NON-ITERATIVE NUMERICAL SOLUTION OF BOUNDARY-VALUE PROBLEMS *)

by MORTON A. HYMAN

Fulbright Scholar, Laboratorium voor Aero- en Hydrodynamica der Technische Hogeschool, Delft

Summary

A method is presented for the numerical solution of boundary-value problems. The method is applicable to arbitrary regions in any number of dimensions. The technique of solution is non-iterative and appears well-adapted to use with high-speed computing machines.

§ 1. *Introduction.* The following problem is of great importance in both pure and applied mathematics. Given a closed region R (in two or more dimensions, simply or multiply-connected) and an elliptic partial differential equation $E(u; x, y, z, \ldots.)$ where u is the dependent and $x, y, z, \ldots.$ the independent variables (which need not be cartesian). One supposes that u (and/or several of its derivatives) are given over the closed boundary of region R and determine a unique solution of equation E throughout R. The problem is to find this solution and, from the standpoint of the applications, find it in a practical manner. Indeed, if a particular boundary distribution be given, one would like to find the *numerical value* of the solution, usually at a very large number of points inside the region. Such a large volume of numbers calls strongly for the use of a high-speed computing machine; in devising a technique of solution one should bear in mind the pecularities of such machines, which differ considerably in their abilities and methods of operation from human computers.

§ 2. *General remarks.* There are two widely used methods of attacking such a boundary-value problem as that described above:

*) Research supported by the U.S. Naval Ordnance Laboratory and the Office of Naval Research. Presented to the American Mathematical Society in April, 1950.

analytical and numerical. For a certain small (but non-trivial) class of regions and equations, it is possible to give an explicit formula for the "exact solution" of the boundary-value problem. Such a formula may be quite complicated, involving special functions and infinite processes, and therefore very difficult to reduce to numbers when the solution is required at many points. From a computational point of view, the use of an explicit formula does have two advantages: the solution can be obtained with any desired precision, and no lengthy, large-scale iteration is required. For the wide class of problems where an explicit formula for the solution cannot be obtained or where such a formula is too formidable, one attacks numerically. The usual procedure is to cover the region R with a grid of points and replace the differential equation by a difference equation. Insisting that this difference equation must be satisfied at every grid point interior to R, one obtains a high-order system of simultaneous algebraic equations (linear if the original differential equation is linear) which nowadays are usually solved by successive approximations [1] [2] [3]. To program (i.e., plan) such a large-scale iterative calculation on a computing machine is a considerable task. Moreover, the machine must remember many numbers and, unless convergence is rapid, calculating time will be an important factor. The difficulties of complicated programming and large memory requirements are also present when a computing machine is used to evaluate an involved analytical formula. Besides its generality (it can be applied to any equation over any region) the numerical method has one great advantage: it usually asks the computer, human or machine, to carry out arithmetic operations of only the very simplest sort. This paper will describe a procedure for solving boundary-value problems which appears to be of rather wide applicability, and to meet the practical requirements sketched in the first paragraph. This procedure is non-iterative, asks the computing machine to carry out only simple arithmetic operations, and does not require complicated programming or large memory capacity.

§ 3. *Laplace's equation over a rectangle.* We shall illustrate the suggested procedure by first considering a rather special case: the solution of Laplace's equation over a rectangle. Extensions (to more than two dimensions, to more general equations, and to more general regions) will be discussed later.

We shall always assume that our regions and boundary distributions are sufficiently non-pathological so that a unique solution of the differential boundary-value problem exists.

§ 4. Consider a rectangle with sides L and D. We cover it with a rectangular grid: $L = (N + 1)\,\Delta x$, $D = (M + 1)\,\Delta y$ (M and N are integers; Δx and Δy are independent.

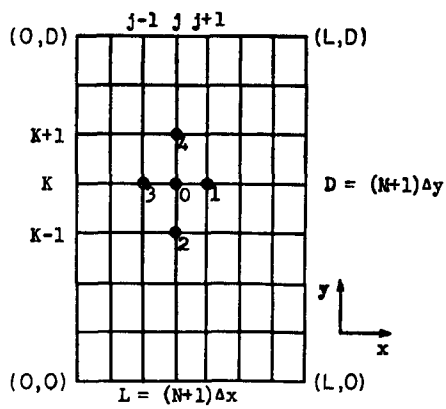

Fig. 1. Computation over a rectangular domain.

$u(x, y)$ must satisfy the difference equation

$$\frac{u(x + \Delta x, y) - 2u(x, y) + u(x - \Delta x, y)}{\Delta x^2} + \frac{u(x, y + \Delta y) - 2u(x, y) + u(x, y - \Delta y)}{\Delta y^2} = 0$$

or

$$\frac{u_4 - 2u_0 + u_2}{\Delta y^2} + \frac{u_1 - 2u_0 + u_3}{\Delta x^2} = 0, \tag{1}$$

where we have replaced Laplace's equation,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

by its simplest possible difference representation.

$u_{jk} \equiv u(j\,\Delta x, k\,\Delta y)$ must satisfy equation (1) at every interior point ($j = 1, 2, \ldots, N$; $k = 1, 2, \ldots, M$), and take on prescribed values (those for the differential problem) at all boundary points.

Various authors [4] [5] [6] [7] have shown that the solution u of this *difference* problem will converge to the solution u^* of the corresponding *differential* problem as $\Delta x, \Delta y \to 0$ independently. Therefore, by using a sufficiently fine grid, we can make the truncation error $(u^* - u)$ entirely negligible. Moreover, changing the boundary distribution at a finite number of points does not change the solution approached in the limit at interior points; we can therefore set $u = 0$ at the four corners of the rectangle. Equation (1) does not call for any of the four corner values of u; however we might want to use a less simple difference representation for Laplace's equation which would yield an equally small truncation error with a coarser grid; the more involved difference equation might call for the corner values.

