8).

Lemma 3 can now be used to prove the following theorem:

Theorem 1 *Let p, j and N be integers such that $p \geq 0$ and $1 \leq j \leq N$. Let $0 \leq t \leq 1, t \neq t_j$. Assume $\gamma(t)$ is not a corner of Γ. Then*

$$\begin{aligned}\left|\frac{d^p\tilde{u}_j(t)}{dt^p}\right| &= \left|\frac{d^p\left(\int_{s=0}^1 K_n(\gamma(s),\gamma(t))\psi_j(s)\,ds\right)}{dt^p}\right| \\ &\leq \frac{C(p)C(\psi)\,R_\Gamma(p)}{N\left[\min_{s\in\operatorname{supp}\psi_j}|||t-s|||\right]^{p+1}},\end{aligned}$$

where $R_\Gamma(p)$ is defined by (8), $C(p)$ depends only on p, and $C(\psi)$ depends only on ψ.

Proof: Since

$$\begin{aligned}
\left|\frac{d^p \tilde{u}_j(t)}{dt^p}\right| &\leq \int_{s \in \operatorname{supp} \psi_j} \left|\frac{d^p}{dt^p}\left(h_s(t)\right)\right| |\psi_j(s)|\, ds \\
&\leq \left(\sup_{0\leq s\leq 1} |\psi_j(s)|\right) \int_{s \in \operatorname{supp} \psi_j} \left|\frac{d^p}{dt^p}\left(h_s(t)\right)\right| ds,
\end{aligned}$$

we can apply Lemma 3 to the above equation, provided that $t \notin \operatorname{supp} \psi_j$. Then, noting that $\sup_{0\leq s\leq 1} |\psi_j(s)| = \sup_{0\leq s\leq 1} |\psi(s)|$, we have

$$\left|\frac{d^p \tilde{u}_j(t)}{dt^p}\right| \leq C(p) \left(\sup_{0\leq s\leq 1} |\psi(s)|\right) R_\Gamma(p) \int_{s \in \operatorname{supp} \psi_j} \frac{1}{|||t-s|||^{p+1}}\, ds.$$

Let $2c_\psi \sup_{0\leq s\leq 1} |\psi(s)|$ be denoted as $C(\psi)$. If $t \notin \operatorname{supp} \psi_j$, and $\operatorname{supp} \psi_j = \frac{2c_\psi}{N}$ we have

$$\begin{aligned}
\left|\frac{d^p \tilde{u}_j(t)}{dt^p}\right| &= \left|\frac{d^p \left(\int_{s=0}^1 K_n(\gamma(s), \gamma(t)) \psi_j(s)\, ds\right)}{dt^p}\right| \\
&\leq \frac{C(p) C(\psi)\, R_\Gamma(p)}{N \left[\min_{s \in \operatorname{supp} \psi_j} |||t-s|||\right]^{p+1}}.
\end{aligned}$$

□

3.2 Bounds on $v_j(t)$ and its derivatives

In this subsection, we will apply the same ideas used to obtain the bounds for $\left|\frac{d^p u_j(t)}{dt^p}\right|$ in Theorem 1 to the function

$$\begin{aligned}
\left|\frac{d^p v_j(t)}{dt^p}\right| &= \left|\frac{d^p \left(\int_{s=0}^1 K(\gamma(s), \gamma(t)) \psi_j(s)\, ds\right)}{dt^p}\right| \\
&= \left|\int_{s=0}^1 \frac{d^p \left(K(\gamma(s), \gamma(t))\right)}{dt^p} \psi_j(s)\, ds\right|
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\left|\int_{s=0}^{1} \frac{d^p}{dt^p}\left(\ln\|\gamma(t)-\gamma(s)\|\right)\psi_j(s)\,ds\right| \\
&\leq \frac{1}{4\pi}\int_{s=0}^{1}\left|\frac{d^p}{dt^p}\left(\ln\langle\gamma(t)-\gamma(s),\gamma(t)-\gamma(s)\rangle\right)\right|\left|\psi_j(s)\right|\,ds.
\end{aligned}
$$

In this case we will initially concentrate our efforts on the term

$$
\left|\frac{d^p}{dt^p}\left(\ln\langle\gamma(t)-\gamma(s),\gamma(t)-\gamma(s)\rangle\right)\right|,
$$

for $p \geq 1$. In case $p = 1$, for $0 \leq t \leq 1$, $s \in \operatorname{supp}\psi_j$ we have

$$
\frac{d}{dt}\ln\langle\gamma(t)-\gamma(s),\gamma(t)-\gamma(s)\rangle = \frac{2\langle\gamma'(t),\gamma(t)-\gamma(s)\rangle}{\langle\gamma(t)-\gamma(s),\gamma(t)-\gamma(s)\rangle}. \tag{9}
$$

Observe that this is essentially the same form as $u_j(t)$, except that the numerator contains $\gamma'(t)$ instead of $n(\gamma(s))$. To obtain higher derivatives of the kernel function, we now differentiate (9). Denoting the numerator of (9) as f and the numerator as g, we now have for $k \geq 0$

$$
f^{(k)} = \sum_{l=0}^{k} D_{l,k}\langle\gamma^{(l+1)}(t),(\gamma(t)-\gamma(s))^{k-l}\rangle,
$$

for some constants $D_{l,k}$. Thus, reintroducing the quantities $r_0, r_1, \ldots$, and using similar arguments as in the last section, for $k \geq 0$ we have

$$
|f^{(k)}| \leq C(k)\max\{r_{\lambda_1}r_{\lambda_2} : \lambda_1+\lambda_2 = k+1\}.
$$

From this we deduce that

$$
\left|\frac{f^{(k)}g^{(l_1)}g^{(l_2)}\cdots g^{(l_p)}}{(g)^{p+1}}\right| \leq C(p)\max\{\frac{r_{\lambda_1}\cdots r_{\lambda_{p+1}}}{r_0^{p+1}} : \lambda_1+\cdots+\lambda_{p+1} = p+1\}.
$$

Now we reintroduce $R_\Gamma(p)$ defined by (8), and we recall that the above expression is actually an expression for a term of the $(p+1)$st derivative of the kernel function. We can therefore conclude that

$$
\left|\frac{d^p}{dt^p}\left(\ln\langle\gamma(t)-\gamma(s),\gamma(t)-\gamma(s)\rangle\right)\right| \leq C(p)\frac{R_\Gamma(p)}{\chi_\Gamma |||s-t|||^p}.
$$

Continuing the arguments from last section leads to the following theorem:

Theorem 2 *Let p, j, and N be integers such that $p \geq 1$ and $1 \leq j \leq N$. Let $0 \leq t \leq 1$, and assume $\gamma(t)$ is not a corner of Γ. Then*

$$\left|\frac{d^p v_j(t)}{dt^p}\right| \leq \frac{C(p)C(\psi)R_\Gamma(p)}{N\chi_\Gamma\left[\min_{s\in\operatorname{supp}\psi_j}|||t-s|||\right]^p}$$

where $R_\Gamma(p)$ is defined by (8), $C(p)$ depends only upon p and $C(\psi)$ depends only on the basis function ψ, but not upon N or the geometry of the domain.

e have just shown that the functions $u_j(t)$ and $v_j(t)$ have as many derivatives as γ does except in the support of ψ_j, and as a result all its derivatives are piecewise bounded, except at a finite number of points. Here the derivatives have a jump discontinuity as well.

3.3 Lower bounds on the discretization of $u_j(t)$

In this subsection we will obtain lower bounds on the 2-norm of the columns of the system matrix A taken from the matrix U defined in §2.1 by (3). Let us denote the discretization of $u_j(t)$ by $\mathbf{u}^j$. Note that

$$\begin{aligned}
\|\mathbf{u}^j\|^2 &= \sum_{i=1}^{N}(\mathbf{u}_i^j)^2 \\
&\geq (\mathbf{u}_j^j)^2 \\
&\geq (u_j(t_j))^2 \\
&\geq \left(\int_{s=0}^{1} K_n(\gamma(t_j),\gamma(s))\psi_j(s)\,ds - c(\gamma(t_j))\right)^2,
\end{aligned}$$

where $c(\gamma(t_j))$ measures the angle of Γ at $\gamma(t_j)$. Since we are assuming that Γ has no cusps, there is a minimum nonzero value for $c(\gamma(t)), 0 \leq t \leq 1$. Also, it can be shown that the first term in parentheses in the expression for $\mathbf{u}^j$ above goes to 0 as $N \to \infty$ like $O(\frac{1}{N})$. So, for N sufficiently large,

$$\left|\int_{s=0}^{1} K_n(\gamma(t_j),\gamma(s))\psi_j(s)\,ds\right| \leq \frac{|c(\gamma(t_j))|}{2},$$

so that

$$\|\mathbf{u}^j\|^2 \geq \frac{(c(\gamma(t_j)))^2}{4} > 0.$$

We have just proved the following theorem:

Theorem 3 *Let N and j be integers such that $1 \leq j \leq N$. Let $\mathbf{u}^j$ be a vector of length N, obtained from discretizing $u_j(t)$ given by*

$$u_j(t) = \int_{s=0}^{1} K_n(\gamma(t), \gamma(s))\psi_j(s)\,ds - c(\gamma(t)),$$

at the points $\frac{i-1}{N}$, $1 \leq i \leq N$. Then for N sufficiently large,

$$\|\mathbf{u}^j\| \geq \frac{|c(\gamma(t_j))|}{2},$$

where $\|\mathbf{u}^j\|$ represents the 2-norm of the vector $\mathbf{u}^j$.

3.4 Lower bounds on the discretization of $v_j(t)$

We begin by defining the N-vector $\mathbf{v}^j$ to be the discretization of $v_j(t)$, which is also the jth column of the matrix V defined in §2.1 by (3). In this subsection, we are interested in obtaining a lower bound on $\|\mathbf{v}^j\|$. Recall that

$$v_j(t_i) = \mathbf{v}_i^j = \frac{1}{2\pi}\int_{s=0}^{1} \ln(\|\gamma(t_i) - \gamma(s)\|)\psi_j(s)\,ds.$$

For a fixed j, let us restrict our attention to those i that satisfy

$$|||s - t_i||| < \frac{1}{\|\gamma'\|}, \tag{10}$$

for all $s \in \text{supp}\psi_j$. Recall, also, that $\|\gamma'\|$ is a constant. In order for (10) to be satisfied, we must have

$$\max\{|||t_j - \frac{c_\psi}{N} - t_i|||, |||t_j + \frac{c_\psi}{N} - t_i|||\} < \frac{1}{\|\gamma'\|},$$

or

$$\max\{(j - i - c_\psi) \bmod N, (j - i + c_\psi) \bmod N\} < \frac{N}{\|\gamma'\|}.$$

That is, i must be an element of the set $\mathcal{I}$ defined by

$$\begin{aligned}\mathcal{I} \;=\; \Big\{ & j - c_\psi - \lfloor\frac{N}{\|\gamma'\|}\rfloor, j - c_\psi - \lfloor\frac{N}{\|\gamma'\|}\rfloor + 1, \ldots, j, \\ & \ldots, j + c_\psi + \lfloor\frac{N}{\|\gamma'\|}\rfloor - 1, j + c_\psi + \lfloor\frac{N}{\|\gamma'\|}\rfloor \Big\} \bmod N.\end{aligned}$$

