MATHEMATICS OF COMPUTATION, VOLUME 30, NUMBER 135
JULY 1976, PAGES 433–468

On the Numerical Solution of Helmholtz's Equation by the Capacitance Matrix Method

By Włodzimierz Proskurowski and Olof Widlund*

Abstract. In recent years the usefulness of fast Laplace solvers has been extended to problems on arbitrary regions in the plane by the development of capacitance matrix methods. The solution of the Dirichlet and Neumann problems for Helmholtz's equation is considered. It is shown, that by an appropriate choice of the fast solver, the capacitance matrix can be generated quite inexpensively. An analogy between capacitance matrix methods and classical potential theory for the solution of Laplace's equation is explored. This analogy suggests a modification of the method in the Dirichlet case. This new formulation leads to well-conditioned capacitance matrix equations which can be solved quite efficiently by the conjugate gradient method. A highly accurate solution can, therefore, be obtained at an expense which grows no faster than that for a fast Laplace solver on a rectangle when the mesh size is decreased.

1. Introduction. In the last decade several very fast methods have been developed for the direct solution of the quite special systems of linear algebraic equations which arise when Laplace's or Helmholtz's equation,

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

is solved by standard finite-difference methods on certain simple regions in the plane. The best known of these methods are due to Hockney [23], [25] and Buneman [5]; see also Fischer, Golub, Hald, Leiva and Widlund [15] and Section 6 of this paper, in which a so-called Fourier-Toeplitz method is developed. For further studies of such methods see Dorr [14] and Buzbee, Golub and Nielson [8].

All these methods can be regarded as efficient computer implementations of the separation of variables method. That technique can only be used for regions which, after a possible change of independent variables, are rectangular and for differential operators of a special form. Typical examples are Laplace's and Helmholtz's equations in Cartesian coordinates on rectangular regions equipped with boundary conditions which do not change type along any of the sides of the rectangle. Another example is provided by the same equations on circular regions for which the use of polar coordinates is appropriate. Similar restrictions are imposed on the corresponding discrete problems. For a discussion of the special structure which is needed to allow for the use of these methods, see Widlund [46]. We note that for a linear system of equations

Received February 2, 1976.
AMS (MOS) subject classifications (1970). Primary 65N20; Secondary 65F05, 65F10.
* The work of this author was supported by ERDA, Contract E(11-1)-3077, at the Courant Institute of Mathematical Sciences, New York University, and in part, by the National Science Foundation, Grant GP–35747, and by the Swedish Science Research Council, (NFR), Grant 2711–012.

with n^2 unknowns arising from Laplace's equation on a square region using n uniformly distributed mesh points in each direction, only $n^2(1 + O(1/n))$ storage locations are needed, while the operation count is $Cn^2\log_2 n(1 + o(1))$. Here C is a constant less than ten.

These fast methods can also be used for elliptic problems on regions, or with boundary conditions, which do not allow for separation of variables, provided that the differential operator allows for separation of variables on a different region or with different boundary conditions, see Buzbee, Dorr, George and Golub [7], Buzbee and Dorr [6] and Hockney [26]. This last condition amounts to a requirement that the differential operator can be split into the sum of commuting operators, each of which contains derivatives with respect to only one variable.

We now briefly describe how a problem can be treated on a bounded region Ω which does not allow for separation of variables. We will limit our discussion throughout to two independent variables. The region Ω is imbedded in either a rectangle, an infinite parallel strip or the whole plane. A uniform mesh is imposed on the enlarged region. An expanded linear system of algebraic equations is derived which has a reducible matrix; see Section 3. This new matrix contains the matrix of our original problem as an irreducible component. The resulting matrix is a rank p modification of a problem which allows for the use of a fast solver. Here p denotes the number of boundary mesh points of the original finite-difference equation. For two-dimensional problems we thus have a value of p of the order n. The problem can be solved with the aid of the Woodbury formula or one of its variants, see Buzbee, Dorr, George and Golub [7]. Similar ideas are also discussed in George [18], Hockney [26] and Martin [35]. It is interesting to note that Hockney and Martin present their algorithms in terms of capacitance matrices, an idea borrowed from potential theory, rather than in terms of the Woodbury formula, which can be regarded merely as an algebraic identity. We note that Bückner (see Todd [45]), has pointed out a relation between potential theory and the method of tearing which is closely related to the Woodbury formula. For a further discussion of previous work see Section 8.

The work reported in this paper on the solution of the interior Dirichlet and Neumann problems grew out of an observation of a formal analogy between the Woodbury formula and a classical solution formula for the Neumann problem for Laplace's equation as presented in Courant-Hilbert [12], Garabedian [17] and Petrowsky [41]. In this potential theoretical approach, which goes back to Neumann and Fredholm, an Ansatz is made in terms of a single layer potential. The charge density is then found by solving a Fredholm integral equation of the second kind. This operator has a simple zero eigenvalue, but it is bounded and has a bounded inverse on a subspace of codimension one. The close analogy between the integral operator and the capacitance matrix which appears in the Woodbury formula suggests that the family of these matrices might be uniformly bounded and have uniformly bounded inverses on subspaces of codimension one. For such uniformly well-conditioned problems, iterative methods might compete successfully with Gaussian elimination. This conjecture has been borne out in practice. We have, therefore, chosen to solve the $p \times p$ linear system of equations

by the conjugate gradient method. The motivation for the choice of this method in this context is detailed in Section 7. The use of this method results in considerable savings compared to previous implementations of the capacitance matrix method in cases where the number of variables is large and only one or a few problems are solved for a given region Ω. Our interest in the conjugate gradient method began when we searched for an explanation for the rapid convergence reported by George [18] with his so-called iterative imbedding algorithms; see further Section 8.

When we turn to the Dirichlet problem, we find that the use of the Woodbury formula corresponds to the attempt to solve the original differential equation in terms of a single-layer potential. That approach is known to give rise to a Fredholm integral equation of the first kind and thus an ill-posed problem. While it is possible to prove that the resulting capacitance matrix is nonsingular for any finite value of p, the ill-posedness of the continuous problem is reflected in a growth of the condition number of the capacitance matrices when p increases. In our experiments, as reported in Section 9, we found that the condition numbers can become large and, also, that they vary rather erratically from case to case. This might lead to a loss of accuracy of the algorithm as previously implemented. The conjugate gradient method also becomes quite uneconomical in this context.

But for the continuous Dirichlet problem, it is known that the proper Ansatz is a layer of dipoles. Changing our Ansatz to a finite-difference analog of a dipole layer, the capacitance matrices, in our experience, are again quite well-conditioned and the conjugate gradient method performs very satisfactorily. We note that this successful Ansatz falls outside the algebraic framework of Buzbee, Dorr, George and Golub [7]. Our treatment also differs from theirs in the Neumann case for Laplace's equation in that we allow the capacitance matrix to become singular.

A potential theory for discrete Poisson problems is developed in Sections 3–5. While our theory is mostly formal it has served very well as a guide in choosing a proper Ansatz for the Dirichlet problem and certain scale factors. A rigorous mathematical analysis of our method is given in certain cases in the thesis of Shieh [42].

Another problem in the use of the method as previously developed is the considerable effort expended in generating the capacitance matrix. A solution has been outlined briefly in Widlund [46]. Taken together these improvements result in an operation count of $Cn^2 \log_2 n(1 + o(1))$ compared to $Cpn^2 \log_2 n(1 + o(1))$ for the algorithm of Buzbee, Dorr, George and Golub [7]. The basic idea involved in the fast generation of the capacitance matrix is the use of translation invariance which can be achieved by imposing a periodicity condition as a boundary condition for the problem for which the fast Poisson solver is applied. The absence of a boundary in such cases makes the discrete fundamental solution very simple. The matrix representing it is a circulant; and therefore, by knowing one of its columns, we know the entire matrix. We note that the same observation was made independently by Peskin [40] and used in his calculations of the blood flow in the heart.

In our numerical experiments, reported in Section 9, we have used the Fourier-Toeplitz method for the separable problems. Other fast Laplace or Helmholtz solvers

could also have been chosen. While the experiments reported in this paper have been carried out using the five-point formula and the well-known Shortley-Weller approximation to the boundary conditions, we believe that our version of the capacitance matrix method should work equally well for other difference schemes. A series of experiments with a high order method suggested by Heinz-Otto Kreiss has been carried out. The number of conjugate gradient iterations and the over all performance of the algorithm is virtually unchanged. Results from these experiments and a description of the method will be given in Pereyra, Proskurowski and Widlund [39]. It should also be noted that a fast Helmholtz solver can be used to a great advantage in the iterative solution of more general elliptic problems; see Concus and Golub [10] and Bartels and Daniel [2]. Similarly, the use of fast Helmholtz solvers for transonic flow calculations has been advocated by Jameson [29] and Martin [36], [37]. The first author has carried out preliminary work to modify our method to handle eigenvalue problems. Dianne P. O'Leary and the second author are exploring the extension of a variant of our method in three dimensions.

This paper has previously appeared as an ERDA-NYU report. That version includes a FORTRAN program for nonnegative values of the constant c, a description of and a guide to the use of our program.

Acknowledgements. The authors want to thank J. Cooley for providing us with a fine fast Fourier transform program and C. Paige and M. Saunders for an excellent conjugate gradient routine. The authors also want to thank P. Concus, G. Dahlquist, A. George, G. Golub, O. Hald, H.-O. Kreiss, D. Martin, D. P. O'Leary, S. Osher, C. Peskin, V. Pereyra and A. Shieh for their interest and encouragement.

2. Certain Results from Classical Potential Theory. We give only a brief review of a few results of potential theory. For a detailed exposition see the references given in the previous section. We begin by defining the potential V resulting from a charge distribution ρ on a boundary curve $\partial\Omega$,

$$V(x) = (1/\pi)\int_{\partial\Omega}\rho(\xi)\log(1/r)\,ds(\xi).$$

We will assume throughout that the boundary $\partial\Omega$ is sufficiently smooth. Here $x = (x_1, x_2)$, $\xi = (\xi_1, \xi_2)$ and $r^2 = (x_1 - \xi_1)^2 + (x_2 - \xi_2)^2$. We note that $(1/2\pi)\log(1/r)$ is a fundamental solution of Poisson's equation, i.e.

$$-\Delta(1/2\pi)\log(1/r) = \delta(r),$$

where $\delta(r)$ is the delta function. Similarly, the potential W of a dipole density μ on $\partial\Omega$ is defined by

$$W(x) = (1/\pi)\int_{\partial\Omega}\mu(\xi)[(\partial/\partial n_\xi)\log(1/r)]\,ds(\xi).$$

We adopt the convention that the normal direction of $\partial\Omega$ is towards the interior of the region Ω in which we want to solve our problem. We note that the normal derivative in the formula for W is taken with respect to the Greek variables. Denote by V^- the limit of V when x approaches $\partial\Omega$ from the interior and by V^+ the corresponding

exterior limit, etc. By using a Green's identity, one can show that V and $\partial W/\partial n$ are continuous across the boundary $\partial\Omega$, while $\partial V/\partial n$ and W satisfy jump conditions. Thus,

$$V^+ = V^-,$$
$$\partial V^{(\mp)}/\partial n = (\mp)\rho + (1/\pi)\int_{\partial\Omega}\rho\,[(\partial/\partial n_x)\log(1/r)]\,ds,$$
$$W^{(\mp)} = (\pm)\mu + (1/\pi)\int_{\partial\Omega}\mu[(\partial/\partial n_\xi)\log(1/r)]\,ds,$$

and

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

The Neumann and Dirichlet problems can be reduced to Fredholm integral equations. For the interior Neumann problem, we make the Ansatz

$$u(x) = (1/2\pi)\iint_\Omega f(\xi)\log(1/r)\,d\xi + (1/\pi)\int_{\partial\Omega}\rho(\xi)\log(1/r)\,ds(\xi)$$
$$= u_S(x) + V(x),$$

for the solution of

$$-\Delta u = f, \qquad x \in \Omega,$$
$$\partial u/\partial n = g, \qquad x \in \partial\Omega.$$

The first term, u_S, is called a space potential term. The boundary condition is satisfied by choosing ρ such that

$$\partial V^-/\partial n = -\rho + (1/\pi)\int_{\partial\Omega}\rho\,[(\partial/\partial n_x)\log(1/r)]\,ds$$
$$= g - (\partial/\partial n)u_S\Big|_{\partial\Omega}. \tag{2.1}$$

This equation can be written as $(I - K)\rho = \tilde{g}$, where K is a compact operator defined by the integral above. The equation is a Fredholm integral equation of the second kind and thus a well-posed problem. It has a zero eigenvalue and is solvable if $\tilde{g}$ has a zero mean value. This condition follows from the standard Fredholm theory. It can also be understood as the condition that in steady state heat flow, the energy liberated by heat sources in Ω must be balanced by the heat flow across the boundary.