§ 5. The solution of (1) which vanishes at the four corners but which is otherwise arbitrary at the boundary grid points may be written as the sum of two solutions, each of which vanishes on two sides and takes the prescribed boundary values on the other two sides. Thus we need only seek a solution of (1) which satisfies the boundary conditions

$$u(0, k\,\Delta y) = u(L, k\,\Delta y) = 0, \tag{2.1}$$

$$u(j\,\Delta x, 0) = F_j,\ u(j\,\Delta x, 0) = G_j, \tag{2.2}$$

$$(j = 1, 2, \ldots, N;\ k = 0, 1, 2, \ldots; M + 1).$$

All such solutions coincide at points of the grid. By direct trial (or more systematically, by separation of variables) we verify that equation (1) has solutions of the form

$$u(x, y) = (\sin \beta x)\, \lambda_1^{\frac{y}{\Delta y}} \quad \text{and} \quad (\sin \beta x)\, \lambda_2^{\frac{y}{\Delta y}},$$

where β is any real number,

$$\lambda_1 = \frac{1}{\lambda_2} = \mu + \sqrt{\mu^2 - 1}, \tag{2.3}$$

$$\mu = 1 + 2r^2 \sin^2 \frac{\beta \Delta x}{2}, \quad r \equiv \frac{\Delta y}{\Delta x}.$$

From (2.1) we get

$$\beta = \frac{n\pi}{L}, \text{ where } n = 0, 1, 2, 3, \ldots.$$

Now in the sum

$$\sum_{n=1}^{N^*} \sin \frac{n\pi x}{L}\left[P_n\lambda_{1n}^{\frac{y}{\Delta y}} + Q_n\lambda_{2n}^{\frac{y}{\Delta y}}\right],\quad N^* \geqslant N,$$

we can choose the $2N^*$ constants P_n and Q_n so as to satisfy the $2N$ conditions (2.2). Since solutions with different values of N^* have the same values at all points of the grid, we choose the simplest expression, setting $N^* = N$.

§ 6. Using a notation previously explained, we write our solution of (1) in the form

$$u_{jk} = \sum_{n=1}^{N} \sin \frac{jn\pi}{N+1}\,[P_n\lambda_{1n}^k + Q_n\lambda_{2n}^k]. \tag{3}$$

This is a matrix equation:

$$U = C\Phi, \tag{4}$$

where

$$U = \| u_{jk} \|,$$

$$C = \| c_{jn} \| = \left\| \sin \frac{jn\pi}{N+1} \right\|,$$

$$\Phi = \| \varphi_{nk} \| = \| P_n\lambda_{1n}^k + Q_n\lambda_{2n}^k \|,$$

$$(j, n = 1, 2, \ldots, N;\ k = 0, 1, 2, \ldots, M+1).$$

The symmetric matrix C has the delightful property of being, to within a constant, its own inverse:

$$C^{-1} = \frac{2}{N+1}\,C \quad \text{or} \quad \sum_{s=1}^{N} C_{rs}C_{st} = \begin{cases} 0 & r \neq t, \\ \dfrac{N+1}{2} & r = t, \end{cases}$$

as is well-known (see Willers [14]), p. 335). Equation (4) is then easily solved:

$$\Phi = C^{-1}U = \frac{2}{N+1}\,CU, \tag{4'}$$

or

$$P_n\lambda_{1n}^k + Q_n\lambda_{2n}^k = \frac{2}{N+1}\sum_{j=1}^{N} u_{jk} \sin \frac{jn\pi}{N+1}.$$

Now, from (2.2), $\quad u_{j0} = F_j, \quad u_{j\,M+1} = G_j.$

It follows that

$$P_n + Q_n = \frac{2}{N+1} \sum_{j=1}^{N} F_j \sin \frac{jn\pi}{N+1} = A_n, \tag{5}$$

$$P_n \lambda_{1n}^{M+1} + Q_n \lambda_{2n}^{M+1} = \frac{2}{N+1} \sum_{j=1}^{N} G_j \sin \frac{jn\pi}{N+1} = B_n,$$

$$(n = 1, 2, 3, \ldots, N)$$

where A_n, B_n are the respective coefficients of the n'th harmonic in the finite Fourier expansions of F_j and G_j. Equations (5) are easily solved to give

$$P_n = -\frac{A_n \lambda_{2n}^{M+1} - B_n}{\lambda_{1n}^{M+1} - \lambda_{2n}^{M+1}}, \qquad Q_n = \frac{A_n \lambda_{1n}^{M+1} - B_n}{\lambda_{1n}^{M+1} - \lambda_{2n}^{M+1}}. \tag{6}$$

If M is large and A_n and B_n are of about the same order of magnitude (the conditions which usually hold), then to a good approximation,

$$P_n \simeq \frac{B_n}{\lambda_{1n}^{M+1}} = B_n \lambda_{2n}^{M+1}, \qquad Q_n \simeq A_n. \tag{6'}$$

The matrix Φ is now completely known and the solution of equation (1) subject to boundary conditions (2) is given by (4) or (more explicitly) by (3). We have replaced the inversion of a matrix of order MN (the number of interior points at which (1) must be satisfied) by the trivial inversion of N matrices of order 2. An essential step was the replacement of the set of base vesters $[\delta_n]$ by the set $[c_n]$, where δ_n and c_n are the n'th columns, respectively, of the unit N'th order matrix and of the matrix C. Both basic sets are orthogonal, but the set $[c_n]$ has the additional advantage that coordinates with respect to it are uncoupled as we move forward from line to line ($k = 0, 1, 2, \ldots, M + 1$) in equation (1). In other words, the set of vectors $[c_n]$ are eigenvectors of the operator which carries line k into line $k + 1$. Previous investigators (see Rosenbloom [7]) have used the set of base vectors $[\delta_n]$ and have been led to the inversion of matrices of order N rather than order 2, as here. We remark finally that the reciprocal character of the matrix formulas

$$F = CA, \qquad A = \frac{2}{N+1} CF$$

is very convenient in checking numerial calculations.