The cardinality of $\mathcal{I}$ is $2(c_\psi + \lfloor\frac{N}{\|\gamma'\|}\rfloor) + 1 \equiv 2(c_\psi + M_\mathcal{I}) + 1$. By taking N sufficiently large, we can guarantee that $\mathcal{I}$ is nonempty. Let us denote the interval $(t_{j-c_\psi-M_\mathcal{I}}, t_{j+c_\psi+M_\mathcal{I}})$ by I. As usual, all subscripts are to be taken modulo N. In order to avoid a more complicated analysis, we will assume, for this section only, that ψ (and hence ψ_j) is a nonnegative function. Then, for any $s \in \mathrm{supp}\psi_j$ and for any $t \in I$, by the mean value theorem we have

$$\|\gamma'\| \geq \frac{\|\gamma(s) - \gamma(t)\|}{|||s-t|||},$$

so that

$$\|\gamma(s) - \gamma(t)\| \leq \|\gamma'\| \, |||s-t|||.$$

Since $t \in I$, we can apply (10) to deduce that

$$\|\gamma(s) - \gamma(t)\| < 1.$$

As a result, for any $i \in \mathcal{I}$

$$\begin{aligned}
\left|\mathbf{v}_i^j\right| &= \frac{1}{2\pi}\left|\int_{s\in\mathrm{supp}\psi_j} \ln(\|\gamma(s)-\gamma(t)\|)\psi_j(s)\,ds\right| \\
&\geq \frac{1}{2\pi}\left|\int_{s\in\mathrm{supp}\psi_j} \ln(\|\gamma'\|\,|||s-t_i|||)\psi_j(s)\,ds\right| \\
&\geq \frac{1}{2\pi}\inf_{s\in\mathrm{supp}\psi_j}\{|\ln(\|\gamma'\|\,|||s-t_i|||)|\}\int_{s\in\mathrm{supp}\psi_j}\psi_j(s)\,ds \\
&\geq \frac{1}{2\pi N}\min\left\{\left|\ln(\|\gamma'\|\,|||t_j+\frac{c_\psi}{N}-t_i|||)\right|, \left|\ln(\|\gamma'\|\,|||t_j-\frac{c_\psi}{N}-t_i|||)\right|\right\}.
\end{aligned}$$

Thus, for N sufficiently large, we see that

$$\begin{aligned}
\|\mathbf{v}^j\|^2 &= \sum_{i=1}^{N}(\mathbf{v}_i^j)^2 \\
&\geq \sum_{i\in\mathcal{I}}(\mathbf{v}_i^j)^2 \\
&\geq \frac{1}{4\pi^2N^2}\sum_{i\in\mathcal{I}}\Bigg[\min\Bigg\{\left|\ln(\|\gamma'\|\,|||t_j+\frac{c_\psi}{N}-t_i|||)\right|, \\
&\qquad\qquad \left|\ln(\|\gamma'\|\,|||t_j-\frac{c_\psi}{N}-t_i|||)\right|\Bigg\}\Bigg]^2
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{1}{4\pi^2 N^2} \sum_{i=c_\psi}^{M_{\mathcal{I}}} \left[\ln(i \frac{\|\gamma'\|}{N}) \right]^2 \\
&\geq \frac{1}{4\pi^2 N^2} \sum_{i=c_\psi}^{\frac{1}{2}\frac{N}{\|\gamma'\|}} \left[\ln(i \frac{\|\gamma'\|}{N}) \right]^2 \\
&\geq \frac{1}{4\pi^2 N^2} \sum_{i=\frac{1}{4}\frac{N}{\|\gamma'\|}}^{\frac{1}{2}\frac{N}{\|\gamma'\|}} \left[\ln(i \frac{\|\gamma'\|}{N}) \right]^2 \\
&\geq \frac{1}{4\pi^2 N^2} \int_{\frac{1}{4}}^{\frac{1}{2}} (\ln(x))^2 \frac{N}{\|\gamma'\|} \, dx \\
&\geq \frac{C}{N\,\|\gamma'\|},
\end{aligned}
$$

where C is a constant independent of the geometry of Γ. We have just proven the following theorem:

Theorem 4 *Let N and j be integers such that $1 \leq j \leq N$. Let $\mathbf{v}^j$ be a vector of length N, obtained from discretizing $v_j(t)$ given by*

$$
v_j(t) = \int_{s=0}^{1} K(\gamma(t), \gamma(s))\psi_j(s)\, ds,
$$

at the points $\frac{i-1}{N}$, $1 \leq i \leq N$. Then for N sufficiently large, $\exists$ a constant C independent of the geometry of the domain, Γ, such that

$$
\|\mathbf{v}^j\|^2 \geq \frac{C}{N\,\|\gamma'\|}.
$$

4 Wavelets

Unless otherwise specified, we will continue to use the definitions and notations introduced in this section throughout the rest of the paper.

4.1 The discrete wavelet transform

In this subsection we will define what we shall term the *discrete wavelet transform* of a function $f \in \mathcal{L}_2(\mathrm{I\!R})$, which involves creating a basis for $\mathcal{L}_2(\mathrm{I\!R})$

out of wavelets. We are interested in the compact orthonormal wavelets that give rise to a basis for $\mathcal{L}_2(\mathrm{IR})$, constructed by Daubechies [9], and [10]. These are defined as follows: consider the functions φ, and w on IR which satisfy

$$\varphi(x) = \sqrt{2}\sum_{k=1}^{2M} h_k\,\varphi(2x-k+1),$$

and

$$w(x) = \sqrt{2}\sum_{k=1}^{2M} g_k\,\varphi(2x-k+1),$$

where

$$g_k = (-1)^{k-1} h_{2M-k+1}, \qquad k = 1, 2, \ldots, 2M,$$

and φ is normalized by

$$\int_{-\infty}^{\infty} \varphi(x)\,dx = 1.$$

The *(Daubechies) wavelet coefficients of order* M, $\{h_k\}_{k=1}^{2M}$, (alternatively $\{g_k\}_{k=1}^{2M}$) are chosen so that the functions

$$w_k^{\nu}(x) = \sqrt{2^{\nu}}\,w(2^{\nu}x-k+1), \qquad \nu, k \in \mathcal{Z} \tag{11}$$

form an orthonormal basis of $\mathcal{L}^2(\mathrm{IR})$. Another condition imposed upon the h_k's is that the first $M-1$ moments of w should vanish, i.e.,

$$\int_{-\infty}^{\infty} w(x)\,x^p\,dx = 0, \qquad p = 0, 1, \ldots, M-1. \tag{12}$$

The function φ is referred to as the *(Daubechies) scaling function of order* M, and the function w is referred to as the *(Daubechies) wavelet of order* M. M itself is referred to as the *wavelet order*. Daubechies showed that for each integer M, there exist wavelet coefficients of order M that guarantee the existence of such scaling functions and wavelets. In fact, the scaling and wavelet functions are compact and the support of w and φ lie in the interval $[0, 2M-1]$.

Given a function $f(x)$ in $\mathcal{L}_2(\mathrm{IR})$, the *discrete wavelet transform of* f is obtained as follows:

$$f(x) = \sum_{\nu,k\in\mathcal{Z}} \hat{d}_k^{\nu}\,w_k^{\nu}(x),$$

where

$$\hat{d}_k^\nu = \int_{-\infty}^{\infty} f(x)\, w_k^\nu(x)\, dx \tag{13}$$

for all integers ν and k. The elements $\hat{d}_k^\nu$ will be referred to as the *kth discrete wavelet representation of f of order ν*. Alternatively, for any integer β, we can express $f(x)$ as

$$f(x) = \sum_{i=-\infty}^{\infty} \hat{s}_i^\beta\, \varphi_i^\beta(x) + \sum_{\nu=\beta}^{\infty} \sum_{i=-\infty}^{\infty} \hat{d}_i^\nu\, w_i^\nu(x). \tag{14}$$

Here $\hat{s}_k^\beta$ is the *kth discrete scaling representation of f of order β* defined via

$$\hat{s}_k^\beta = \int_{-\infty}^{\infty} f(x)\, \varphi_k^\beta(x)\, dx, \tag{15}$$

where the translations and dilations of φ, namely φ_k^β are defined in a manner similar to that of equation (11):

$$\varphi_k^\beta(x) = \sqrt{2^\beta}\varphi(2^\beta x - k + 1), \tag{16}$$

for any integers β and k. In the expression (14), $\{\varphi_i^\beta\}_{i\in\mathcal{Z}}$ are orthogonal.

4.2 Periodizing the discrete wavelet transform

In the last section we described the wavelet transform for functions in $\mathcal{L}_2(\mathsf{IR})$. Recall that our functions are defined on $[0, 1]$ only. Indeed, because the boundary Γ is a closed curve, our functions u_j and v_j are most easily understood as periodic functions on IR (so that $u_j(x) = u_j(x + k)$ for all integers k). It turns out that the discrete wavelet transform for such a function has a special property: all the wavelet representation entries above order 0 vanish. Thus, it turns out that the wavelet transform of a periodic function can be represented as follows: there one scaling function entry, namely, (14) in the case that $\beta = 0, k = 0$. Then the remaining entries in the representation are given by (12) in the case that $\nu \geq 0$ and $1 \leq k \leq 2^\nu$. This is elucidated further in the next subsection.

4.3 The fast wavelet transform

One usually thinks of wavelets as presented in §4.1, that is, as functions that give rise to a basis for $\mathcal{L}^2(\mathrm{IR})$. For our purposes we view wavelets primarily as functions that give rise to a linear map from IR^N to IR^N. We consider the the same orthonormal wavelets as introduced in §4.1 above, but apply them numerically as outlined by Beylkin, Coifman, and Rokhlin (BCR) [5].

We begin by introducing more notation and terminology, in addition to what we will reuse from §4.1. Let α be a positive integer, and set $N = 2^\alpha$. Our goal in this section is to define the *fast wavelet transform*, a linear map $W : \mathrm{IR}^N \longmapsto \mathrm{IR}^N$. With this in mind, let $\mathbf{f} \in \mathrm{IR}^N$ be a given vector, which we view as being obtained from a function $f(s), s \in [0,1]$, discretized at N equally spaced points $0, \frac{1}{N}, \ldots, \frac{N-1}{N}$. We will speak of the function $f(s)$ as being the *parent function of* $\mathbf{f}$. We will refer to the ith element of $\mathbf{f}$ as $\mathbf{f}_i$. Let the wavelet order M be a fixed integer parameter (in our case $1 \le M \le 10$) and let $\{h_i\}_{i=1}^{2M}$ be the wavelet coefficients of order M. Finally, let $s_k^\alpha = \mathbf{f}_k$ for $k = 1, \ldots, N$. We now define the *scaling representations of* $\mathbf{f}$ *at level* ν recursively via the formula

$$s_k^\nu = \sum_{i=1}^{2M} h_i \, s^{\nu+1}_{i+2k-2 \bmod 2^{\nu+1}} \qquad k = 1, \ldots 2^\nu, \tag{17}$$

and we similarly define the *wavelet representations of* $\mathbf{f}$ *at level* ν by

$$d_k^\nu = \sum_{i=1}^{2M} g_i \, s^{\nu+1}_{i+2k-2 \bmod 2^{\nu+1}} \qquad k = 1, \ldots, 2^\nu. \tag{18}$$

Here the *wavelet level* ν runs from $\alpha - 1$ down to 0. We will always express the wavelet level by Greek letters.