If we now attempt to use the same single layer Ansatz for the Dirichlet problem, we obtain a Fredholm integral equation of the first kind. It has the form

$$(1/\pi)\int_{\partial\Omega}\rho\log(1/r)\,ds = g - u_S\Big|_{\partial\Omega}$$

and is an ill-posed problem. The kernel is now symmetric, a fact which might initially appeal to a numerical analyst.

The Ansatz

$$u(x) = (1/2\pi)\iint_\Omega f\log(1/r)\,d\xi + (1/\pi)\int_{\partial\Omega}\mu[(\partial/\partial n_\xi)\log(1/r)]\,ds$$
$$= u_S(x) + W(x),$$

provides a correct approach to the interior Dirichlet problem. If g now denotes the Dirichlet data, we obtain

$$W^- = \mu + (1/\pi)\int_{\partial\Omega}\mu[(\partial/\partial n_\xi)\log(1/r)]\,ds = g - u_S\Big|_{\partial\Omega} = \tilde{g}. \tag{2.2}$$

This is a well-posed problem of the form $(I + K^T)\mu = \tilde{g}$, where K^T is the adjoint of the compact operator K introduced above.

We remark that the exterior Neumann and Dirichlet problems similarly give rise to the equations

$$(I + K)\rho = \tilde{g} \quad \text{and} \quad (I - K^T)\mu = \tilde{g},$$

respectively. This is important in developing the existence and uniqueness theory for the elliptic problems by using the Fredholm theory of compact operators.

We conclude this section by remarking that the fundamental solution $(1/2\pi)\log(1/r)$ can be replaced by any other fundamental solution of the Laplace operator. In particular, we could have used the Green's function for a problem which is periodic in one or two directions. The periods must then be chosen such that the region Ω is a subset of the interior of a fundamental region of periodicity; i.e., we must use a wide enough infinite parallel strip or large enough rectangle. This observation is quite important when we implement our ideas computationally. This theory also generalizes, in a straightforward way, to Helmholtz's equation with a constant $c \neq 0$.

3. The Imbedding of Discrete Poisson Problems. In the next three sections we will develop a similar, formal potential theory for discrete Poisson problems. For the purpose of exposition only, we limit our discussion to a case with the same uniform mesh size in the two coordinate directions and to Laplace's equation, i.e. $c = 0$. The treatment of that case is actually more complicated than that of $c > 0$. Our method can also be used for negative values of c. The theoretical problems arising from the singularity of the operators for certain negative values of c are quite similar to those of the case under consideration.

The Laplace operator is replaced by a finite-difference approximation such as the five-point formula. The fundamental solution $(1/2\pi)\log(1/r)$, used in Section 2, will then be replaced by the discrete Green's function for the surface of an infinite cylinder, the surface of a torus or the whole plane. Very efficient and accurate Poisson and Helmholtz solvers exist for such problems; see Section 6. We will denote by B the matrix representing the discrete Laplacian $-h^2\Delta_h$, employing undivided differences, for the extended region and boundary conditions of our choice.

We decompose the set of mesh points into three disjoint sets Ω_h, $\partial\Omega_h$ and $(C\Omega)_h$. The set Ω_h is the set of interior mesh points; i.e., each of its members has all its relevant neighbors in the open set Ω. The remaining mesh points in Ω constitute $\partial\Omega_h$, the set of irregular mesh points, while the set $(C\Omega)_h$ contains all the remaining, the exterior, mesh points. We call the exterior and the interior points regular mesh points. Our numerical method produces values of a mesh function even for the points in $(C\Omega)_h$. These values are largely arbitrary, a useless by-product from the fast Poisson solver. It is likewise necessary to provide some largely arbitrary extension of the data to the set $(C\Omega)_h$. We note that the formulas for the continuous problem, given in the previous section, can be interpreted as having resulted from the extension by zero of the right-hand side f.

For all regular mesh points, i.e. those in $\Omega_h \cup (C\Omega)_h$, we use the basic discretization formula of the Laplace operator which we have already selected. That formula cannot be used for the irregular mesh points, i.e. those in $\partial\Omega_h$; because we must introduce an approximation of the boundary conditions; and we must also assure that our solution on $\Omega_h \cup \partial\Omega_h$ is independent of the extended values of the data and the solution on $(C\Omega)_h$. This will be achieved by eliminating from the discrete Laplacian, centered at an irregular mesh point, the values of the solution at its exterior neighbors. Thus, in the Dirichlet case, we will combine the discrete Laplacian with interpolation formulas which approximate the boundary condition, while in the Neumann case we will use difference approximations to the normal derivative for the same purpose. Examples with full details will be given below. The resulting linear equations corresponding to the irregular mesh points should be multiplied by appropriate scale factors. Because the choice of these factors is quite important for our theory and the performance of the conjugate gradient method, it will be discussed at some length below.

We will denote by A the matrix corresponding to our difference equations on the entire mesh. The equations and unknowns are ordered in the same way as the regularly structured problem given by the matrix B. The following representations of the matrix A are direct consequences of the fact that the rows of A and B are the same for all regular mesh points. For the Neumann problem we write $A = B - UV^T$, while we use the notation $A = B + UZ^T$ for the Dirichlet case. The matrices U, V and Z have p columns, where p is the number of points in the set $\partial\Omega_h$. The matrix U represents an extension operator. It maps any mesh function defined only on $\partial\Omega_h$ into a function on all mesh points. It retains the values of the mesh function on $\partial\Omega_h$ and makes the remaining values equal to zero. Its transpose, U^T, is a trace operator; i.e., U^T maps any mesh function defined for all mesh points into its restriction to $\partial\Omega_h$. The rows of V^T, and $-Z^T$, are simply the difference between the corresponding rows of B and A. One can therefore regard V^T and $-Z^T$ as compact representations of $B - A$ from which the zero rows corresponding to the regular mesh points have been deleted. We note that V^T and Z^T depend on the scaling chosen for the irregular rows of A.

In our construction of the matrix A the couplings to exterior mesh points were eliminated for the irregular rows. If we choose a suitable permutation matrix P we see that A is a reducible matrix

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

The submatrix A_{11} is the matrix for the linear system of equations which we set out to solve. It is easy to see from the structure of the matrix A that the solution on $\Omega_h \cup \partial\Omega_h$ is independent of the solution and data on $(C\Omega)_h$. The second diagonal block A_{22} can be interpreted as a finite-difference approximation to a Dirichlet problem on $C\Omega$. It is therefore easy to verify that A_{22} is nonsingular if our basic difference approximation is of positive type. Similarly, any convergent difference approximation to the

interior Dirichlet problem will give rise to a nonsingular A_{11} and thus a nonsingular matrix A. For the Neumann case we assume that the row sums of A_{11}, and hence of V^T, vanish and that the matrix A_{11} has a simple zero eigenvalue. It is then easy to see that the matrix A also will have a simple zero eigenvalue.

4. A Capacitance Matrix Method for the Dirichlet Case. We will now describe our method for dealing with the system of equations

$$Au = (B + UZ^T)u = F.$$

Let $Au - F \equiv r$ define the residual vector r. We wish to find a vector u so that the components of r vanish for all points in $\Omega_h \cup \partial\Omega_h$. The restriction of any such u to $\Omega_h \cup \partial\Omega_h$ provides the solution of the original difference equations. Guided by the continuous analog,we try the Ansatz,

$$u = GF + 2G\hat{V}D\mu.$$

The vector GF corresponds to the space potential part of the solution and satisfies $BGF = F$. The operator G thus plays the same role as the integral operator defined by the fundamental solution of the continuous problem. Any constant belongs to the nullspace of B and the choice of G is therefore not unique. Our particular choice of G corresponds closely to the standard choice for the continuous problem. We remark in particular that GF can be defined for all vectors F. For details see Section 6.

The second term in our Ansatz is a discrete analog of the potential from a dipole density. The vector μ has p components and is determined by solving a system of linear equations which we will derive below. The square matrix D is diagonal and contains scale factors. The choice of D and the scaling of the irregular rows of A is directed towards making the discrete problem close to the potential theory of Section 2. The mesh function $\hat{V}D\mu$ should vanish for all $x \in \Omega_h$. It is the sum of p mesh functions each representing a discrete dipole. The simplest construction of such discrete dipoles, which is the one we have used exclusively in our numerical experiments, is given in Figure 1. We have thus chosen to work with undivided differences. With the choice of weights as in our figure $\hat{V}^T v$ is an approximation of an undivided normal difference of the form $h_\alpha(\partial v/\partial n) + o(h)$. Here $h_\alpha = h/\cos\alpha$ and α is the angle between the normal through the irregular mesh point and the closest coordinate axis.

We now use our Ansatz and compute the residual vector,

$$Au - F = (B + UZ^T)(GF + 2G\hat{V}D\mu) - F = (2\hat{V}D + 2UZ^TG\hat{V}D)\mu + UZ^TGF. \tag{4.1}$$

From the properties of U and $\hat{V}$, it follows that the correct difference equations are satisfied for all $x \in \Omega_h$ and any vector μ. To derive a linear system of equations for the vector μ we multiply Eq. (4.1) by the trace operator U^T. It is easy to verify that $U^TU = I_p$ and also, for our choice of $\hat{V}$, that $U^T\hat{V} = I_p$. Here I_p is the $p \times p$ identity matrix. We thus obtain

$$(2D + 2Z^TG\hat{V}D)\mu = -Z^TGF. \tag{4.2}$$

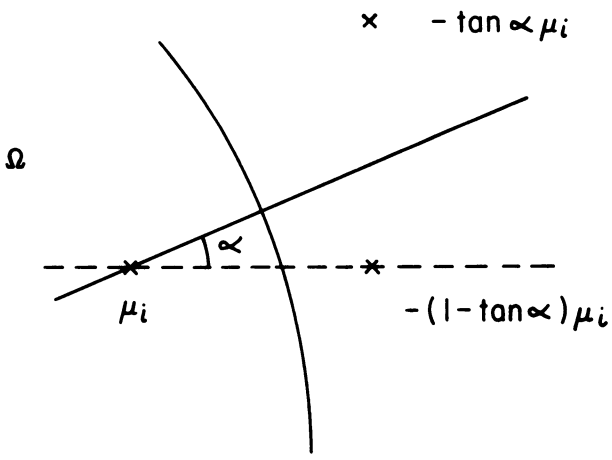

FIGURE 1

A discrete dipole with the scaling matrix $D = I$. We always assume that the boundary of the region and the mesh size is such that the two points with negative coefficients fall in the complement of Ω. The two exterior mesh points lie on opposite sides of the normal. One of these points is at a distance h and the other at a distance $\sqrt{2}h$ from the irregular mesh point.

The solution μ of Eq. (4.2) will make the components of the residual vector vanish for $x \in \partial\Omega_h$. The residual vector will in general not be equal to zero for all $x \in (C\Omega)_h$. As we saw in Section 3, this is of no importance because the restriction of the vector u to $\Omega_h \cup \partial\Omega_h$ will still provide a solution of the original problem. The matrix on the left-hand side of Eq. (4.2) is the capacitance matrix C.