§ 7. We now observe that it is unnecessary to use formulas (3) or (4) to calculate the solution u_{jk} at each of the MN interior grid points. We may write equation (1) in the form

$$u_4 = (2u_0 - u_2) - r^2(u_1 - 2u_0 + u_3), \qquad r \equiv \frac{\Delta y}{\Delta x}, \tag{7}$$

and it is clear from inspection of fig. 1 that the solution of any line $(k + 1)$ can be expressed very simply in terms of the solution on the two preceding lines $(k, k - 1)$. u_{j0} is given; from formula (3), we may compute u_{j1}, and then using formula (7) "step-ahead" the solution in the remainder of the region. This "stepping-ahead" process (or "marching", as it was picturesquely termed by Richardson [9]) has been effectively used for many years in the machine solution of initial-value problems (where the governing differential equation is hyperbolic or parabolic). The process has apparently never been applied to boundary-value problems for two reasons:

(*a*) the absence of a technique for transferring data readily from remote boundaries to the initial surface;

(*b*) the essential *instability* of a stepping-ahead process when applied to an equation which calls naturally for data given on a closed boundary.

However, the idea of "stepping-off" the solution of an elliptic equation is probably old, though scarcely mentioned in the literature; see Runge [15]).

§ 8. *Error growth.* By the instability of a stepping-ahead process, we mean that an error (either a computer's blunder or a round-off error) introduced at any step is magnified in subsequent steps. Such matters are discussed by setting-up the variational equation, easily obtained from the equation governing the stepping-ahead process [10]). If the governing equation is linear and homogeneous, the variational equation will have the same form, and the error will have the same law of growth as the general error-free solution. Thus if our governing equation be (1) or (7), the errors will also satisfy this equation and their growth will be given by formula (3). We see that the n'th harmonic of the error consists of two parts

$$\left(P_n \sin \frac{n\pi x}{L}\right) \lambda_{1n}^k \text{ and } \left(Q_n \sin \frac{n\pi x}{L}\right) \lambda_{2n}^k;$$

these parts are multiplied respectively by

$$\lambda_{1n} \text{ and } \lambda_{2n}$$

every time we move forward one line. Here

$$\lambda_{1n} = \frac{1}{\lambda_{2n}} = \mu_n + \sqrt{\mu_n^2 - 1},$$

$$\mu_n = 1 + 2r^2 \sin^2\left(\frac{n}{N+1}\frac{\pi}{2}\right),$$

$$\left(r \equiv \frac{\Delta y}{\Delta x} > 0, \quad n = 1, 2, \ldots, N\right).$$

Clearly $\lambda_{1n} > 1$ for all n and thus an error of any frequency fed into the calculation must in general grow. An upper limit on λ_{1n} which is closely approached for large N is

$$\lambda^* = (1 + 2r^2) + \sqrt{(1 + 2r^2)^2 - 1}.$$

For

$$r = 0.1, \quad \lambda^* = 1.2210,$$
$$r = 1.0, \quad \lambda^* = 5.8284,$$
$$r = 10.0, \quad \lambda^* = 401.9975.$$

1 step at $r = 10$ equals 10 steps at $r = 1.0$ and 100 steps at $r = 0.1$. However

$r = 0.1$:	$(1.2210)^{10} = 7.3653$	$(1.2210)^{100} = 4.7\ (10^8)$
$r = 1.0$:	$(5.8284)^1 = 5.8284$	$(5.8284)^{10} = 4.5\ (10^7)$
$r = 10.0$:		$(402.00)^1 = 4.0\ (10^2)$

and it is clear that in order to go a given distance in the y-direction, it is better from the standpoint of error growth to use a big rather than a small r. However, the most useful difference solutions are those in which Δx and Δy are equal or nearly so throughout any subregion, i.e., r is 1 or very close to it. Moreover, the convergence of the solution of the difference problem to that of the differential problem is presumably chiefly dependent on the magnitude of the larger of the two increments Δx, Δy; hence, from the standpoint of convergence, very little is gained by using a Δy which is appreciably larger or smaller than Δx. We add one final remark: the general solution of (1) over a rectangle requires the superposition of two special solutions which step-ahead respectively in the x- and y-directions and hence have reciprocal r's (see § 5). Therefore, if we intend to step-off these solutions separately and afterwards add them (rather than vice-

versa, which leads to only one stepping-ahead traverse of the region) we should clearly use an r close to 1 to cut down error growth. For all these reasons, then, we believe that $\Delta y/\Delta x$ should be as close to 1 as is permitted by the dimensions of the rectangle without using too fine a grid.

