MATHEMATICS OF COMPUTATION, VOLUME 33, NUMBER 147
JULY 1979, PAGES 849–879

Capacitance Matrix Methods for the Helmholtz Equation on General Three-Dimensional Regions

By Dianne P. O'Leary* and Olof Widlund**

Abstract. Capacitance matrix methods provide techniques for extending the use of fast Poisson solvers to arbitrary bounded regions. These techniques are further studied and developed with a focus on the three-dimensional case. A discrete analogue of classical potential theory is used as a guide in the design of rapidly convergent iterative methods. Algorithmic and programming aspects of the methods are also explored in detail. Several conjugate gradient methods are discussed for the solution of the capacitance matrix equation. A fast Poisson solver is developed which is numerically very stable even for indefinite Helmholtz equations. Variants thereof allow substantial savings in primary storage for problems on very fine meshes. Numerical results show that accurate solutions can be obtained at a cost which is proportional to that of the fast Helmholtz solver in use.

1. Introduction. It is well known that highly structured systems of linear algebraic equations arise when Helmholtz's equation

$$-\Delta u + cu = f, \qquad c = \text{constant}, \tag{1.1}$$

is discretized by finite difference or finite element methods using uniform meshes. This is true, in particular, for problems on a region Ω which permits the separation of the variables. Very fast and highly accurate numerical methods are now readily available to solve separable problems at an expense which is comparable to that of a few steps of any simple iterative procedure applied to the linear system; see Bank and Rose [2], [3], Buneman [5], Buzbee, Golub and Nielson [8], Fischer, Golub, Hald, Leiva and Widlund [16], Hockney [24], [26], Swarztrauber [50], [51], Swarztrauber and Sweet [52], [53] and Sweet [54]. Adopting common usage, we shall refer to such methods as fast Poisson solvers.

The usefulness of these algorithms has been extended in recent years to problems on general bounded regions by the development of capacitance matrix, or imbedding, methods; see Buzbee and Dorr [6], Buzbee, Dorr, George and Golub [7], George [19], Hockney [25], [27], Martin [35], Polozhii [40], Proskurowski [41], [42], [43], Proskurowski and Widlund [44], [45], Shieh [46], [47], [48] and Widlund [57]. We refer to Proskurowski and Widlund [44] for a discussion of this development up to the

Received October 20, 1978.

AMS (MOS) subject classifications (1970). Primary 65F10, 65N20.

*The work of this author was supported in part by a National Science Foundation Grant MCS 76-06595 at the University of Michigan.

**The work of this author was supported by ERDA, Contract No. EY-76-C-02-3077.

beginning of 1976. All of the numerical experiments reported in those papers were carried out for regions in the plane. Strong results on the efficiency of certain of these methods have been rigorously established through the excellent work of Shieh [46], [47], [48]. Algorithms similar to those which we shall describe have recently been implemented very successfully for two-dimensional regions by Proskurowski [42], [43] and Proskurowski and Widlund [45]. In that work, a new fast Poisson solver, developed by Banegas [1], has been used extensively; see Section 5. We note that the performance of computer programs implementing capacitance matrix algorithms depends very heavily on the efficiency of the fast Poisson solver, and if properly designed, they can be easily upgraded by replacing that module when a better one becomes available.

In this paper, we shall extend the capacitance matrix method to problems in three dimensions. The mathematical framework, using discrete dipole layers in the Dirichlet case, is an extension of the formal discrete potential theory developed in Proskurowski and Widlund [44]. We note that these algorithms must be quite differently designed in the three-dimensional case. As in two dimensions the fast Poisson calculations strongly dominate the work. The number of these calculations necessary to meet a given tolerance remains virtually unchanged when the mesh size is refined. We have developed a FORTRAN program for Cartesian coordinates and the Dirichlet problem, which turns out to be technically more demanding than the Neumann case. This program has been designed to keep storage requirements low. The number of storage locations required is one or two times N, the number of mesh points in a rectangular parallelepiped in which the region is imbedded, plus a modest multiple of p, the number of mesh points which belong to the region Ω and are adjacent to its boundary. A further substantial reduction of storage can be accomplished for very large problems by using the ideas of Banegas [1]; see further Section 5.

In the second section, we discuss the imbedding idea. Following a review of classical potential theory, we derive our capacitance matrix methods in Section 3. Section 4 focuses on algorithmic aspects which are of crucial importance in the development of fast, reliable and modular computer code. We solve the capacitance matrix equations by conjugate gradient methods. These methods, originally used in a similar context by George [19], are reviewed in that section. We also discuss how spectral information and approximate inverses of the capacitance matrices can be obtained and used at a moderate cost in computer time and storage. The fast Poisson solver which is used in our program is described in Section 5. It is numerically stable even for negative values of the coefficient c of the Helmholtz operator. Finally, we give details on the organization of our computer program and results from numerical experiments. These tests were designed to be quite severe, and the method has proved efficient and reliable.

Our program has been checked by the CDC ANSI FORTRAN verifier at the Courant Mathematics and Computing Laboratory of New York University. It has been run successfully on the CDC 6600 and the DEC 11/780 VAX at the Courant Institute, a CDC 7600 at the Lawrence Berkeley Laboratory and the Amdahl 470V/6 at the University of Michigan.

2. Discrete Helmholtz Problems and Imbedding.

2.1. *The Imbedding of the Discrete Problem.* In this section, we shall discuss how discretizations of the problem

$$-\Delta u + cu = f \quad \text{on } \Omega,$$

with a boundary condition and data given on $\partial\Omega$, can be imbedded in problems for which fast Poisson solvers can be used. In the second subsection, we describe in detail how these ideas apply to the finite difference scheme which we have used in our numerical experiments.

The efficiency of capacitance matrix methods depends on the choice of appropriate finite difference and finite element meshes. Interior parts of the mesh should be made regular in the sense that the linear equations at the corresponding mesh points match those of a fast Poisson solver. We denote the set of these mesh points by Ω_h where h is a mesh width parameter. The set of the remaining, irregular mesh points is denoted by $\partial\Omega_h$. These points are typically located on or close to the boundary $\partial\Omega$ and the discrete equations associated with them are computed from local information on the geometry of the region. For efficiency, the number of unknowns associated with the points in $\partial\Omega_h$ should be kept small, since the equations and other information required at the regular mesh points are inexpensive to generate and can be stored in a very compact form.

If we work in Cartesian coordinates it is natural to imbed our open, bounded region Ω in a rectangular parallelepiped and to use a rectangular mesh. Other choices which permit the separation of the variables on the larger region, can equally well be chosen. On the larger region a mesh suitable for a fast Poisson solver is introduced which coincides with the regular part of the mesh previously introduced for the region Ω. The position of the larger region relative to Ω is largely arbitrary but when using discrete dipoles (see Section 3), we need a layer of exterior mesh points, one mesh width thick, outside of $\Omega_h \cup \partial\Omega_h$. We shall use some or all of the discrete equations at exterior mesh points to expand our original linear system into one which is of the same size as the one which is solved by the fast Poisson solver. The set of mesh points corresponding to these equations is denoted by $C\Omega_h$.

Before we describe how these larger systems of equations are derived, we shall show by two examples how these sets of mesh points can be constructed. We first consider a Dirichlet problem solved by a classical finite difference scheme on a rectangular mesh. The values of the approximate solution are sought at the mesh points which belong to Ω. The discretization of the Helmholtz operator on the larger region induces, for each mesh point, a neighborhood of points used by its stencil. A mesh point in Ω belongs to Ω_h if and only if all its relevant stencil neighbors are in Ω, and $\partial\Omega_h$ is the set of the remaining mesh points in Ω. The set $C\Omega_h$ is the set of all mesh points which belong to the complement of Ω. It, thus, includes any mesh point which is on the boundary $\partial\Omega$.

As a second example, consider a Neumann problem for Laplace's equation in two dimensions solved by a finite element method with piecewise linear trial functions. The region is approximated by a union of triangles using a regular triangulation, based on a uniform mesh, in the interior of the region. The set Ω_h will then correspond to the set of equations which are not affected by the particular geometry of the region. Values of the discrete solution are also sought at the vertices on the boundary. These points normally fail to lie on a regular mesh. They belong to $\partial\Omega_h$ together with certain mesh points which are close to the boundary. Each irregular point can be assigned to a close-by mesh point of the regular mesh which covers the larger region; and we then define $C\Omega_h$ as the set of remaining, exterior mesh points. There are a number of permissible ways in which this assignment can be made. Similar constructions can be carried out for higher order accurate finite element methods; see Proskurowski and Widlund [45] for further details.

Let us write the expanded linear system in the form

$$Au = b, \tag{2.1}$$

where u is the vector of values of the discrete solution at the mesh points and the components of b are constructed from the function f and the data given on $\partial\Omega$. By construction, our formulas for the interior and irregular mesh points do not involve any coupling to exterior mesh points, and the matrix is therefore reducible, i.e. there exists a permutation matrix P such that

$$P^TAP = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}.$$

The block matrix A_{11} represents the approximation of the problem on $\Omega_h \cup \partial\Omega_h$. It is clear from the structure of this system that the restriction of the solution of the system (2.1) to this set is independent of the solution and the data at the exterior points. Our methods also produce values of a mesh function for the points of $C\Omega_h$ but they are largely arbitrary and useless. Similarly, we must provide some extension of the data to the set $C\Omega_h$, but the performance of the algorithms is only marginally affected by this choice.

Let B denote the matrix representation of the operator obtained by using the basic discretization at all the mesh points. Only those rows of A and B which correspond to the irregular mesh points differ provided the equations and unknowns are ordered in the same way. We can, therefore, write

$$A = B + UZ^T,$$

where U and Z have p columns, with p equal to the number of elements of the set $\partial\Omega_h$. It is convenient to choose the columns of U to be unit vectors in the direction of the positive coordinate axes corresponding to the points of $\partial\Omega_h$. The operator U is then an extension operator which maps any mesh function, defined only on $\partial\Omega_h$, onto a function defined on all mesh points. The values on $\partial\Omega_h$ are retained while all the remaining values are set equal to zero. The transpose of U, U^T, is a restriction, or trace, operator which maps any mesh function defined everywhere onto its restriction to $\partial\Omega_h$. The matrix Z^T can, with this choice of U, be regarded as a compact representation of

$A - B$, obtained by deleting the zero rows corresponding to the equations for the mesh points in $\Omega_h \cup C\Omega_h$. It is important to note that Z and U are quite sparse, a reflection of the sparsity of A and B.

In Sections 3 and 4, we shall discuss efficient and stable ways of solving the linear system (2.1).

2.2. *The Shortley-Weller Scheme.* We shall now discuss the finite difference scheme which has been used in our numerical experiments to solve the Dirichlet problem and also describe how the necessary information on the geometry of the boundary is handled.

The second order accurate Shortley-Weller formula (see Collatz [9, Chapter 5.1] or Forsythe and Wasow [17, Section 20.7]) can be understood as the sum of three point difference approximations for the second derivative with respect to each of the three independent variables. The value at the nearest mesh neighbor in each positive and negative coordinate direction is used unless this neighbor belongs to the set $C\Omega_h$. In that case the Dirichlet data at the point of intersection of the mesh line and the boundary is used.