By analogy to the continuous case we expect the capacitance matrix C to be nonsingular and in our numerical experiments we have consistently found C to be very well conditioned. When we attempt to prove that C is nonsingular we find that the existence of a nontrivial p-vector ϕ such that $C\phi = 0$ implies that the restriction of $AG\hat{V}D\phi$ to $\Omega_h \cup \partial\Omega_h$ must vanish identically. Hence, since A_{11} is nonsingular, the mesh function $G\hat{V}D\phi$ must be zero for all $x \in \Omega_h \cup \partial\Omega_h$. Conversely, if there exists a nontrivial ϕ such that $G\hat{V}D\phi$ is identically zero on $\Omega_h \cup \partial\Omega_h$, it follows from the block structure of PAP^T that $C\phi = U^T AG\hat{V}D\phi = 0$. We have not been able to exclude the existence of such a vector ϕ. In our opinion tools from mathematical analysis are required to resolve this problem completely. A proof that C is nonsingular when B and A are nonsingular, e.g. when $c > 0$, is given in Section 8.

We would like to be able to interpret Eq. (4.2) as a discretization of the appropriate Fredholm integral equation of the second kind. A simple discretization of such an equation would be based on a numerical quadrature rule of the form

$$\mu_j + 2\sum_i ((\partial/\partial n_\xi)\Gamma)_{ji}\mu_i h_i = \tilde{g}_j, \tag{4.3}$$

where $\Gamma(x, \xi)$ is the fundamental solution of the continuous problem, h_i is a local mesh length and the values of μ_i are approximate values of the dipole density μ. In the case

of zero Dirichlet data $\tilde{g} = -u_S$. By comparing right-hand sides we see that it is natural to choose a scaling of the irregular rows of A such that the row sums of Z^T equal one. Such a scaling is always possible for the Dirichlet case; see further the discussion at the end of this section.

We next turn our attention to the left-hand sides of Eqs. (4.2) and (4.3). We must then determine the local mesh length h_i. The distance between consecutive irregular mesh points will of course vary in a highly irregular way and it therefore seems appropriate to use a value of h_i which represents a local average. Let us consider a right triangle with a slightly curved hypotenuse constructed from a segment of the boundary of length $\sqrt{h}$. The average distance between the boundary mesh points close to this segment is then $h/\cos\alpha + o(h)$, where $\alpha \in [0, \pi/4]$ is the smallest angle between the normal of $\partial\Omega$ and a coordinate axis. We will adopt $h/\cos\alpha$ as our choice of h_i.

Away from the diagonal the elements of $2Z^T G\hat{V}$ behave like $2h_i(\partial/\partial n_\xi)\Gamma + o(h)$. This follows from our choice of h_i, Z^T and $\hat{V}$ and the fact that divided differences of the discrete fundamental solution G converge to the corresponding derivatives of the continuous fundamental solution $\Gamma(x, \xi)$ for $x \neq \xi$, see Thomée [44]. Therefore, we can simply adopt the choice $D = I_p$. The elements of the capacitance matrix C close to the diagonal present particular problems. While all off-diagonal elements of the matrix resulting from a numerical quadrature as in formula (4.3) are $O(h)$, this is not true in general for the elements of the capacitance matrix close to the diagonal. This is a consequence of the irregular positions of the points of $\partial\Omega_h$. Arthur Shieh [42] has therefore suggested that we should consider all elements of the capacitance matrix C corresponding to the influence between pairs of points within a distance of $\sqrt{h}$ as an entity. The resulting band matrix E_h should then ideally be a consistent approximation to the identity operator.

We will not pursue this difficult subject further at this time. We refer to Shieh [42] for a detailed theoretical analysis. However, we would like to point out that while our description of the algorithms so far closely reflects our actual implementation of the method we see several alternatives. Thus the discrete dipole layer could be constructed differently by allocating a fraction of the weights μ_i to nearby discrete dipoles located entirely outside $\Omega_h \cup \partial\Omega_h$. This would change the leading term of the capacitance matrix while the elements far away from the diagonal would be virtually unchanged. Variations of the elements of D are also possible, while retaining a local average of one, without spoiling our comparison between Eqs. (4.2) and (2.2).

In our experiments we have always used the five-point formula for the regular mesh points and combined it with the well-known Shortley-Weller approximation for the irregular points, see Collatz [9, Chapter 5.1], or Forsythe-Wasow [16, Section 20.9]. Thus, if $x \in \partial\Omega_h$ has its eastern and southern neighbors in $(C\Omega)_h$, we obtain

$$
\begin{aligned}
(4.4) \qquad & -(2/(1+\delta_1))u_w + (2/\delta_1 + 2/\delta_2 + ch^2)u_c - (2/(1+\delta_2))u_n \\
& \qquad = h^2 f_c + (2/\delta_1(1+\delta_1))u_e + (2/\delta_2(1+\delta_2))u_s.
\end{aligned}
$$

The scheme is of positive type. The parameter $\delta_1 = h_1/h$ where h_1 is the distance,

along a mesh line parallel with the x_1-axis, of the mesh point x and the boundary $\partial\Omega$. Hence, $\delta_1 \in (0, 1]$. The Dirichlet data at this point on $\partial\Omega$ is denoted by u_e. The values of δ_2 and u_s are defined similarly. If only the southern neighbor of x belongs to $(C\Omega)_h$, δ_1 is set equal to one and the u_e-term is moved to the left-hand side of the equation. We note that we can derive this formula by using quadratic polynomial extrapolation to find auxiliary values of u at the exterior mesh points which are next neighbors of x. These values are then eliminated by using the five-point formula. We note that we can regard any other boundary approximation from the same point of view. The sum of the extrapolation coefficients must always be nonzero and we can therefore safely assume that we will be able to scale each irregular row of A, making each row sum equal to one.

Now let $c = 0$. If we apply the recipe of scaling suggested above, we find that Eq. (4.4) should be multiplied by

$$(2/\delta_1(1 + \delta_1) + 2/\delta_2(1 + \delta_2))^{-1}.$$

If only the southern neighbor is exterior, the scale factor is $\delta_2(1 + \delta_2)/2$. It can easily be shown that the diagonal elements of the resulting matrix A take values in the interval $[1, 4]$. In many of our experiments we have used a scaling which makes all the diagonal elements of A equal to 4. This choice is close to the recipe suggested by our formal theory, and it results in a performance which is often slightly better. The use of formula (4.4) in its unscaled form results in a much poorer performance; see further Section 9.

5. A Capacitance Matrix Method for the Neumann Case. In the Neumann case our system of equations is written as

$$Au = (B - UV^T)u = F.$$

It is solvable if and only if the right-hand side F is orthogonal to the left eigenvector of A which corresponds to the zero eigenvalue. By assumption this eigenvalue is simple. By using the reducible structure of A, described in Section 3, it is easy to see that the mesh function corresponding to this eigenvector must vanish for all $x \in (C\Omega)_h$. The right-hand side F is therefore consistent regardless of its values on $(C\Omega)_h$ if the data is already consistent on $\Omega_h \cup \partial\Omega_h$.

Guided by the continuous analog, we try the Ansatz $u = GF + 2GUD\rho$. The p-vector ρ corresponds to the boundary charge distribution of the continuous case. The extension operator U was introduced in Section 3 while the diagonal matrix D contains certain scale factors.

Computing the residual vector, we obtain

$$(5.1)\quad Au - F = (B - UV^T)(GF + 2GUD\rho) - F = (2UD - 2UV^TGUD)\rho - UV^TGF.$$

Because of the factor U, the residuals are zero for all $x \in \Omega_h \cup (C\Omega)_h$. The vector ρ is determined by requiring the residuals to vanish on $\partial\Omega_h$. We therefore multiply Eq. (5.1) by U^T and obtain

$$(2D - 2V^T GUD)\rho = V^T GF. \tag{5.2}$$

We note that our formula is virtually identical to the Woodbury formula; see Householder [28].

The matrix of Eq. (5.2) is the capacitance matrix C of the Neumann problem. It is singular because if it were not our procedure would provide a solution for any data. It is also easy to verify that the zero eigenvalue of C is simple. Let us assume that ϕ_1 and ϕ_2 were two linearly independent eigenvectors of the capacitance matrix C corresponding to the zero eigenvalue. Then

$$V^T GUD\phi_i = D\phi_i, \qquad i = 1, 2,$$

and $GUD\phi_i$, $i = 1, 2$, must be nonzero vectors. These vectors are, by a similar argument, linearly independent. A direct calculation shows that $GUD\phi_i$, $i = 1, 2$, are eigenvectors of A with a zero eigenvalue. We have thus arrived at a contradiction because, by assumption, the zero eigenvalue of A is simple.

The right-hand side $V^T GF$ of Eq. (5.2) is consistent if F is consistent for the original problem $Au = F$. Let ϕ^T be the left eigenvector of C corresponding to the zero eigenvalue. It can readily be shown that $\psi^T = \phi^T V^T G$ is the corresponding eigenvector of A. But $\psi^T F = 0$ implies $\phi^T(V^T GF) = 0$.

We will show in Section 7 that the singularity of C will cause little difficulty when we solve for ρ by the conjugate gradient method. Any solution ρ of Eq. (5.2) will result in a zero residual for all mesh points and thus a solution of the original problem.

We now attempt to choose the scaling of the irregular rows of A and the diagonal matrix D so that Eq. (5.2) appears, as closely as is possible, as a discretization of the integral equation

$$\rho - 2\int[(\partial/\partial n_x)\Gamma]\rho\, ds = \tilde{g}. \tag{5.3}$$

Our arguments here are very similar to those in Section 4. The right-hand side $\tilde{g}$ is by Eq. (2.1) equal to the normal derivative of the space potential if the Neumann data is equal to zero. It follows from our discussion in Section 3 that the rows of V^T are linear combinations of undivided difference approximations of the normal derivative and the undivided discrete Laplacian. The mesh function $V^T v$ is therefore essentially of the form $\beta h\, \partial v/\partial n + o(h)$ where the factor $\beta(x)$ is nonzero. We therefore scale the irregular rows of A to make $\beta = 1$. This choice makes the right-hand side of (5.2) a consistent approximation to that of Eq. (5.3) except for an unimportant factor h.

Adopting this choice of scaling of the rows of A we next examine the elements of the capacitance matrix which are far from the diagonal. They are seen to behave like $2h(\partial/\partial n_x)\Gamma + o(h)$ which suggests that an appropriate choice of the elements of the matrix D is $1/\cos\alpha$. Here, $\alpha \in [0, \pi/4]$ is the angle between the normal direction and the closest coordinate axis. In our experiments we have simply used a scaling $D = I_p$.

As in the Dirichlet case we could modify our Ansatz in an attempt to make our discrete model more flexible. We could thus spread certain fractions of the charges ρ_i

to nearby exterior mesh points. This approach mainly affects the elements of the capacitance matrix which are on and close to the diagonal. We have not tried any such procedures numerically. An alternative Ansatz of this form will lead to difficulties similar to those in the Dirichlet case. We have thus been unable to show that the right-hand side of the equation replacing (5.2) is consistent except in the version of the method which we have used consistently in our experiments.

We will now describe the scheme which we used in our computations. Our point of departure is the scheme (4.4) for the Dirichlet problem. The values u_e and u_s must be eliminated by using some approximation to the normal derivative. Denote by $\gamma_1 \in [0, \pi/2)$ the angle between the normal direction at the eastern point and the x_1-axis. We have used the formula

$$(5.4) \qquad (1 - \delta_1 \tan\gamma_1)u_c + \delta_1 \tan\gamma_1 u_n - u_e = (\delta_1 h/\cos\gamma_1)(\partial u/\partial n)_e.$$

Similarly, denoting by γ_2 the angle between the x_2-axis and the normal direction at the southern point, we obtain

$$(5.5) \qquad (1 - \delta_2 \tan\gamma_2)u_c + \delta_2 \tan\gamma_2 u_w - u_s = (\delta_2 h/\cos\gamma_2)(\partial u/\partial n)_s.$$

By combining formulas (4.4), (5.4) and (5.5), we obtain

$$\begin{aligned} &-2(1/(1+\delta_1) + \tan\gamma_2/(1+\delta_2))u_w \\ &\quad + (2(1+\tan\gamma_1)/(1+\delta_1) + 2(1+\tan\gamma_2)/(1+\delta_2) + ch^2)u_c \\ &\quad - 2(1/(1+\delta_2) + \tan\gamma_1/(1+\delta_1))u_n \\ &= h^2 f_c - (2h/((1+\delta_1)\cos\gamma_1))(\partial u/\partial n)_e \\ &\quad -(2h/((1+\delta_2)\cos\gamma_2))(\partial u/\partial n)_s. \end{aligned}$$

When we modify the formula in the obvious way for the case of only one, a southern, exterior neighbor, we obtain

$$\begin{aligned} &-(1 + 2\tan\gamma_2/(1+\delta_2))u_w \\ &\quad + (2 + 2(1+\tan\gamma_2)/(1+\delta_2) + ch^2)u_c - (2/(1+\delta_2))u_n - u_e \\ &= h^2 f_c - 2h/(\cos\gamma_2(1+\delta_2))(\partial u/\partial n)_s. \end{aligned}$$

We note that this scheme is of positive type.