§ 9. Choosing $r = 1$, then, we calculate the maximum possible growth of a single error; since

$$(5.8284)^{13.06} = 10^{10},$$

we see that in stepping-ahead thirteen lines, we can corrupt as many as ten decimal places ahead of the place where the error enters. It should be remarked that the actual error present at any point is the sum of many errors of different amplitudes, frequencies, and growth-rates, so that the description of growth in the preceding section is in general too dire, based as it is on only a single error of the fastest growing frequency possible. We reserve for another paper a detailed discussion of single and cumulative errors in step-wise processes; such an analysis is facilitated by expansion of each line of errors in terms of eigenvectors of the stepping-ahead operator (see § 6). In the present case of elliptic boundary-value problems the cumulative error appears to grow rapidly, in a manner indicated by the simple calculation above. However, it is very important to note that *the true (error-free) difference solution usually grows as rapidly as the error*, since both are given by formula (3). Thus the tendency in general will be for the *relative error* to become and remain constant as we step-ahead; the number of correct significant figures after many steps should differ by perhaps one or two from the number of correct significant figures with which one *begins* (see § 11). This is clearly seen from Table I: considering the first two lines ($k = 0, 1$) to be *exact*, nineteen more lines were stepped-off using formula (7) with $r = 1$, and retaining systematically eight significant figures. The exact values, calculated from formula (3), are given for comparison; the relative error is in all cases of the order $10^{-5}\%$, which is exactly what would be expected from a priori comparison of $\bar{P}_1$, $\bar{P}_2$, $\bar{P}_3$ (see §§ 11, 13). Noting that the number of correct significant figures remains nearly constant, while the decimal point migrates greatly, it seems desirable (but of course, not essential) to carry out the stepping-ahead process using a *floating-decimal-point* calculating machine.

TABLE I

$k \backslash j$	0	1	2	3	4
0	0.000 0000	2.000 0000	3.000 0000	4.000 0000	0.000 0000
	0.000 0000	2.000 0000	3.000 0000	4.000 0000	0.000 0000
1	0.000 0000	6.785 7160	1.001 7859(1)	7.785 7160	0.000 0000
	0.000 0000	6.785 7160	1.001 7859(1)	7.785 7160	0.000 0000
5	0.000 0000	1.177 8591(2)	1.982 6769(2)	2.147 8591(2)	0.000 0000
	0.000 0000	1.177 8591(2)	1.982 6769(2)	2.147 8591(2)	0.000 0000
10	0.000 0000	2.202 3728(4)	—6.257 0348(4)	9.224 9728(4)	0.000 0000
	0.000 0000	2.202 3735(4)	—6.257 0357(4)	9.224 9735(4)	0.000 0000
15	0.000 0000	1.718 2771(8)	—2.781 8599(8)	2.226 7123(8)	0.000 0000
	0.000 0000	1.718 2773(8)	—2.781 8602(8)	2.226 7126(8)	0.000 0000
20	0.000 0000	7.470 4734(11)	—1.082 4813(12)	7.838 5798(11)	0.000 0000
	0.000 0000	7.470 4743(11)	—1.082 4814(12)	7.838 5806(11)	0.000 0000

Lines $k = 0$, 1 are assumed exact (correct to an infinite number of significant figures).
On top: exact solution, rounded to 8 significant figures, obtained using formula (3).
On bottom: solution obtained by stepping-ahead from lines $k = 0, 1$, keeping always 8 significant figures.
2.202 3726(3) is an abbreviation for $2.202\,3726(10^3)$.

§ 10. Thus the *numerical* solution of the elliptic difference equation (call it N) *) is "close" to the *exact* solution of the difference equation (Δ) *), in the sense that they agree to a large number of significant figures. Moreover, the proofs in [4] [5] [6] [7] show that, for a sufficiently fine grid, Δ will be arbitrarily close to D, the exact solution of the *differential* equation. Since N and D are close, we will have satisfactorily solved the differential equation by our numerical methods. Thus the usual terror inspired by *instability* does not seem justified in such elliptic problems. The fear is justified for parabolic problems, because (as pointed out in [10]) instability or error growth is often accompanied in such problems by *non-convergence* (i.e., Δ does not approach D as the mesh is made finer and finer).

§ 11. We have remarked that the number of correct significant figures is nearly preserved as we step-ahead. This refers to *the number of correct significant figures in the amplitudes of the various frequencies*. We now shall see how such significant figures can be lost at the beginning of stepping-ahead, and how this loss may be avoided. We have seen that each frequency β gives rise to two waves,

$$P\lambda^k \sin \beta x \quad \text{and} \quad Q\lambda^{-k} \sin \beta x,$$

*) We use the notation (N, Δ, D) introduced in [10].

which respectively increase and decrease exponentially as y (or k) increases. P and Q are so chosen that

$$P + Q = A$$

and

$$P\lambda^{M+1} + Q\lambda^{-(M+1)} = B,$$

where A, B are the amplitudes of the harmonic β on the first and last lines (respectively, $k = 0$ and $k = M + 1$). If A, B are of the same order of magnitude (this is usually the case, or can be managed, as we shall see) and M is large, both P and $P\lambda$ will be very small in comparison with Q and $Q\lambda^{-1}$. Let S be the amplitude on the second line,

$$P\lambda + Q\lambda^{-1} = S,$$

calculated from the P and Q just found. In stepping-ahead, we tacitly assume the two "starting" lines (here the first and second lines) to be exact, while actually they are accurate only to a certain number of significant figures. That is, our process *implicitly* solves the equations

$$\overline{P} + \overline{Q} = A$$
$$\overline{P}\lambda + \overline{Q}\lambda^{-1} = S$$

for $\overline{P}$ and $\overline{Q}$, and uses $\overline{P}$ and $\overline{Q}$ for stepping-ahead. But it is clear from these equations and the remark that the first term in each case is very small compared to the second, that while $\overline{Q}$ will have about as many correct significant figures as A or S, $\overline{P}$ will have considerably fewer correct significant figures. By "correct" significant figures of an approximate number we mean simply the number of significant figures in which that number, considered as an infinite decimal, agrees with the exact number it is approximating.