As an example, suppose that the mesh spacings in the x, y and z directions are all equal to h. Consider an irregular mesh point, with indices (i, j, k), which has two exterior neighbors in the x direction and one in the positive y direction. Let δ_{-x}, δ_{+x} and δ_{+y} be the distances to the boundary, in the respective coordinate directions, measured in units of the mesh size h; and let g_{-x}, g_{+x} and g_{+y} be the Dirichlet data at the corresponding points on the boundary $\partial\Omega$. Then our approximation to $-\Delta u + cu = f$ at this irregular point is

$$(2/(\delta_{+x}\delta_{-x}) + 2/\delta_{+y} + 2 + ch^2)u_{ijk} - (2/(1 + \delta_{+y}))u_{i,j-1,k} - u_{ij,k+1} - u_{ij,k-1}$$

$$= h^2 f_{ijk} + (2/(\delta_{+x}^2 + \delta_{+x}\delta_{-x}))g_{+x} + (2/(\delta_{-x}^2 + \delta_{+x}\delta_{-x}))g_{-x}$$

$$+ (2/(\delta_{+y}^2 + \delta_{+y}))g_{+y}.$$

At the regular points the formula reduces to a simple seven point approximation.

The Shortley-Weller formula has a matrix of positive type. This permits the use of the classical error estimates based on a discrete maximum principle, as in the references given above. The only information required on the geometry of the region is the coordinates of the irregular mesh points and the distances along the mesh lines from each such point to the boundary. This appears to be close to the minimum information required by any method with more than first order accuracy. See Proskurowski and Widlund [44], Pereyra, Proskurowski and Widlund [39] and Strang and Fix [49] for more details. This geometrical information is also sufficient to construct higher order accurate approximations to the Helmholtz equation, as in Pereyra, Proskurowski and Widlund [39] where a family of methods suggested by Kreiss is developed. These methods have proven quite effective for two-dimensional problems but their usefulness is limited by the requirement that each irregular mesh point must have several interior mesh neighbors along each mesh line. This requirement is met by shifting the region and refining the mesh if necessary. Although this is practical in two dimensions, it is much more difficult for three-dimensional regions.

We are free to scale the rows of the matrix A which correspond to the irregular mesh points. The choice of scaling is important since it affects the rate of convergence of our iterative method. Based on the analysis given in the next section, the experience in the two-dimensional case (see Proskurowski and Widlund [44]) and our numerical experiments, we have chosen to make all diagonal elements of A equal to one.

3. Potential Theory and Discrete Dipoles.

3.1. *The Continuous Case.* In this section, we shall give a brief survey of certain results of classical potential theory and also develop an analogous, formal theory for the discrete case. We shall mainly follow the presentation of Garabedian [18] when discussing the continuous case, specializing to the case of $c = 0$. A discrete, formal theory has previously been developed by Proskurowski and Widlund [44] but our presentation in subsections 3.2–3.4 will be more complete in several respects.

We first introduce the volume, or Newton, potential

$$u_V(x) = (1/4\pi) \int_{R^3} f(\xi)/r \, d\xi, \tag{3.1}$$

where $x = (x_1, x_2, x_3)$, $\xi = (\xi_1, \xi_2, \xi_3)$ and

$$r = ((x_1 - \xi_1)^2 + (x_2 - \xi_2)^2 + (x_3 - \xi_3)^2)^{1/2}.$$

We note that $(1/4\pi)(1/r)$ is a fundamental solution of the operator $-\Delta$, i.e.,

$$-\Delta u_V = f.$$

A single layer potential, with a charge density ρ, is given by

$$V(x) = (1/2\pi) \int_{\partial\Omega} \rho(\xi)/r \, d\sigma \tag{3.2}$$

and a double layer potential, with a dipole moment density μ, by

$$W(x) = (1/2\pi) \int_{\partial\Omega} \mu(\xi)(\partial/\partial\nu_\xi)(1/r) \, d\sigma. \tag{3.3}$$

Here ν denotes the normal of the boundary $\partial\Omega$ directed towards the interior of Ω. By V^+ and V^-, we denote the limits of V when the boundary is approached from the outside and inside, respectively, and similar notations are also used for the limits of W. The functions V and W are real analytic functions in the complement of $\partial\Omega$. By using a Green's formula one can establish that V and $\partial W/\partial\nu$ are continuous and that jump conditions hold for $\partial V/\partial\nu$ and W; see Garabedian [18, Chapter 9]. Thus, for a region with a smooth boundary,

$$V^+ = V^-,$$

$$\partial V^{(\mp)}/\partial\nu = (\mp)\rho + (1/2\pi) \int_{\partial\Omega} \rho(\partial/\partial\nu_x)(1/r) \, d\sigma,$$

$$W^{(\mp)} = (\pm)\mu + (1/2\pi) \int_{\partial\Omega} \mu(\partial/\partial\nu_\xi)(1/r) \, d\sigma,$$

$$\partial W^+/\partial\nu = \partial W^-/\partial\nu.$$

With the aid of these relations the Neumann and Dirichlet problems can be reduced to

Fredholm integral equations. For the interior Neumann problem,

$$-\Delta u = f \quad \text{in } \Omega,$$

$$\partial u/\partial \nu = g_N \quad \text{on } \partial\Omega,$$

we make the Ansatz,

$$u(x) = u_V(x) + V(x).$$

The boundary condition is satisfied by choosing ρ such that

$$\begin{aligned} \partial V^-/\partial\nu &= -\rho + (1/2\pi) \int_{\partial\Omega} \rho(\partial/\partial\nu_x)(1/r)\,d\sigma \\ &= g_N - (\partial/\partial\nu)u_V|_{\partial\Omega} = \widetilde{g}. \end{aligned} \tag{3.4}$$

This equation can be written as $(I - K)\rho = -\widetilde{g}$, where K is a compact operator defined by the formula above. It is a Fredholm integral equation of the second kind with a simple zero eigenvalue. Since K is compact in L^2 the integral operator $I - K$ is bounded in L^2 and it has an inverse of the same form on a space of codimension one. Equation (3.4) is solvable if $\widetilde{g}$ is orthogonal to the left eigenfunction of $(I - K)$ corresponding to the zero eigenvalue. In this case this simply means that $\widetilde{g}$ should have a zero mean value. By using the same Ansatz for the exterior Neumann problem, we obtain an integral equation with the operator $I + K$.

If we use the same single layer Ansatz for the interior Dirichlet problem, with data g_D, we get an integral equation of the first kind,

$$(1/2\pi) \int_{\partial\Omega} \rho/r\,d\sigma = g_D - u_V|_{\partial\Omega}.$$

This operator does not have a bounded inverse in L_2. The use of an analogous Ansatz for the discrete Dirichlet problem gives rise to capacitance matrices which become increasingly ill-conditioned as the mesh is refined.

The Ansatz

$$u(x) = u_V(x) + W(x),$$

which employs a double layer potential, leads to a Fredholm integral equation of the second kind,

$$\begin{aligned} W^- &= \mu + (1/2\pi) \int_{\partial\Omega} \mu(\partial/\partial\nu_\xi)(1/r)\,d\sigma \\ &= g_D - u_V|_{\partial\Omega}. \end{aligned} \tag{3.5}$$

The integral operator is now $I + K^T$, where K^T is the transpose of the operator introduced when solving the Neumann problems. We shall obtain well-conditioned capacitance matrices when using a discrete analogue of this approach.

The close relationship between the integral equations for the interior Dirichlet and exterior Neumann problems is used to establish the solvability of the Dirichlet problem; see Garabedian [18, Chapter 10]. A similar argument is given in subsection 3.3 for a discrete case.

The integral operator K is not symmetric except for very special regions. Nevertheless, it has real eigenvalues; see, e.g. Kellogg [32, p. 309]. For future reference, we also note that there exist variational formulations of the Fredholm integral equations given in this section; see Nedelec and Planchard [37]. It can be shown that the mapping defined by the single layer potential V is an isomorphism from $H^{-1/2}(\partial\Omega)/P_0$ to the subspace of $H^1(\Omega)/P_0$ of weak solutions of Laplace's equation. Here $H^1(\Omega)$ is the space of functions with square integrable first distributional derivatives, $H^{1/2}(\partial\Omega)$ the space of traces of $H^1(\Omega)$, $H^{-1/2}(\partial\Omega)$ the space dual to $H^{1/2}(\partial\Omega)$, and P_0 the space of constants. By substituting the single layer potential into the standard variational formulation of the interior Neumann problem and using a Green's formula, an alternative formulation is obtained. The resulting bilinear form is coercive on $H^{-1/2}(\partial\Omega)/P_0$ and is equivalent to Eq. (3.4).

Before we turn to the discrete problems, we note that, in the theory just developed, the function $(1/4\pi)(1/r)$ can be replaced by other fundamental solutions of the Laplace operator. In particular, we can use a Green's function for a rectangular parallelepiped in which the region Ω is imbedded. The theory can also be extended, in a straightforward way, to Helmholtz's equation with a nonzero coefficient c.

3.2. *Discrete Potential Theory.* We now return to the solution of $Au = b$, (Eq. (2.1)) with $A = B + UZ^T$. Guided by the theory for the continuous case, we shall develop two algorithms, one suitable for the Neumann and the other for the Dirichlet case.

We shall assume that B is invertible. This is not a very restrictive assumption since we have a great deal of freedom to choose the boundary conditions on the larger region.

We recall from subsection 2.1 that the columns of U were chosen to be unit vectors corresponding to the irregular mesh points. If we order the points of Ω_h first, followed by those of $\partial\Omega_h$ and $C\Omega_h$, we can obtain the representation,

$$U = \begin{pmatrix} 0 \\ I \\ 0 \end{pmatrix},$$

where I is a $p \times p$ identity matrix. Let us, in analogy to the continuous case, make the Ansatz

$$u = G\tilde{b} + GWs, \tag{3.6}$$

where the vector s has p components, G is the inverse of B, and W has the form

$$W = \begin{pmatrix} 0 \\ W_2 \\ W_3 \end{pmatrix}.$$

The operator G plays a role very similar to that of a fundamental solution for the continuous problem. The second term GWs corresponds to a single or double layer potential. For additional flexibility, we have introduced the mesh function $\tilde{b}$ which coincides with b except possibly at the irregular points of $\partial\Omega_h$. In particular, if the Helmholtz equation has a zero right-hand side, we can often choose $\tilde{b} = 0$, eliminating the first term of the Ansatz. To arrive at an equation for the vector s, we calculate the residual,

$$b - Au = b - (B + UZ^T)(G\widetilde{b} + GWs) = (b - \widetilde{b}) - UZ^TG\widetilde{b} - (I + UZ^TG)Ws.$$

From the form of $\widetilde{b}$, U, and W, we have the following result:

LEMMA 3.1. *The residuals for the system* (2.1) *corresponding to the points of* Ω_h *are zero for any choice of the vector* s *in* (3.6). *If the matrix* W_3 *is zero, they also vanish at all points of* $C\Omega_h$.