Also, given both the wavelet representations of $\mathbf{f}$ at level ν and the scaling representations of $\mathbf{f}$ at level ν we can obtain the scaling representations of $\mathbf{f}$ at level $\nu + 1$ as follows:

$$\begin{aligned} s_{2k}^{\nu+1} &= \sum_{i=1}^{M} h_{2i} \, s^\nu_{k-i+1 \bmod 2^\nu} + \sum_{i=1}^{M} g_{2i} \, d^\nu_{k-i+1 \bmod 2^\nu} \\ s_{2k-1}^{\nu+1} &= \sum_{i=1}^{M} h_{2i-1} \, s^\nu_{k-i+1 \bmod 2^\nu} + \sum_{i=1}^{M} g_{2i-1} \, d^\nu_{k-i+1 \bmod 2^\nu} \end{aligned} \tag{19}$$

for $k = 1, \ldots, 2^\nu$, and for $\nu = 0, \ldots, \alpha - 1$.

We are now in a position to define the fast wavelet transform of $\mathbf{f}$, namely the N-vector $W\mathbf{f}$ which will henceforth be denoted by $\mathbf{w}$, by

$$W\mathbf{f} \equiv \mathbf{w} = \{s_1^0, d_1^0, d_1^1, d_2^1, \ldots, d_{2^{\alpha-2}}^{\alpha-2}, d_1^{\alpha-1}, d_2^{\alpha-1}, \ldots, d_{2^{\alpha-1}}^{\alpha-1}\}.$$

The fact that (19) is the inverse of (17) and (18) is a property of the values $\{h_i\}_{i=1}^{2M}$ (see Daubechies [10]). In fact, (19) is the orthogonal transformation of (17) and (18), and as a result, with the proper scaling, the fast wavelet transform W is an orthonormal matrix so that $W^{-1} = W^T$ where W^T is the transpose of W.

As can be seen from the formulae (17) and (18), if $s_k^{\nu+1}$ were known for $k = 1, 2, \ldots, 2^{\nu+1}$, then s_k^ν could be computed numerically with only $2M$ multiplications and $2M - 1$ additions for each $k = 1, \ldots, 2^\nu$, and similarly for the d_k^ν's. Hence $\mathbf{w}$ could be computed with a total of $(2M)(N-1)$ multiplications and $(2M-1)(N-1)$ additions, or an $O(N)$ cost per transform. Also, it is easy to see that given $\mathbf{w}$, one can retrieve $\mathbf{f}$ (by applying (19) repeatedly) at an $O(N)$ cost as well.

5 Bounds on the fast wavelet transform for the BEM matrix

In this section we will state and prove some theorems regarding the size of the fast wavelet representations of the columns of the BIEM matrix A defined in §2.1 by (3) and (4). Initially, in §5.1, we will phrase the main theorem in terms of a general vector $\mathbf{f}$ and parent function $f(t)$, but the vectors and functions we have in mind are those given by (5) and (6). In §5.2 we will apply this theorem to those specified vectors and functions. Unless otherwise stated, all subscripts are taken to be modulo N.

5.1 General theorems

As usual, we will denote the fast wavelet transform of $\mathbf{f}$, $W\mathbf{f}$, by $\mathbf{w}$. The main theorem gives a bound on the wavelet representations (d_k^ν) of a vector $\mathbf{f}$ in terms of the derivatives of the parent function $f(t)$, for $k = 1, 2, \ldots, 2^\nu$ and for $\nu = \alpha - 1, \alpha - 2, \ldots, 0$, for some given α. However, in order to prove this theorem, we will need the following lemma:

Lemma 4 *Let M (wavelet order) and α (number of wavelet levels) be given. Let $p : \mathbb{R} \mapsto \mathbb{R}$ be a polynomial of degree $k \le M-1$. Then the function*

$$q(x) = \sum_{i=-\infty}^{\infty} p(i)\varphi_i^\alpha(x)$$

is a polynomial in x also of degree k (exactly). Here

$$\varphi_i^\alpha(x) = \sqrt{2^\alpha}\varphi(2^\alpha x - k + 1),$$

and $\varphi(x)$ is the scaling function of order M as defined by (16) in §4.1.

Proof: Let $r : \mathbb{R} \mapsto \mathbb{R}$ be a polynomial of degree $k \le M$. We can express the polynomial $r(x)$ as

$$r(x) = \sum_{i=-\infty}^{\infty} a_i \varphi_i^\alpha(x),$$

where the coefficients a_i are independent of x, and the above equation holds for all $x \in \mathbb{R}$. Since $r(x) \notin \mathcal{L}_2(\mathbb{R})$, this sum is interpreted pointwise. Its existence follows because $r(x) \in \mathcal{L}_2(\mathbb{R})$ when restricted to any finite interval.

Let us concentrate on the coefficients a_i for now. We know that

$$\begin{aligned} a_i &= \int_{-\infty}^{\infty} r(x)\varphi_i^\alpha(x)\,dx \\ &= \int_{-\infty}^{\infty} r(x)\varphi_0^\alpha(x - 2^{-\alpha}i)\,dx, \end{aligned}$$

so that we can define the function

$$a(y) = \int_{-\infty}^{\infty} r(x)\varphi_0^\alpha(x-y)\,dx.$$

We wish to show that if $r(x)$ is a polynomial of degree $k \le M-1$ then $a(y)$ is also a polynomial of degree k. Since the set $\left\{1, x, x^2, \ldots, x^{M-1}\right\}$ constitutes a basis for the space of all polynomials of degree $\le M-1$ defined on $\mathbb{R}$, we can, without loss of generality, consider $r(x) = x^k$. Then

$$\begin{aligned} a(y) &= \int_{-\infty}^{\infty} x^k \varphi_0^\alpha(x-y)\,dx \\ &= \int_{-\infty}^{\infty} (x+y)^k \varphi_0^\alpha(x)\,dx \\ &= \sum_{l=0}^{k} \frac{k!}{l!\,(k-l)!} y^l \int_{-\infty}^{\infty} x^{k-l}\varphi_0^\alpha(x)\,dx, \end{aligned}$$

and because φ_0^α is bounded and has compact support, and because the leading coefficient satisfies:

$$\int_{-\infty}^{\infty} \varphi_0^\alpha(x)\, dx = \frac{1}{\sqrt{2^\alpha}},$$

$a(y)$ is a polynomial in y of degree k. Thus, there exists a linear map $T : V \mapsto V$ defined by $T(r) = a$, where r and a are defined above and where V is the space of all polynomials of degree $\leq M-1$ defined on $\mathbb{R}$. We have also shown that T is of full rank, that is, T is invertible. As a result, given a polynomial $p(y)$ of degree $k \leq M-1$, there exists a polynomial $q(x)$ of degree k such that

$$p(y) = \int_{-\infty}^{\infty} q(x)\varphi_0^\alpha(x-y)\, dx,$$

and

$$q(x) = \sum_{i=-\infty}^{\infty} p(i)\varphi_i^\alpha(x),$$

as required.

□

We are now in a position to state and prove our main theorem of this subsection.

Theorem 5 *Let M (wavelet order) and α (number of wavelet levels) be given. Let $\mathbf{f} \in \mathbb{R}^N$, be a given vector, and let $f(t)$ be its corresponding parent function, i.e. $\mathbf{f}_i = f\left(\frac{i}{N}\right)$ for $i = 1, 2, \ldots, N$. Assume that $f \in C^M[0,1]$, with*

$$\lim_{t\to 0^+} \frac{d^k f(t)}{dt^k} = \lim_{t\to 1^-} \frac{d^k f(t)}{dt^k}, \qquad k = 0, 1, \ldots, M.$$

That is, the periodic extension of f should be an element of $C^M(\mathbb{R})$. Then we can obtain the following upper bound on the entries of the fast wavelet transform of $\mathbf{f}$, namely the elements of $\mathbf{w}$:

$$|d_k^\nu| \leq \frac{\sqrt{2^\alpha}}{(2^\nu)^M\sqrt{2^\nu}} C(M) \sup_{x\in\mathbb{R}}\left(|w(x)|\right) \sup_{x\in\mathbb{R}}\left(|\varphi(x)|\right) \sup_{\eta\in I_k^\nu}\left(\left|f^{(M)}(\eta)\right|\right),$$

where

$$C(M) = \frac{(4M-1)\,(2M-1)^{M+1}}{M!},$$

and I_k^ν *is the interval*

$$\left(\frac{2^{\alpha-\nu}(k-1)-2M+1}{2^\alpha}, \frac{2^{\alpha-\nu}(k-1)+2^{\alpha-\nu}(2M-1)}{2^\alpha}\right),$$

for $k = 1, 2, \ldots, 2^\nu$ *and for* $\nu = \alpha-1, \alpha-2, \ldots, 0$.

Proof: It can be shown that if we define the function $\bar{f} : \mathbb{R} \mapsto \mathbb{R}$ by

$$\bar{f}(x) = \sum_{i=1}^{2^\alpha} \mathbf{f}_i \varphi_i^\alpha(x),$$

then the fast wavelet representations of the vector $\mathbf{f}$ are related to the discrete wavelet representations of $\bar{f}(x)$ in the following way:

$$d_k^\nu = \int_{-\infty}^{\infty} \bar{f}(x) w_k^\nu(x)\, dx,$$

where φ_i^α and w_k^ν are defined by (11) and (16) in §4.1. Substituting our expression for $\bar{f}$ into the expression for d_k^ν yields

$$d_k^\nu = \int_{-\infty}^{\infty} \sum_{i=1}^{2^\alpha} \mathbf{f}_i \varphi_i^\alpha(x) w_k^\nu(x)\, dx.$$

Let us restrict our attention to $k = 1, 2, \ldots, 2^\nu$ and to $\nu = \alpha-1, \alpha-2, \ldots, 0$. If we apply Taylor's Theorem to the parent function $f(t)$, expanding about the point $\frac{k^\star-1}{2^\alpha}$ and evaluating at the point $\frac{i-1}{2^\alpha}$, we see that

$$f(x_i) = \mathbf{f}_i = \sum_{j=0}^{M-1} c_j (i-k^\star)^j + r_i,$$

where for $j = 0, 1, \ldots, M-1$

$$c_j = \frac{1}{(2^\alpha)^j\, j!} f^{(j)}\left(\frac{k^\star-1}{2^\alpha}\right),$$

and

$$r_i = \frac{(i-k^\star)^M}{(2^\alpha)^M\, M!} f^{(M)}(\eta_{i,k^\star}),$$

for some $\eta_{i,k^\star}$ between $\frac{i-1}{2^\alpha}$ and $\frac{k^\star-1}{2^\alpha}$. Here we choose

$$k^\star = 2^{\alpha-\nu-1}(2(k-1)+2M-1)+1$$

so that $\frac{k^\star-1}{2^\alpha}$ is the nodal point at the center of suppw_k^ν, $1 \le k \le 2^\nu$. All subscripts will be taken modulo N from here on, unless otherwise specified. Inserting the expansion for $\mathbf{f}_i$ into the above expression for d_k^ν yields

$$\begin{aligned} d_k^\nu &= \int_{-\infty}^{\infty} \sum_{i=1}^{2^\alpha} \left(\sum_{j=0}^{M-1} c_j (i-k^\star)^j + r_i \right) \varphi_i^\alpha(x) w_k^\nu(x)\, dx \\ &= \int_{-\infty}^{\infty} \sum_{i=1}^{2^\alpha} \left(\sum_{j=0}^{M-1} c_j (i-k^\star)^j \right) \varphi_i^\alpha(x) w_k^\nu(x)\, dx \\ &\quad + \int_{-\infty}^{\infty} \sum_{i=1}^{2^\alpha} r_i \varphi_i^\alpha(x) w_k^\nu(x)\, dx. \end{aligned}$$