At any line k the value obtained by stepping-ahead is

$$\overline{P}\lambda^k + \overline{Q}\lambda^{-k},$$

plus accumulated round-off errors which, as we have seen, do not much affect the number of correct significant figures. For k small, $\overline{P}\lambda^k$ contributes negligibly to the sum, but for k large, this term is dominant and

$$\overline{P}\lambda^k + \overline{Q}\lambda^{-k}$$

will agree with the exact amplitude

$$P\lambda^k + Q\lambda^{-k}$$

only to as many significant figures as $\overline{P}$ agrees with P. Thus the error which may rise to swamp us is not the growth of round-off error, but the inaccuracy implicit in our initial data, which inaccuracy makes its presence felt ever more strongly as we step-ahead.

§ 12. To minimize this inaccuracy, we chose the two lines from which we step-ahead as the lines where $\overline{P}\lambda^k$ and $\overline{Q}\lambda^{-k}$ are of comparable magnitude (are "balanced"); $\overline{P}$ and $\overline{Q}$, evaluated at the "balancing lines", will each have about as many correct significant figures as we wish to employ in our stepping-ahead. If the values prescribed on opposite sides of the rectangle are of the same order of magnitude, we choose two lines at the middle of the region and step-off in both directions. If there are a great many lines between opposite sides, it will be desirable (see also § 13) to insert *several* pairs of very accurate "starting lines" (using formula (3)) and step-off the entire region piecewise. Beginning at or near the middle, one would step-ahead perhaps ten or twenty lines, then compute two more "starting lines", step-ahead again, compute two starting lines, step-ahead, etc. The stepping-ahead process can be checked and the error growth

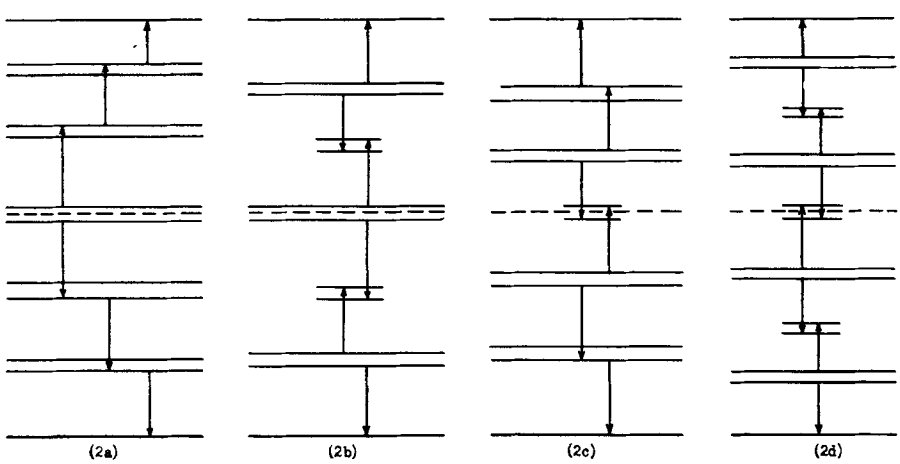

Fig. 2. Alternative stepping-ahead routines.

"starting" lines, computed from formula (3).
adjacent lines where steppings-ahead overlap.
balancing position.

It is not intended to imply that the balancing position need be in the center or that the number of lines between opposite sides ($k = 0, 1, 2, \ldots, M + 1$) is the same in all four figures.

monitored by making the last two lines computed by stepping-ahead coincide with the two starting lines which begin the next stage of calculation; this also checks the evaluation of formula (3). The pattern of calculation just described is indicated schematically in Figure 2*a*; three other patterns which possess some advantage are given in Figures 2*b*, 2*c*, 2*d*. We remark that the balancing position will be usually (but not always) near the middle of the region; moreover the distance between pairs of starting lines need not (and if we wish to minimize the number of starting lines while keeping the tolerable error the same should not) be constant.

§ 13. It should be observed that, whereas by starting in the middle we have balanced the amplitudes $\overline{P}\lambda_1^k$, $\overline{Q}\lambda_1^{-k}$ of the two waves generated by each frequency, the amplitudes of different frequencies $\overline{P}_1\lambda_{11}^k$, $\overline{P}_2\lambda_{12}^k$, $\overline{P}_3\lambda_{13}^k$, ..., and $\overline{Q}_1\lambda_{11}^{-k}$, $\overline{Q}_2\lambda_{12}^{-k}$, $\overline{Q}_3\lambda_{13}^{-k}$, ..., are not of the same order of magnitude at the middle. We conclude, by an argument similar to that used before, that $\overline{P}_1$, $\overline{P}_2$, $\overline{P}_3$, ... will differ in their number of correct significant figures, as will $\overline{Q}_1$, $\overline{Q}_2$, $\overline{Q}_3$, ... However, the resulting loss as we step ahead does not appear to approach in magnitude the loss we have just avoided. We may diminish (often considerably) the present loss by the device of subtracting out at the very beginning the constant part of the solution — i.e., the part obtained by averaging the boundary values. This device serves the double purpose of

(*a*) diminishing the amplitudes of the lower frequencies which decrease at the smallest exponential rates in passing to the middle of the region;

(*b*) making the average of boundary values on opposite faces approximately the same, thus in general justifying our use of the middle for balancing.

By subtracting out the constant part of the solution, we are able to obtain more accurately the variable part, a matter of considerable importance if we intend to differentiate (via differencing) the numerical solution.