We now demand that the residuals vanish on the set $\partial\Omega_h$

$$0 = U^T(b - Au) = U^T(b - \widetilde{b}) - Z^TG\widetilde{b} - U^TAGWs.$$

This gives us a system of p equations

$$Cs = U^TAGWs = (U^TW + Z^TGW)s = U^T(b - \widetilde{b}) - Z^TG\widetilde{b}, \tag{3.7}$$

where C is the capacitance matrix. We ignore the residuals on the set $C\Omega_h$, since the extension of the data to this set is largely arbitrary. It follows from the reducible structure of A that if the capacitance matrix C is nonsingular the restriction of the mesh function u, given by Eq. (3.6), solves the discrete Helmholtz equation. We shall now discuss two choices of the matrix W and study the invertibility of the resulting matrices.

For a Neumann problem, our choice of W should correspond to a single layer Ansatz. We, therefore, choose $W = U$ and note that the capacitance matrix $C_N = U^TAGU$ is then the restriction of AG to the subspace corresponding to the set $\partial\Omega_h$. Using Eqs. (3.6) and (3.7), we find,

$$u = G\widetilde{b} - GU(U^TAGU)^{-1}(Z^TG\widetilde{b} - U^T(b - \widetilde{b})).$$

This is, for $\widetilde{b} = b$, the well-known Woodbury formula; see Householder [29]. For completeness, we give a proof of the following result.

THEOREM 3.1. *The capacitance matrix* C_N *is singular if and only if the matrix* A *is singular. For* $\widetilde{b} = b$ *the equation* (3.7) *fails to have a solution if and only if* b *does not lie in the range of* A.

PROOF. Let ϕ be a nontrivial element of the null space of C_N. Then, since $C_N = I + Z^TGU$, the vector

$$Z^TGU\phi = -\phi$$

is nonzero; and therefore, $GU\phi$ cannot vanish identically. But $AGU\phi = UC_N\phi = 0$; and therefore, A is singular. Let now ψ belong to the null space of C_N^T and assume that

$$\psi^T(Z^TGb) = (\psi^TZ^TG)b \neq 0.$$

Then b does not belong to the range of A since

$$A^TG^TZ\psi = (B^T + ZU^T)G^TZ\psi = ZC_N^T\psi = 0.$$

Finally, given data for Eq. (2.1), which does not belong to the range of A, Eq. (3.7) cannot be solvable since otherwise Eq. (3.6) would provide a solution of Eq. (2.1).

The Woodbury formula is popular for computation, especially when the rank p of $A - B$ is small. In our application, p is usually very large, often exceeding 1000.

This precludes the computation and storage of the dense, nonsymmetric matrix C_N. We must, therefore, solve the $p \times p$ linear system,

$$C_N s = U^T(b - \widetilde{b}) - Z^T G\widetilde{b}, \tag{3.8}$$

by an iterative method which does not require the explicit calculation of the elements of C_N; see further Section 4. We see from Eq. (3.6) that in addition to solving the system (3.8), we need only to solve at most two simple Helmholtz problems on the entire mesh in order to complete the calculation of the solution u. Our main task is, therefore, the efficient solution of Eq. (3.8).

The efficiency of the iterative solution of Eq. (3.8) depends crucially on the distribution of the singular values of C_N. The choice $W = U$ is suitable for Neumann problems, since it is based on a single layer Ansatz; but it gives rise to increasingly ill-conditioned capacitance matrices if applied to Dirichlet problems.

An alternative to the Woodbury formula gives well-conditioned capacitance matrices for the Dirichlet problem. We shall specialize to a case of a uniform rectangular mesh; cf. subsection 2.2. Our choice of W should correspond to a double layer potential. Let $W = VD$, where D is a square diagonal matrix of nonzero scale factors and each column of V represents a discrete dipole of unit strength associated with an irregular .mesh point. The solution to our problem is then

$$u = G\widetilde{b} - GVD(U^T AGVD)^{-1}(Z^T G\widetilde{b} - U^T(b - \widetilde{b})),$$

and the capacitance matrix is $C_D = U^T AGVD$.

We would like to construct the discrete dipoles by placing a positive unit charge at an irregular mesh point and a negative unit charge at another point located on the exterior normal through the irregular point. Since the data for the fast Poisson solver must be given at mesh points only, we instead divide this negative charge and place it on three mesh points. As an example, consider an irregular mesh point with indices (i, j, k), for which the exterior normal through this mesh point lies in the positive octant. Let the distances, measured in units of the mesh size, to the boundary along the three positive coordinate axes be δ_{+1}, δ_{+2} and δ_{+3}, respectively. Let further $0 < \delta_{+1} \leqslant \delta_{+2} \leqslant \delta_{+3}$. We find the first of the three mesh points for the negative charges by moving in the positive x_1 direction, the direction of the smallest distance, to the point $(i + 1, j, k)$. The weight for this point is $-(1 - \delta_{+1}/\delta_{+2})$. We then proceed in the x_2 direction, the direction of the medium distance, to the point $(i + 1, j + 1, k)$ which is given the weight $-(\delta_{+1}/\delta_{+2} - \delta_{+1}/\delta_{+3})$; and we finally go to the point $(i + 1, j + 1, k + 1)$ which is given the weight $-\delta_{+1}/\delta_{+3}$. We note that all these are nonpositive and that their sum equals -1. Assuming that the boundary $\partial\Omega$ is smooth enough, we find by expanding the expression $V^T v$ in a Taylor series, that it equals $h_\delta(\partial v/\partial \nu) + o(h)$, where

$$h_\delta = h\delta_{+1}(\delta_{+1}^{-2} + \delta_{+2}^{-2} + \delta_{+3}^{-2})^{1/2}. \tag{3.9}$$

For future reference, we note that the area, A_δ, of the triangle with vertices at the intersections of the boundary and the mesh lines through the irregular mesh point is

$$A_\delta = (h^2/2)\delta_{+1}\delta_{+2}\delta_{+3}(\delta_{+1}^{-2} + \delta_{+2}^{-2} + \delta_{+3}^{-2})^{1/2}.$$

For a region with a smooth boundary none of the mesh points used in the discrete dipole construction belong to the set Ω_h provided that the mesh is fine enough. We shall assume that this condition is satisfied and reject any problem which violates it. For an irregular mesh point which, along the same mesh line, is within h of the boundary in both the positive and negative directions, we use the smaller distance of the two in the dipole construction, resolving a tie in an arbitrary way.

3.3. *The Invertibility of the Matrix C_D.* An attempt to prove that C_D is nonsingular, modeled strictly on the proof of Theorem 3.1, is not successful and some additional ideas must be introduced. The proof of the following theorem is in an important part due to Arthur Shieh.

THEOREM 3.2. *Assume that the discrete Helmholtz problem is uniquely solvable, that $c \geqslant 0$, and that the matrix B is of positive type. Assume further that any mesh function of the form $GU\psi$ takes on a maximum or a minimum. Then the capacitance matrix C_D is invertible.*

Remark. The last assumption of this theorem is of course always satisfied if the number of mesh points is finite. It must be verified for fast solvers on regions with an infinite number of points; cf. Section 5.

Proof. We begin as in our proof of Theorem 3.1. To simplify our notations, we choose $D = I$. Suppose that there exists an eigenvector ϕ such that $C_D\phi = U^TAGV\phi = 0$. The mesh function $AGV\phi$, therefore, vanishes on $\partial\Omega_h$ and by Lemma 3.1, it also vanishes on Ω_h. Since the discrete problem represented by the matrix A_{11} is uniquely solvable, the mesh function $GV\phi$ vanishes for all $x \in \Omega_h \cup \partial\Omega_h$. Conversely, if there exists a nontrivial vector ϕ such that $GV\phi$ is identically zero on $\Omega_h \cup \partial\Omega_h$, then by the reducible structure of A, $C_D\phi = 0$.

To conclude, we must prove that there exists no nontrivial discrete dipole potential which vanishes identically on $\Omega_h \cup \partial\Omega_h$. We shall work with a very primitive approximation of the Dirichlet problem, since the particular choice of the rows of A corresponding to the points of $\partial\Omega_h$ is of no importance in this context and also use a simple approximation of an exterior Neumann problem. After a suitable symmetric permutation, which we suppress in order to simplify our notations, we write the discrete Helmholtz operator on the entire mesh in the form

$$\begin{pmatrix} B_{11} & B_{12} & 0 \\ B_{21} & B_{22} & B_{23} \\ 0 & B_{32} & B_{33} \end{pmatrix}.$$

Here the subscripts 1, 2 and 3 refer to the interior, irregular and exterior mesh points, respectively. Our interior Dirichlet problem is simply chosen so that

$$\tilde{A}_D = \begin{pmatrix} B_{11} & B_{12} & 0 \\ 0 & I & 0 \\ 0 & B_{32} & B_{33} \end{pmatrix}.$$

The dipole capacitance matrix is then

$$\tilde{C}_D = G_{22}V_2 + G_{23}V_3,$$

where a discrete dipole layer is written as

$$V\mu = \begin{pmatrix} 0 \\ V_2 \\ V_3 \end{pmatrix} \mu.$$

The matrices G_{ij}, $i, j = 1, 2, 3$, are the blocks of the inverse of B. The exterior Neumann problem is approximated by

$$\tilde{A}_N = \begin{pmatrix} B_{11} & B_{12} & 0 \\ 0 & V_2^T & V_3^T \\ 0 & B_{32} & B_{33} \end{pmatrix}.$$

Using a single layer Ansatz, the capacitance matrix becomes

$$\tilde{C}_N = V_2^TG_{22} + V_3^TG_{32}.$$

By the symmetry of the operator G, we obtain

$$\tilde{C}_D^T = \tilde{C}_N;$$

cf. the continuous case. By the arguments given in the proof of Theorem 3.1 the matrix $\tilde{C}_N$ is invertible if

$$\tilde{A}_NGU\psi = 0$$

only for $\psi = 0$. Let $c = 0$. Since, by assumption, $GU\psi$ attains an extremal value and $\tilde{A}_N$ clearly satisfies a discrete maximum principle, we can conclude that $GU\psi$ is a constant and that then $BGU\psi = U\psi = 0$. This argument can easily be modified for the case of $c > 0$; and the proof is, therefore, concluded.

We note that the assumptions of this theorem, except for the invertibility of the matrix A_{11}, were used solely to prove that the null spaces of $\tilde{A}_N$ and B coincide. We also note that one of the arguments given in a similar context in Proskurowski and Widlund [44] is incorrect. The proof given above can be modified to give rather crude, but still quite useful estimates of the condition number of the matrix C_D, see Shieh [48].

3.4. *The Choice of Scale Factors.* The capacitance matrix equation (3.7) is solved by iterative methods; and it is, therefore, quite important to use a suitable scaling of the variables and the equations. When choosing the scaling, we shall be guided by an interpretation of Eq. (3.7) as approximations of the well-conditioned continuous problems (3.4) and (3.5). We shall only discuss the Dirichlet case, since a discussion of the Neumann problem adds little new, and also specialize to the case when $c = 0$.