However, by Lemma 4, the function

$$\sum_{j=0}^{M-1} c_j (i-k^\star)^j \varphi_i^\alpha(x)$$

is a polynomial in x of degree no more than $M-1$, and so this integral vanishes because of the vanishing moment properties of wavelets (§4.1, (12)), leaving us with

$$\begin{aligned} d_k^\nu &= \int_{-\infty}^{\infty} \sum_{i=1}^{2^\alpha} r_i \varphi_i^\alpha(x) w_k^\nu(x)\, dx \\ &= \frac{1}{(2^\alpha)^M\, M!} \sum_{i=1}^{2^\alpha} (i-k^\star)^M f^{(M)}(\eta_{i,k^\star}) \int_{-\infty}^{\infty} \varphi_i^\alpha(x) w_k^\nu(x)\, dx \\ &= \frac{\sqrt{2^{\alpha+\nu}}}{(2^\alpha)^M\, M!} \sum_{i=1}^{2^\alpha} (i-k^\star)^M f^{(M)}(\eta_{i,k^\star}) \\ &\quad \int_{-\infty}^{\infty} \varphi(2^\alpha x - i + 1) w(2^\nu x - k + 1)\, dx. \end{aligned}$$

At each wavelet level ν, the only φ_i^α whose support intersects the support of w_k^ν are in the range

$$2^{\alpha-\nu}(k-1) - 2M + 1 \le i \le 2^{\alpha-\nu}(k-1) + 2^{\alpha-\nu}(2M-1).$$

Here we are taking the upper and lower bounds modulo N. Therefore,

$$
\begin{aligned}
d_k^\nu &= \frac{\sqrt{2^{\alpha+\nu}}}{(2^\alpha)^M\, M!} \sum_{i=k^\star-(2^{\alpha-\nu-1}+1)(2M-1)}^{k^\star+2^{\alpha-\nu-1}(2M-1)} (i-k^\star)^M f^{(M)}(\eta_{i,k^\star}) \\
&\qquad \int_{-\infty}^{\infty} \varphi(2^\alpha x - i + 1) w(2^\nu x - k + 1)\, dx,
\end{aligned}
$$

and so

$$
\begin{aligned}
|d_k^\nu| &\leq \frac{\sqrt{2^{\alpha+\nu}}}{(2^\alpha)^M\, M!} \sum_{i=k^\star-(2^{\alpha-\nu-1}+1)(2M-1)}^{k^\star+2^{\alpha-\nu-1}(2M-1)} |(i-k^\star)|^M \left|f^{(M)}(\eta_{i,k^\star})\right| \\
&\qquad \int_{-\infty}^{\infty} |\varphi(2^\alpha x - i + 1) w(2^\nu x - k + 1)|\; dx \\
&\leq \frac{\sqrt{2^{\alpha+\nu}}}{(2^\alpha)^M\, M!} \sup_{\eta\in I_k^\nu} \left(\left|f^{(M)}(\eta)\right|\right) \sum_{i=k^\star-(2^{\alpha-\nu-1}+1)(2M-1)}^{k^\star+2^{\alpha-\nu-1}(2M-1)} |(i-k^\star)|^M \\
&\qquad \frac{2M-1}{2^\alpha} \sup_{x\in\mathbb{R}}(|\varphi(x)|) \sup_{x\in\mathbb{R}}(|w(x)|) \\
&\leq \frac{\sqrt{2^\alpha}}{\sqrt{2^\nu}\,(2^\nu)^M} C(M) \sup_{\eta\in I_k^\nu} \left(\left|f^{(M)}(\eta)\right|\right) \sup_{x\in\mathbb{R}}(|\varphi(x)|) \sup_{x\in\mathbb{R}}(|w(x)|),
\end{aligned}
$$

where

$$
C(M) = \frac{(4M-1)\,(2M-1)^{M+1}}{M!},
$$

and I_k^ν is the interval

$$
\left(\frac{k^\star - 2^{\alpha-\nu-1}(2M-1) - 2M}{2^\alpha}, \frac{k^\star + 2^{\alpha-\nu-1}(2M-1) - 1}{2^\alpha}\right).
$$

$\square$

Note that if the parent function f does not satisfy the smoothness criteria of Theorem 5 at some point $x \in [0,1)$, we can still apply the theorem to those d_k^ν such that f is M times differentiable on the interval I_k^ν, because the proof is local in nature. We will use this fact in the next subsection, where we apply Theorem 5 to the functions $u_j(t)$ and $v_j(t)$.

5.2 Applying the theorems to the BEM matrix A

In this subsection, we will combine the results from §3 and §5.1 to obtain bounds on the fast wavelet transform of the vectors $\mathbf{u}^j$ and $\mathbf{v}^j$ defined in §3.3 and §3.4. We will begin by bounding the vector $W\mathbf{u}^j$, the fast wavelet transform of $\mathbf{u}^j$. The method used in this section is to define a notion of *valid* entries of $W\mathbf{u}^j$, that is, indices (ν, k) such that d_k^ν has a small magnitude. We first bound the 2-norm of the valid entries of $W\mathbf{u}^j$, showing that they are very small with respect to $\|\mathbf{u}^j\|$ (i.e., small with respect to $\|W\mathbf{u}^j\|$). Then we show that almost all the elements of $W\mathbf{u}^j$ are valid. Thus, if the valid entries are dropped, then a sparse vector that is a good approximation to $W\mathbf{u}^j$ will result. The same ideas carry over to $W\mathbf{v}^j$.

Let α be a given positive integer and set $N = 2^\alpha$. Let $1 \le j \le N$ be fixed. The first step in obtaining bounds on the fast wavelet transform of $\mathbf{u}^j$ is to obtain bounds on each element of the wavelet transform of $\mathbf{u}^j$ using Theorem 5 of §5. To do this we need to replace the generic function f used in Theorem 5 by the parent function of $\mathbf{u}^j$, $u_j(t)$. Recall that the parent function of $\mathbf{u}^j$ is given by

$$u_j(t) = \int_{s=0}^1 K_n(\gamma(s), \gamma(t))\psi_j(s)\, ds - \Delta(t - t_j)c(\gamma(t_j)).$$

By applying the bounds for $\frac{d^M u_j(t)}{dt^M}$ ($t \ne t_j$, $\gamma(t)$ not a corner) obtained in §3.1, namely

$$\left|\frac{d^M u_j(t)}{dt^M}\right| \le \frac{C(M)C(\psi)R_\Gamma(M)}{N\left[\min_{s\in \operatorname{supp}\psi_j} |||t-s|||\right]^{M+1}},$$

to the result of Theorem 5 we obtain

$$\begin{aligned} |d_k^\nu| \quad \le \quad & \frac{\sqrt{2^\alpha}}{(2^\nu)^M\sqrt{2^\nu}\,2^\alpha} \cdot \sup_{t\in I_k^\nu}\left[\frac{1}{\left[\min_{s\in \operatorname{supp}\psi_j} |||t-s|||\right]^{M+1}}\right] \cdot \\ & C(M)C(\psi)R_\Gamma(M) \sup_{x\in \mathbb{R}}\left(|w(x)|\right) \sup_{x\in \mathbb{R}}\left(|\varphi(x)|\right), \end{aligned} \qquad (20)$$

where $R_\Gamma(M)$ is defined by (8). This equation holds provided that k is chosen such that there are no corners in $\{\gamma(t) : t \in I_k^\nu\}$.

Let us confine our attention to those $t \in [0, 1)$ that are sufficiently far from $\operatorname{supp}(\psi_j)$ and from a corner. Recall that ψ_j is the jth basis function (see §2.1)and that ψ_j has small constant support ($\frac{2c_\psi}{N}$) around t_j. In order to ensure that I_k^ν does not meet the support of ψ_j, we restrict k to the set $\{1,\ldots,2^\nu\} - V_\nu$, where

$$V_\nu = \left\{ \left\lfloor \frac{2^\nu(j-c_\psi)}{2^\alpha} \right\rfloor - 2M + 2 - \hat{k}, \ldots, \left\lceil \frac{2^\nu(j+c_\psi)}{2^\alpha} \right\rceil + 1 + \hat{k} \right\} \bmod 2^\nu. \tag{21}$$

Here $\hat{k}$ is a constant nonnegative integer to be determined later. Note that V_ν depends not only on ν but also on j, M, c_ψ etc. Let the corners of the domain be $\gamma(\tau_1), \ldots, \gamma(\tau_\rho)$, where ρ is the total number of corners. For each corner define a set $C_{i,\nu} = \{k : \tau_i \in I_k^\nu\}$.

We will define the pair (k,ν) to be *valid* if

$$k \in \{1, \ldots 2^\nu\} - (V_\nu \cup C_{1,\nu} \cup \cdots \cup C_{\rho,\nu}).$$

Then for valid (k,ν), we can replace

$$\sup_{\substack{t \in I_k^\nu \\ t \notin \mathrm{supp}(\psi_j)}} \left[\frac{1}{\left\{ \min_{s \in \mathrm{supp}\, \psi_j} |||t-s||| \right\}^{M+1}} \right]$$

by

$$\left[\max\left\{ \frac{1}{|||\frac{k-1}{2^\nu} - \frac{j+c_\psi+2M-1}{2^\alpha}|||}, \frac{1}{|||\frac{k+2M-2}{2^\nu} - \frac{j-c_\psi}{2^\alpha}|||} \right\} \right]^{M+1}.$$

Substituting the above expression into (20), and changing the supremum of $t \in I_k^\nu$ to the supremum of $t \in [0,1]$ yields, for valid (k,ν)

$$|d_k^\nu| \;\le\; \frac{C(M)C(\psi)R_\Gamma(M)}{(2^\nu)^M \sqrt{2^\nu}\sqrt{2^\alpha}} \sup_{x \in \mathbb{R}} (|w(x)|) \sup_{x \in \mathbb{R}} (|\varphi(x)|) \tag{22}$$

$$\left[\max\left\{ \frac{1}{|||\frac{k-1}{2^\nu} - \frac{j+c_\psi+2M-1}{2^\alpha}|||}, \frac{1}{|||\frac{k+2M-2}{2^\nu} - \frac{j-c_\psi}{2^\alpha}|||} \right\} \right]^{M+1}. \tag{23}$$

We are interested in squaring the term $|d_k^\nu|$ given by (22) above, and summing the result over all valid (k,ν). In the summation over k, the only factor depending on k is the last, leading to a summation

$$\sum_{\text{valid}\, k} \left[\max \left\{ \frac{1}{|||\frac{k-1}{2^\nu} - \frac{j+c_\psi+2M-1}{2^\alpha}|||}, \frac{1}{|||\frac{k+2M-2}{2^\nu} - \frac{j-c_\psi}{2^\alpha}|||} \right\} \right]^{2M+2}.$$

