THE USE OF DISCRETE GREEN'S FUNCTIONS IN THE NUMERICAL SOLUTION OF POISSON'S EQUATION

BY

J. M. BERGER AND G. J. LASHER[1]

1. Introduction

The numerical solution of Poisson's equation in two dimensions has, with the advent of large-scale computers, attracted a great deal of interest in recent years. Most of the literature concerns itself with the convergence of the finite difference equation approximation [1] or with iterative methods of approaching the numerical solution [2], although some discussion has been given by Hyman [3] on the exact solutions of Poisson's difference equation.

This paper presents a practical method for directly obtaining the exact solution of the usual difference equation that approximates the solution of Poisson's equation. The method, as developed here, applies to problems defined on a rectangle or rectangular strip, and therefore also to problems defined on regions which may be conformally mapped onto a rectangle or rectangular strip. It consists of the following procedure: (*a*) deriving an expression for the exact discrete Green's function satisfying the required boundary conditions, (*b*) evaluating this function numerically, and (*c*) applying this function to obtain the desired solution of the difference equation.

This method is currently being used on the IBM 704 computer to obtain the solutions of certain time dependent vector field problems. These problems require the solution of a two-dimensional Poisson's equation at each time step.

In general, the use of the Green's function method is most advantageous when many solutions of the difference equation are required with the same boundary conditions, but with different inhomogeneous parts. A specific advantage of the Green's function method, as compared with iterative methods, is that a definite high degree of accuracy of the solution of the difference equation is obtained in a predetermined sequence of arithmetic operations. An additional advantage is that the method may be employed to obtain the solution in only that portion of the total region for which it is desired. On the other hand, in many cases, particularly those problems for which modest accuracy suffices, the iterative methods will yield an approximate solution after a fewer total number of arithmetic operations.

The Green's function is the inverse of a matrix whose elements are given by the coefficients of the difference equation and the boundary conditions, and

Received March 29, 1957; received in revised form October 15, 1957.

[1] The authors are indebted to Dr. H. H. Goldstine of the Institute for Advanced Study for his aid and advice. We wish also to thank Miss Maureen Kane for programming the problem for the IBM 704. This work was supported, in part, by the Air Force.

the method could be derived and applied in matrix language without mention of the concept of a Green's function. However, the use of the Green's function concept and the consequent manipulation of difference operators in analogy with the methods of partial differential equations seems to greatly simplify the derivation and exposition of the method. In particular, it leads naturally to the full exploitation of the special properties of the system of difference equations. Practically, the use of a matrix inversion routine, such as one based on the method of von Neumann and Goldstine [4], would use much more computer time, as indicated in Section 7, and result in a much greater round-off error than the numerical evaluation of the expressions derived in this paper, as demonstrated in Section 5.

In Section 2, a brief discussion is given of the origin of the discrete problems solved in Section 4. A general mathematical description of the discrete Green's function method and its relation to matrix methods is given in Section 3. In Section 4, the Green's functions are derived for two particular sets of boundary conditions. An expression for an upper bound of the round-off error in the numerical evaluation of a Green's function is given in Section 5. This is then shown to be of much lower order than the upper bound obtained by von Neumann and Goldstine [4] for the error in the inversion of a general matrix by an elimination method. An extension of the method to include the problems in which the solution achieves an arbitrary function on the boundary is derived in Section 6. In Section 7, there is presented a refinement of the method which reduces the number of necessary arithmetic operations, and which allows the use of different lattice spacings in the various portions of the region.

2. Origin of the discrete problems

The work reported here arose from the necessity of obtaining the solution of a two-dimensional set of vector field equations defined over an annular region. The solution of the vector equations is derivable from the solution of Poisson's equation in the given region, and the resultant problem is further simplified by a conformal mapping of the annular region onto a rectangular strip with periodicity along the strip.[2] That is, the problem was reduced to the solution of

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\varphi(x, y) = f(x, y), \tag{2.1}$$

with the periodic condition

$$\varphi(x, Y + y) = \varphi(x, y) \tag{2.2}$$

in the region $0 < x < X$.

[2] A procedure alternate to performing the conformal mapping is to difference the original Poisson's equation in polar coordinates over a lattice of points determined by the corners of curvilinear squares. The resulting difference equation would be identical to that found by the procedure pursued in the text.

Two distinct problems were considered, corresponding to the boundary conditions at $x = 0, X$: (a) the solution is constant on the boundaries with one of the constants determined by an integral condition on the normal derivative over one boundary, that is,

$$\varphi(0, y) = 0, \qquad \varphi(X, y) = C, \tag{2.3}$$

where C is determined by the condition

$$\int_0^Y \frac{\partial}{\partial x}\varphi(x, y)\Big|_{x=0} dy = 0; \tag{2.3a}$$

and (b) the normal derivative in the x direction vanishes on the boundaries, that is,

$$\frac{\partial}{\partial x}\varphi(x, y)\Big|_{x=0} = \frac{\partial}{\partial x}\varphi(x, y)\Big|_{x=X} = 0, \tag{2.4}$$

thus defining the solution, $\varphi(x, y)$, only to within an additive constant.