The scaling of C_D is carried out by choosing the matrix D and the row sums of U^TA or equivalently the row sums of Z^T. It is easy to see that these are strictly positive in the special case considered in subsection 2.2 and that this property holds for any other consistent approximation of the Dirichlet problem for Laplace's equation. We shall now show that it is appropriate to choose $D = I$ and to make the row sums of Z^T equal to two.

With this choice of D the first term of the capacitance matrix C_D equals U^TV; see (3.7). In the typical case where all the mesh points corresponding to the negative weights belong to $C\Omega_h$, $U^TV = I$. When we turn to the other term, we first note that it can be shown, by elementary arguments, that with the choice of scaling of the matrix B consistent with the formulas in subsection 2.2, $h^{-1}G$, regarded as a mesh function, approximates $\Gamma(x, \xi)$, a fundamental solution of the Laplace operator. In subsection 3.2, we have interpreted V^T as a difference operator in the normal direction. We find that $(hh_\delta)^{-1}Z^TGV$ formally converges to $2\partial\Gamma/\partial\nu_\xi$, since the operator Z^T is a local difference operator with a combined weight equal to two; see (3.9). By using finite difference theory or by studying the discrete fundamental solution directly, we can show that this convergence is pointwise for any $x \neq \xi$. See Shieh [46] or Thomée [55]. We note, however, that this convergence fails to be uniform. See further discussion below.

We want to interpret the vector $Z^TGV\mu$ as a numerical quadrature approximation of the corresponding term

$$2\int_{\partial\Omega} \partial\Gamma/\partial\nu_\xi \mu \, d\sigma \tag{3.10}$$

of a Fredholm integral equation similar to Eq. (3.5). We note that the factor 2 is appropriate since the function $(1/2\pi)(1/r)$ appearing in that equation is twice a fundamental solution of the Laplace operator. To verify that our choice of scalings gives a formally convergent approximation, we must consider the density of the discrete dipoles and the area elements to be assigned to them. Since the distances between the dipoles vary in a highly irregular way, we shall consider local averages over patches of the boundary with a diameter on the order of $\sqrt{h}$. Over an area of that size the direction of the normal can be regarded as a constant. We shall specialize to the case discussed in subsection 3.2, in which the discrete dipoles were introduced, and use the same notations. In the patch considered there is then one irregular mesh point within a distance of h to the boundary along any mesh line through the patch parallel to the x_1-axis. The area A_δ, previously computed, should, therefore, be compared with the area $(h^2/2)\delta_{+2}\delta_{+3}$ of the other relevant face of the polyhedron with vertices at the irregular point and the intersections of the mesh lines and the boundary. Each dipole should, therefore, be assigned the weight,

$$\delta_{+1}\delta_{+2}\delta_{+3}(\delta_{+1}^{-2} + \delta_{+2}^{-2} + \delta_{+3}^{-2})^{1/2}/\delta_{+2}\delta_{+3} = \delta_{+1}(\delta_{+1}^{-2} + \delta_{+2}^{-2} + \delta_{+3}^{-2})^{1/2} = h_\delta/h.$$

Combining these observations, we see that $Z^TGV\mu$ formally converges to the integral (3.10).

It is natural to ask if the singular values of C_D converge to those of the integral operator. This is not in general the case, a fact intimately related to the nonuniform distribution of the irregular mesh points. The study of this question is of very considerable difficulty. Following Shieh [46], [47], [48], let

$$C_D = B_h + K_h,$$

where B_h represents the coupling between irregular mesh points which are within $\sqrt{h}$ of each other. With the scaling introduced above, K_h converges pointwise to the correct integral operator. However, the operator B_h is not in general a formally convergent approximation of the identity operator, but for certain important finite difference schemes and general plane regions Shieh [46], [47], [48] has been able to show that the spectral condition number of B_h can be bounded independently of h. These results, combined with the crude estimates of the spectral condition number C_D mentioned in the previous subsection, suffice to show that the number of conjugate gradient steps required for a specific decrease of the error grows only in proportion to $\log(1/h)$. See also Proskurowski [41], [42], [43], Proskurowski and Widlund [44], [45] and Section 6 of this paper for numerical evidence.

4. Capacitance Matrix Algorithms.

4.1. *The Generation of the Capacitance Matrix.* We have previously pointed out that the central problem in our work is the efficient solution of Eq. (3.7). In this section, we shall examine various alternatives.

We shall first consider the cost of computing the capacitance matrices $C_N = U^TAGU$ and $C_D = U^TAGV$, respectively. These are $p \times p$ dense nonsymmetric matrices where p is the number of variables associated with the set $\partial\Omega_h$. Since the matrices U^TA, U^T and V^T have only a few nonzero elements per row, the computation of an individual element of C_N or C_D requires only a modest number of arithmetic operations if the elements of G are known. Since the order of G is at least as large as the number of mesh points in $\Omega_h \cup \partial\Omega_h$, the computation and storage of all its elements is out of the question. Alternatively, columns of C_N or C_D can be computed one at a time using the fast solver once per column of GU or GV. For problems in three dimensions the cost would be enormous.

The number of arithmetic operations can be reduced drastically by using a device described already in Widlund [56]. The separable problem can be made periodic or the larger region can otherwise be chosen without a boundary. In the absence of a boundary, the problem becomes translation invariant in the sense that the solution at any mesh point, due to a single point charge at another mesh point, depends only on the difference of the coordinates of the two mesh points. One use of the fast Poisson solver, with a discrete delta function as data, provides one column of the matrix of G. By this observation, all elements of G are then easily available from this one solution. Given a column of G, the entire capacitance matrix can then be found at an expense which grows in proportion to p^2. This cost is thus of the same order of magnitude as the evaluation of a numerical quadrature approximation of the integral equations of the classical potential theory (see, for example, (3.5)) employing a comparable number of quadrature points. At an expense of $p^3/3$ multiplications and additions, a triangular factorization of the capacitance matrix can be computed by Gaussian elimination. The solution of the capacitance matrix equation (3.7) can then be found at an additional expense of p^2 additions and multiplications.

If the capacitance matrix is available, the equation (3.7) can also be solved by iterative methods at an expense of between p^2 and $2p^2$ additions and multiplications per step; see further Proskurowski and Widlund [44]. When using an iterative method of this kind, the elements of the capacitance matrix can either be stored, possibly on a secondary mass storage device, or they can be regenerated whenever they are needed.

In two dimensions the number of irregular mesh points typically grows only in proportion to $N^{1/2}$ while in three dimensions the growth is proportional to $N^{2/3}$. Many problems in the plane can be solved satisfactorily using a value of p which is less than 200, but in three dimensions values of p in excess of 1000 occur even for quite coarse meshes. We must, therefore, find alternative algorithms which do not require the storage or direct manipulation of the large capacitance matrices unless we are willing to accept a very substantial number of arithmetic operations and the use of out of core storage devices.

To put the methods discussed so far in some perspective, we compare them with known results on symmetric Gaussian elimination methods applied to standard finite difference problems in two and three dimensions. For problems in two dimensions Hoffman, Martin and Rose [28] have shown that the number of nonzero elements of the triangular factors must grow at least in proportion to $N \log_2 N$. George [20] has designed such optimal methods and also has shown that at least $N^{3/2}$ multiplications and additions are required to carry out the factorization step. The corresponding best bounds for three-dimensional problems are on the order of $N^{4/3}$ and N^2, respectively; see Eisenstat [13], Eisenstat, Schultz and Sherman [14].

We shall now demonstrate that we can compute the product of a capacitance matrix and any vector t at a much smaller expense. In the next subsections, we shall show how such products can be used in efficiently solving Eq. (3.7) by iterative methods. We note that in their original form these ideas are due to George [19]. We shall specialize this discussion to the discrete dipole case, $C_Dt = U^TAGVt$, but similar remarks can be made for the discrete Neumann problem.

We first note that the generation of the mesh function Vt can be carried out using only on the order of p operations on a three-dimensional array initialized to zero. The fast Poisson solver is then applied to give GVt, and only on the order of p operations are then needed to obtain $C_Dt = U^TA(GVt)$. Similarly C_D^Tt can be obtained, if so desired, by using a factored form of the matrix. The sparse matrices U^TA and V can be computed from the coordinates of the irregular points and other local information on the geometry of the region using only on the order of p arithmetic operations. Since it is inexpensive to generate these matrices, we can choose to recompute their nonzero elements whenever they are needed but they could also be stored at a cost of on the order of p storage locations.

We remark that when U^TAGVt is computed from GVt only a small fraction of the values of this mesh function is needed. Similarly the vector Vt is very sparse. This has inspired the development of fast Poisson solvers which exploit the sparsity inherent in problems of this kind; see further discussion in Section 5.

4.2. *The Use of the Standard Conjugate Gradient Method.* We shall first review some material on conjugate gradient methods and then discuss their use in solving Eq. (3.7).

Let $Mv = c$ be a linear system of equations with a symmetric, positive definite matrix M. The kth iterate v_k of the conjugate gradient method can then be characterized as the minimizing element for the problem,

$$\min_{v-v_0\in S^{(k)}} \frac{1}{2} v^T M v - v^T c. \tag{4.1}$$

Here $S^{(k)}$ is the subspace spanned by the first k elements of the Krylov sequence, r_0, $Mr_0, M^2r_0, \ldots$, where $r_0 = c - Mv_0$ is the initial residual and v_0 is the initial guess. See further Hestenes and Stiefel [23] or Luenberger [34].

The kth iterate is thus of the form

$$v_k = v_0 + P_{k-1}(M)r_0,$$

where P_{k-1} is some polynomial of degree $k - 1$. The quadratic form in (4.1) differs from the error functional

$$E(v_k) = ½(v_k - v)^T M(v_k - v),$$

only by an irrelevant constant term. Here v is the exact solution. The optimality result (4.1) and an expansion of the initial error $v_0 - v$ in the eigenvectors of M easily leads to the estimate

$$E(v_k) \leqslant \min_{P_{k-1}} \max_{\lambda\in\sigma(M)} (1 - \lambda P_{k-1}(\lambda))^2 E(v_0), \tag{4.2}$$

where $\sigma(M)$ is the spectrum of M. See further Daniel [12], Kaniel [31] or Luenberger [34]. This inequality remains valid if eigenvalues corresponding to modes absent from the initial error are ignored when forming the maximum in (4.2). This is important since it allows us the use of the method and the estimate for semidefinite problems if the data and initial guess lie in the range of the operator.

From inequality (4.2) and a special construction of the polynomial P_{k-1} in terms of Chebyshev polynomials, the estimate

$$E(v_k) \leqslant (2(1 - 1/\kappa)^k/((1 + 1/\sqrt{\kappa})^{2k} + (1 - 1/\sqrt{\kappa})^{2k}))^2 E(v_0) \tag{4.3}$$

is easily obtained; see references given above. Here κ is the spectral condition number of the operator M. When this ratio κ of eigenvalues of M is computed, we can again ignore eigenvalues corresponding to modes which are absent from the initial error.

A convenient way of implementing the conjugate gradient algorithm is as follows: Let v_0 be an initial guess. Compute

$$r_0 = c - Mv_0 \tag{4.4}$$

and set $p_0 = r_0$.