Because we are summing over all valid k, we can bound this factor by

$$\begin{aligned} 2\sum_{k=\hat{k}}^{2^{\nu-1}} \left(\frac{2^\nu}{k}\right)^{2M+2} &\leq 2(2^\nu)^{2M+2} \sum_{k=\hat{k}}^{2^{\nu-1}} \frac{1}{k^{2M+2}} \\ &\leq 2(2^\nu)^{2M+2} \sum_{k=\hat{k}}^{\infty} \frac{1}{k^{2M+2}} \\ &\leq 2(2^\nu)^{2M+2} \int_{x=\hat{k}-1}^{\infty} \frac{dx}{x^{2M+2}} \\ &\leq \frac{2(2^\nu)^{2M+2}}{(2M+1)(\hat{k}-1)^{2M+1}}, \end{aligned}$$

so long as $\hat{k} \geq 2$. In the sum $\sum_{\text{valid}\,\nu,k} (d_k^\nu)^2$, the factor involving ν is thus of the form

$$\frac{(2^\nu)^{2M+2}}{(2^\nu)^{2M+1}\, 2^\alpha} = \frac{(2^\nu)}{2^\alpha}.$$

Summing this result over all ν yields 1. Absorbing the factor $\frac{2}{2M+1}$ into $C(M)$ in (22), yields

$$\sum_{\text{valid}\,\nu,k} (d_k^\nu)^2 \leq \frac{(C(M)C(\psi)R_\Gamma(M))^2}{(\hat{k}-1)^{2M+1}} \sup_{x\in\mathbb{R}} (|w(x)|)^2 \sup_{x\in\mathbb{R}} (|\varphi(x)|)^2 .$$

Let us define $\hat{\mathbf{w}}$ to hold the "small" elements of the fast wavelet transform of $\mathbf{u}^j$. That is, $\hat{\mathbf{w}}$ is the N-vector that is equal to d_k^ν for entries when (k,ν) is valid, and is equal to zero otherwise. Then, combining the above calculations and the fact that

$$\|\mathbf{u}^j\| \geq \frac{|c(\gamma(t_j))|}{2},$$

(see Theorem 3.3, §3.3), we obtain the following result:

$$\frac{\|\hat{\mathbf{w}}\|^2}{\|\mathbf{u}^j\|^2} \leq \frac{1}{(\hat{k}-1)^{2M+1}} \frac{(C(M)C(\psi)R_\Gamma(M))^2}{(c(\gamma(t_j)))^2} \sup_{x\in\mathbb{R}} (|w(x)|)^2 \sup_{x\in\mathbb{R}} (|\varphi(x)|)^2 .$$

Thus, given $\epsilon > 0$, for N sufficiently large, we can choose $\hat{k}$ large enough (of the order $\epsilon^{-\frac{1}{M+1/2}}$) so that $\|\hat{\mathbf{w}}\| < \epsilon \|\mathbf{u}^j\|$. Here, the choice of $\hat{k}$ depends upon the wavelet chosen ($w(x)$), the scaling function ($\varphi(x)$), the wavelet order (M), the basis chosen ($\psi(x)$), and the domain, but not upon N. Also, because

$$
\begin{aligned}
\sum_{\text{valid}\,\nu,k} 1 &\geq \sum_{\nu=0}^{\alpha-1} [2^\nu - |V_\nu| - |C_{1,\nu}| - \cdots - |C_{\rho,\nu}|] \\
&\geq \sum_{\nu=0}^{\alpha-1} \left[2^\nu - 2^\nu(2c_\psi)/2^\alpha - 2M - 2 - 2\hat{k} - 2\rho(2M-1)\right] \\
&\geq \sum_{\nu=0}^{\alpha-1} \left[2^\nu - 2M - 2 - 2\hat{k} - 2\rho(2M-1)\right] - 2c_\psi \\
&= 2^\alpha - \alpha[2M + 2\hat{k} + 2\rho(2M-1) + 2] - 2c_\psi - 1 \\
&= N - (\log_2 N)[2M + 2\hat{k} + 2\rho(2M-1) + 2] - 2c_\psi - 1,
\end{aligned}
$$

the vector $\hat{\mathbf{w}}$ has $N - O(\log N)$ elements. As a result, if we define the *sparse fast wavelet transform* of the vector $\mathbf{u}^j$ as $\mathbf{w} - \hat{\mathbf{w}}$, then this vector is sparse, having $O(\log N)$ elements, and is such that $\|W\mathbf{u}^j - (\mathbf{w} - \hat{\mathbf{w}})\| \leq \epsilon \|W\mathbf{u}^j\|$ for any given $\epsilon > 0$. We have just proven the following theorem:

Theorem 6 *Let M (wavelet order), α (number of wavelet levels), and $\epsilon > 0$ be given. Let j be an integer such that $1 \leq j \leq N$. Let $\mathbf{u}^j$ be a vector of length N, obtained from discretizing $u_j(t)$ given by*

$$u_j(t) = \int_{s=0}^{1} K_n(\gamma(t), \gamma(s)) \psi_j(s)\, ds - c(\gamma(t)),$$

at the points $\frac{i-1}{N}$, $1 \leq i \leq N$. Then, provided that α(or N) is sufficiently large, $\hat{\mathbf{w}}$ has at least $N - O(\log N)$ nonzero elements, and $\|\hat{\mathbf{w}}\| \leq \epsilon \|\mathbf{u}^j\|$, where $\hat{\mathbf{w}}$ is a vector of length N containing either zeros or the corresponding entries of $\mathbf{w}$ (if they are "small" enough).

Remark: We are thinking of N and ϵ as independent parameters. In some applications it may make sense to regard ϵ as a function of N, for instance, $\epsilon = cN^{-q}$. In this case, the wavelet transform still yields a sparse column, but the number of nonzero entries in $\mathbf{w} - \hat{\mathbf{w}}$ is $O(N^\beta \log N)$ rather than $O(\log N)$. Here, β depends on q and M.

We can perform a similar analysis for $\mathbf{v}^j$ by applying the bounds for $\frac{d^M v_j(t)}{dt^M}$ obtained in Theorem 2, §3.2 to Theorem 5. We start with the inequality

$$
\begin{aligned}
|d_k^\nu| \quad \leq \quad & \frac{\sqrt{2^\alpha}}{(2^\nu)^M \sqrt{2^\nu}\, 2^\alpha} \sup_{t \in I_k^\nu} \left[\frac{1}{\left[\min_{s \in \operatorname{supp} \psi_j} |||t-s|||\right]^M} \right] \\
& C(M) C(\psi) R_\Gamma(M) \sup_{x \in \mathbb{R}} \left(|w(x)|\right) \sup_{x \in \mathbb{R}} \left(|\varphi(x)|\right) / \chi_\Gamma .
\end{aligned}
$$

We say that (ν, k) is valid if $k \in \{1, \ldots, 2^\nu\} - (V_\nu \cup C_{1,\nu} \cup \cdots \cup C_{\rho,\nu})$, where V_ν and $C_{i,\nu}$ were defined in the last section. We will use a new choice for $\hat{k}$ specified below.

Summing the squares of d_k^ν over all valid (ν, k), the only factor that depends on k is

$$
\sum_{\text{valid}\, k} \left[\max \left\{ \frac{1}{|||\frac{k-1}{2^\nu} - \frac{j+c_\psi+2M-1}{2^\alpha}|||}, \frac{1}{|||\frac{k+2M-2}{2^\nu} - \frac{j-c_\psi}{2^\alpha}|||} \right\} \right]^{2M} .
$$

Because we are summing over all valid k, we can bound this term by

$$
\begin{aligned}
2 \sum_{k=\hat{k}}^{2^{\nu-1}} \left(\frac{2^\nu}{k} \right)^{2M} \quad & \leq \quad 2 (2^\nu)^{2M} \sum_{k=\hat{k}}^{2^{\nu-1}} \frac{1}{k^{2M}} \\
& \leq \quad 2 (2^\nu)^{2M} \int_{x=\hat{k}-1}^{\infty} \frac{dx}{x^{2M}} \\
& = \quad \frac{2 (2^\nu)^{2M}}{(\hat{k}-1)^{2M-1} (2M-1)} .
\end{aligned}
$$

For the sum $\sum_{\text{valid}\, \nu,k} (d_k^\nu)^2$, the factor involving ν is thus of the form

$$
\frac{(2^\nu)^{2M}}{(2^\nu)^{2M+1}\, 2^\alpha} = \frac{1}{2^\nu\, 2^\alpha} .
$$

Summing this result over all ν yields a factor of $2^{-\alpha} = 1/N$. This leads to an equation for $\mathbf{v}^j$ similar to (22):

$$
\sum_{\text{valid}\, \nu,k} (d_k^\nu)^2 \leq \frac{(C(M) C(\psi) R_\Gamma(M))^2}{\chi_\Gamma^2 \cdot 2^\alpha (\hat{k}-1)^{2M-1}} \sup_{x \in \mathbb{R}} \left(|w(x)|\right)^2 \sup_{x \in \mathbb{R}} \left(|\varphi(x)|\right)^2 ,
$$

where $R_\Gamma(M)$ is defined by (8). Let us redefine $\hat{\mathbf{w}}$ to hold the "small" entries of the fast wavelet transform of $\mathbf{v}^j$. That is, $\hat{\mathbf{w}}$ is the N-vector that is equal to the fast wavelet transform of $\mathbf{v}^j$ when the corresponding index (k,ν) is valid, and is equal to zero otherwise. Then, combining the above calculations and the fact that

$$\|\mathbf{v}^j\|^2 \geq \frac{C}{N\|\gamma'\|},$$

(see Theorem 4, §3.4), we obtain the following result:

$$\frac{\|\hat{\mathbf{w}}\|^2}{\|\mathbf{v}^j\|^2} \leq \frac{(C(M)C(\psi)R_\Gamma(M))^2}{(\hat{k}-1)^{2M-1}\chi_\Gamma^2} \cdot \sup_{x\in\mathbb{R}} (|w(x)|)^2 \sup_{x\in\mathbb{R}} (|\varphi(x)|)^2 \cdot \|\gamma'\|$$

Thus, given $\epsilon > 0$, for N sufficiently large we can choose $\hat{k}$ large enough (of the order $\epsilon^{-\frac{1}{M-1/2}}$) so that $\|\hat{\mathbf{w}}\| < \epsilon\|\mathbf{v}^j\|$. Here, the choice of $\hat{k}$ again depends upon the wavelet chosen ($w(x)$), the scaling function ($\varphi(x)$), the wavelet order (M), the basis chosen ($\psi(x)$), and the domain, but not upon N.

The analysis of the number of valid (ν,k) is identical to the analysis in the last subsection. The conclusion is that the vector $\hat{\mathbf{w}}$ has $N - O(\log N)$ elements. As a result, if we define the sparse fast wavelet transform of the vector $\mathbf{v}^j$ as $\mathbf{w}-\hat{\mathbf{w}}$, then this vector is sparse, having $O(\log N)$ elements, and is such that $\|W\mathbf{v}^j - (\mathbf{w}-\hat{\mathbf{w}})\| \leq \epsilon\|W\mathbf{v}^j\|$, for any given $\epsilon > 0$. We have just proven the following theorem:

Theorem 7 *Let M (wavelet order), α (number of wavelet levels), and $\epsilon > 0$ be given. Let j be an integer such that $1 \leq j \leq N$. Let $\mathbf{v}^j$ be a vector of length N, obtained from discretizing $v_j(t)$ given by*

$$v_j(t) = \int_{s=0}^{1} K(\gamma(t),\gamma(s))\psi_j(s)\,ds,$$

at the points $\frac{i-1}{N}$, $1 \leq i \leq N$. Then, provided that α(or N) is sufficiently large, $\hat{\mathbf{w}}$ has at least $N - O(\log N)$ nonzero elements, and $\|\hat{\mathbf{w}}\| \leq \epsilon\,\|\mathbf{u}^j\|$, where $\hat{\mathbf{w}}$ is a vector of length N containing either zeros or the corresponding elements of $\mathbf{w}$ (if they are "small" enough).