§ 14. We present three calculations which illustrate and support the foregoing remarks. Table II shows the results of applying the process to a 5 × 5 square, stepping-off in turn from the left and top sides, and then adding; the calculation was performed twice, car-

TABLE II

$k \backslash j$	0	1	2	3	4
0	0.000 0000	1.000 0000	1.000 0000	1.000 0000	0.000 0000
	0.000 0000	1.000 0000	1.000 0000	1.000 0000	0.000 0000
1	2.000 0000	2.000 0000	8.000 000	26.000 001	62.000 009
	2.000 0000	2.000 0000	8.000 0001	26.000 0006	62.000 0032
2	3.000 0000	— 3.000 000	2.999 999	32.999 996	98.999 983
	3.000 0000	— 3.000 0001	2.999 9998	32.999 9991	98.999 9952
3	4.000 0000	—20.000 000	—26.000 000	4.000 001	88.000 009
	4.000 0000	—20.000 0002	—25.999 9999	4.000 0008	88.000 0032
4	0.000 0000	—54.999 999	—91.000 000	—79.000 001	0.000 0000
	0.000 0000	—55.000 0008	—91.000 0000	—78.999 9992	0.000 0000

On top: numerical solution, carrying eight significant figures (no guard figures).
On bottom: numerical solution, carrying seven decimal places.

TABLE III

$k \backslash j$	0	1	2	3	4
0	0.0000 0000	6.0000 0297(4)	—9.0000 0363(4)	4.0000 0216(4)	0.0000 0000
1	0.0000 0000	1.0330 3168(4)	—1.9951 4298(4)	4.9713 3080(3)	0.0000 0000
7	0.0000 0000	—3.5218 3628(1)	—5.2725 5056(1)	—3.7202 4376(1)	0.0000 0000
8	0.0000 0000	—1.7696 1667(1)	—2.5694 5728(1)	—1.8230 3406(1)	0.0000 0000
9	0.0000 0000	—9.8717 3115	—1.4126 2782(1)	—1.0024 3523(1)	0.0000 0000
10	0.0000 0000	—7.6644 7973	—1.0914 4566(1)	—7.7407 9027	0.0000 0000
11	0.0000 0000	—9.8717 3117	—1.4126 2782(1)	—1.0024 3522(1)	0.0000 0000

Boundary distribution: at $j = 0, 4$: identically zero
at $k = 0,20$: exactly $6(10^4)$, $-9(10^4)$, $4(10^4)$.

Line $k = 10, 11$ were calculated correct to nine significant figures, and the solution stepped off, carrying one guard figure (i.e., ten columns of a digital decimal calculator were used). In the results given above, the tenth column has been rounded away.

rying, respectively, eight significant figures and seven decimal places; the exact solution may be obtained by rounding to the nearest integer the numbers shown; the boundary distribution imposed was not averaged. Tables III and IV show the same potential problem solved respectively by stepping-off from the middle and from one side. It is supposed that a ten-place machine with floating decimal point is being used; we round the starting data to nine significant figures and use the tenth column for a guard figure. Starting from the middle (lines $k = 10, 11$) it is seen that we lose about 2.3 significant figures in advancing ten steps (Table III). A simple calculation shows that by assuming lines 10 and 11 to be exact

$$\overline{P}_1 = \overline{Q}_1, \quad \overline{P}_2 = \overline{Q}_2, \quad \overline{P}_3 = \overline{Q}_3$$

have respectively 8.9, 7.4, 6.8 correct significant figures. Starting from one side and stepping-ahead (Table IV), we see that after eight steps almost all significance is lost and the numbers computed are useless. In general, there is a rapid loss of correct significant figures from both the left and right of computed numbers, as we move from the boundary toward the center of the region. The loss on the left is by cancellation, and the loss on the right is caused by round-off error feed-in and build-up. In the example summarized in Table IV, however, no round-off errors are fed-in by the stepping-ahead calculations, because the numerical work calls only for additions, subtractions, and multiplications by the integer 4, with no overflow of the registers necessitating round-off on the right.

TABLE IV

$k \backslash j$	0	1	2	3	4
0	0.0000 0000	6.0000 0000(4)	−9.0000 0000(4)	4.0000 0000(4)	0.0000 0000
1	0.0000 0000	1.0330 3108(4)	−1.9951 4229(4)	4.9713 2697(3)	0.0000 0000
7	0.0000 0000	−3.53812 (1)	−5.25341 (1)	−3.73103 (1)	0.0000 0000
8	0.0000 0000	−1.85057 (1)	−2.46947 (1)	−1.88348 (1)	0.0000 0000
9	0.0000 0000	−1.39468 (1)	−8.9043	−1.33342 (1)	0.0000 0000
10	0.0000 0000	−2.83772 (1)	+1.63583 (1)	−2.55975 (1)	0.0000 0000
11	0.0000 0000	−1.159202 (2)	+1.283122 (2)	−1.054140 (2)	0.0000 0000

Boundary distribution: at $j = 0, 4$: identically zero.
at $k = 0, 20$: exactly $6(10^4), -9(10^4), 4(10^4)$.

Lines $k = 0, 1$ were calculated correct to nine significant figures, and the solution stepped-off, carrying one guard figure (i.e., ten columns of a digital decimal calculator were used). In the results given above, the tenth column has been rounded away. Notice how very rapidly correct significant figures are lost on the left by cancellation.