For $k = 0, 1, 2, \ldots$:

Update the solution and the residual by

$$v_{k+1} = v_k + \alpha_k p_k, \qquad r_{k+1} = r_k - \alpha_k M p_k, \tag{4.5}$$

where

$$\alpha_k = r_k^T r_k / p_k^T M p_k \tag{4.6}$$

provides the minimum of the error functional along the search direction p_k.

Compute a new M-conjugate search direction by

$$p_{k+1} = r_{k+1} + \beta_k p_k, \tag{4.7}$$

where

$$\beta_k = r_{k+1}^T r_{k+1} / r_k^T r_k. \tag{4.8}$$

We note that the use of this algorithm requires no a priori information on the spectrum of M. By a standard result, the residual vectors r_k are mutually orthogonal; see Luenberger [34].

In order to use this algorithm to solve the Dirichlet problem, we first form the normal equations equivalent to Eq. (3.7) and obtain,

$$C_D^T C_D s = C_D^T(-Z^T G\widetilde{b} - U^T(\widetilde{b} - b)).$$

We expect that the new matrix $C_D^T C_D$ will still be quite well conditioned.. The product of it and an arbitrary vector can be obtained by the methods described in subsection 4.1.

In our experience the inequality (4.3) gives realistic bounds for Helmholtz problems with nonnegative values of c. If a negative value of c is chosen so that the discrete Helmholtz operator is almost singular, the capacitance matrix must have at least one small singular value. By analogy with the continuous case, we however expect that there will only be a few such values, well separated from the rest of the spectrum. Bounds, much improved in comparison with (4.3), can therefore be obtained from inequality (4.2) by constructing polynomials which vanish at the isolated small eigenvalues of M and are small over the interval containing the rest of the spectrum. A similar idea was used by Hayes [21], who proved that the conjugate gradient algorithm is superlinearly convergent when applied to a Fredholm integral equation of the second kind. See Widlund [57] and Proskurowski and Widlund [44] for further discussion. Such arguments are also central in the work of Shieh [47]. He was able to prove that all except a fixed number of singular values of certain capacitance matrices for problems in the plane lie in a fixed interval while the remaining few are no closer than Kh^q, K and q constants, from the origin. A construction of polynomials as indicated above leads to a bound for the number of iterations required to obtain a prescribed reduction of the error. This bound grows only in proportion to $\log(1/h)$.

The algorithm described in this section can equally well be used for the capacitance matrix equation (3.8).

4.3. *An Alternative Conjugate Gradient Algorithm for Neumann Problems.* We shall now describe an alternative conjugate gradient method, which can be used with the single layer Ansatz for discrete Helmholtz problems with positive semidefinite symmetric coefficient matrices. It has the advantage that a normal equation formulation of the capacitance matrix equation can be avoided and the cost per step is, therefore, reduced by a factor two. That such a reduction is possible is not immediately apparent since the continuous analogue of the capacitance matrix is a nonsymmetric operator. The search for a method of this kind was inspired by the variational formulation of the Fredholm integral equations mentioned in subsection 3.1. This algorithm has recently been implemented successfully by Proskurowski and Widlund [45] for a finite element approximation of the two-dimensional Neumann problem.

Consider the solution of a linear system of the form

$$\widetilde{A}x = b,$$

where $\widetilde{A}$ is a positive semidefinite, symmetric operator. We make the Ansatz

$$x = \widetilde{G}y,$$

where $\widetilde{G}$ is a suitable, strictly positive definite symmetric operator. A new variable is now introduced by $z = \widetilde{G}^{1/2}y$, and the resulting equation is multiplied by $\widetilde{G}^{1/2}$

$$\widetilde{G}^{1/2}\widetilde{A}\widetilde{G}^{1/2}z = \widetilde{G}^{1/2}b.$$

The new operator is symmetric, positive semidefinite while $\widetilde{A}\widetilde{G}$, in general, fails to be symmetric. The standard conjugate gradient algorithm is applied to this transformed system, and the final algorithm is then obtained by returning to the variable y.

Carrying out this substitution, we find that the formulas given in subsection 4.2 must be modified in two respects:

Replace the operator M by $\widetilde{A}\widetilde{G}$ when calculating the residuals by formulas (4.4) and (4.5).

In the calculation of the parameters α_k and β_k, in formulas (4.6) and (4.8), replace the inner products $r_k^Tr_k$ and $p_k^TMp_k$ by $r_k^T\widetilde{G}r_k$ and $p_k^T\widetilde{G}\widetilde{A}\widetilde{G}p_k$, respectively.

The error estimates (4.2) and (4.3) apply in this case. The relevant spectrum is now that of the operator $\widetilde{A}\widetilde{G}$.

In our application $\widetilde{A}$ is the operator corresponding to the discretization of the Helmholtz problem on the original region Ω, and $\widetilde{G}$ the restriction of the operator G to the set $\Omega_h \cup \partial\Omega_h$. No extension of the operator $\widetilde{A}$ to a larger region is necessary. If the right-hand side b vanishes on the set Ω_h, then so will the vector y, since the solution x can be expressed as a discrete single layer potential. The iteration can, therefore, be organized using only vectors with p components. A version of the algorithm has been designed which requires only one application of operator $\widetilde{G}$ in each step. For details see Proskurowski and Widlund [45].

In our problem the possibility of using the sparsity of the vectors y_k gives this algorithm an advantage over the generalized conjugate gradient algorithm considered by Concus, Golub and O'Leary [10] and others; see also Hestenes [22]. Their algorithm is obtained from ours by using the iterates $x_k = \widetilde{G}y_k$. The vectors x_k fail to be sparse in our applications.

4.4. *Estimates of the Singular Values and Approximate Inverses of Capacitance Matrices.* We have previously pointed out that the residuals r_k of the conjugate gradient method are orthogonal. By combining Eqs. (4.5) and (4.7), eliminating the vectors p_k, we obtain

$$\begin{aligned} Mr_0 &= -(1/\alpha_0)r_1 + (1/\alpha_0)r_0, \\ Mr_k &= -(1/\alpha_k)r_{k+1} + (1/\alpha_k + \beta_{k-1}/\alpha_{k-1})r_k - (\beta_{k-1}/\alpha_{k-1})r_{k-1}. \end{aligned} \tag{4.9}$$

Let $R^{(k)}$ be a matrix with its k columns chosen as the normalized residual vectors. Using the definition of the parameter β_k, the equations (4.9) can be rewritten as

$$MR^{(k)} = R^{(k)}J^{(k)} - (\sqrt{\beta_{k-1}}/(\alpha_{k-1}|r_k|))r_k e_k^T.$$

Here e_k is a unit vector in the direction of the positive' kth coordinate direction and $J^{(k)}$ the symmetric, tridiagonal matrix,

$$J^{(k)} = \begin{pmatrix} 1/\alpha_0 & -\sqrt{\beta_0}/\alpha_0 & \\ -\sqrt{\beta_0}/\alpha_0 & (1/\alpha_1 + \beta_0/\alpha_0) & -\sqrt{\beta_1}/\alpha_1 \\ \ddots & \ddots & \ddots \end{pmatrix}.$$

Using the orthogonality of the residuals, we find that

$$J^{(k)} = R^{(k)T}MR^{(k)},$$

i.e. $J^{(k)}$ is a matrix representation of the restriction of the operator M to the space spanned by the vectors $r_0, \ldots, r_{k-1}$. This space can easily be shown to be the same as the Krylov subspace $S^{(k)}$ which was defined in subsection 4.2. See further Engeli, Ginsburg, Rutishauser and Stiefel [15].

We shall exploit these facts in two ways. Approximations of the eigenvalues of M are obtained from the eigenvalues of $J^{(k)}$. The eigenvalues of $J^{(k)}$ interlace those of $J^{(k+1)}$ and improved estimates of the largest and smallest eigenvalues of M and a lower bound for its condition number are, therefore, obtained in each step. This procedure is in fact a variant of a well-known eigenvalue algorithm due to Lanczos [33]. The extreme eigenvalues of $J^{(k)}$ often converge quite rapidly. See for example, Kaniel [31] and Paige [38]. In our problems we quickly obtain realistic estimates of the condition number of M. This idea has proven a very useful tool in the development of our algorithms, in particular when different scalings of the capacitance matrices were tested. The cost of computing the eigenvalues of $J^{(k)}$ is very moderate and grows no faster than k^2.

The analogy between the capacitance matrices and the Fredholm integral operators of the second kind inspired an attempt to compute and use approximate inverses of these matrices of the form of an identity operator plus a low rank operator. The information contained in the matrices $J^{(k)}$ and $R^{(k)}$ was used as follows. We suppose that these matrices have been retained from a previous problem with the same coefficient matrix but with different data. The component $R^{(k)}t_0$ of the new solution in the space $S^{(k)}$ can then be computed inexpensively by solving the tridiagonal system,

$$J^{(k)}t_0 = R^{(k)T}(M\hat{v}_0 - \hat{c}),$$

where $\hat{v}_0$ and $\hat{c}$ are the initial guess and the data for the new problem, respectively. We can then start the conjugate gradient iteration from the initial point $\hat{v}_0 - R^{(k)}t_0$. This procedure requires $kp + 2k - 1$ additional storage locations. The computational cost is modest since the improved initial guess essentially only requires the calculation of k inner products of length p and the linear combination $R^{(k)}t_0$. The same improved initial guess could also be obtained by using a variable metric algorithm for the first set of data, with the identity matrix as a first approximation of the Hessian, and then using the updated Hessian in the calculation of the second solution. See Broyden [4], Huang [30] and Myers [36]. We note that our method clearly retains only the minimum of necessary information to obtain the projection of the new solution on $S^{(k)}$.

5. Fast Poisson Solvers in Three Dimensions. In this section, we shall describe several variants of a Fourier-Toeplitz method for the discrete Helmholtz equation on a region for which the variables can be separated. We use a Fourier transformation for two of the three variables and solve the tridiagonal linear systems of equations, which result from this change of basis, by a Toeplitz method. See Fischer, Golub, Hald, Leiva and Widlund [16] and Proskurowski and Widlund [44] for descriptions of similar algorithms for two-dimensional problems. As shown by Proskurowski [43], for problems in two dimensions, the execution time of a well-written code of this kind can compare quite favorably with those of good programs implementing other better known methods. We also note that Wilhelmson and Ericksen [58] have presented strong evidence which shows that methods based on Fourier analysis should be chosen for problems in three dimensions. Our methods are designed so that we can guarantee a very high degree of numerical stability for all values of the coefficient c, positive or negative.