We have succeeded in proving that the BEM matrix A as defined by (4) in §2.1 contains $N^2 - O(N\log N)$ "small" elements, and making this matrix

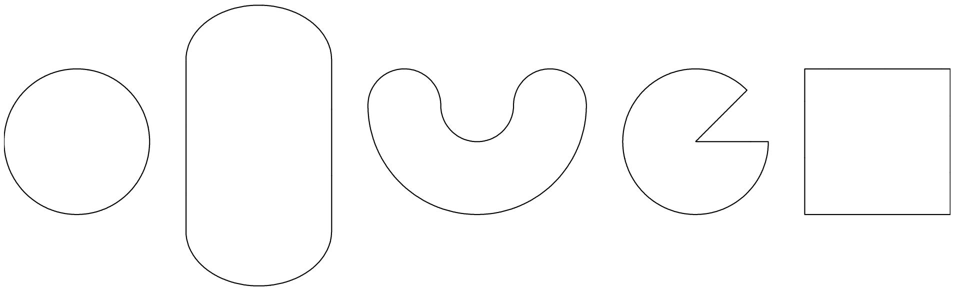

Figure 1: The domain shapes reported on in our computational experiments.

sparse as outlined in §1.2 yields a sparse ($O(N \log N)$ entries) $N \times N$ matrix. Moreover, the 2-norm of every column of this sparse matrix can be as close to the 2-norm of corresponding column of the original dense matrix as desired, for sufficiently large N.

6 Numerical Computations

In this section we present some numerical results to support the theorems presented in §5. The numerical computations that were carried out were of two types. In §6.1 we are interested in how well the sparse fast wavelet transform of A (denoted A'' see §1.2) compares with A in solving a general matrix equation $A\mathbf{x} = \mathbf{b}$. In §6.2 we are interested in how sparse A'' actually is. In each case the numerical computations were performed using the same matrices, obtained using the following procedure: for each of the domains in Figure 1, we formed the $N \times N$ system matrix A ((4), §2.1) for varying N (nodal points), M (wavelet orders) and ϵ (tolerance levels), providing Neumann boundary data at half the nodes and Dirichlet boundary data at the other half of the nodes. All tests were performed in MATLAB 4.1 on a Sun 4/670 MP Sparc Workstation. MATLAB, an interactive software package for numerical computation, is a trademark of The Mathworks, Inc.

6.1 Numerical approximation of A'' to A

In this subsection we are interested in comparing the two vectors $\mathbf{x}$ and $\mathbf{x}_w$, obtained by solving the two $N \times N$ systems $A\mathbf{x} = \mathbf{b}$ and $A''\mathbf{x}_w = W\mathbf{b}$ respectively, for various values of N, M and ϵ. Here, as mentioned above,

A is the BEM matrix obtained using half Neumann boundary data and half Dirichlet boundary data, A'' is the sparse fast wavelet transform of A, and $W\mathbf{b}$ is the fast wavelet transform of the fixed random vector $\mathbf{b}$. The sparse fast wavelet transform of A was performed as follows: the boundary element matrix A was formed as outlined in §6. For a given wavelet order M, the fast wavelet transform of A, WA was calculated. For a chosen tolerance, ϵ, we went through the matrix WA on a column by column basis, zeroing out first the smallest element of each column, then the second smallest element in the column, and so on, until $\|A''(:,j) - A(:,j)\| \geq \epsilon\|A(:,j)\|$, for all j, $1 \leq j \leq N$. In practice, one would most likely drop elements of the column $WA(:,j)$ that are less than $\epsilon'\|A(:,j)\|$ in magnitude for some arbitrary ϵ', but we have used this more complicated procedure to test our theory directly. Here, $A(:,j)$ represents the jth column of the matrix A, and $A''(:,j)$ denotes the jth column of the sparse version of the matrix WA. Both of the systems ($A\mathbf{x} = \mathbf{b}$ and $A''\mathbf{x}_w = W\mathbf{b}$) were solved using Gaussian elimination with pivoting. The relative error between $\mathbf{x}$ and $\mathbf{x}_w$,

$$\mathrm{err}(\mathbf{x}, \mathbf{x}_w) \equiv \frac{\|\mathbf{x} - \mathbf{x}_w\|}{\|\mathbf{x}\|}, \tag{24}$$

was calculated and plotted against N, for a variety of N values ($N = 2^k, k = 1, 2, \ldots, 10$). We performed these computations for fast wavelet transforms with $M = 2$ and with $M = 6$. For each different fast wavelet transform, we took ϵ to be $10^{-3}, 10^{-5}$, and 10^{-7}. We used a piecewise linear basis as our choice of $\{\psi_j\}_{j=1}^N$, where $\mathrm{support}(\psi_j(s)) = [t_{j-1 \bmod N}, t_{j+1 \bmod N}]$. The computational results for the domains C and E are plotted in Figures 2 and 3. The key to these figures is given in Table 1. The computational results for the remaining domains were similar to the plotted result and hence are not shown here.

Note that examining the relative error between $\mathbf{x}$ and $\mathbf{x}_w$ is equivalent to examining the relative error between the vectors $\mathbf{x}$ and $\mathbf{y}$, where $\mathbf{y}$ solves the system $\hat{A}\mathbf{y} = \mathbf{b}$, and where $\hat{A}$ is a matrix obtained from A via the formula $\hat{A} = W^{-1}A''$. Recall that $A'' = WA + E$ where $\|E\| \leq \epsilon\|WA\|$. Since W is an orthogonal matrix, this implies that the solution $\mathbf{x}_w$ described above satisfies $(A + E')\mathbf{x}_w = \mathbf{b}$, where $\|E'\| \leq \epsilon\|A\|$. In general, this implies that $\|\mathbf{x} - \mathbf{x}_w\|/\|\mathbf{x}\|$ is bounded above by $\epsilon\kappa(A)$, where κ denotes the condition number (see [12]). Because our perturbation E has special structure, this bound could be an overestimate. Unfortunately, we do not know of any analysis in

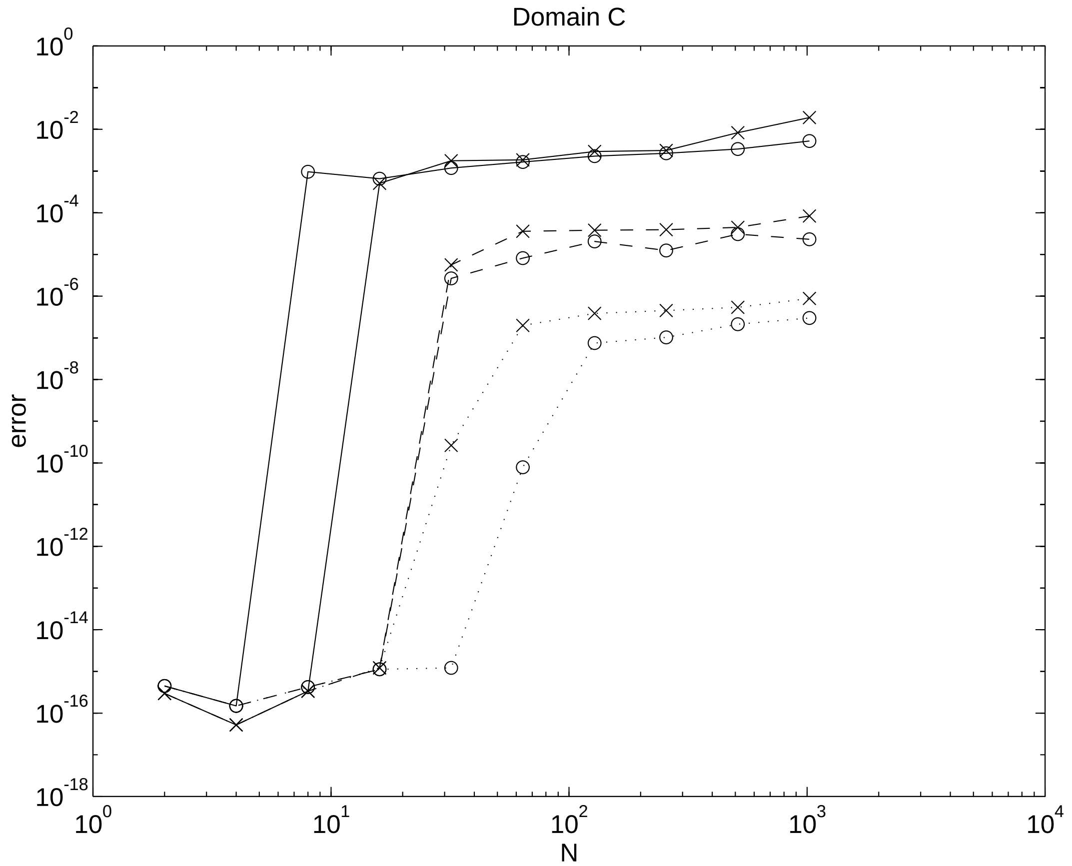

Figure 2: Plot of $\mathrm{err}(\mathbf{x}, \mathbf{x}_w)$ *vs.* N for Domain C. For key see Table 1.

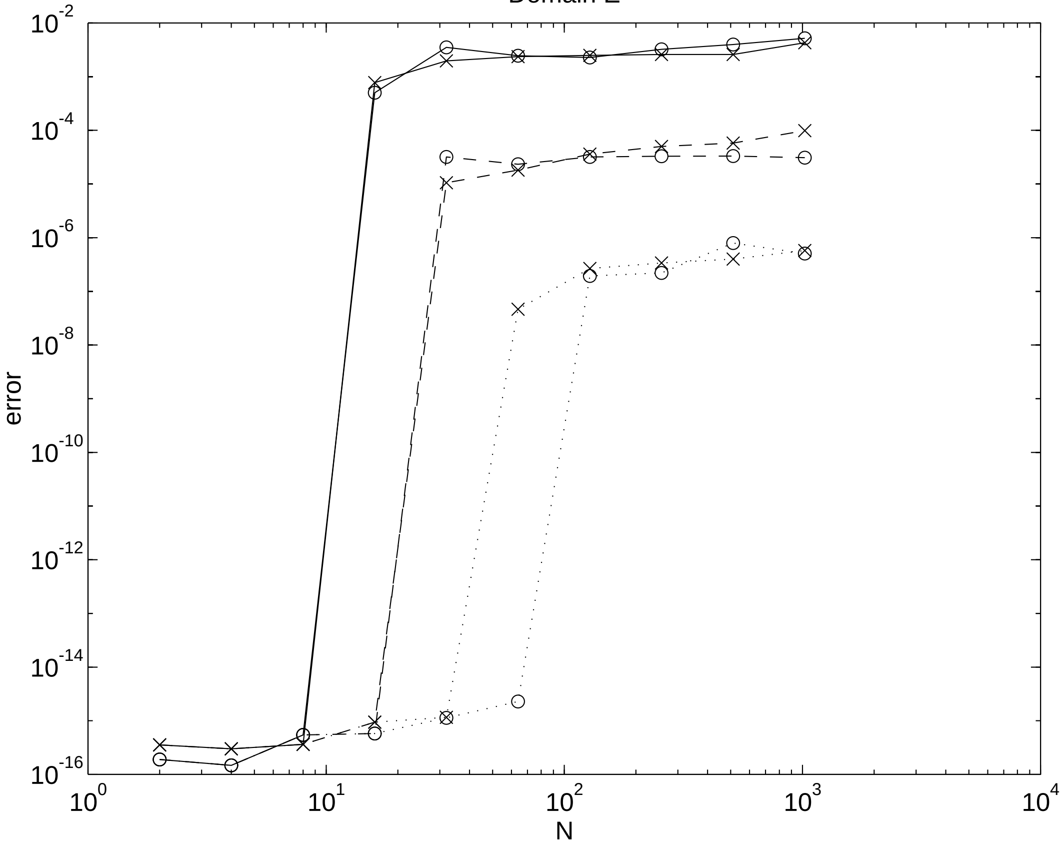

Figure 3: Plot of $\mathrm{err}(\mathbf{x}, \mathbf{x}_w)$ *vs.* N for Domain E. For key see Table 1.