According to our recipe above, the scale factors for the irregular rows of A are

$$d = (2/((1+\delta_1)\cos\gamma_1) + 2/((1+\delta_2)\cos\gamma_2))^{-1}$$

and

$$d = \cos\gamma_2(1+\delta_2)/2,$$

respectively. One can easily show that this results in diagonal elements which vary between one and three. Here, just as in the Dirichlet case we have frequently used a scaling which makes the diagonal elements equal to four, in which case we are still quite close to the scaling suggested by our discussion above.

6. The Fourier-Toeplitz Method and the Fast Generation of the Capacitance Matrix. We will now describe the variant of the Fourier-Toeplitz method which we have used in our numerical experiments to solve the Helmholtz equation on the entire mesh. We will also show how the capacitance matrix is generated in our program at a modest expense. We have previously indicated that it is important to achieve translation invariance. We should therefore choose a region without a boundary. In the original paper on the Fourier-Toeplitz method (see Fischer, Golub, Hald, Leiva and Widlund [15]), the Dirichlet and the doubly periodic problems on rectangles and the Dirichlet problem on an infinite parallel strip were discussed in detail. Of these, the doubly periodic case, which amounts to solving the problem on the surface of a torus, would have been an appropriate choice. However, we have instead chosen an infinite parallel strip with a periodicity condition, i.e. a problem on the surface of an infinite cylinder. This variant of the Fourier-Toeplitz method is somewhat faster than the doubly periodic case because an intermediary step employing the Woodbury formula is eliminated. More importantly, our choice allows for a more satisfactory resolution of a special problem associated with cases where the operator B is singular. A third alternative would be a method of Hockney's [25, p. 178], which allows for the calculation of the solution of Helmholtz's equation on the entire plane. However, that method requires four times as much storage and also more arithmetic operations and is not therefore appropriate in our present context. Hockney's idea could however prove quite useful should we attempt to compute the solution of exterior problems without a prior transformation of the region into a bounded set.

The solution of the five-point discrete Helmholtz problem on an infinite strip parallel to the x_1-axis gives rise to a block-tridiagonal linear system of equations of infinite order,

$$\begin{pmatrix} \ddots & \ddots & \ddots & & & & \\ \ddots & \ddots & \ddots & \ddots & & 0 & \\ & -I & A_c & -I & & & \\ & & -I & A_c & -I & & \\ & & & -I & A_c & -I & \\ 0 & & & & \ddots & \ddots & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ \vdots \\ u^{(-1)} \\ u^{(0)} \\ u^{(1)} \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \vdots \\ f^{(-1)} \\ f^{(0)} \\ f^{(1)} \\ \vdots \\ \vdots \end{pmatrix}.$$

The vectors $u^{(j)}$ and $f^{(j)}$ represent the values of the approximate solution and the data at the mesh points on $x_1 = jh$. We will always assume that the vectors $f^{(j)}$ vanish except for a finite number of indices. The matrix A_c is an $n \times n$ circulant,

$$A_c = \begin{pmatrix} 4+ch^2 & -1 & 0 & & 0 & -1 \\ -1 & 4+ch^2 & -1 & & 0 & 0 \\ & & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ 0 & 0 & & -1 & 4+ch^2 & -1 \\ -1 & 0 & & 0 & -1 & 4+ch^2 \end{pmatrix}.$$

For convenience, we assume that n is even. The matrix A_c then has the normalized eigenvectors $(1/n)^{1/2}(1, 1, \ldots, 1)^T$ and $(1/n)^{1/2}(1, -1, \ldots, -1)^T$ corresponding to the simple eigenvalues $2 + ch^2$ and $6 + ch^2$, respectively, and $(n-2)/2$ double eigenvalues $4 + ch^2 - 2\cos(2\pi l/n)$, $l = 1, 2, \ldots, (n-2)/2$, with the eigenvectors $\phi_{\mathrm{I}}^{(l)}$ and $\phi_{\mathrm{II}}^{(l)}$. These are given by

$$\phi_{\mathrm{I},k}^{(l)} = (2/n)^{1/2} \sin(kl2\pi/n),$$
$$\text{and} \qquad k = 0, 1, \ldots, n-1, \quad l = 1, 2, \ldots, (n-2)/2.$$
$$\phi_{\mathrm{II},k}^{(l)} = (2/n)^{1/2} \cos(kl2\pi/n),$$

A change of basis corresponding to the diagonalization of A_c can be carried out inexpensively by using the Fast Fourier Transforms if n has many prime factors; see Cooley, Lewis and Welch [11].

This change of basis, cf. Fischer, Golub, Hald, Leiva and Widlund [15], results in a block-tridiagonal system with the matrix

$$\begin{pmatrix} \cdot & \cdot & \cdot & & & & & \\ & \cdot & \cdot & \cdot & & & & 0 \\ & & -I & D_c & -I & & & \\ & & & -I & D_c & -I & & \\ & & & & -I & D_c & -I & \\ & & & & & \cdot & \cdot & \cdot \\ & 0 & & & & & \cdot & \cdot & \cdot \end{pmatrix}.$$

The diagonal matrix D_c contains the eigenvalues of A_c. A permutation of rows and columns of this matrix which preserves symmetry and groups the lth equation of each block together into one block results in a matrix which is the direct sum of n-tridiagonal Toeplitz matrices Λ_l, $l = 0, 1, \ldots, n-1$, of infinite order. Here

$$\Lambda_l = \begin{pmatrix} \cdot & \cdot & \cdot & & & & & \\ & \cdot & \cdot & \cdot & & & & 0 \\ & & -1 & \lambda_l & -1 & & & \\ & & & -1 & \lambda_l & -1 & & \\ & & & & -1 & \lambda_l & -1 & \\ & & & & & \cdot & \cdot & \cdot \\ & 0 & & & & & \cdot & \cdot & \cdot \end{pmatrix},$$

where λ_l is an eigenvalue of A_c. We will first restrict our attention to the case $c > 0$. All these matrices are then positive definite. From our assumptions concerning the data f, we see that the corresponding right-hand side, derived by using the Fast Fourier Transform on the vectors $f^{(j)}$, vanishes for $j < N_-$ and $j > N_+$ with N_- and N_+ appropriately chosen; we can therefore write

$$-\hat{u}_{j-1}^{(l)} + \lambda_l \hat{u}_j^{(l)} - \hat{u}_{j+1}^{(l)} = 0 \quad \text{for } j < N_- \text{ and } j > N_+. \tag{6.1}$$

This homogeneous difference equation has the solution const μ_l^j + const μ_l^{-j} where $\mu_l = \lambda_l/2 + (\lambda_l^2/4 - 1)^{1/2} > 1$. Imposing the natural free space boundary condition, i.e.

requiring that the solution remains bounded for all j, we find that

$$\hat{u}_j^{(l)} = \hat{u}_{N_+}^{(l)} \mu_l^{(N_+ - j)}, \qquad j \geqslant N_+,$$

$$\hat{u}_j^{(l)} = \hat{u}_{N_-}^{(l)} \mu_l^{(j - N_-)}, \qquad j \leqslant N_-. \tag{6.2}$$

By using Eqs. (6.1) and (6.2) for $j = N_+ + 1$ and $j = N_- - 1$ we obtain

$$\mu_l \hat{u}_{N_-}^{(l)} - \hat{u}_{N_-+1}^{(l)} = \hat{f}_{N_-}^{(l)} \quad \text{and} \quad -\hat{u}_{N_+-1}^{(l)} + \mu_l \hat{u}_{N_+}^{(l)} = \hat{f}_{N_+}^{(l)}.$$

We can therefore set up a finite linear system of equations,

$$\begin{pmatrix} \mu_l & -1 & 0 & & & & 0 \\ -1 & \lambda_l & -1 & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & -1 & \lambda_l & -1 \\ 0 & & & & & -1 & \mu_l \end{pmatrix} \begin{pmatrix} \hat{u}_{N_-}^{(l)} \\ \hat{u}_{N_-+1}^{(l)} \\ \cdot \\ \cdot \\ \cdot \\ \hat{u}_{N_+-1}^{(l)} \\ \hat{u}_{N_+}^{(l)} \end{pmatrix} = \begin{pmatrix} \hat{f}_{N_-}^{(l)} \\ \hat{f}_{N_-+1}^{(l)} \\ \cdot \\ \cdot \\ \cdot \\ \hat{f}_{N_+-1}^{(l)} \\ \hat{f}_{N_+}^{(l)} \end{pmatrix}.$$

This system can be solved very easily by an LU-decomposition because, as is easily verified,

$$\begin{pmatrix} \mu_l & -1 & & & & & 0 \\ -1 & \lambda_l & -1 & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & & -1 & \lambda_l & -1 \\ 0 & & & & & -1 & \mu_l \end{pmatrix}$$

$$= \begin{pmatrix} 1 & & & & & & 0 \\ -\mu_l^{-1} & 1 & & & & & \\ & -\mu_l^{-1} & 1 & & & & \\ & & \cdot & \cdot & & & \\ & & & \cdot & \cdot & & \\ & & & & -\mu_l^{-1} & 1 & \\ 0 & & & & & -\mu_l^{-1} & 1 \end{pmatrix} \begin{pmatrix} \mu_l & -1 & & & & & 0 \\ & \mu_l & -1 & & & & \\ & & \mu_l & -1 & & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \\ 0 & & & & & \mu_l & -1 \\ & & & & & & \sigma_l \end{pmatrix}.$$

Here $\sigma_l = \mu_l - \mu_l^{-1}$. We note that except for one element in the second factor these bidiagonal matrices are Toeplitz matrices. The solution of the five-point discrete Helmholtz equation is thus obtained by Fast Fourier Transforms in the x_2-direction, by the solution of n very special tridiagonal systems of equations and by an inverse Fast Fourier Transform for each value x_1 for which we need the solution. We note that our method imposes no restriction on the number of mesh points used along the strip.