We shall consider the solution of the Helmholtz equation

$$-\Delta u + cu = f$$

on the unit cube, $0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1, 0 \leqslant z \leqslant 1$. Periodicity conditions are imposed on the data and the solution by

$$f(x + 1, y, z) = f(x, y + 1, z) = f(x, y, z)$$

and

$$u(x + 1, y, z) = u(x, y + 1, z) = u(x, y, z)$$

and a homogeneous Dirichlet condition is used at $z = 0$,

$$u(x, y, 0) = 0.$$

We also assume that $f(x, y, 0) = 0$. An additional boundary condition is required at $z = 1$ and will be introduced below after a Fourier transformation step. Our methods provide an extension of the solution to all positive values of z. The homogeneous condition at $z = 0$ also allows us to extend the solution and the data to negative values of z by making them odd functions,

$$f(x, y, -z) = -f(x, y, z) \quad \text{and} \quad u(x, y, -z) = -u(x, y, z).$$

When necessary, we extend the data $f(x, y, z)$ by zero for $|z| > 1$. In our experience, an alternative extension, which brings the data more gradually to zero, offers no benefits in our application.

We shall discuss in detail only the seven-point difference approximation and, to simplify our notations, we shall use the same uniform mesh size h in the three coordinate directions. We shall also, without loss of generality, concentrate on the case when $n = 1/h$ is an even number. The discrete Helmholtz problem can be written as

$$(6 + h^2c)u_{ijk} - u_{i+1,jk} - u_{i-1,jk} - u_{i,j+1,k} - u_{i,j-1,k} - u_{ij,k+1} - u_{ij,k-1} = h^2f_{ijk}.$$

The same periodicity and boundary conditions are used for these difference equations.

It is well known that the undivided second centered difference operator, operating on periodic functions, has the normalized eigenfunctions

$$(1/n)^{1/2}(1, 1, \dots, 1)^T \quad \text{and} \quad (1/n)^{1/2}(1, -1, \dots, -1)^T$$

corresponding to the simple eigenvalues 0 and 4, respectively, and the $(n - 2)/2$ double eigenvalues $2 - 2\cos(2\pi l/n)$, $l = 1, 2, \dots, (n - 2)/2$, with the eigenfunctions

$$\begin{aligned} \Phi^{(l)}_{\mathrm{I},k} &= (2/n)^{1/2} \sin(kl2\pi/n), \\ & \qquad\qquad\qquad\qquad k = 0, 1, \dots, n - 1. \\ \Phi^{(l)}_{\mathrm{II},k} &= (2/n)^{1/2} \cos(kl2\pi/n), \end{aligned}$$

The change of basis resulting in the diagonalization of the centered difference operator can be carried out inexpensively by a fast Fourier transform if n has many prime factors; see for example, Cooley, Lewis and Welsh [11].

We choose to work with a partial Fourier transform, transforming with respect to the two variables x and y. The resulting operator can then be represented as the direct sum of n^2 tridiagonal Toeplitz matrices which will be of infinite order if we consider the problem for all positive values of z. The diagonal elements of each of these matrices are equal to one of the numbers,

$$\lambda_{l,m} = 6 + ch^2 - 2\cos(2\pi l/n) - 2\cos(2\pi m/n), \qquad l, m = 0, 1, \dots, n/2,$$

and the off-diagonal elements equal -1.

Thus, these tridiagonal systems of equations can be represented by difference equations,

$$-\hat{u}_{k+1} + \lambda\hat{u}_k - \hat{u}_{k-1} = h^2\hat{f}_k. \tag{5.1}$$

Here $\lambda = \lambda_{l,m}$ and $\hat{f}_k$ and $\hat{u}_k$ are values at $z = kh$ of the appropriate components of the partial Fourier transform of the mesh functions f and u. Since $f(x, y, z) \equiv 0$ for $z > 1$, $\hat{f}_k = 0$ for $k > n$. Once all the components of $\hat{u}$ have been computed, the solution u can be found for the desired values of z by an inverse fast Fourier transform. It is well known that the fast Fourier transform algorithm is very stable.

We solve the tridiagonal systems of equations by two different methods.

Case 1. If $|\lambda| \geqslant 2$, we use a special simple factorization of the matrix into triangular factors. We must first choose the additional boundary condition at $z = 1$. For $k > n$ the difference equation (5.1) is homogeneous and for $|\lambda| > 2$ its solution has the form

$$\hat{u}_k = A\mu^k + B\mu^{-k}.$$

Here A and B are constants and $\mu = \lambda/2 + (\lambda^2/4 - 1)^{1/2}$ and μ^{-1} are the roots of the characteristic equation. We note that $|\mu| > 1$. It is natural to make $A = 0$ since the solution will then decay as $k \longrightarrow +\infty$. This is equivalent to the boundary condition $\hat{u}_{n+1} = \mu^{-1}\hat{u}_n$, and the equation at $z = 1$ reduces to $\mu\hat{u}_n - \hat{u}_{n-1} = h^2\hat{f}_n$. The resulting $n \times n$ tridiagonal matrix can be written as

$$\begin{pmatrix} \mu & -1 & & & \\ -1 & \lambda & -1 & & \\ & -1 & \lambda & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & \lambda \end{pmatrix}.$$

We have ordered the unknowns in order of decreasing indices $(\hat{u}_n, \dots, \hat{u}_1)$ and used the homogeneous Dirichlet condition at $z = 0$ to obtain the last row of the matrix. This matrix has a most convenient factorization, as the product of two bidiagonal Toeplitz matrices

$$\begin{pmatrix} 1 & & & & \\ -\mu^{-1} & 1 & & & \\ & -\mu^{-1} & 1 & & \\ & & \ddots & \ddots & \\ & & & -\mu^{-1} & 1 \end{pmatrix} \begin{pmatrix} \mu & -1 & & & \\ & \mu & -1 & & \\ & & \mu & -1 & \\ & & & \ddots & \ddots \\ & & & & \mu \end{pmatrix}.$$

The linear systems can, therefore, be solved by using very simple two term recursion procedures which are highly stable since $|\mu| > 1$. The same procedure also works well for the case when $|\lambda| = 2$.

Case 2. If $|\lambda| < 2$, the roots of the characteristic equation fall inside the unit circle; and we can use the three term recursion formula (5.1) to compute $\hat{u}_k$ in a stable way. Before we can use this marching procedure, we need to find a value of $\hat{u}_1$ to provide a second initial value in addition to $\hat{u}_0 = 0$. This can be done by using the formula

$$\hat{u}_j = \sum_{k=1}^{n} \frac{\sin(|j+k|\phi) - \sin(|j-k|\phi)}{2 \sin \phi} h^2 \hat{f}_k,$$

which can easily be verified to give a solution of the difference equation. Here $\phi = \arccos(\lambda/2)$. For $j = 1$, we find the simple formula,

$$\hat{u}_1 = \sum_{k=1}^{n} \frac{\sin((k+1)\phi) - \sin((k-1)\phi)}{2 \sin \phi} h^2 \hat{f}_k = \sum_{k=1}^{n} \cos(k\phi) h^2 \hat{f}_k.$$

There are other solutions of the difference equation (5.1), but the present choice gives the same solution in the limit case $|\lambda| = 2$ as the method developed for Case 1. We, therefore, obtain a solution of the Helmholtz problem which is a continuous function of the parameter c. We also note that by our choice of boundary conditions, instability has been avoided for all values of the parameter c.

The method requires $n^3(1 + o(1))$ storage locations and, if n is a power of two, on the order of $n^3(\log_2 n + 1)$ arithmetic operations.

Although quite efficient this algorithm does not fully exploit the structure of our problem. During the conjugate gradient iteration the mesh functions representing the right-hand sides of the Helmholtz equation vanish except at mesh points used for the construction of the discrete single or dipole layers. Similarly during this main part of the calculation, we need the solution only at the points of the stencils of the irregular mesh points. Thus, on any line parallel to a coordinate axes only a few source and target points have to be considered.

We shall now briefly describe a method due to Banegas [1]. For large problems the direct and inverse Fourier transforms with respect to one of the variables can be carried out more economically by computing inner products of sparse vectors and the basis vectors of the new coordinate system. The fast Fourier transform should be used for the second variable because after the first Fourier transform step the arrays will no longer be sparse. The main advantage of this variant is that it can be implemented using only a two-dimensional work array if the necessary information on the coordinates and values of the source and target points is stored elsewhere. Only on the order of $N^{2/3}$ storage locations are, therefore, required for the main iteration. See Banegas [1] and Proskurowski [42] for more details and a discussion of the use of a similar algorithm for Helmholtz problems in two dimensions. The three-dimensional algorithm has not yet been implemented. The savings in storage would not show dramatically for problems in three dimensions unless a million words of storage is available.

The calculation of the space potential terms and the final solution can also be carried out without using arrays with n^3 elements. See Proskurowski [42] for a design of a third variant of a Fourier-Toeplitz method. It requires access to all elements of the right-hand side twice but no intermediary results need to be written on secondary storage devices. The primary storage requirement can be reduced drastically at an expense of a modest increase of the computational work.

We conclude this section by proving a result needed in connection with Theorem 3.2. We restrict ourselves to $z \geqslant 0$ and assume, as in that theorem, that $c \geqslant 0$.

THEOREM 5.1. *Let f have its support in* $0 < z \leqslant 1$, *and let* $c \geqslant 0$. *The mesh function* $u = Gf$, *defined by the Fourier-Toeplitz method of this section, takes on a maximum or a minimum.*

Proof. We first consider the case of $c > 0$. By construction all modes of the solution decay as $z \longrightarrow \infty$. The conclusion then follows since we need to consider only a finite subset of the mesh.

For $c = 0$, we partition the solution into two parts, $u = u_0 + u_1$. The function u_0 corresponds to the lowest frequency for which $\lambda = 2$. It is easy to see that u_0 depends only on z and that it reduces to a linear function for $z > 1$. u_1 has a zero average for each z and decays as $z \longrightarrow \infty$. If u_0 is an unbounded function, the conclusion easily follows. If u_0 is constant for $z > 1$, u takes on a maximum and a minimum on that set since any nontrivial u_1 changes sign for each z and decays as $z \longrightarrow \infty$. If the maximum and minimum of u on $0 < z \leqslant 1$ are also considered, an extremal value of u on $z > 0$ can be found.

6. Implementation of the Algorithm and Numerical Results.

6.1. *The Program in Outline.* We have implemented a capacitance matrix algorithm for the three-dimensional Helmholtz equation as a FORTRAN program. The Shortley-Weller approximation of the Dirichlet boundary condition described in subsection 2.2 is used, and a normal equation form of the capacitance matrix equation is solved by using the conjugate gradient method described in subsection 4.2. Discrete dipoles are used as in subsection 3.2.

In designing the program, clarity and ease of modification have been prime objectives with efficiency in execution time and storage important but secondary. The program has been successfully checked by the CDC ANSI FORTRAN verifier on the CDC 6600 at the Courant Institute. No machine dependent constants are used.

The program is fully documented in a Courant Institute technical report, so we give only an outline of the program here.

The main subroutine HELM3D is the only subroutine with which the user needs to have direct contact. The geometric information necessary to describe the region, the data for the differential equation, scratch storage space and convergence tolerances are passed to this routine.

The coordinates of the irregular mesh points, altogether 3(IP1 + IP2) integer values, are needed. Here IP1 is the number of irregular points with at most one neighbor on or outside the boundary in each coordinate direction, and IP2 is the number of remaining irregular points.

The signed distances from the irregular mesh points to the boundary in the x, y and z directions, 3IP1 + 6IP2 real values, are also required.