the literature for the condition number of a boundary element matrix with mixed boundary conditions. Accordingly, we can judge the relative error in $\mathbf{x}$ only by computational experiments. In fact, this behavior is essentially what we observe from the figures. Any apparent dependence upon N appears as a result of the relatively small values of N used. Nevertheless, even with these small values of N, we can observe the order ϵ relative error between $\mathbf{x}$ and $\mathbf{y}$. Also, as is evidenced by the computational results, the relative error between $\mathbf{x}$ and $\mathbf{x}_w$ does not depend upon the wavelet order M, and just upon ϵ as mentioned above. Note also, that the relative error between $\mathbf{x}$ and $\mathbf{x}_w$ is similar, both qualitatively and quantitatively, for domains with and without corners. As can be expected, for small N (eg., $N < 128$), the matrix WA cannot be made sufficiently sparse, and as a result A'' is dense and the relative error between $\mathbf{x}$ and $\mathbf{x}_w$ is approximately machine precision. We can observe from the results that as ϵ decreases, we need a larger value of N in order for the matrix A'' to differ from WA, or to significantly alter $\mathbf{x}_w$ from $\mathbf{x}$.

6.2 Sparsity of A''

In this subsection, we present some numerical results to support the theorems presented in §5. We will see that while the theorems guarantee their stated results for sufficiently large N, their results are realized numerically for a relatively small value of N. The numerical computations that were carried out were as follows: the boundary element matrix A was formed as outlined in §6. For a given wavelet order M and tolerance ϵ, the sparse fast wavelet transform of A, A'' was calculated as outlined in §6.1. Let us denote a column of A'' that arises from the matrix U (see §2.1 by (3)) by $\mathbf{u}''$. Similarly, let us denote any column of A'' that is associated with the matrix V by $\mathbf{v}''$. For domains C and E, we plotted the average number of elements of $\mathbf{u}''$ against N. We also plotted the average number of elements in $\mathbf{v}''$ against N for domains C and E. These plots are presented in Figures 4 and 5 (for the sparsity of $\mathbf{u}''$) and Figures 6 and 7 (for the sparsity of $\mathbf{v}''$). We performed these same computations for fast wavelet transforms with $M = 2$ and with $M = 6$. For each different fast wavelet transform, we took ϵ to be $10^{-3}, 10^{-5}$, and 10^{-7}. We chose N to take the values $2^k, k = 1, 2, \ldots, 10$. We used a piecewise linear basis as our choice of $\{\psi_j\}_{j=1}^N$, where support$(\psi_j(s)) = [t_{j-1 \bmod N}, t_{j+1 \bmod N}]$. The key to all these Figures is presented in Table 1. The computational

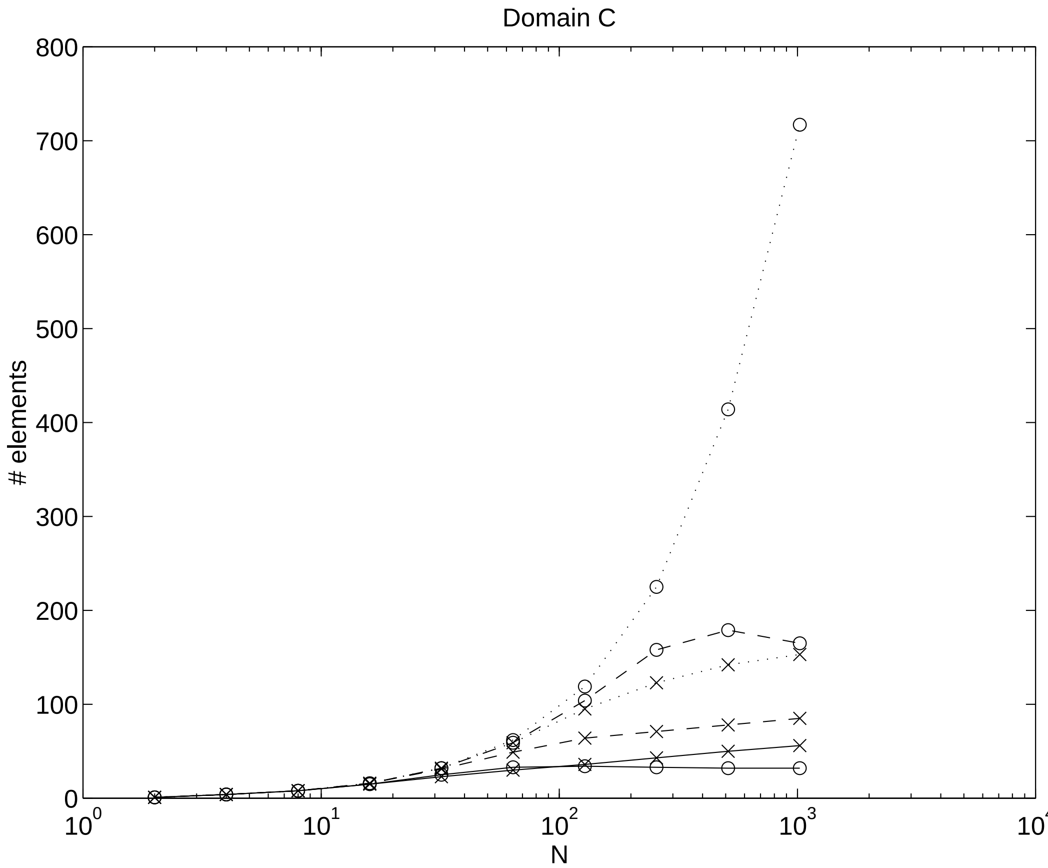

Figure 4: Plot of the average number of elements in $\mathbf{u}''$ *vs.* N for Domain C. For key see Table 1.

results for the remaining domains were similar to the plotted result and hence are not shown here.

A few general trends are evident from the test results corresponding to Figures 4 through 7. First, except for the few computations that have yet to settle down (specifically $\{M, \epsilon\} = \{2, 10^{-7}\}$), the $O(\log N)$ sparsity that the theorems prove for sufficiently large N, is evident for all domains, even for N as small as 128. Also, as expected, for a given ϵ, there are fewer nonzero elements in the vectors $\mathbf{u}''$ and $\mathbf{v}''$ corresponding to higher values of M. This is because the values of $\hat{k}$ derived in §5 will be smaller as M is larger. The price paid by using larger wavelets (larger values of M) appears

Table 1: Key for Figures 2 through 7

$\{M, \epsilon\}$	point style	line style
$\{2, 10^{-3}\}$	o	solid
$\{2, 10^{-5}\}$	o	dashed
$\{2, 10^{-7}\}$	o	dotted
$\{6, 10^{-3}\}$	x	solid
$\{6, 10^{-5}\}$	x	dashed
$\{6, 10^{-7}\}$	x	dotted

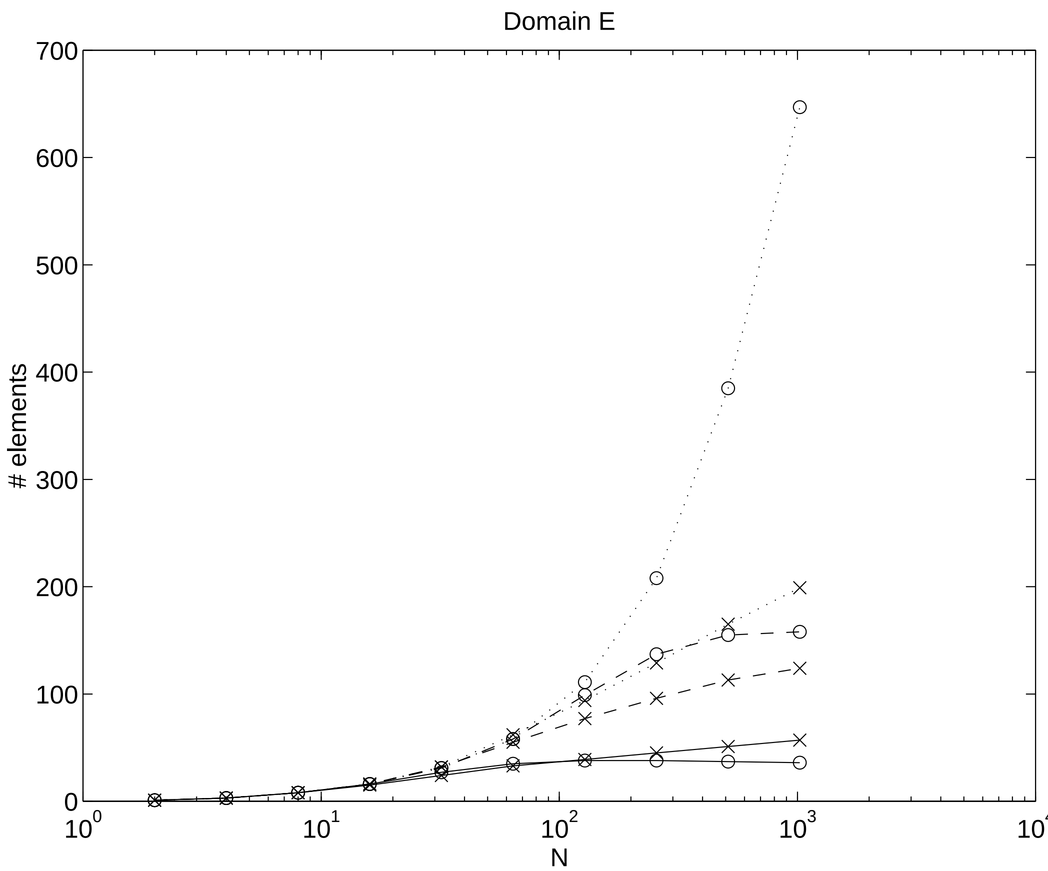

Figure 5: Plot of the average number of elements in $\mathbf{u}''$ *vs.* N for Domain E. For key see Table 1.