For convenience, the central difference operator, Δ_n^2, is defined as

$$\Delta_n^2 \psi_n = \psi_{n+1} - 2\psi_n + \psi_{n-1}\,. \tag{2.5}$$

The solution of the differential equation (2.1) is approximated at discrete points by the solution of the difference equation

$$(\Delta_k^2 + \Delta_l^2)\varphi_{k,l} = f_{k,l}\,. \tag{2.6}$$

The discrete points are given by $x = kh$, $y = lh$, where h is the lattice spacing, and k, l are integers. The physical boundaries of the strip were taken at $k = \frac{1}{2}$, $k = K + \frac{1}{2}$. The boundary conditions then become, for equation (2.2), the periodic condition

$$\varphi_{k,l+L} = \varphi_{k,l}\,. \tag{2.7}$$

Equations (2.3) and (2.3a), for case (a) become

$$\varphi_{0,l} + \varphi_{1,l} = 0, \qquad \varphi_{K,l} + \varphi_{K+1,l} = C, \tag{2.8}$$

with C determined by

$$\sum_{l=0}^{L-1} (\varphi_{1,l} - \varphi_{0,l}) = 0; \tag{2.8a}$$

and equation (2.4), for case (b),

$$\varphi_{1,l} - \varphi_{0,l} = \varphi_{K+1,l} - \varphi_{K,l} = 0. \tag{2.9}$$

In the following sections, the Green's functions are derived for these two problems.

3. Outline of the discrete Green's function method

The partial difference equation (2.6) may be solved by direct analogy with

the method of the Green's function in partial differential equations. A discrete Green's function, $G_{k,l;k',l'}$, is defined to satisfy the difference equation

$$(\Delta_k^2 + \Delta_l^2)G_{k,l;k',l'} = \delta_{kk'}\,\delta_{ll'}\,, \tag{3.1}$$

where δ_{ij} is the Kronecker delta. The function $G_{k,l;k',l'}$ is then completely determined by the requirement that it satisfy the same boundary conditions as are satisfied by the desired solution. The Green's function, thus defined, yields the desired solution, $\varphi_{k,l}$, by

$$\varphi_{k,l} = \sum_{k',l'} G_{k,l;k',l'}\, f_{k',l'}\,. \tag{3.2}$$

The function $G_{k,l;k',l'}$ may be expressed as a linear combination of a complete orthonormal set of the eigenfunctions of the operator $(\Delta_k^2 + \Delta_l^2)$. That is, if $\psi_{k,l;j,m}$ is defined to satisfy: (a) the difference equation

$$(\Delta_k^2 + \Delta_l^2 - \lambda_{j,m})\psi_{k,l;j,m} = 0, \tag{3.3}$$

where j, m are indices labeling the distinct eigenfunctions; (b) the orthonormality condition

$$\sum_{k,l} \psi^*_{k,l;j,m}\,\psi_{k,l;j',m'} = \delta_{jj'}\,\delta_{mm'}\,, \tag{3.3a}$$

which implies

$$\sum_{j,m} \psi^*_{k,l;j,m}\,\psi_{k',l';j,m} = \delta_{kk'}\,\delta_{ll'}\,; \tag{3.3b}$$

and (c) the appropriate boundary conditions; then the Green's function is given by

$$G_{k,l;k',l'} = \sum_{j,m} (1/\lambda_{j,m})\psi^*_{k,l;j,m}\,\psi_{k',l';j,m}\,. \tag{3.4}$$

It is highly desirable to perform analytically at least one of the sums occurring in the expression for the Green's function, equation (3.4), in order to avoid the accumulation of large round-off errors in the numerical evaluation. Two lemmas are derived in the appendix that are useful for that purpose, and these are applied in the next section.

If a complete set of boundary conditions is combined with the difference equations (3.1), the result may be considered as the KL dimensional matrix equation

$$DG = I, \tag{3.5}$$

where the matrix D is formed from the coefficients of the difference equations as modified by the boundary conditions; the matrix G is formed from the elements of the Green's function, and I is the unit matrix. The method outlined above for obtaining G, stated in matrix language, is equivalent to determining the orthogonal similarity transformation formed from the eigenvectors of D, which diagonalizes D. The inverse of this transformation operating upon the inverse of the diagonal form of D yields the Green's function as in equation (3.4). The necessary manipulations of the matrix D may be done by

operations with submatrices of D. This decomposition of D into submatrices corresponds to the separation of variables of the difference equation (3.1) which is used in the next section. The determination of the eigenvectors of the triply diagonal submatrices of D is equivalent to the solution of the one-dimensional difference equations obtained by the separation of variables in the next section. Thus, the solution of the difference equation may be obtained by purely matrix methods. However, it seems clear that the language adopted in this paper leads more easily and directly to the final result.