We now turn to the case of $c = 0$. The tridiagonal matrix which corresponds to the smallest eigenvalue of A_c will then have the form

$$\Lambda_0 = \begin{pmatrix} \cdot & \cdot & \cdot & & & & 0 \\ & \cdot & \cdot & \cdot & & & \\ & & -1 & 2 & -1 & & \\ & & & -1 & 2 & -1 & \\ & & & & \cdot & \cdot & \cdot \\ 0 & & & & & \cdot & \cdot & \cdot \end{pmatrix},$$

while all the others remain positive definite. Guided by the continuous problem, we introduce a fundamental solution. We choose the function $-|j-k|/2$. We note that we could add any linear function and still obtain a fundamental solution. With our choice

$$\hat{u}_j^{(0)} = -\sum(|j-k|/2)\hat{f}_k^{(0)}. \tag{6.3}$$

We note further that no assumption of the form $\Sigma_k \hat{f}_k^{(0)} = 0$ is needed. This is particularly important when we generate a representation of the operator G by using the data $f_k^{(j)} = \delta_{jl}\delta_{km}$. The formula (6.3) can be used to evaluate two consecutive values of $\hat{u}_j^{(0)}$, and we can then find the remaining values by employing the simple stable recursion

$$\hat{u}_{j+2}^{(0)} = 2\hat{u}_{j+1}^{(0)} - \hat{u}_j^{(0)} - \hat{f}_{j+1}^{(0)}.$$

There remains the case of $c < 0$. At least one of the tridiagonal matrices Λ_l will then have diagonal elements $\lambda_l < 2$. By imposing a mild restriction on ch^2, we assume that $|\lambda_l| < 2$. It is easy to verify that $-\sin(|j-k|\phi_l)/2 \sin\phi_l$, where $2\cos\phi_l = \lambda_l$, is a fundamental solution. With this choice

$$\hat{u}_j^{(l)} = -\sum_k (\sin(|j-k|\phi_l)/2\sin\phi_l)\hat{f}_k^{(l)}, \qquad \lambda_l = 2\cos\phi_l. \tag{6.4}$$

We can again use a stable recursion formula

$$\hat{u}_{j+1}^{(l)} = \lambda_l \hat{u}_j^{(l)} - \hat{u}_{j-1}^{(l)} - \hat{f}_j^{(l)},$$

to find the remaining values of $\hat{u}_j^{(l)}$, once two consecutive values of $\hat{u}_j^{(l)}$ have been computed. We note that the Chebyshev polynomials of the second kind are defined by

$$U_n(x) = \sin((n+1)\theta)/\sin\theta, \qquad x = \cos\theta,\ n \geqslant 0,$$

and that they satisfy the recursion formula

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \qquad n \geqslant 1,$$
$$U_0(x) = 1, \qquad U_1(x) = 2x.$$

We can therefore combine this formula with formula (6.4) to compute the required two values of $\hat{u}_j^{(l)}$.

We now turn our attention to the generation of the capacitance matrices. The matrices U, V, $\hat{V}$ and Z are very sparse with a fixed upper bound on the number of nonzero elements in any row and column. It therefore follows that the number of arithmetic operations required to generate a capacitance matrix grows only as const $p^2(1 + o(1))$

if complete information on the operator G is available. To obtain this information we solve the Helmholtz equation for $u^{(0,0)}$ with the data given by $f_k^{(j)} = \delta_{j0}\delta_{k0}$. In our program we take advantage of the very special form of the data which virtually eliminates the need for the first Fast Fourier Transform step. If we denote by $u^{(j,k)}$ the solution resulting from a single unit charge at the point (j, k), we see, by the translation invariance, that $u_{l,m}^{(j,k)} = u_{l-j,m-k}^{(0,0)}$. We also use the fact that the solution $u^{(0,0)}$ satisfies

$$u_{l,m}^{(0,0)} = u_{-l,m}^{(0,0)} \quad \text{and} \quad u_{l,m}^{(0,0)} = u_{l,n-m}^{(0,0)}.$$

Complete information on G can therefore easily be obtained from the mesh function $u^{(0,0)}$ which is generated by using a simplified fast Helmholtz solver only once.

7. The Conjugate Gradient Method. In this section we will briefly describe the conjugate gradient method and those of its properties which make it particularly attractive for solving the capacitance matrix equations. We have used this method extensively as an iterative method for the linear systems of Eqs. (4.2) and (5.2). For a detailed theory of the conjugate gradient method see Daniel [13], Hestenes [21], Hestenes and Stiefel [22], Kaniel [31], Lanczos [33] and Luenberger [34]. In our experiments we have used an implementation of the method, SYMMLQ, developed by Paige and Saunders [38].

We again denote the capacitance matrix by C. It is a $p \times p$, nonsymmetric and dense matrix. The conjugate gradient method requires a symmetric coefficient matrix and we therefore work with the corresponding normal equations $Qv = b$ where $Q = C^TC$. In each conjugate gradient step we must multiply the matrix Q by a vector u. We never compute the matrix Q in our program because that would require $p^2(p + 1)/2$ multiplications. Instead, we first compute $w = Cu$ and then $Qu = C^Tw$. Each step of the conjugate gradient method requires $2p^2(1 + o(1))$ multiplications.

We denote by v_* an exact solution of our system $Qv = b$ and by v_0 the initial guess. In the first step of the algorithm the residual $r_0 = b - Qv_0$ is computed. This is the steepest descent direction and it is chosen as the first search direction p_1. We then find the minimum v_1 of the quadratic form $\frac{1}{2}v^TQv - v^Tb$ along the line $v = v_0 + sp_1$. The search direction p_k, $k \geqslant 1$, is chosen as that linear combination of the residual $r_{k-1} = b - Qv_{k-1}$ and the previous search direction p_{k-1}, which makes the vectors p_i Q-conjugate, i.e. $p_k^TQp_i = 0$ for $i < k$. The kth approximation v_k is the minimum of the quadratic form along the line $v = v_{k-1} + sp_k$. No a priori information about the spectrum of Q is needed. This is an advantage compared with the Chebyshev and certain other iterative methods. The iteration is terminated when the l_2-norm of the residual r_k becomes less than a certain tolerance. For a detailed description of the termination criterion of the conjugate gradient method which we have used in our experiments, see Paige and Saunders [38].

It is easy to verify that the iterates satisfy

$$v_k = v_0 + P_{k-1}(Q)r_0, \tag{7.1}$$

where P_{k-1} is a polynomial of degree $k-1$. It is known (see for example Luenberger [34]) that of all iterative methods which satisfy the relation (7.1) the conjugate gradient method is optimal in the sense that

$$E(v) = \tfrac{1}{2}(v - v_*)^T Q(v - v_*)$$

is minimal. Denote by λ_i, ϕ_i, $i = 1, \ldots, p$, the eigenvalues and eigenvectors of Q and let $v_0 - v_* = \Sigma \alpha_i \phi_i$ be an eigenvector expansion of the initial error. It then follows easily from the optimality of the conjugate gradient method that

$$E(v_k) \leqslant \max_{\lambda_i} (1 - \lambda_i P_{k-1}(\lambda_i))^2 E(v_0), \tag{7.2}$$

for any choice of a polynomial P_{k-1} of degree $k-1$. This inequality remains valid if the maximum is taken only over those λ_i for which $\alpha_i \neq 0$. The conjugate gradient method will therefore, in the absence of round-off, give an exact solution in at most q steps where $q \leqslant p$ is the number of distinct eigenvalues of Q for which a coefficient α_i differs from zero. It can be shown by using the inequality (7.2) and a special choice of the polynomial P_{k-1} that

$$\begin{aligned} E(v_k) &\leqslant (2(1 - 1/\kappa))^k / ((1 + 1/\sqrt{\kappa})^{2k} + (1 - 1/\sqrt{\kappa})^{2k}))^2 E(v_0) \\ &\leqslant 4((1 - 1/\sqrt{\kappa})/(1 + 1/\sqrt{\kappa}))^{2k} E(v_0); \end{aligned} \tag{7.3}$$

see for example Daniel [13]. Here κ is the condition number of Q.

We have seen in Section 5 that Eq. (5.2) has a singular matrix when $c = 0$. It is, therefore, of interest to consider the performance of the conjugate gradient algorithm in a singular case. It is easy to show that the right-hand side b of our normal equations $Qv = b$ is orthogonal to any eigenvector ϕ which corresponds to a zero eigenvalue. Starting from a zero initial guess all the iterates v_k will also be orthogonal to ϕ. In the absence of round-off the entire computation will proceed in a subspace orthogonal to ϕ. The estimate (7.3) is then valid with κ redefined as the ratio between the largest and smallest strictly positive eigenvalues of Q. Our experience with the conjugate gradient method for singular and almost singular capacitance matrices has been excellent; see Section 9.

As we pointed out in Section 4, we have not been able to prove that the capacitance matrix in Eq. (4.2) is nonsingular when $c = 0$. As a safeguard we calculate the residual for the original nonsymmetric linear system of equations (4.2) and print a warning when its norm is larger than expected. We have found no Dirichlet problems with the constant $c \geqslant 0$ for which there are indications that the capacitance matrix equation (4.2) fails to have a solution. This feature of our program is of course also useful when the constant c is negative. We note that the solution of the least squares problems with the conjugate gradient method has been considered from a theoretical point of view by Kammarer and Nashed [30].

In our experiments we have found that the estimate (7.3) is realistic for Dirichlet problems. For such problems the capacitance matrices are typically very well conditioned. We have also observed impressive rates of convergence for Neumann problems

with very small values of c. By a continuity argument such problems must have large condition numbers. To explain the success of the conjugate gradient method in these cases we must consider the entire spectral distribution. The results can be understood in terms of the estimate (7.2) once we realize that the small singular values of C, i.e. the positive square root of the eigenvalues of C^TC, are isolated and that the singular values cluster around a point well removed from the origin.

We recall that our choice of a double layer Ansatz for the Dirichlet problem was based on the conjecture that the singular values of the capacitance matrices would be distributed similarly to those of the integral operators of the classical potential theory. In order to illuminate this point we consider a case for which the singular values of the integral operators are known. Kantorovich and Krylov [32, p. 135], have shown that for ellipsoidal regions and Laplace's equation the dipole Ansatz for the Dirichlet problem leads to a Fredholm integral operator of the second kind with the singular values

$$1 + \gamma^j, \quad j = 0, 1, \ldots, \quad \text{and} \quad 1 - \gamma^j, \quad j = 1, 2, \ldots .$$

Here $\gamma = (a - b)/(a + b)$ where a and b are the half axes of the ellipse. The Neumann problem, with the single layer Ansatz leads to the singular values

$$1 - \gamma^j, \quad j = 0, 1, \ldots, \quad \text{and} \quad 1 + \gamma^j, \quad j = 1, 2, \ldots .$$

We note that in both cases there is a very pronounced cluster of the spectrum at the point one. Very rapid convergence can therefore be anticipated from the estimate (7.2) with an appropriate choice of a polynomial.

Hayes [20] has shown that the conjugate gradient method converges superlinearly for Fredholm integral equations of the second kind. We have verified this result by numerical experiments with diagonal coefficient matrices. In our applications to the capacitance matrix method, we normally fail to observe superlinear convergence. The overall structure of the spectrum with a pronounced cluster around the point one is however inherited from the continuous problem, see Section 9; and as we have previously noted, the rate of convergence is, therefore, often far better than the estimate (7.3).

The single layer Ansatz for the continuous Dirichlet problem on the interior of an ellipse gives the singular values

$$\gamma^j, \quad j = 0, 1, \ldots .$$

A similar cluster of singular values very close to the origin can be observed in experiments with the analogous discrete Ansatz. This explains the slow convergence and erratic behavior of the conjugate gradient method when the Woodbury formula is applied to the Dirichlet problem; see Section 9.

In our program we also have an option of using Gaussian elimination for Eqs. (4.2) and (5.2) if the capacitance matrix is nonsingular. The factorization step of Gaussian elimination requires $p^3/3(1 + o(1))$ multiplicative operations. The conjugate gradient method produces highly accurate solutions in a number of steps which remains virtually constant when p is increased. The conjugate gradient method therefore essentially requires only const p^2 operations. The constant depends on the region, the value of c and the boundary condition. In our experience, its variation is not very large.

Hence, the conjugate gradient method should be used for large enough p if only one or a few sets of data are used for the same problem. For experimental evidence see Section 9.