§ 15. *Poisson's equation over a rectangle.* So far we have been discussing the solution of Laplace's equation. But our techniques are easily applied to Poisson's equation:

$$u_{xx} + u_{yy} = F(x, y). \tag{8}$$

As before, we replace the differential equation by a difference equation (see fig. 1)

$$\frac{u_4 - 2u_0 + u_2}{\Delta y^2} + \frac{u_1 - 2u_0 + u_3}{\Delta x^2} = G_0, \tag{9}$$

or

$$u_4 = (2u_0 - u_2) - r^2(u_1 - 2u_0 + u_3) + G_0 \Delta y^2 \tag{10}$$

where $r \equiv \Delta y/\Delta x$ as before, and $G_0 \equiv G(x_0\, y_0)$.

Equation (9) must hold at every interior point of the mesh ($j = 1, 2, \ldots, N$; $k = 1, 2, \ldots, M$). Let G_{jk} denote $G(j\Delta x, k\Delta y)$; then we easily obtain the double sine expansion

$$G_{jk} = \sum_{n=1}^{N} \sum_{m=1}^{M} \gamma_{nm} \sin \frac{\pi jn}{N+1} \sin \frac{\pi km}{M+1}, \tag{11}$$

where

$$\gamma_{nm} = \frac{4}{(M+1)(N+1)} \sum_{j=1}^{N} \sum_{k=1}^{M} G_{jk} \sin \frac{\pi jn}{N+1} \sin \frac{\pi km}{M+1}. \tag{11'}$$

As before we can put these relations in matrix form:

$$G = C\gamma D, \tag{12}$$

$$\gamma = C^{-1}GD^{-1} = \frac{4}{(M+1)(N+1)} CGD, \tag{12'}$$

where

$$G = \| G_{jk} \|, \qquad \gamma = \| \gamma_{nm} \|,$$

and the matrices

$$C = \left\| \sin \frac{\pi jn}{N+1} \right\|, \qquad D = \left\| \sin \frac{\pi Km}{M+1} \right\|$$

are used in solving the related homogeneous (Laplace's) equation. Setting

$$\omega_{nm} = -\frac{\Delta y^2}{4} \frac{\gamma_{nm}}{\left[r^2 \sin^2 \frac{\pi n}{2(N+1)} + \sin^2 \frac{\pi m}{2(M+1)} \right]} \tag{13}$$

we see, by direct substitution, that

$$W_{jk} = \sum_{n=1}^{N} \sum_{m=1}^{M} \omega_{nm} \sin \frac{\pi jn}{N+1} \sin \frac{\pi km}{M+1} \tag{14}$$

or

$$W = C\,\omega\,D \tag{14'}$$

satisfies equation (9) at all interior mesh points and vanishes identically on the boundary. We next obtain an explicit solution of Laplace's equation for the prescribed boundary values, in the manner described previously:

$$V_{jk} = I + V'_{jk} + V''_{jk},$$

where I is the constant solution, and V'_{jk} and V''_{jk} are, respectively,

the "horizontal" and "vertical" solutions for the rectangle. It is clear that

$$U_{jk} = V_{jk} + W_{jk} \tag{15}$$

is the solution we seek, because

(*a*) it satisfies equation (9) at every interior mesh point;

(*b*) it takes on the prescribed boundary values.

Our procedure assumes a certain amount of regularity for the function $G(x, y)$. If G possesses irregularities of certain sorts which are not first removed, the procedure here followed may not yield properly convergent (or sufficiently rapidly convergent) results as the mesh dimensions shrink to zero. In such cases, other methods have to be employed, somewhat more complicated than the simple methods described here.

§ 16. We now proceed as before: using the explicit solution (15), we compute several pairs of lines in the interior of the region and step-off the solution in between these lines using formula (10) and analogous expressions. Note that we may adopt any of several procedures to obtain

$$I + V'_{jk} + V''_{jk} + W_{jk}$$

which vary in the amount of work they require and their accuracy:

(*a*) $$I + (V'_{jk}) + (V''_{jk}) + (W_{jk}),$$

(*b*) $$I + (V'_{jk} + V''_{jk}) + (W_{jk}),$$

(*c*) $$I + (V'_{jk}) + (V''_{jk} + W_{jk}),$$

(*d*) $$I + (V'_{jk} + V''_{jk} + W_{jk}),$$

etc.; the quantity inside any pair of parentheses is computed by a single stepping-ahead traverse of the region, set up from the explicit formulas available for V'_{jk}, V''_{jk}, W_{jk}. Our choice of procedure is based on convenience and the stability considerations of §§ 8–13. Procedure (*a*) is probably the most accurate, usually, and procedure (*d*) the quickest. One should always be careful about combining quantities of different orders of magnitude in setting-up a stepping-ahead calculation. It seems best, in general, to add on I at the end of the calculation, particularly if it differs in order of magnitude from the other numbers entering.

§ 17. *Extensions of the technique.* Let us summarize the principal features of our technique, after which it will be clear what the possible extensions are.

(*a*) We first by some means obtain an explicit solution (ES) of the difference equation satisfying prescribed conditions at the boundary mesh points. If the equation (which may be non-homogeneous) has a simple structure and the region is simple geometrically, it will be possible to obtain (ES) by direct analogy with the techniques which are known to succeed for the corresponding *differential* boundary-value problem — such as separation of variables, generalized Fourier expansions, integral representations, etc. — each suitably modified for application to difference equations. For all those differential problems, with various equations and regions, which have been solved explicitly, we may expect to obtain an explicit solution of the corresponding difference problem. Thus we may expect to solve Laplace's and Poisson's equations not only over rectangles, but these and somewhat more complicated equations over the interior and exterior of circles, rings, ellipses, etc.; indeed, we may go to higher dimensions without difficulty, solving over parallelopipeds, spheres, ellipsoids, etc. It should be remarked that problems in more than two dimensions are almost entirely beyond the scope of present-day iteration methods (A l l e n has recently made a preliminary attack using relaxation methods [11]); from the computing machine point of view, the number storage required with more than two independent variables is enormous.