4. Evaluation of the Green's functions

In further analogy with the solution of partial differential equations, the difference equation for the eigenfunctions, (3.3), may be solved by the separation of variables, that is, under the assumption

$$\psi_{k,l;j,m} = X_{k,m}\, Y_{l,j}\,. \tag{4.1}$$

The eigenfunctions are obtained from the solutions of

$$(\Delta_k^2 - \nu_m)X_{k,m} = 0, \qquad (\Delta_l^2 - \gamma_j)Y_{l,j} = 0, \tag{4.2}$$

where

$$\nu_m + \gamma_j = \lambda_{j,m}\,. \tag{4.3}$$

The orthonormal set of solutions satisfying equations (4.2) and the boundary conditions (2.7) and (2.8), with C set equal to zero, are given by

$$\begin{aligned} X_{k,m} &= (2/K)^{1/2} \sin(\pi m/K)(k - \tfrac{1}{2}), && m = 1, 2, \cdots, K-1, \\ X_{k,K} &= (1/K)^{1/2} \sin \pi(k - \tfrac{1}{2}), \\ Y_{l,j} &= (1/L)^{1/2} \exp(2\pi i l j/L), && j = 0, 1, \cdots, L-1, \end{aligned} \tag{4.4}$$

with the eigenvalues

$$\lambda_{j,m} = 2\cos 2\pi j/L + 2\cos \pi m/K - 4. \tag{4.5}$$

The Green's function is obtained by substitution of equations (4.4) and (4.5) into equation (3.4). The sum over j in the resultant expression may be performed with the aid of Lemma I (cf. Appendix I), and the final expression for the Green's function is

$$\begin{aligned} G^0_{k,l;k',l'} = \frac{2}{K}\Bigg\{ &\sum_{m=1}^{K-1} S_{l-l'}(r_m) \sin\frac{m\pi}{K}(k - \tfrac{1}{2}) \\ &\cdot \sin\frac{m\pi}{K}(k' - \tfrac{1}{2}) + (-1)^{k+k'} S_{l-l'}(r_K)\Bigg\}, \end{aligned} \tag{4.6}$$

where

$$S_n(r_m) = -\frac{r_m^{(L/2-n)} + r_m^{-(L/2-n)}}{(a_m^2 - 4)^{1/2}(r_m^{L/2} - r_m^{-L/2})}, \tag{4.7}$$

and

$$r_m = \tfrac{1}{2}(a_m + (a_m^2 - 4)^{1/2}); \qquad a_m = 4 - 2\cos(\pi m/K).$$

In order to obtain the Green's function, $G^a_{k,l;k',l'}$, satisfying the sum condition (2.8a), as well as conditions (2.7) and (2.8), a solution of the homogeneous difference equation is added to $G^0_{k,l;k',l'}$. That is,

$$G^a_{k,l;k',l'} = G^0_{k,l;k',l'} + A_{k'} \cdot (k - \tfrac{1}{2}), \tag{4.8}$$

where $A_{k'}$ is determined, by substitution of (4.8) into the sum condition, to be

$$A_{k'} = (1/KL)(K - k' + \tfrac{1}{2}), \tag{4.9}$$

which completes the derivation of the Green's function for case (a).

The difference equation for the Green's function, equation (3.1), has no nontrivial solutions which satisfy the conditions of case (b): (2.7) and (2.9). Correspondingly, solutions of Poisson's difference equation exist only when $\sum_{k,l} f_{k,l} = 0$. These are the result of a condition exactly analogous to the integral condition for continuous variables:

$$\int_A \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \varphi(x, y)\, dx\, dy = \oint \vec{\nabla}\varphi(x, y) \cdot \vec{n}\, dl, \tag{4.10}$$

where A is a two-dimensional region, $\vec{n}$ is the unit normal to the boundary of A, and dl is the line increment along the boundary. The discrete analogy to (4.10) is a sum condition on any function, $\varphi_{k,l}$, of two discrete variables:

$$\sum_S (\Delta_k^2 + \Delta_l^2)\varphi_{k,l} = \sum \text{(Normal differences of } \varphi_{k,l} \text{ on the boundary of } S), \tag{4.11}$$

where S is any set of points in the region.