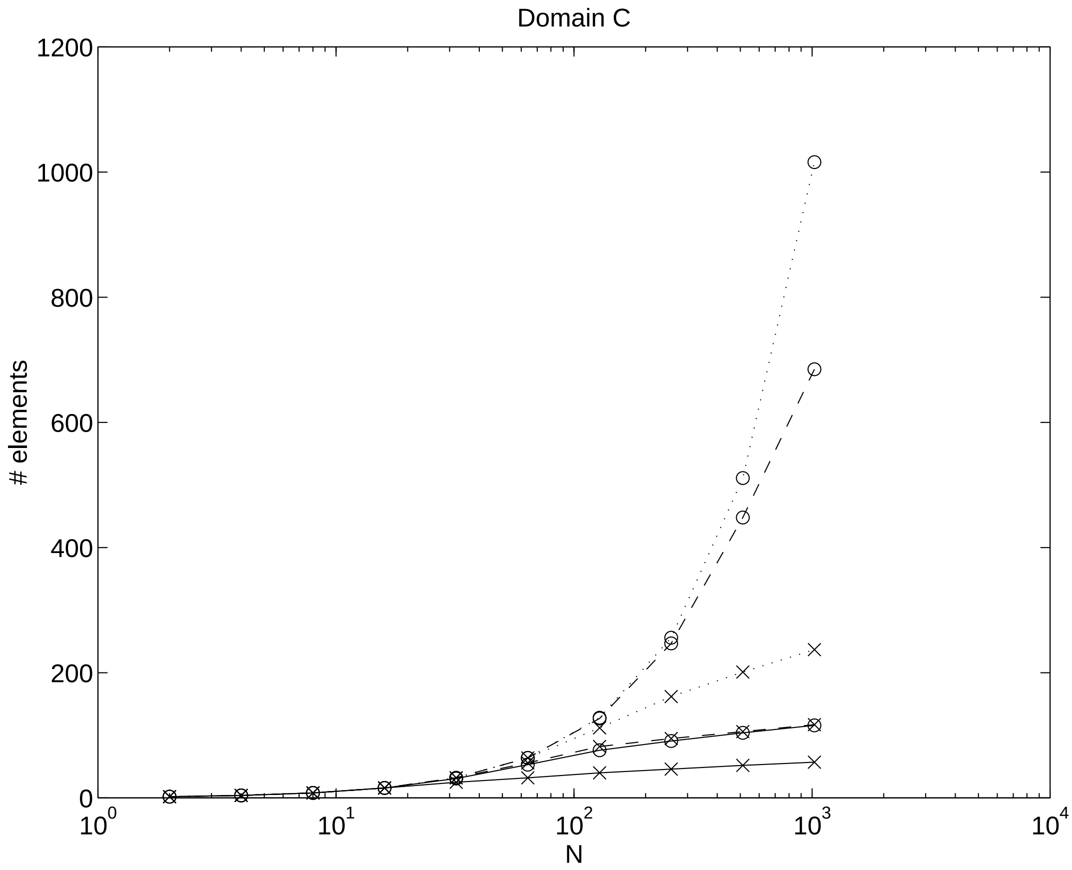

Figure 6: Plot of the average number of elements in $\mathbf{v}''$ *vs.* N for Domain C. For key see Table 1.

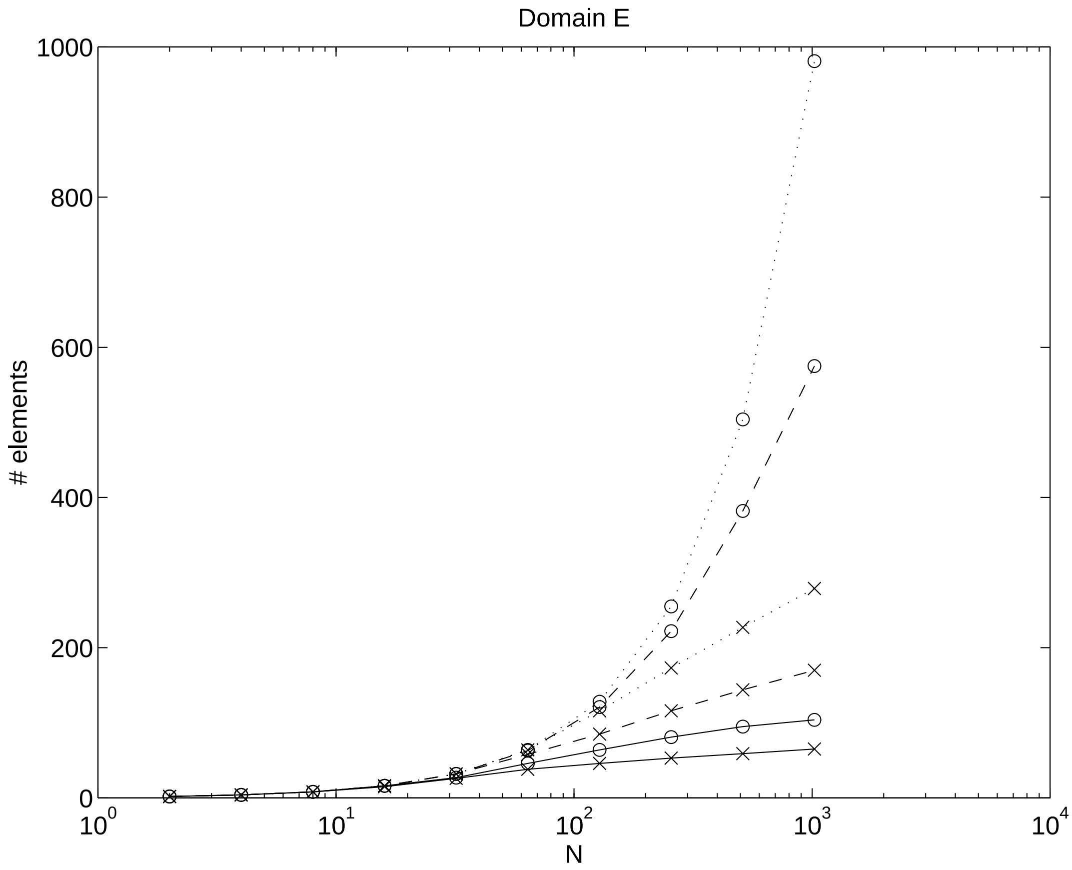

Figure 7: Plot of the average number of elements in $\mathbf{v}''$ *vs.* N for Domain E. For key see Table 1.

as larger constants appearing in Theorems 6 and 7, and as a cost increase in numerically calculating the fast wavelet transform of a vector. Recall that if we think of M as a variable, the cost of a fast wavelet transform of a vector of length N is $O(MN)$. Note that for a given domain, M and ϵ, the vector $\mathbf{v}''$ is slightly less sparse than the corresponding vector $\mathbf{u}''$, especially for the case $M = 2, \epsilon = 10^{-6}$. Also, as a general rule, as the domain becomes more complex, the number of nonzero entries in each sparse (transformed) vector increases (for a fixed N). In particular, as more corners are added, the sparsity goes down; this is predicted by our analysis. However, for the domains in question, each of these vectors is still very sparse, especially as N increases. Thus, for computational purposes, the theorems seem to hold true in all respects for all but very small N.

7 Conclusions

Applying a standard BIEM to Laplace's equation in two dimensions, with any type of consistent boundary conditions yields a coefficient matrix made up of columns of U and V as defined by (3). Theorems 5 through 7 show that if we apply the fast wavelet transform to the columns of this coefficient matrix, we can obtain a matrix that has at most $O(N \log N)$ non-negligible elements in total. Thus, as mentioned in §1 we can approximate the transformed matrix by zeroing out the "small" entries and obtain a sparse coefficient matrix that can be exploited. The penalty for the sparseness of the approximation matrix is paid in the change of the norm of each individual column of the original matrix, which for sufficiently large N, is negligible.

While Theorems 6 and 7 above require a sufficiently large number of boundary element points (N) in order to have the 2-norm of the approximate wavelet transformed vector closely match the 2-norm of the original vector, experimentally the number of boundary element points does not need to be very large for the results of the theorems to hold. That is, even for moderate size vectors ($N \approx 128$), the approximate wavelet transform is sparse, and has a 2-norm close to the 2-norm of the original vector (the vectors in question coming from the columns of A).

References

[1] K. A. Atkinson and I. G. Graham, *Iterative solution of linear systems arising from the boundary integral method,* SIAM J. on Scientific and Statistical Computing, 13 (1992), pp. 694–722.

[2] B. K. Alpert, *Construction of simple multiscale bases for fast matrix operations,* in Wavelets and Their Applications, M. B. Ruskai et. al (ed.), Jones and Bartlett, Boston, 1992.

[3] B. K. Alpert, *Sparse representation of smooth linear operators,* Yale Research Report YALEU/DCS/RR-814, 1990.

[4] B. Alpert, G. Beylkin, R. Coifman, V. Rokhlin, *Wavelets for the fast solution of second-kind integral equations,* Yale Research Report YALEU/DCS/RR-837, 1990.

[5] G. Beylkin, R. Coifman, and V. Rokhlin, *Fast wavelet transforms and numerical algorithms* I, Comm. Pure and Applied Math. XLIV (1991), pp. 141–183.

[6] A. Brandt and A. A. Lubrecht, *Multilevel matrix multiplication and fast solution of integral equations,* J. Comp. Phys. 90 (1990), pp 348–370.

[7] C. A. Brebbia (ed.), *Topics in Boundary Element Research*, Springer-Verlag, Berlin, 1984.

[8] F. X. Canning, *Sparse approximations for solving integral equations with oscillatory kernels,* SIAM J. on Scientific and Statistical Computing, 13 (1992), pp. 71–87.

[9] I. Daubechies, *Orthonormal bases of compactly supported wavelets,* Comm. Pure and Applied Math. XLI (1988), pp. 909–996.

[10] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[11] G. David, *Singular Integrals on Curves and Surfaces*, Springer, Berlin, 1991.

[12] G. H. Golub and C. F. Van Loan, *Matrix Computations*, John Hopkins University Press, Baltimore, 1989.

[13] A. Greenbaum, L. Greengard, and A. Mayo, *On the numerical solution of the biharmonic equation in the plane*, Physica D, 60 (1992), pp. 216–225.

[14] J. L. Hess and A.O.M. Smith, *Calculation of potential flow about arbitrary bodies,* in Progress in Aeronautical Sciences, vol 8, D. Küchemann, ed., Pergamon press, London, 1967, pp. 1–138.

[15] S. Jaffard, *Wavelet methods for fast resolution of elliptic problems,* in SIAM J. Numerical Analysis, 29 (1992), pp. 965–986.

[16] M. A. Jaswon, *Integral equation methods in potential theory, I,* Proc. Royal Soc. Ser. A 275 (1963), pp. 23–32.

[17] Y. Meyer, *Wavelets and operators,* in Analysis at Urbana, vol. 1, London Math. Soc. Lecture Notes Ser. 137 (1989), E. Berkson, N. T. Peck and J. Uhl (eds.), pp 256–365.

[18] V. Rokhlin, *Rapid solution of integral equations of classical potential theory,* J. Comp. Phys. 60 (1985), pp. 187–207.

[19] H. Schippers, *Theoretical and practical aspects of multigrid methods in boundary element calculations,* in Topics in Boundary Element Research, vol. 3, Computational Aspects, C. A. Brebbia (ed.), Springer-Verlag, Berlin, 1984.

[20] G. Strang *Wavelets and dilation equations: A brief introduction,* SIAM Review 31 (1989), pp 614–627.

[21] G. T. Symm, *Integral equation methods in potential theory, II,* Proc. Royal Soc. Ser. A 275 (1963), pp. 33–46.

[22] G. T. Symm, *Preconditioning for boundary integral equations,* SIAM J. Matrix Anal App. 12 no. 4 (1992), pp. 905–925.

[23] Y. Yan, *Sparse preconditioned iterative methods for dense linear systems,* Proceedings of the Copper Mountain Conference on Iterative Methods, Univ. of Colorado at Denver, April, 1992.