The data is entered by using four real arrays. The values of the inhomogeneous term f at the mesh points are stored in a three-dimensional array of dimension NX × NY × NZ where NX, NY and NZ are the number of mesh points in the different coordinate directions in the rectangular parallelepiped in which the region is embedded. Values of this mesh function can be set arbitrarily at mesh points on or outside of the boundary. The boundary data, i.e. the values of the solution at the points where mesh lines cross the boundary, are stored in three one-dimensional arrays requiring 3IP1 + 6IP2 real words of storage.

In total two real three-dimensional arrays of dimension NX × NY × NZ and eleven one-dimensional arrays are used. One of the one-dimensional arrays is real and of dimension max (IP1 + 2IP2, NX × NZ, NY × NZ). The remaining four integer and six real arrays are of length IP1 + 2IP2. The need for array space could be decreased by, among other things, packing the coordinates of the irregular points into one array. If f is zero, one of the three-dimensional arrays is eliminated simply by not dimensioning it in the calling program. In the general case, this second array could be kept on a secondary storage device with very little degradation in the performance of the program. For a discussion of further possible reduction of array space, see Section 5.

The conjugate gradient iteration is controlled by two input parameters NIT, the maximum number of iterations allowed, and EPS, a tolerance for the norm of the residual.

Upon termination the approximate solutions of the Helmholtz and capacitance matrix equations and the residual of the capacitance matrix equation are available. The values of the three-dimensional array containing the solution at mesh points on or outside of the boundary are useless by-products of the calculation. The capacitance matrix solution can be refined, if so desired, by additional calls of HELM3D using current values of the dipole strength and the residual.

A sample driver is provided in our program to illustrate the use of the HELM3D subroutine. We note that we have found it relatively convenient to describe our regions in terms of inequalities.

HELM3D calls other subroutines to set up the right-hand side and solves the capacitance matrix equation. It is the only subroutine which needs to be modified in order to incorporate the singular value estimates or the accumulation of an approximate inverse discussed in subsection 4.4. The right-hand side of the capacitance matrix equation is calculated by the subroutine BNDRY. The subroutines BNDRY, UTAMLT and UTATRN, all related to the finite difference formulas near the boundary, must be changed if a different approximation of the boundary condition is to be implemented. The two subroutines VMULT and VTRANS depend on the discrete dipole construction. Single layer versions of these subroutines should be written if the program is modified to solve the Neumann problem.

The fast Poisson solver of Section 5 is implemented in subroutine CUBE. It uses two FFT subroutines RFORT and FORT provided by Dr. W. Proskurowski, who has modified code written by Dr. J. Cooley.

The product of the capacitance matrix C_D and an arbitrary vector is formed by calling the subroutines VMULT, CUBE and UTAMLT. Similarly, the product of C_D^T and a vector is formed by using UTATRN, CUBE and VTRANS.

The system also has an error checking module, HELMCK. This subroutine checks that enough storage space has been allocated, that the indices of the irregular points are within range, that no irregular points are missing or listed twice and that the discrete dipoles point out of the region.

One of the three-dimensional arrays, w, is used when checking the geometric information for self-consistency. For each irregular point the corresponding element of w is set to indicate $\partial\Omega_h$ after a check that this point has not been previously marked as irregular or exterior. The current values of w at the six neighbors of the point are checked for consistency by using the distances to the boundary which are given as data. Appropriate elements of w are then set to indicate that these points belong to $\Omega_h \cup \partial\Omega_h$ or $C\Omega_h$.

Each line of points of the three-dimensional array begins at an outside point. In a second stage, we march across each line, setting w to indicate $C\Omega_h$ until an indicator of Ω_h (signalling an error) or $\partial\Omega_h$ is encountered. We proceed along the line, setting w elements to indicate Ω_h whenever appropriate, until we leave the region via a point of $\partial\Omega_h$. In this way an array is created which could be used to display the subsets Ω_h, $\partial\Omega_h$ and $C\Omega_h$ graphically. We then use this array and the data on the distances to the boundary to check that no dipole charge falls on an interior mesh point; see subsection 3.2. Finally, we make sure that no interior mesh point has an exterior neighbor.

Our code could be modified to perform these checks locally, without using a three-dimensional array.

The execution time could be reduced in several ways. In the current program the coefficients for the difference equation at the irregular mesh points and the dipole weights are recomputed every time they are used. Storage of these elements would save time. The subroutine CUBE can be replaced by a faster Poisson solver. Overhead in subroutine calls could be reduced through the use of COMMON.

6.2. *Numerical Experiments.* Extensive numerical experiments have been carried out with our program on the CDC 6600 at the Courant Institute and the Amdahl 470V/6 at the University of Michigan. Dr. W. Proskurowski has also kindly run some problems on a CDC 7600 at the Lawrence Berkeley Laboratory. We report in detail only on experiments carried out on the CDC 6600 using a FTN, OPT = 2, compiler and no more than 50000 words of storage for the arrays. In our experience, the program runs about six times faster on a CDC 7600.

The runs reported have been made for problems with the solutions $x^2 + y^2 + 2z^2$ and $x^2 + y^2 - 2z^2$, but extensive experiments with other types of data make us confident that the performance of our algorithm is virtually independent of the right-hand side. The efficiency of our method as a highly specialized linear equation solver can easily be studied for these simple solutions since there is no truncation error. For the finest meshes, we consider only homogeneous problems, i.e. $f \equiv 0$, in order to save one three-dimensional array. The initial guess is always chosen to be zero.

The parameter EPS is used in the stopping criterion of the conjugate gradient algorithm. The iteration is terminated when the Euclidean norm of the residual of the capacitance matrix equation drops below EPS × $\sqrt{\mathrm{IP}}$ where IP = IP1 + IP2. The condition number of $C_D^T C_D$, $\kappa(C_D^T C_D)$, is estimated by using ideas from subsection 4.4 and the TQL1 subroutine of EISPACK. The time required for this calculation is included in the tables.

Three regions have been used in these experiments and the results are reported in Tables 1–3. The smallest recorded times for the execution of the fast Poisson solver are .055, .432 and 2.757 seconds for 8 × 8 × 9, 16 × 16 × 17 and 32 × 32 × 24 points, respectively.

TABLE 1

Radius of sphere	.360		.424			.424			.447	
Number of interior and irregular points	93		1357			7556			8796	
Number of irregular points, IP	66		438			1522			1698	
NX × NY × NZ	8 × 8 × 9		16 × 16 × 17			32 × 32 × 24			32 × 32 × 24	
Condition number, $\kappa(C_D^T C_D)$	14.7		39.7			56.7			77.2	
Tolerance, EPS	.1E-2	.1E-5	.1E-2	.1E-5	.1E-8	.1E-2	.1E-5	.1E-8	.1E-2	.1E-5
Number of iterations	5	9	7	15	22	8	17	26	8	17
Maximum error	.403E-2	.936E-5	.314E-1	.167E-5	.596E-8	.384E-1	.367E-4	.262E-7	.584E-1	.548E-4
Total execution time in seconds	.952	1.58	9.03	17.9	25.9	58.5	117	173	58.8	116
Percentage of time spent using the fast Poisson solver	65.0	68.8	72.7	76.1	76.8	81.1	84.3	85.5	80.1	83.4

Experiments with spherical regions centered at (.5,.5,.5) with c = 0.

When we examine the tables, we note the very modest growth in the number of iterations when the size of the problem increases. The stability of our method is further illustrated by the very accurate solutions obtained when the tolerance EPS is chosen to be very small.

TABLE 2

Number of interior and irregular points	2050			10464		
Number of irregular points, IP	1000			3172		
NX × NY × NZ	16 × 16 × 17			32 × 32 × 24		
Condition number, $\kappa(C_D^T C_D)$	602			554		
Tolerance, EPS	.1E-2	.1E-5	.1E-8	.1E-2	.1E-5	.1E-8
Number of iterations	13	23	32	13	23	35
Maximum error	.258E-1	.325E-4	.377E-7	.517E-1	.554E-3	.805E-7
Total execution time in seconds	19.7	33.1	45.5	101	171	255
Percentage of time spent using the fast Poisson solver	62.3	63.3	63.5	75.3	77.1	77.6

Experiments with c = 0 and a cube with a sphere cut out, $0.1 \leq x \leq 0.9$, $0.1 \leq y \leq 0.9$, $0.1 \leq z \leq 0.9$ and $x^2 + y^2 + z^2 \geq (0.2)^2$.

The experiments of Table 3 require some further comments. Faster methods are of course available for rectangular regions. This region has been chosen since the eigenvalues of the discrete Laplace operator are known explicitly. We note that when *c* is

TABLE 3

The constant c	100			0			-34.892	-52.238	-77.91		-205.5
Condition number $\kappa(C_D^T C_D)$	2.39			27.1			42.2	6.07E+6	4.35E+3		8.78E+5
Tolerance, EPS	.1E-3	.1E-5	1E-11	.1E-3	.1E-5	.1E-11	.1E-7	.1E-7	.1E-5	.1E-11	-
Number of iterations	4	6	15	8	12	23	22	42	47	66	200
Maximum error	.121E-2	.233E-4	.140E-10	.433E-2	.177E-4	.201E-10	.371E-6	.124E-6	.343E-4	.372E-10	.995E-5
Execution time in seconds	6.76	9.43	20.2	11.4	16.4	30.0	27.9	52.3	60.0	85.6	259

Experiment with the region $0.125 \le x \le 0.875$, $0.125 \le y \le 0.875$ and $0.125 \le z \le 0.875$ and different values of c and EPS. The number of interior and irregular points is 1331 and IP, NX, NY and NZ are 602, 16, 16 and 17, respectively. Between 70.3 and 74.1% of the execution time is used by the fast Poisson solver.

large and positive, as in the application of our method to the solution of a parabolic equation by an implicit method, the convergence is extremely rapid. In such applications an excellent initial guess is also normally available. Negative values of c lead to more difficult problems. The smallest eigenvalue of the operator is $\lambda_{\min} = 52.337926 \ldots$ and another eigenvalue is equal to $205.78497 \ldots$. The values 34.892 and 77.91 approximate $(2/3)\lambda_{\min}$ and the average of the two smallest eigenvalues, respectively. The problems which are almost singular or indefinite are very ill-conditioned. However, only a few eigenvalues of $C_D^T C_D$ are very small, and the conjugate gradient method is still relatively successful; see further discussion in Proskurowski and Widlund [44].

Using the approximate inverse idea of subsection 4.4, improved initial approximations for the discrete dipole strengths have been obtained for a series of problems on a spherical region. To illustrate the performance of this method, we consider the problem of Table 1 with 1357 unknowns. The tolerance EPS was chosen to be .1E-4 and 14 iterations were required. Eight vectors were saved from this run and used to construct an initial approximation of the discrete dipole layer for two problems with solutions drastically different from the previous one. For these subsequent problems only nine iterations were required to reach a comparable accuracy. In implementing this method, precautions must be taken to insure that round-off does not contaminate the computation. The orthogonality of the residual vectors should be monitored and vectors and parameters computed after loss of orthogonality must be discarded. With careful implementation, this can be a very effective technique and can lead to substantial savings when many problems are to be solved for the same region.