8. **Previous Work on Capacitance Matrix Methods.** In this section we will examine some earlier work on capacitance matrix methods and also derive estimates of the condition number of certain capacitance matrices. A method of this type is described very briefly in a paper by Hockney [24], who credits Oscar Buneman for the idea. A detailed description of this method, its implementation in Fortran and a listing of a program is given in Hockney [26]. The problems considered there differ in certain ways from those considered in this paper. The five-point discrete Laplacian is used on a uniform mesh in a rectangular region. A large variety of boundary conditions, all of separable type, are allowed on the sides of the rectangle. A number of electrodes are introduced in the interior and on the boundary of the rectangle. Each is represented by a straight line segment to which one or several mesh points are assigned. Prescribed average values of the potential on each such set of mesh points are obtained by using a capacitance matrix calculation. If the length of a line segment is chosen to be zero, an electrode will consist of a single mesh point. Although primarily designed for other applications, Hockney's program can be used for the solution of an interior Dirichlet problem for Laplace's equation. The capacitance matrix is then the restriction of the solution operator, for the problem of our choice on the rectangle, to the set of irregular mesh points. Hockney's capacitance matrices are always symmetric, positive definite, and storage and time is saved by using the Cholesky algorithm. Hockney's method thus corresponds to a single-layer Ansatz for the Dirichlet problem. In a typical application of his program the solution is required for the entire rectangle, a fact which seems to preclude the use of discrete dipoles; cf. Section 4. It would not seem possible to use his program without modifications to obtain a more than first order accurate solution of elliptic finite-difference equations on a general interior region. We doubt that the symmetry of the capacitance matrix can be maintained in higher order cases. See discussion below.

The capacitance matrices are generated at the expense of p calls of the subroutine which gives the solution on the rectangle. Here p is the number of electrodes. The cases of Neumann and periodic boundary conditions on the rectangle, which make the Laplace operator singular, are discussed; and the related difficulty is well resolved. The capacitance matrices are symmetric even in these cases. By carefully examining his fast Poisson solver, Hockney finds a short-cut for the treatment of so-called boundary electrodes, which are located on two parallel sides of the rectangle. This observation leads to the use of two capacitance matrices for the boundary electrodes and the remaining electrodes, respectively. Once the Cholesky factors of the capacitance matrix for the boundary electrodes are computed, the correction due to boundary electrodes is handled quite inexpensively, while the Poisson problem on the rectangle is solved only once.

The report contains detailed information on the accuracy and execution time of the program when run on an IBM 360/91.

The reports by Buzbee, Dorr, George and Golub [7] and George [18] appeared almost at the same time as Hockney's paper. Their results are formulated in matrix language. In both papers the imbedding idea is discussed generally. Focusing on the first of these papers, we find that the Ansatz $u = B^{-1}F + B^{-1}UW\rho$ is used when the matrix B is nonsingular. Here B is the operator of our choice for the rectangle, W a $p \times p$ nonsingular matrix and U the extension operator introduced in Section 3. A choice of $W = I$ gives the Woodbury formula. The case of a singular matrix B is treated in a way different from Hockney's and ours. A nonsingular capacitance matrix results in that case if and only if A is nonsingular. An operation count for this method has already been given in Section 1.

A symmetric capacitance matrix is obtained when this single-layer Ansatz is used for the Dirichlet problem if no irregular mesh point has a neighbor outside the closure of Ω and $W = I$. We see no practical possibility of choosing a W which will retain the symmetry of the capacitance matrix when a more accurate boundary approximation is used. Symmetry is also lost for the Neumann problem even for the simplest boundary.

We refer to Bramble and Hubbard [3] for a proof that first order accuracy is obtained if a general boundary of a Dirichlet problem is approximated by a set of mesh points. We also note that, close to the boundary, approximations of the gradient of the solution obtained from such a scheme cannot be expected to converge, cf. Bramble and Hubbard [4]. This is in contrast to schemes with a more accurate treatment of the boundary condition for which asymptotic error expansions exist. From these expansions the second order accuracy of difference approximations to the first and at times higher order derivatives can be obtained; cf. Pereyra, Proskurowski and Widlund [39]. It should also be noted that the use of isoparametric finite element methods (cf. Strang and Fix [43]), enjoy an increasing popularity for second order elliptic problems. In these methods the boundary is approximated quite accurately by piecewise polynomials of degree two or higher. The success of such methods undoubtedly reflects the importance of an accurate representation of the boundary in many applications.

One can of course also use a nonsymmetric capacitance matrix in the Buzbee, Dorr, George and Golub [7] method. However, we see no reason to prefer their single-layer Ansatz for the Dirichlet problem to our double-layer Ansatz. Only a marginally lower operation count will result from their choice except in a case where no irregular mesh point has a neighbor in the complement of the closure of Ω, or an only first order accurate approximation is accepted. In such cases the savings in operations result almost entirely from the use of the Cholesky method instead of Gaussian elimination for the factorization of the capacitance matrix.

Buzbee, Dorr, George and Golub [7] also show how capacitance matrix techniques can be used to handle problems on unions of rectangles as well as problems with piecewise constant coefficients.

A number of numerical results are reported in their paper for runs on a CDC 6600. Certain of these experiments were later rerun on a CDC 7600 using improved programs. A gain in speed of roughly a factor three is reported in a paper by Buzbee and Dorr [6]. Some of these results can be found in Table 8. The main emphasis of their paper is the

application of a capacitance matrix method to a finite-difference approximation of rectangular clamped plate problems. Their method is an improved version of an algorithm derived by Golub [19]. The variables can be separated for this plate problem once the boundary conditions on two opposite sides of the plate have been suitably modified. The order of the capacitance matrix, p, is therefore twice the number of mesh points across the plate. To ensure flexibility, the matrix decomposition method (cf. Buzbee, Golub and Nielson [8]), is used as a fast solver. This method allows an arbitrary number of mesh points in both directions, but it requires on the order of n^3 operations. An efficient implementation of this method still allows very fast execution. When generating the capacitance matrix, Buzbee and Dorr take advantage of the sparsity of the data vectors and the fact that the solutions are needed only at relatively few points. This computation therefore requires only on the order of pn^2 operations compared to pn^3 needed for a straightforward implementation of the capacitance matrix method using the matrix decomposition method for a general value of n. We see no possibility of reducing this operation count substantially by a particular choice of n unless the problem is modified to allow for the use of translation invariance; cf. Section 6. Buzbee and Dorr also exploit obvious symmetries of the rectangular plate problem when computing the capacitance matrix. In their experiments with Laplace's equation they take advantage of a choice of $n = 2^k$, k integer, by using a faster Laplace solver than the matrix decomposition method.

Once the capacitance matrix and its Cholesky factors have been found, the solution of the clamped plate problem requires the usual two calls of the fast solver subroutine and the use of the backsolving part of the Cholesky algorithm for a matrix of order p.

George's [18] paper contains a very interesting idea. George notes that a residual vector for the capacitance matrix equation $C\rho = b$ can be calculated at the expense of one call of the fast Laplace solver subroutine even if the elements of C have not been computed. If the capacitance matrix is positive definite symmetric, the corresponding linear system can therefore be solved by the conjugate gradient or the Davidon-Fletcher-Powell method. George reports rapid convergence for these methods and that the number of iterations required is proportional to $\sqrt{p}$. We note that, in the absence of rounding errors, the conjugate gradient and Davidon-Fletcher-Powell methods produce the same approximate solutions ρ_k from the same initial guess if the first approximation of C^{-1} in the Davidon-Fletcher-Powell algorithm is the identity matrix. The performance of the two methods is also reported to be virtually identical.

The Dirichlet problem for Laplace's equation is considered in detail. It is imbedded in a Dirichlet problem on a rectangular region, a problem with only strictly positive eigenvalues. A single-layer Ansatz is used and a symmetric capacitance matrix is obtained for simple boundaries or by using a crude approximation of the boundary condition; see discussion above. The capacitance matrix is then the restriction of the solution operator for the rectangle to the set of irregular mesh points. An attractive feature of this method is that the error in the boundary values is easily obtained in each step.

The error elsewhere is, by the maximum principle, bounded by the error on the boundary. This observation is used to obtain a most convenient stopping criterion.

The capacitance matrix C is a principal minor of the symmetric matrix B^{-1} in this case. Its condition number $\kappa(C)$, induced by the Euclidean vector norm, is therefore bounded by $\kappa(B)$ which is of the order n^2. George's numerical evidence is consistent with a conjecture that the condition number of C grows linearly with p. We have computed the condition numbers of certain families of capacitance matrices obtained from a single-layer Ansatz for Dirichlet problems and have found a linear growth with p; see Section 9. In that section we also report on cases where the single-layer Ansatz leads to an excessive number of conjugate gradient iterations. It should be noted that we have consistently worked with the normal equations $C^TC\rho = C^Tb$. It would be interesting to know how the conjugate gradient method performs when used directly for the nonsymmetric linear system of equations obtained by using a single-layer Ansatz and the Shortley-Weller approximation for a general region.

We will now give an estimate of $\kappa(C)$ in some more general cases. We will only treat cases in which both A and B are nonsingular. If we ignore the factors 2 and D, we can write the capacitance matrix C as $U^TAB^{-1}U$ for a single-layer Ansatz. We need an upper bound for

$$\max|Cx|_2 = \max|U^TAB^{-1}Ux|_2, \qquad |x|_2 = 1,$$

and a lower bound for

$$\min|Cx|_2 = \min|U^TAB^{-1}Ux|_2, \qquad |x|_2 = 1.$$

Such bounds are provided by

$$\max|U^TAB^{-1}y|_2, \qquad |y|_2 = 1, \quad \text{and} \quad \min|U^TAB^{-1}y|_2, \qquad |y|_2 = 1,$$

respectively, which in turn can be estimated by

$$|A|_2 \cdot |B^{-1}|_2 \quad \text{and} \quad 1/(|A^{-1}|_2|B|_2).$$

Thus,

$$\kappa(C) \leqslant \kappa(A)\kappa(B).$$

Both A and B typically have condition numbers on the order of h^{-2}. The upper bound for $|C|_2$ can be improved by the observation that in this case, hC is a quadrature-like discretization of the Fredholm integral equation of the first kind in Section 2. By using the techniques of Shieh [42], we can prove that $|C|_2$ is on the order of h^{-1} and the condition number $\kappa(C)$ will, therefore, grow no faster than const h^{-3} or const p^3.

Estimates of this type do not enable us to distinguish between a Neumann problem with $c = 1$, for which the capacitance matrices in our experience are uniformly well-conditioned, and a single-layer Ansatz for a Dirichlet problem. We will now show, by an example provided by Ole Hald, that there is little hope of improving this method.

Consider two tridiagonal matrices

$$B = \begin{pmatrix} 1 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}.$$

We note that A is a rank one perturbation of B which results when a discrete Neumann condition at one end of an interval is replaced by a discrete Dirichlet condition. The order of the capacitance matrix is, therefore, one; and the condition number equals one.

The inverse of B can easily be computed because $B = LL^T$, where

$$L = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & -1 & 1 \end{pmatrix}.$$

A straightforward computation shows that

$$AB^{-1} = I + \begin{pmatrix} n & (n-1) & \cdots & 1 \\ & & 0 & \end{pmatrix} = I + uv^T,$$

where $u^T = (1, 0, \ldots, 0)$ and $v^T = (n, n-1, \ldots, 1)$. To find the singular values, we consider

$$(I + uv^T)(I + vu^T) = I + W\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}W^T,$$

where W is a matrix with the two columns $v + (v^Tv/2)u$ and u. All singular values of AB^{-1} except two are thus equal to one, and we have only to find the remaining two. These coincide with the eigenvalues of

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (v + (v^Tv/2)u,\ u)^T(u,\ v + (v^Tv/2)u).$$

A direct calculation reveals that the smallest eigenvalue is on the order of $1/n$ while the largest is on the order of n^3. The condition number of AB^{-1}, therefore, grows

quadratically with n. This example shows that the estimate of the condition number of the restriction of the operator AB^{-1} to a subspace in terms of the condition number of AB^{-1} can be very crude.

In order to prove that the capacitance matrices introduced in Section 4 are nonsingular for $c > 0$, we first note that $C = U^T AB^{-1}\hat{V}$, where we again ignore the factors 2 and D. It is easy to show that $\hat{V} = PLP^T U$, where L is a unit lower triangular matrix and P a permutation matrix. By using the same arguments as in the single-layer case, we find that

$$\kappa(C) \leqslant \kappa(A)\kappa(B)\kappa(L)$$

and that C, thus, is nonsingular. This argument is again quite crude and fails to reveal the great power of the double-layer Ansatz.