Taking the region S to be the rectangle ($1 \leqq k \leqq K, 1 \leqq l \leqq L$), the boundary condition (2.9) and the periodic condition (2.7) state that the right-hand side of (4.11) vanishes for the desired solution $\varphi_{k,l}$ of Poisson's equation, and hence also for the appropriate Green's function $G^b_{k,l;k',l'}$. The difference equation for the Green's function, equation (3.1), must therefore be modified so that the sum of the inhomogenous part over the region vanishes. That is, equation (3.1) becomes

$$(\Delta_k^2 + \Delta_l^2)G^b_{k,l;k',l'} = \delta_{kk'}\,\delta_{ll'} - 1/KL. \tag{4.12}$$

In the same manner as before, the eigenfunctions are obtained as

$$\begin{aligned} X_{k,m} &= (2/K)^{1/2} \cos(\pi m/K)(k - \tfrac{1}{2}), \qquad m = 1, 2, \cdots, K - 1, \\ X_{k,0} &= (1/K)^{1/2} \\ Y_{l,j} &= (1/L)^{1/2} \exp(2\pi i l j/L), \qquad j = 0, 1, \cdots, L - 1, \end{aligned} \tag{4.13}$$

with the eigenfunctions given by the same expression, equation (4.5), as before, but with the range of the index m from 0 to $(K - 1)$, thus yielding a zero eigenvalue. The Green's function $G^b_{k,l;k',l'}$ is derived from these functions using the same expression, equation (3.4), except that the term including

the zero eigenvalue is omitted.[3] The resultant Green's function, after performing the sum over j, is

$$G^b_{k,l;k',l'} = \frac{2}{K}\left\{\sum_{m=1}^{K-1} S_{l-l'}(r_m)\cos\frac{\pi m}{K}(k-\tfrac{1}{2})\cos\frac{\pi m}{K}(k'-\tfrac{1}{2}) + \sigma_{l-l'}\right\}, \tag{4.14}$$

$$\sigma_n = -\frac{(L-1)(L+1)}{12L} - \frac{n(L-1)}{2L} + \frac{n(n-1)}{2L},$$

where $S(r_m)$ and r_m are as previously defined.

5. Effect of round-off error in the numerical evaluation of a Green's function

This section consists of an analysis of the effect of round-off error for the digital computer evaluation of a Green's function assuming a reasonable arithmetic procedure. The upper bound derived for this error is then compared with that obtained by von Neumann and Goldstine for the inversion of a general matrix by an elimination method [4]. An estimate of the range of magnitude of the elements of the Green's function is made to determine the expected relative error of the computed elements.

The analysis is done only for the Green's function $G^0_{k,l;k',l'}$ of equations (4.6) and (4.7). The formulas for Green's functions satisfying other boundary conditions differ only by the addition of a small number of terms whose contribution to the total round-off error is negligible. It is assumed that the integer L is greater than or equal to K. In the alternative case, $K > L$, the analytic sum using the lemmas of Appendix I could be performed over the index m instead of j, thus reversing the roles of K and L in the error analysis.

The error analysis of this section applies to computations by a floating-point machine or a set of floating-point subroutines. In the course of the analysis, it is convenient to distinguish between generated and propagated errors in the result of an arithmetic operation. Generated error refers to truncation and round-off error committed in the operation referred to; that is, it is the error in the result of an operation if the numbers operated upon are assumed exact. Propagated error refers to that part of the error in the result of an operation which arises from the errors in the numbers operated upon. This separation of total error into two parts is expressed by the equation

$$F(x,y) - \overline{F(\bar{x},\bar{y})} = [F(x,y) - F(\bar{x},\bar{y})] + [F(\bar{x},\bar{y}) - \overline{F(\bar{x},\bar{y})}], \tag{5.1a}$$

where the barred quantities are the truncated digital approximations to the unbarred quantities, and the function $F(x, y)$ represents the arithmetic operation. The two expressions in square brackets represent respectively the

[3] Dr. H. H. Goldstine has pointed out to the authors that the condition, $\sum_{k,l}\varphi_{k,l} = 0$, may be imposed on the solution in case (b). This condition determines the solution completely and suppresses the zero eigenvalue, since its corresponding eigenfunction does not satisfy the new condition.

propagated and generated errors. In terms of relative errors, this equation may be approximated by

$$R[F(x, y)] = \left|\frac{x}{F}\frac{\partial F}{\partial x}\right| R(x) + \left|\frac{y}{F}\frac{\partial F}{\partial y}\right| R(y) + \delta, \tag{5.1b}$$

where $R[F]$, $R(x)$ stand for upper bounds for the relative errors in F and x respectively, and δ is an upper bound for the generated relative error in $F(x, y)$. In the case of the floating-point operations of the IBM 704, it can be shown that the value of $\delta = 2^{-26}$ may be taken when $F(x, y)$ represents one of the elementary arithmetic operations of division, multiplication, addition, or subtraction.[4] The generalization for this value of the generated error is $\delta = \beta^{-s+1}$, where β equals the radix used and s, the number of digits carried in the fraction part of the floating-point number. An expression similar to (5.1b) is also used in the analysis to evaluate error bounds for the case of a function of one variable which is computed by a machine subroutine. Another relation used is that for the upper bound on the absolute error in a computed sum of N numbers:

$$E\left(\sum_{i=1}^{N} x_i\right) = \sum_{i=1}^{N} E(x_i) + N\delta \sum_{i=1}^{N} |x_i|, \tag{5.2}$$

where $E(x)$ stands for an upper bound on the absolute error in x. To simplify the language of this section, the term "error" will be used instead of "an upper bound for the error".