Acknowledgements. The authors want to thank John G. Lewis, Włodzimierz Proskurowski and Arthur Shieh for their interest and help with this project.

Computer Science Department
University of Maryland
College Park, Maryland 20742

Courant Institute of Mathematical Sciences
New York University
New York, New York 10012

1. A. BANEGAS, "Fast Poisson solvers for problems with sparsity," *Math. Comp.*, v. 32, 1978, pp. 441–446.

2. R. E. BANK, "Marching algorithms for elliptic boundary value problems. II: The variable coefficient case," *SIAM J. Numer. Anal.*, v. 14, 1977, pp. 950–970.

3. R. E. BANK & D. J. ROSE, "Marching algorithms for elliptic boundary value problems. I: The constant coefficient case," *SIAM J. Numer. Anal.*, v. 14, 1977, pp. 792–829.

4. C. G. BROYDEN, "Quasi-Newton methods" in *Numerical Methods for Unconstrained Optimization* (W. Murray, Ed.), Academic Press, New York, 1972, pp. 87–106.

5. O. BUNEMAN, *A Compact Non-Iterative Poisson Solver*, Report SUIPR-294, Inst. Plasma Research, Standford University, 1969.

6. B. L. BUZBEE & F. W. DORR, "The direct solution of the biharmonic equation on rectangular regions and the Poisson equation on irregular regions," *SIAM J. Numer. Anal.*, v. 11, 1974, pp. 753–763.

7. B. L. BUZBEE, F. W. DORR, J. A. GEORGE & G. H. GOLUB, "The direct solution of the discrete Poisson equation on irregular regions," *SIAM J. Numer. Anal.*, v. 8, 1971, pp. 722–736.

8. B. L. BUZBEE, G. H. GOLUB & C. W. NIELSON, "On direct methods for solving Poisson's equation," *SIAM J. Numer. Anal.*, v. 7, 1970, pp. 627–656.

9. L. COLLATZ, *The Numerical Treatment of Differential Equations*, Springer-Verlag, Berlin and New York, 1960.

10. P. CONCUS, G. H. GOLUB & D. P. O'LEARY, "A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations," *Sparse Matrix Computations* (J. R. Bunch and D. J. Rose, Eds.), Academic Press, New York, 1976, pp. 309–332.

11. J. W. COOLEY, P. A. W. LEWIS & P. D. WELSH, "The fast Fourier transform algorithm: Programming considerations in the calculation of sine, cosine and Laplace transforms," *J. Sound Vib.*, v. 12, 1970, pp. 315–337.

12. J. W. DANIEL, "The conjugate gradient method for linear and nonlinear operator equations," *SIAM J. Numer. Anal.*, v. 4, 1967, pp. 10–26.

13. S. C. EISENSTAT, "Complexity bounds for Gaussian elimination." (To appear.)

14. S. C. EISENSTAT, M. H. SCHULTZ & A. H. SHERMAN, "Applications of an element model for Gaussian elimination," *Sparse Matrix Computations* (J. R. Bunch and D. J. Rose, Eds.), Academic Press, New York, 1976, pp. 85–96.

15. M. ENGELI, TH. GINSBURG, H. RUTISHAUSER & E. STIEFEL, *Refined Iterative Methods for Computation of the Solution and the Eigenvalues of Self-Adjoint Boundary Value Problems*, Birkhäuser, Basel-Stuttgart, 1959.

16. D. FISCHER, G. GOLUB, O. HALD, C. LEIVA & O. WIDLUND, "On Fourier-Toeplitz methods for separable elliptic problems," *Math. Comp.*, v. 28, 1974, pp. 349–368.

17. G. E. FORSYTHE & W. R. WASOW, *Finite Difference Methods for Partial Differential Equations*, Wiley, New York, 1960.

18. P. R. GARABEDIAN, *Partial Differential Equations*, Wiley, New York, 1964.

19. J. A. GEORGE, *The Use of Direct Methods for the Solution of the Discrete Poisson Equation on Non-Rectangular Regions*, Computer Science Department Report 159, Stanford University, 1970.

20. A. GEORGE, "Nested dissection of a regular finite element mesh," *SIAM J. Numer. Anal.*, v. 10, 1973, pp. 345–363.

21. R. M. HAYES, "Iterative methods of solving linear problems on Hilbert space," *Contributions to the Solution of Systems of Linear Equations and the Determination of Eigenvalues* (O. Taussky, Ed.), Nat. Bur. Standards Appl. Math. Series, Vol. 39, 1954, pp. 71–103.

22. M. R. HESTENES, *The Conjugate Gradient Method for Solving Linear Systems*, Proc. Sympos. Appl. Math., Vol. 6, Amer. Math. Soc., Providence, R. I., 1956, pp. 83–102.

23. M. R. HESTENES & E. STIEFEL, "Methods of conjugate gradients for solving linear systems," *J. Res. Nat. Bur. Standards*, v. 49, 1952, pp. 409–436.

24. R. W. HOCKNEY, "A fast direct solution of Poisson's equation using Fourier analysis," *J. Assoc. Comput. Mach.*, v. 12, 1965, pp. 95–113.

25. R. W. HOCKNEY, "Formation and stability of virtual electrodes in a cylinder," *J. Appl. Phys.*, v. 39, 1968, pp. 4166–4170.

26. R. W. HOCKNEY, "The potential calculation and some applications," *Methods in Computational Physics*, Vol. 9, Academic Press, New York, 1970.

27. R. W. HOCKNEY, POT 4—*A Fast Direct Poisson Solver for the Rectangle Allowing Some Mixed Boundary Conditions and Internal Electrodes*, IBM Research, R. C. 2870, 1970.

28. A. J. HOFFMAN, M. S. MARTIN & D. J. ROSE, "Complexity bounds for regular finite difference and finite element grids," *SIAM J. Numer. Anal.*, v. 10, 1973, pp. 364–369.

29. A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.

30. H. Y. HUANG, "Unified approach to quadratically convergent algorithms for function minimization," *J. Optimization Theory Appl.*, v. 5, 1970, pp. 405–423.

31. S. KANIEL, "Estimates for some computational techniques in linear algebra," *Math. Comp.*, v. 20, 1966, pp. 369–378.

32. O. D. KELLOGG, *Foundations of Potential Theory*, Springer-Verlag, Berlin, Heidelberg, New York, 1967.

33. C. LANCZOS, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Standards*, v. 45, 1950, pp. 255–282.

34. D. G. LUENBERGER, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass., 1973.

35. E. D. MARTIN, "A generalized capacity matrix technique for computing aerodynamic flows," *Internat. J. Comput. Fluids*, v. 2, 1974, pp. 79–97.

36. G. E. MYERS, "Properties of the conjugate gradient and Davidon methods," *J. Optimization Theory Appl.*, v. 2, 1968, pp. 209–219.

37. J.-C. NEDELEC & J. PLANCHARD, "Une méthode variationelle d'éléments finis pour la résolution numérique d'un problème extérieur dans R^3," *R.A.I.R.O.*, v. 7-R, 1973, pp. 105–129.

38. C. C. PAIGE, *The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices*, Ph.D. thesis, London University, Institute of Computer Science, 1971.

39. V. PEREYRA, W. PROSKUROWSKI & O. WIDLUND, "High order fast Laplace solvers for the Dirichlet problem on general regions," *Math. Comp.*, v. 31, 1977, pp. 1–16.

40. G. N. POLOZHII, *The Method of Summary Representation for Numerical Solution of Problems of Mathematical Physics*, Pergamon Press, New York, 1965.

41. W. PROSKUROWSKI, "On the numerical solution of the eigenvalue problem of the Laplace operator by the capacitance matrix method," *Computing*, v. 20, 1978, pp. 139–151.

42. W. PROSKUROWSKI, "Numerical solution of Helmholtz's equation by implicit capacitance matrix methods," *ACM Trans. Math. Software*, v. 5, 1979, pp. 36–49.

43. W. PROSKUROWSKI, *Four* FORTRAN *Programs for Numerically Solving Helmholtz's Equation in an Arbitrary Bounded Planar Region*, Report LBL-7516, Lawrence Berkeley Laboratory, 1978.

44. W. PROSKUROWSKI & O. WIDLUND, "On the numerical solution of Helmholtz's equation by the capacitance matrix method," *Math. Comp.*, v. 30, 1976, pp. 433–468. Appeared also as an ERDA-NYU report.

45. W. PROSKUROWSKI & O. WIDLUND, "A finite element-capacitance method for the Neumann problem for Laplace's equation." (To appear.)

46. A. S. L. SHIEH, *Fast Poisson Solver on Nonrectangular Domains*, Ph.D. thesis, New York University, 1976.

47. A. S. L. SHIEH, "On the convergence of the conjugate gradient method for singular capacitance matrix equations from the Neumann problem of the Poisson equation," *Numer. Math.*, v. 29, 1978, pp. 307–327.

48. A. S. L. SHIEH, "Fast Poisson solvers on general two dimensional regions for the Dirichlet problem," *Numer. Math.*, v. 31, 1979, pp. 405–429.

49. G. STRANG & G. J. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, N. J., 1973.

50. P. N. SWARZTRAUBER, "A direct method for the discrete solution of separable elliptic equations," *SIAM J. Numer. Anal.*, v. 11, 1974, pp. 1136–1150.

51. P. N. SWARZTRAUBER, "The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle," *SIAM Rev.*, v. 19, 1977, pp. 490–501.

52. P. N. SWARZTRAUBER & R. A. SWEET, "The direct solution of the discrete Poisson equation on a disk," *SIAM J. Numer. Anal.*, v. 10, 1973, pp. 900–907.

53. P. SWARZTRAUBER & R. SWEET, *Efficient* FORTRAN *Subprograms for the Solution of Elliptic Partial Differential Equations*, Report NCAR-TN/IA-109, National Center for Atmospheric Research, Boulder, Colorado, 1975.

54. R. SWEET, "A cyclic reduction algorithm for solving block tridiagonal systems of arbitrary dimension," *SIAM J. Numer. Anal.*, v. 14, 1977, pp. 706–720.

55. V. THOMÉE, "On the convergence of difference quotients in elliptic problems," *Numerical Solution of Field Problems in Continuum Physics*, SIAM-AMS Proc., Vol. 2, Amer. Math. Soc., Providence, R. I., 1970, pp. 186–200.

56. O. WIDLUND, "On the use of fast methods for separable finite difference equations for the solution of general elliptic problems," *Sparse Matrices and Their Applications* (D. J. Rose and R. A. Willoughby, Ed.), Plenum Press, New York, 1972, pp. 121–134.

57. O. WIDLUND, *Capacitance Matrix Methods for Helmholtz's Equation on General Bounded Regions*, Proc. Oberwolfach meeting 1976, Springer Lecture Notes in Math., Vol. 631, 1978, pp. 209–219.

58. R. B. WILHELMSON & J. H. ERICKSEN, "The direct solutions for Poisson's equation in three dimensions," *J. Computational Phys.*, v. 25, 1977, pp. 319–331.