An interesting application of a capacitance matrix method to the calculation of the potential flow of an incompressible fluid past a two-dimensional airfoil is given in Martin [35]. The stream function satisfies a zero Dirichlet condition on the airfoil and is expressed, far from the airfoil, in terms of the free stream velocity and an unknown parameter Γ, the circulation. An additional condition, the Kutta condition, allows the determination of Γ. The Dirichlet condition is represented by an interpolation formula which does not seem to be directly related to classical finite-difference techniques. The values of Γ and the appropriate point charges at certain mesh points inside and on the airfoil are computed by solving a capacitance matrix equation.

We also note that Angel and Bellman [1] have explored an imbedding idea in their work on the numerical solution of elliptic problems. A matrix for a linear system of equations, very similar to our matrix A, is obtained by imbedding the original problem in a rectangle. This system is then solved by a block-Gaussian elimination method. The imbedding simplifies the programming, because all blocks of the block-tridiagonal system of equations will be of the same order.

9. Numerical Experiments. In this section we will report on results from a series of numerical experiments which were carried out mainly on the IBM 360/75 computer at the Royal Institute of Technology in Stockholm. In our experiments we have used the program listed in the ERDA-NYU version of this paper and some slightly different versions of that program. Our programs closely reflect the ideas presented in Sections 1–7. We report on runs for different values of the mesh size and the constant c, for Dirichlet and Neumann boundary conditions and for three different regions. Our main purpose throughout has been to study the performance of the capacitance matrix method as a highly specialized linear equation solver. We have therefore worked extensively with problems with no discretization error such as those whose solutions are polynomials of second degree. This simplifies the study of the error originating from the linear equation solver. We are confident after a series of experiments that the performance of the method is virtually independent of the character of the data, the method of extending the data to $(C\Omega)_h$, the width of the strip and the values of the parameters N_- and N_+. Some of these experiments are further discussed below.

TABLE 1

	Neumann Problem		Dirichlet Problem, Double-Layer Ansatz		Dirichlet Problem Single-Layer Ansatz	
λ	p = 32	p = 64	p = 32	p = 64	p = 32	p = 64
0 - 0.1	1	1			0	37
0.1 - 0.2	0	0			17	16
0.2 - 0.3	0	0			7	4
0.3 - 0.4	0	0	1	1	2	2
0.4 - 0.5	1	1	0	0	1	1
0.5 - 0.6	0	0	1	1	1	2
0.6 - 0.7	0	0	0	2	2	0
0.7 - 0.8	0	0	10	24	0	0
0.8 - 0.9	1	7	13	24	0	0
0.9 - 1.0	6	13	1	6	0	0
1.0 - 1.1	11	22	4	4	0	0
1.1 - 1.2	10	18	0	0	0	0
1.2 - 1.3	0	0	0	1	0	0
1.3 - 1.4	0	1	1	0	0	1
1.4 - 1.5	2	1	0	0	0	0
1.5 - 1.6			0	0	1	1
1.6 - 1.7			0	0	0	
1.7 - 1.8			0	1	1	
1.8 - 1.9			0			
1.9 - 2.0			0			
2.0 - 2.1			1			

The distribution of the singular values of capacitance matrices of order p. The domain is a circle and the constant c in the Helmholtz equation is zero.

The first series of experiments was carried out to test the validity of the discrete potential theory. The singular values of certain capacitance matrices were computed by using the subroutines HSBG and ATEIG of the Scientific Subroutine Package. The results are shown in Table 1. These singular values are not directly comparable to those given in Section 7, because we have used the fundamental solution for a strip in our experiments. Other experiments for different regions also showed similar pronounced clusters of the singular values of the capacitance matrices. We note that the capacitance matrices in Table 1 which resulted from the single-layer Ansatz for the Dirichlet problem were fairly well conditioned. As expected, they have clusters of singular values close to the origin. Other cases reported in Table 2 are much more ill conditioned. The erratic behavior of the spectrum of the capacitance matrix when a single-layer Ansatz is used for the Dirichlet problem is clearly seen when we compare the case of $c = 0$ and that of $c = 10^{-6}$. See further the discussion in Section 7. In an earlier experiment the use of the method of Buzbee, Dorr, George and Golub [7] to handle

TABLE 2

		Neumann Problem		Dirichlet Problem Double-Layer Ansatz		Dirichlet Problem Single-Layer Ansatz	
		p = 32	p = 64	p = 32	p = 64	p = 32	p = 64
c = 0	I	10	11	9	11	10	22
	κ	2.9*	2.9*	5.4	4.8	12	20
$c = 10^{-6}$	I	10	11	10	15	17	35
	κ	670	440	5.4	4.8	**	**
$c = 10^{-3}$	I	11	12	10	14	13	28
	κ	67	60	4.9	4.4	500	860
c = 1	I	13	13	9	11	10	21
	κ	2.6	2.4	2.1	2.0	10	17
c = -1	I	14	15	9	12	12	21
	κ	2.9	3.0	2.8	2.6	8.4	14

The number of conjugate gradient iterations I and the spectral condition number $\kappa(C)$ of the capacitance matrices for problems on a circle for different values of the constant c and the number of irregular mesh points p.

* This condition number was computed disregarding a singular value very close to zero, see Section 7.

** In this case the eigenvalue routine returned negative values for the smallest singular values. The positive definiteness of C^TC did not survive the formation of C^TC, in double precision, and the use of the eigenvalue routine.

the case $c = 0$ resulted in a capacitance matrix which had a condition number 161 for $p = 64$ and a circular region. Thus, the use of their method would not appear to resolve the difficulties inherent in the use of the single-layer Ansatz for the Dirichlet problem.

In Tables 2 and 3 we report on a series of experiments designed to assess the efficiency of the conjugate gradient method. The tolerance for the conjugate gradient method was set equal to 10^{-6}. We note that in most applications we would be satisfied with a cruder tolerance.

The results reported in Table 4 indicate that the norm of the residual at the termination of the iteration gives a reliable estimate of the resulting error.

In Table 5 we report on an experiment which shows the importance of a proper scaling of the capacitance matrix. The choice of the scaling recommended in Section 4 resulted in nineteen iterations for the case $c = 0$. We note that the poorly-scaled capacitance matrices are fairly well conditioned. The large number of iterations undoubtedly reflects an unfortunate distribution of the singular values.

TABLE 3

		Neumann Problem			Dirichlet Problem, Double-Layer Ansatz		
		p = 27	p = 56	p = 117	p = 27	p = 56	p = 117
$c = 0$	I	13	15	15	16	16	18
	κ	3.5*	3.5*	**	3.5	4.0	**
$c = 10^{-6}$	I	26	31	38	20	19	22
	κ	3.5*	3.5*	**	3.5	4.0	**
$c = 10^{-3}$	I	20	23	26	16	17	20
	κ	3.5*	3.5*	**	3.3	3.8	**
$c = 1$	I	14	15	16	12	13	15
	κ	4.8	3.3	**	2.4	2.5	**
$c = -1$	I	14	16	16	13	14	15
	κ	3.9	3.6	**	3.5	3.3	**

The number of conjugate gradient iterations I and the condition number $\kappa(C)$ of capacitance matrices for the region displayed in Figure 2 for the different values of c and the number of irregular mesh points p.

* This condition number was computed disregarding a singular value very close to zero, see Section 7.

** Not available.

Experiments were carried out for the domain of Figure 2 to study the influence of the data on the performance of the algorithm. Dirichlet problems with the constant $c = 1$ were used and the tolerance for the conjugate gradient method was 10^{-6}. For a problem with a second degree polynomial solution, twelve iterations were required for $p = 27$ and fifteen iterations were needed for $p = 117$. The same problems with random data required thirteen and sixteen iterations, respectively.

The classical Fredholm theory is directly applicable only to regions with smooth enough boundaries. Experiments with square regions, with the sides of the square at 45° angle to the coordinate axes seem to indicate that our program will not require a more than normal number of iterations for domains with corners. We note, however, that such problems often fail to have sufficiently smooth solutions. This can make the discretization error of the finite-difference approximation very large.

In Diagram 1 we compare the rate of convergence predicted by the first part of inequality (7.3) and that actually observed in two experiments.

In Table 6 we report the CPU-time for our program when run on an IBM 360/75 using a FORTRAN H-level 2.1 optimizing compiler. Variations of measurements of up to 10% have been observed. We report the minimum recorded of several runs.

The required storage, excluding the program itself, is $(p + 22)p + 2n^2$ in our

TABLE 4

	Neumann Problem		Dirichlet Problem, Double-Layer Ansatz		
	p = 27	p = 56	p = 27	p = 56	p = 56
Norm of residual on exit from SYMMLQ	$2.3 \cdot 10^{-7}$	$4.0 \cdot 10^{-7}$	$6.8 \cdot 10^{-7}$	$6.0 \cdot 10^{-7}$	$4.7 \cdot 10^{-2}$
Maximum norm of the error	$1.2 \cdot 10^{-6}$	$1.4 \cdot 10^{-6}$	$6.0 \cdot 10^{-7}$	$9.6 \cdot 10^{-7}$	$8.4 \cdot 10^{-2}$
Normalized l_2-norm of the error	$2.8 \cdot 10^{-7}$	$4.0 \cdot 10^{-7}$	$1.9 \cdot 10^{-7}$	$2.2 \cdot 10^{-7}$	$2.0 \cdot 10^{-2}$
Number of iterations	14	15	12	13	3

A comparison of the norm of the residual of the conjugate gradient algorithm and the over all error for problems on the domain of Figure 2 and with $c = 1$. The normalized l_2-norm of the error is the l_2-norm divided by the square root of the number of points in $\Omega_h \cup \partial\Omega_h$.

```
434-DEKEN1    PAGE     21
HELMHOLTZ EQN.WITH DIRICHL. BOUNDARY CONDITIONS
LU+CU=F IN REGION
U=G        ON BOUNDARY

C= 0.0
NUMBER OF POINTS IN RECT.REGION= 16* 16
NUMBER OF BOUNDARY POINTS= 27

-LOG(|ERROR|)

 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 010 8 9 8 7 0 0 0 0 0
 0 0 0 0 0 0 9 8 8 8 8 8 0 0 0 0
 0 0 0 0 0 8 8 8 8 8 8 8 9 0 0 0
 0 0 0 0 0 8 8 8 8 8 8 9 8 0 0 0
 0 0 0 0 0 9 8 8 8 8 8 8 9 0 0 0
 0 0 0 0 0 8 8 8 8 8 8 8 9 0 0 0
 0 0 0 0 0 8 8 8 8 8 8 8 8 0 0 0
 0 0 0 0 0 8 8 8 8 8 8 0 0 0 0 0
 0 0 0 0 0 8 8 8 8 8 0 0 0 0 0 0
 0 0 0 0 0 0 8 8 8 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 8 8 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

NO.POINTS MAX & L2 - NORMS OF THE ERROR
  68     3.16924D-08     1.15157D-08
```

FIGURE 2

The last page of output from a run with our program. Points with a zero value represent exterior mesh points. The strictly positive numbers indicate the number of correct decimal digits, rounded to the nearest integer, obtained at the particular mesh points.

TABLE 5

	Scaling according to formula (4.4)		Scaling making diagonal element equal to $4 + ch^2$	
Constant c	0	1	0	1
Number of iterations with tolerance ACCY = 10^{-6}	> 56	34	16	13
Condition number of C	10	11.5	4.0	2.5

A comparison on the effect of scaling of the capacitance matrix on the performance of the conjugate gradient method for Dirichlet problems on the domain of Figure 2. The number of irregular points was $p = 56$.