The absolute error, $E(G)$, of any element, $G_{k,l;k',l'}$, of the Green's function is given by

$$E(G) \leqq K\delta \sum_{m=1}^{K} |T_m| + \sum_{m=1}^{K} E(T_m) \leqq (K\delta + R_T) \sum_m |T_m|, \tag{5.3}$$

where

G, any element of the Green's function $= \sum_m T_m$,
$E(T_m)$ is absolute error in each term, T_m,
R_T is upper bound for relative error of T_m for all m.

The upper bound R_T is obtained explicitly by summing the relative errors in each factor of T_m. This bound is derived in detail in Appendix II by considering the steps of the computation. The result, equation (II.9), is

$$R_T = (12L + 42)\delta. \tag{5.4}$$

An upper bound for the sum $\sum_m T_m$ is obtained by noting that

$$\sum_{m=1}^{K} |T_m| \leqq \frac{2}{K} \sum_{m=1}^{K} \frac{1}{(a_m^2 - 4)^{1/2}} \frac{r_m^L + 1}{r_m^L - 1}. \tag{5.5}$$

[4] In the case of subtraction of almost equal numbers of like sign but unequal exponents, the 704 arrives at a result whose *generated* relative error is only 2^{-26} by saving and taking account of the extra digit shifted off in the process of making the exponents of the two numbers equal. In such a case, the *total* error will be much larger than this if there exists any error in the numbers subtracted. This larger error is, of course, contributed by the *propagated* error.

Application of the inequality (II.7) yields the result:

$$(r_m^L + 1)/(r_m^L - 1) < 2, \tag{5.6}$$

where $L \geqq K$ has been assumed. The expression (II.3) indicates that

$$(a_m^2 - 4)^{1/2} > 4 \sin(\pi m/2K) \geqq 4m/K, \quad \text{for } 1 \leqq m \leqq K. \tag{5.7}$$

The final result is

$$\sum_{m=1}^{K} | T_m | < \sum_{m=1}^{K} 1/m < 1 + \ln K. \tag{5.8}$$

If we combine this result with that for R_T, equation (5.4), the error of an element of the Green's function becomes

$$E(G) = (\ln K + 1)(K + 12L + 42)\delta. \tag{5.9}$$

The term $K \ln K$ of (5.9) arises from the numerical sum over the index m. The desirability of performing the sum over j analytically is indicated by the fact that a double numerical sum over both m and j would result in an error term of order $KL \ln K$.

The error bound derived by von Neumann and Goldstine for the inversion of a general symmetric matrix with positive eigenvalues by a numerical elimination method is

$$| BD - I | \leqq 14.24(\lambda/\mu)n^2\beta^{-s}, \tag{5.10}$$

where D is the given matrix; B, the numerically computed inverse; I, the unit matrix; λ/μ, the ratio of greatest to least eigenvalue of D; n, the dimension of D; $| BD - I |$, the upper bound of the matrix $BD - I$, that is, its largest eigenvalue; and β^{-s} is as previously defined. For the matrix defined by the difference equation, λ/μ can be determined from equation (4.5). To highest order in K

$$\lambda/\mu \cong 8K^2/\pi^2. \tag{5.11}$$

Specializing the result of von Neumann and Goldstine, equation (5.10), to the case treated here, with $n = KL$, one obtains

$$| BD - I | = | ED | < 12K^4L^2\delta/\beta, \tag{5.12}$$

where E is an error matrix defined by $B = D^{-1} + E$.

The comparable result for the method of this paper may be obtained from the error bound $E(G)$, given by equation (5.9), and the relation

$$| ED | \leqq \text{Norm}(ED) = [\sum_{i,j=1}^{n} (\sum_{k=1}^{n} E_{ik} D_{kj})^2]^{1/2}, \tag{5.13}$$

where $E_{ik} \leqq E(G)$ by definition of $E(G)$, and $\sum_{k=1}^{n} | D_{kj} | = 8$ from the difference equation (3.1). The result is

$$\begin{aligned} | ED | &\leqq 8nE(G) \\ &= 8KL(12L + K)(\ln K)\delta + \text{terms of lower order in } K, L. \end{aligned} \tag{5.14}$$

Comparison of the error bounds (5.14) and (5.12) demonstrates the expected advantage gained by computing the inverse of D by the method of this paper. The ratio of the two error estimates, i.e. of the order of $K^3/\ln K$, indicates that there exists a range of values of K, L, and δ for which the solution of the problem would be impossible by the elimination method, but feasible by use of the expressions for the Green's functions derived in this paper. There may well exist, however, other special numerical inversion methods which take advantage of the sparseness or the special form of the matrix D, and which also yield a smaller error.

To appreciate the practical importance of errors of magnitude, $E(G)$, it is necessary to compare them to the magnitudes of the elements of $G^0_{k,l;k',l'}$. The elements of greatest magnitude are those for which $l = l'$, $k = k'$, that is, the value of the solution at the source. An estimate for the magnitude of these elements is given by the upper bound, (5.8) for $\sum_m | T_m |$. The elements of least magnitude are those for $k = 1$, $k' = K$, and $| l - l' | = L/2$, that is, those elements at points most distant from the source. In these cases, the terms of the sum over m have alternating signs and rapidly decreasing magnitudes. The first term is then a fair estimate for the sum, giving

$$G^0_{1,l;K,l-L/2} \cong \pi/(4K^2 \sinh (L\pi/2K)). \tag{5.15}$$

By comparing (5.8), the magnitude of the larger elements, with the value for $E(G)$, (5.9) gives the result that the relative error in these elements is of order L. The relative error in the smallest elements of G^0 are of order $LK^2 \sinh (L\pi/2K)$, which may be very large if the ratio L/K is large. However, if the Green's function is applied to the solution of a problem having sources distributed throughout the domain of solution, the contribution of these very small elements will be unimportant. In our cases (a) and (b), the additional terms added to the Green's function to satisfy the boundary conditions make higher order contributions to the smaller elements which have the effect of reducing the order of their relative errors.

6. Arbitrary function as a boundary condition

The Green's function method is extended here to include those cases in which the solution must attain some arbitrary set of values on the boundary. A function, $H_{k,l;p,q}$, is defined where the indices p and q specify only boundary points. This function satisfies Laplace's difference equation within a rectangular region bounded by the lines $k = 0, K$ and $l = 0, L$, and vanishes at all boundary points but one, $k = p$, and $l = q$, where it has the value unity. That is,

$$(\Delta_k^2 + \Delta_l^2) H_{k,l;p,q} = 0 \tag{6.1}$$

for k, l on the boundary: $H_{k,l;p,q} = \delta_{kp}\, \delta_{lq}$. Then the solution of Poisson's difference equation, $\varphi_{k,l}$, is given by

$$\varphi_{k,l} = \sum_{k',l'} G_{k,l;k',l'}\, f_{k',l'} + \sum_{p,q} H_{k,l;p,q}\, \Phi_{p,q}, \tag{6.2}$$

where $G_{k,l;k',l'}$ is the discrete Green's function which vanishes on the boundaries, $f_{k',l'}$ is the inhomogeneous term of Poisson's difference equation, and $\Phi_{p,q}$ is the value of the solution desired at the boundary points p, q. The function $H_{k,l;p,q}$ may be obtained as a linear combination of the solutions $\psi_{k,l;j}$ of Laplace's difference equation which vanish on three of the boundaries. For example, for $p = K$, and q any value between 0 and L, the appropriate solutions are

$$\psi_{k,l;j} = \sin(\pi l/L)j \sinh \beta_j k, \tag{6.3}$$

where β_j is determined by

$$\cosh \beta_j = 2 - \cos(\pi/L)j.$$

The solution for $H_{k,l;p,q}$ in terms of these functions, is given by

$$H_{k,l;K,q} = \frac{2}{L} \sum_{j=1}^{L-1} \frac{\sin(\pi/L)qj \sin(\pi/L)lj \sinh \beta_j k}{\sinh \beta_j K}. \tag{6.4}$$

The solutions for the other three boundaries may be obtained by simple changes of variables in equation (6.4).

7. Computational use of the Green's function method

The direct application of the method requires the computation of at most K^2L^2 numbers which constitute the complete Green's function, and the permanent storage of these on magnetic tape. As shown in Section 5, the round-off error in this computation has been greatly reduced by performing analytically one of the sums in the expression for the Green's function. For example, the Green's functions, (4.6) and (4.14), for the case $K = 11$, $L = 100$, have been evaluated on the IBM 704. These results satisfied the difference equation for the Green's function to seven of the eight decimal digits carried by the machine. The symmetries of the boundary conditions, in general, reduce the number of distinct values required to specify the Green's function. The Green's functions (4.6) and (4.14) depend only on $l - l'$, k, k', as a result of the periodic condition (2.7), and thus only $K^2L/2$ values are required. Thus, the number of multiplications required to evaluate a complete Green's function is of the order of K^3L, since the three factors, S and the two sine functions, occurring in each term of the Green's function, may be computed initially for all terms. Therefore, the largest number of multiplications arises simply from the multiplication of these factors to form the terms of the Green's function. In the absence of the symmetries noted above, of the order of K^3L^2 multiplications would be required. Both of these results may be compared with the K^3L^3 multiplications required by the straightforward application of the matrix inversion method of von Neumann and Goldstine [4].