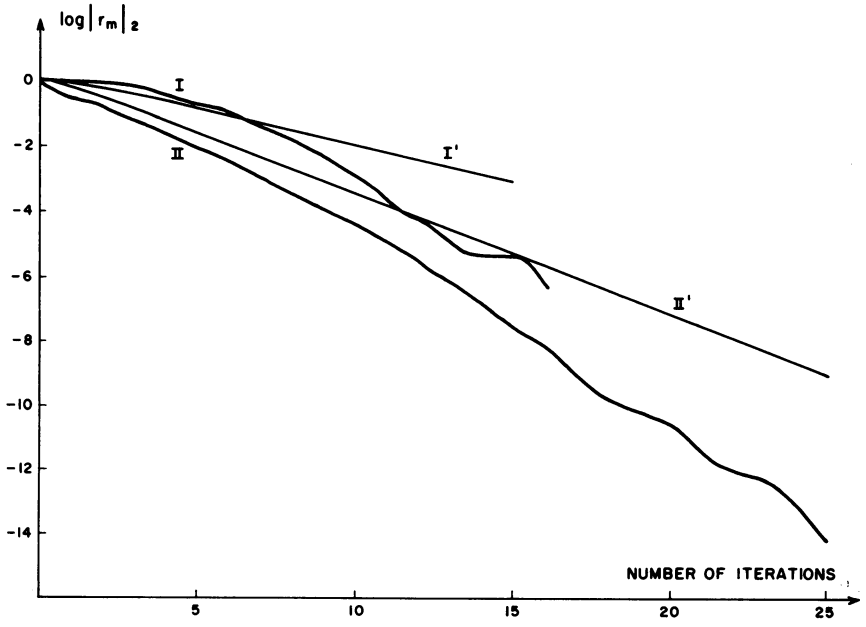

DIAGRAM 1

The rate of convergence of the conjugate gradient method for the capacitance matrix equation for Dirichlet problems. The curve I is for the constant $c = 0$ and the curve II for $c = 1$. The curves I′ and II′ are based on computed values of the condition numbers and formula (7.3).

implementation. The number of irregular points p grows linearly with n. In the case reported in Table 6 p is approximately equal to $2n$. We note that the Helmholtz solver contributes less than 20% to the total execution time in these cases. We therefore ignore the term proportional to $n^2 \log_2 n$ in the estimate of the execution time and conjecture that the time grows as const p^2 if the conjugate gradient variant is used. When we test

TABLE 6

Number of irregular mesh points p	27	56	117
Total number of mesh points	16 × 16	32 × 32	64 × 64
Time for Helmholtz solver SOLVE	0.04	0.15	0.73
Time to generate capacitance matrix in the Neumann case	0.06	0.26	1.15
In the Dirichlet case	0.16	0.70	3.06
Factorization time for Gaussian elimination	0.08	0.55	5.11
Backsolving time for Gaussian elimination	< 0.01	0.03	0.12
Time to reach 10^{-6} tolerance by using the conjugate gradient method	0.26	0.96	4.08
Time to generate the problem and other overhead*	0.18	0.53	2.11
Total time using conjugate gradient method in the Neumann case	0.50	1.75	7.34
In the Dirichlet case	0.60	2.19	9.55
Total time with Gaussian elimination, Dirichlet case	0.36	1.81	10.62
Total time to solve an additional problem using the Gaussian elimination option	0.16	0.46	1.77

CPU-time in seconds for problems on the region of Figure 2 on an IBM 360/75.

* The time to read in data and to print the results is not included.

this conjecture using the two right most columns of Table 6 we find that the ratio of the squares of the values of p equals 4.36, while the execution time grows by a factor of 4.20 for the Neumann problem and by a factor of 4.36 for the Dirichlet problem.

In order to assess the efficiency of our program we compare the CPU-time of our Helmholtz solver with those of other solvers on a 128 ×128 rectangular mesh as they are given in Hockney [27]. The results are given in Table 7. All times refer to runs on the IBM 360/75 computer using an H-level FORTRAN compiler. We note that our program is slower than the others, but that this is of little importance in the present context.

This and other parts of our code could undoubtedly be made faster. Our current program is a modification of a program developed primarily to test the validity of the discrete potential theory; relatively little attention was paid to its speed of execution.

A comparison is made in Table 8 between the running times of two versions of

TABLE 7

POT 1, Hockney's method	2.55
XYPOIS, Buneman's method	3.41
ODDEVN, A. George's implementation of Buneman's method	3.18
SOLVE, our Fourier-Toeplitz solver	3.55

CPU-time in seconds for Poisson solvers on IBM 360/75 computers. H-level FORTRAN compilers are used and the mesh is 128 × 128. The data on Hockney's and Buneman's methods are obtained from Hockney [27].

TABLE 8

	Neumann Problem		Dirichlet Problem	
	n = p = 32	n = p = 64	n = p = 32	n = p = 64
Our current program on an IBM 360/75. Conjugate gradient option with ACCY = 10^{-6}	0.95	3.36	1.14	4.02
An older version of our program on an IBM 360/75. Conjugate gradient option with ACCY = 10^{-6}	–	–	1.72	6.19
The older version on a CDC 7600 with the RUN 76 compiler. Conjugate gradient option with ACCY = 10^{-6}	–	–	0.348	1.43
With Gaussian elimination option	–	–	0.338	1.45
Buzbee and Dorr's Fortran program, as reported in [6], on a CDC 7600	–	–	0.365	2.65

CPU-time in seconds. The region in our experiments was a circle.

our program and those reported in Buzbee and Dorr [6]. The runs with an older version of our program on the CDC 7600 were kindly carried out for us by Dr. Victor Pereyra while he was visiting the Lawrence Berkeley Laboratory. When the parameters n and p were doubled the execution time of our program grew only by a factor between

3.52 and 4.10, while those of Buzbee and Dorr grew by a factor 7.26 or roughly as the second and third powers of p, respectively. This is as predicted by the operation counts. We, therefore, expect our program to be even more competitive for larger problems. We also note that the ratio p/n is only one in these experiments while a more typical value would be about three. In such cases with relatively few boundary points the Gaussian elimination option is just about as efficient as the conjugate gradient option. The experiments reported by Buzbee and Dorr are for quite simple geometries which would permit a simplification of our program.

Computer Science Department
Royal Institute of Technology
Stockholm 70, Sweden

Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York, New York 10012

1. E. ANGEL & R. BELLMAN, *Dynamic Programming and Parital Differential Equations,* Math. in Sci. and Engineering, vol. 88, Academic Press, New York and London, 1972. MR **50** #11822.

2. R. BARTELS & J. W. DANIEL, "A conjugate gradient approach to nonlinear elliptic boundary value problems in irregular regions," *Conference on the Numerical Solution of Differential Equations* (Dundee, July 1973), Lecture Notes in Math., vol. 363, Springer-Verlag, Berlin and New York, 1974, pp. 1–11.

3. J. H. BRAMBLE & B. E. HUBBARD, "A theorem on error estimation for finite difference analogues of the Dirichlet problem for elliptic equations," *Contributions to Differential Equations,* v. 2, 1963, pp. 319–340. MR **27** #2114.

4. J. H. BRAMBLE & B. E. HUBBARD, "Approximation of derivatives by finite difference methods in elliptic boundary value problems," *Contributions to Differential Equations,* v. 3, 1964, pp. 399–410. MR **29** #4208.

5. O. BUNEMAN, *A Compact Non-Iterative Poisson Solver,* Report SUIPR-294, Inst. Plasma Research, Stanford 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. MR **45** #1403.

8. B. L. BUZBEE, G. H. GOLUB & C. W. NIELSON, "On direct methods for solving Poisson's equations," *SIAM J. Numer. Anal.,* v. 7, 1970, pp. 627–656. MR **44** #4920.

9. L. COLLATZ, *The Numerical Treatment of Differential Equations,* Springer-Verlag, Berlin and New York, 1966.

10. P. CONCUS & G. H. GOLUB, "Use of fast direct methods for efficient numerical solution of nonseparable elliptic equations," *SIAM J. Numer. Anal.,* v. 10, 1973, pp. 1103–1120. MR **49** #6636.

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

12. R. COURANT & D. HILBERT, *Methods of Mathematical Physics.* Vol. I, Interscience, New York, 1953. MR **16**, 426.

13. J. W. DANIEL, "The conjugate gradient method for linear and nonlinear operator equations," *SIAM J. Numer. Anal.,* v. 4, 1967, pp. 10–26. MR **36** #1076.

14. F. W. DORR, "The direct solution of the discrete Poisson equation on a rectangle," *SIAM Rev.,* v. 12, 1970, pp. 248–263. MR **42** #1353.

15. 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.

16. G. E. FORSYTHE & W. R. WASOW, *Finite-Difference Methods for Partial Differential Equations,* Appl. Math. Ser., Wiley, New York, 1960. MR **23** #B3156.

17. P. R. GARABEDIAN, *Partial Differential Equations,* Wiley, New York, 1964. MR **28** #5247.

18. 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.

19. G. H. GOLUB, "An Algorithm for the Discrete Biharmonic Equation," (Unpublished.)

20. 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,* Nat. Bur. Standards Appl. Math. Ser., no. 39, U. S. Government Printing Office, Washington, D. C., 1954, pp. 71–103. MR **16**, 597.

21. M. R. HESTENES, "The conjugate gradient method for solving linear systems," *Proc. Sympos. Appl. Math.,* vol. 6, *Numer. Anal.,* Amer. Math. Soc., Providence, R. I., 1956, pp. 83–102. MR **18**, 824.

22. M. R. HESTENES & E. STIEFEL, "Methods of conjugate gradients for solving linear systems," *J. Res. Nat. Bur. Standards,* v. 49, 1952, pp. 409–436. MR **15**, 651.

23. R. W. HOCKNEY, "A fast direct solution of Poisson's equation using Fourier analysis," *J. Assoc. Comput. Mach.,* v. 12, 1965, pp. 95–113. MR **35** #3913.

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

25. R. W. HOCKNEY, *The Potential Calculation and Some Applications,* Methods in Computational Physics, vol. 9, Academic Press, New York, 1970.

26. 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.

27. R. W. HOCKNEY, *Computers, Compilers and Poisson-Solvers,* Computer Science Department Report, Reading University, 1972.

28. A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis,* Blaisdell, New York, 1964. MR **30** #5475.

29. A. JAMESON, *Accelerated Iteration Schemes for Transonic Flow Calculations Using Fast Poisson Solvers,* N. Y. U. ERDA report C00-3077-82, 1975.

30. W. J. KAMMERER & M. Z. NASHED, "On the convergence of the conjugate gradient method for singular operator equations," *SIAM J. Numer. Anal.,* v. 9, 1972, pp. 165–181.

31. S. KANIEL, "Estimates for some computational techniques in linear algebra," *Math. Comp.,* v. 20, 1966, pp. 369–378. MR **38** #2934.

32. L. V. KANTOROVIČ & V. I. KRYLOV, *Approximate Methods of Higher Analysis,* 3rd ed., GITTL, Moscow, 1950; English transl., Interscience, New York; Noordhoff, Groningen, 1958. MR **13**, 77; **21** #5268.

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. MR **13**, 163.

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. MR **49** #11953.

36. E. D. MARTIN, "Progress in application of direct elliptic solvers for transonic flow computations," *Aeordynamics Analyses Requiring Advanced Computers,* NASA SP-347, 1975. (To appear.)

37. E. D. MARTIN, "A fast semidirect method for computing transonic aerodynamic flows," *Proceedings of the AIAA 2nd Computational Fluid Dynamics Conference,* June 1975. (To appear.)

38. C. C. PAIGE & M. A. SAUNDERS, "Solutions of sparse indefinite systems of equations," *SIAM J. Numer. Anal.,* v. 12, 1975, pp. 617–629.

39. V. PEREYRA, W. PROSKUROWSKI & O. WIDLUND, "On a family of elliptic difference schemes suggested by Heinz-Otto Kreiss." (To appear.)

40. C. S. PESKIN, Personal communication.

41. I. G. PETROVSKIĬ, *Lectures on Partial Differential Equations,* GITTL, Moscow, 1950; English transl., Interscience, New York, 1954. MR **13**, 241; **16**, 478.

42. A. SHIEH, *Fast Poisson Solver on Nonrectangular Domains,* Ph. D. Thesis, New York University, June 1976.

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

44. 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. MR 41 #4834.

45. J. TODD (Editor), *Survey of Numerical Analysis,* McGraw-Hill, New York, 1962. MR 24 #B1271.

46. 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,* Edited by D. J. Rose and R. A. Willoughby, Plenum Press, New York, 1972.