The straightforward substitution of the Green's function into equation (3.2) to obtain the solution of Poisson's difference equation would require K^2L^2 multiplications. It is apparent, however, that one may save almost half of

these operations by calculating the value of the solution only at alternate lattice points from the Green's function, and supplying the missing values through direct use of the difference equation (2.6). A logical extension of this procedure is the solution of a problem by a process of repeated binary division of the region. Consider an $N \times N$ lattice. The first step consists of using the Green's function for the whole region to obtain the solution along the line $k = N/2$. The original problem is now reduced to two independent problems of half the size; i.e. rectangles $N \times N/2$. The solution in each of the new regions may be obtained by the use of the Green's function appropriate to the smaller regions together with the extension described in Section 6, to take account of the arbitrary function along the boundary. This process of dividing the regions may be repeated until the solution is obtained at all desired points. An analysis of this procedure for an $N \times N$ problem, where N is an integral power of two, shows that to highest order in N, only $3N^3$ multiplications are required in contrast to the N^4 multiplications required in the straightforward application of the Green's function method. This procedure would probably be resorted to only in extremely large problems, since it has the obvious drawbacks of complicating the programming and requiring the evaluation of a Green's function for each of the successively smaller regions. However, the principle of this procedure may be advantageously applied to problems in which one desires either the solution in only a small part of the total region, or the use of different lattice spacings in different parts of the region.

Appendix I

It is required to obtain the value of the sum

$$S_n = \frac{1}{L} \sum_{j=0}^{L-1} \frac{\exp (2\pi i/L)jn}{\exp (2\pi i/L)j + \exp (-2\pi i/L)j - a}, \tag{I.1}$$

where a is a real constant and $a \geqq 2$.

This is accomplished by multiplying numerator and denominator by $\exp (2\pi i/L)j$ and decomposing into partial fractions in terms of the factors of the denominator. That is,

$$S_n = \frac{1}{L(r_+ - r_-)} \cdot \sum_{j=0}^{L-1} \exp \frac{2\pi i}{L} j(n+1) \left(\frac{1}{\exp (2\pi i/L)j - r_+} - \frac{1}{\exp (2\pi_i/L)j - r_-} \right), \tag{I.2}$$

where $r_\pm = \frac{1}{2}(a \pm (a^2 - 4)^{1/2})$. Since $r_+ > |\exp (2\pi i/L)j|$, each partial fraction can be expanded in powers of $\{\exp (2\pi i/L)j\}/r_+$.

Inverting the order of summations, and making use of the fact that

$$\sum_{j=0}^{L-1} \exp \frac{2\pi i}{L} js = \begin{cases} 0 & \text{if } s \neq NL, \\ L & \text{if } s = NL, \end{cases} \qquad N = 0, 1, 2 \cdots \tag{I.3}$$

one obtains

$$S_n = \frac{1}{r_+ - r_-}\left(-\frac{1}{r_+}\sum_{m=1}^{\infty} r_+^{-mL+n+1} + \frac{1}{r_-}\sum_{m=1}^{\infty} r_-^{-mL+n+1}\right). \tag{I.4}$$

By summing the geometric series and simplifying, the desired result is obtained:

LEMMA I.

$$S_n = \frac{-1}{(a^2 - 4)^{1/2}} \frac{r_+^{L/2-n} + r_+^{-L/2+n}}{r_+^{L/2} - r_+^{-L/2}}. \tag{I.5}$$

It is required to obtain the value of sum:

$$\sigma_n = \frac{1}{L}\sum_{j=1}^{L-1} \frac{\exp(2\pi i/L)jn}{\exp(2\pi i/L)j + \exp(-2\pi i/L)j - 2}. \tag{I.6}$$

The sum, σ_n, is rewritten as

$$\sigma_n = \frac{1}{L}\operatorname*{Lim}_{\varepsilon\to 0}\left(\sum_{j=0}^{L-1} \frac{\exp(2\pi i/L)jn}{\exp(2\pi i/L)j + \exp(-2\pi i/L)j - 2 - \varepsilon} + \frac{1}{\varepsilon}\right). \tag{I.7}$$

The first term in brackets is just S_n, with $a = 2 + \varepsilon$, and may be evaluated as above. The limit is then taken, with the result:

LEMMA II.

$$\sigma_n = -\frac{(L-1)(L+1)}{12L} + \frac{n(L-1)}{2} - \frac{n(n-1)}{2}. \tag{I.8}$$

Appendix II

It is desired to derive an upper bound for the relative error from round-off in the computation of

$$T_m = \frac{2}{K}\frac{1}{(a_m^2 - 4)^{1/2}} \frac{r_m^{L/2-n} + r_m^{-L/2+n}}{r_m^{L/2} - r_m^{-L/2}} \sin\frac{\pi m}{2K}(2k-1)\sin\frac{\pi m}{2K}(2k'-1), \tag{II.1}$$

where

$$a_m = 4 - 2\cos(\pi m/K) = 2 + 4\sin^2(\pi m/2K),$$
$$r_m = \tfrac{1}{2}(a_m + (a_m^2 - 4)^{1/2}).$$

The derivation is done according to the assumptions of Section 5. The relative errors of each of the factors of T_m are derived and then summed to yield the relative error in T_m.

(1) To compute the factors of the form $\sin(\pi I/2K)$, where I is an integer, the integer J is found exactly, such that

$$-K \leqq J \leqq K \quad \text{and} \quad \sin(\pi J/2K) = \sin(\pi I/2K).$$

From the arithmetic operations, the argument, $\pi J/2K$, has relative error 3δ.

Using the relation (5.1b), the sine then has relative error:

$$R\left(\sin\frac{\pi I}{2K}\right) = \left|\frac{(\pi J/2k)\cos(\pi I/2k)}{\sin(\pi I/2k)}\right| 3\delta + \text{Error in evaluation of sine} \tag{II.2}$$
$$= 3\delta + 5\delta = 8\delta,$$

where it is assumed that the numerical evaluation of the sine introduces a relative error of 5δ.

(2) To compute the factor $(a_m^2 - 4)^{-1/2}$, the expression for $a_m^2 - 4$ is put in the form

$$a_m^2 - 4 = 16\sin^2(\pi m/2K)(1 + \sin^2(\pi m/2K)). \tag{II.3}$$

By applying (II.2),

$$R(\sin^2(\pi m/2K)) = 17\delta,$$

and thus

$$R(a_m^2 - 4) = 29\delta,$$

and

$$R\big((a_m^2 - 4)^{-1/2}\big) = \tfrac{1}{2}29\delta + \delta + \delta < 17\delta, \tag{II.4}$$

where the evaluation of the square root contributes a relative error δ.

(3) To compute the factor, $(r_m^{L/2-n} + r_m^{-L/2+n})/(r_m^{L/2} - r_m^{-L/2})$, a_m is computed with relative error

$$R(a_m) = 19\delta,$$

and thus

$$R(r_m) \leqq 21\delta.$$

The number $r_m^{L/2}$ may be formed by at most $2\ln_2(L/2)$ multiplications, and thus

$$R(r_m^{L/2}) = (L/2)R(r_m) + 2\ln_2(L/2)\delta \leqq 11L\delta. \tag{II.5}$$

From the expressions for $a_m^2 - 4$ it follows that

$$(a_m^2 - 4)^{1/2} > 4\sin(\pi m/2K) > 4/K \quad \text{for } 1 \leqq m \leqq K,$$

and hence that

$$r_m > 1 + 2/K. \tag{II.6}$$

For $L \geqq K \geqq 1$,

$$r_m^{L/2} > 1 + L/K \geqq 2; \tag{II.7}$$

therefore

$$R\left(\frac{r_m^{L/2-n} + r_m^{-L/2+n}}{r_m^{L/2} - r_m^{-L/2}}\right) \leqq R\left(\frac{r_m^{L/2} + r_m^{-L/2}}{r_m^{L/2} - r_m^{-L/2}}\right)$$
$$\leqq \frac{4r_m^L}{r_m^{-L/2} - 1} R(r_m^{L/2}) + 6\delta \leqq \frac{16}{15} R(r_m^{L/2}) + 6\delta \leqq (12L + 6)\delta, \tag{II.8}$$

where we have used relation (5.1b). The quantity 6δ occurring in equation

(II.8) is generated by the elementary arithmetic operations required to form the factor $(r_m^{L/2} + r_m^{-L/2})/(r_m^{L/2} - r_m^{L/2})$ from $r_m^{L/2}$. The subtraction to form the denominator introduces a relative error less than 2δ because of relation (II.7).

Summing the relative errors of the factors, and adding 3δ for the multiplications to form T_m, yield

$$R_T = (12L + 42)\delta. \tag{II.9}$$

References

1. R. Courant, K. Friedrichs und H. Lewy, *Über die partiellen Differenzengleichungen der mathematischen Physik*, Math. Ann., vol. 100 (1928), pp. 32–74.
   J. L. Walsh and David Young, *On the degree of convergence of solutions of the difference equations to the solution of the Dirichlet problem*, J. Math. Phys., vol. 33 (1954), pp. 80–93.
2. David Young, *Iterative methods for solving partial difference equations of elliptic type*, Trans. Amer. Math. Soc., vol. 76 (1954), pp. 92–111.
3. M. A. Hyman, *On the numerical solution of partial differential equations*, Thesis, Technische Wetenschap aan de Technische Hogeschool te Delft.
4. John von Neumann and H. H. Goldstine, *Numerical inverting of matrices of high order*, Bull. Amer. Math. Soc., vol. 53 (1947), pp. 1021–1099.

IBM Research Center
Poughkeepsie, New York