(*b*) Using the explicit solution (ES) obtained in some way, we compute sufficient "starting lines" (varying in number with the order of the difference equation) near the middle of the region (or wherever the different components of the solution are most nearly "balanced") and step-off the solution, a process which involves only simple arithmetic operations. After having taken as many steps forward as permitted by the tolerable loss of significant figures, we compute several more "starting lines" (generally the last ones obtained by stepping-ahead, which checks that operation) and step ahead again, etc. — stepping in general always from the interior of the region toward the boundaries. The distances between groups of starting lines need not all be the same; indeed, given the tolerable loss of significant figures, it seems possible to calculate optimum values for these distances. If the totality and complexity of the

"starting" calculations are not too great, it would seem possible to calculate all the starting data by some relatively slow computing device and do only stepping-ahead on a fast electronic computing machine. In describing how to find the explicit solution, we have split it into parts: the constant part, the non-homogeneous part, the several parts of the homogeneous solution; all these partial solutions can be combined and the whole region traversed only once; or, any combination of partial solutions may be stepped-off separately over the region and later combined; the decision is based on convenience and error control considerations. The process as we have described it is non-iterative. However, it may be thought desirable to employ some iterative techniques analogous to the iterative techniques used in solving ordinary differential equations by stepping-ahead methods — whose purpose is to decrease the order of the truncation error in the solution (thus allowing a larger step with the same or smaller-sized error). These techniques involve computing the next point (or the next line), correcting it several times, then proceeding to the next following point (line), etc.; they are obviously iterations of a much lower order of magnitude in time and number storage required than the techniques ordinarily referred to as iterative in the numerical solution of boundary value problems [12] [13] [14]. For parabolic and hyperbolic partial differential problems, such iterative stepping-ahead techniques have been used by Hartree and by von Neumann (for a fuller discussion, see O'Brien, Hyman, and Kaplan [10]).

§ 18. We would like to add several specific remarks to the above generalities. First of all, it seems desirable to subtract out whenever possible, a so-called "constant part" of the solution, which is usually some convenient average of the boundary values. This constant will in many cases satisfy the homogeneous equation, but, if not, the residue can often be subtracted from the right member. Thus, for example, if $u(x, y)$ satisfies

$$u_{xx} + u_{yy} + 4u = F(x, y)$$

then $v \equiv (u - I)$ satisfies

$$v_{xx} + v_{yy} + 4v = F(x, y) - 4I \equiv G(x, y).$$

Similarly, it is usually wise to subtract out any other simple parts

of the solution whose subtraction at every mesh-point of the region is easily managed. These devices are always applicable if the governing equation is linear.

§ 19. Next we remark that the equation

$$L(u) = G(x, y, z, \ldots, t), \tag{16}$$

where $L(u)$ is a general linear difference operator with constant coefficients and arbitrary order, can be solved explicitly over a rectangle, parallelopiped, etc. (we assume a certain amount of regularity for G). This statement follows from the observation that $L(u) = 0$ admits solutions of the form

$$(\sin \alpha_n x)\,(\sin \beta_m y)\,(\sin \gamma_l z)\,\ldots\,(\lambda_{nml}^{t}\ldots), \quad \text{etc.},$$

where

$$\{\alpha_n\},\ \{\beta_m\},\ \{\gamma_l\},\ \ldots$$

are, as before, finite sequences of constants of the form

$$\alpha_n = \frac{\pi n}{(N+1)\,\Delta x},\quad \beta_m = \frac{\pi m}{(M+1)\,\Delta y},\quad \gamma_l = \frac{\pi l}{(L+1)\,\Delta z},\ \ldots$$

$(n = 0, 1, \ldots, N+1;\ m = 0, 1, \ldots, M+1;\ l = 0, 1, \ldots, L+1; \ldots)$

and $\lambda_{nml}\ldots$ is a known constant depending only on $n, m, l; \ldots$. In three dimensions, for example, we would write

$$u(x, y, z) = v' + v'' + v''' + w,$$

where w is a particular solution of $L(u) = G$ and v' is that solution of $L(u) = 0$ which vanishes on all four faces of the parallelopiped which are parallel to the x-direction. Then

$$v' = \sum_{m=1}^{M}\sum_{l=1}^{L} \sin \beta_m y \sin \gamma_l z\,(P_{ml}\lambda_{1ml}^{\Delta x} + Q_{ml}\lambda_{2ml}^{\Delta x}) \tag{17}$$

and

$$w = \sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{l=1}^{L} (\sin \alpha_n x \sin \beta_m y \sin \gamma_l z)\,\omega_{nml} \tag{18}$$

where P_{ml}, Q_{ml}, ω_{nml} are constants determined without difficulty in a manner strictly analogous to the one already explained. Formulas (17) and (18) are analogous to equations (3) and (14) and can be thrown into matrix forms analogous to equations (4) and (14'). For convenience in this example, we hare implicitly assumed that $L(u)$ is a second-order difference operator and that λ_1 is in all cases distinct from λ_2.

§ 20. We would like to stress our use of the method of separated variables, which provides us with "uncoupled coordinates". As we cross from one boundary to another, the amplitudes of these coordinates change in an easily followed manner, and there is no interaction or influence on each other of the amplitudes of different components. It is this property which enables us to connect up with ease data given on separated surfaces, obtain the undetermined constants in our "explicit" solution so that it becomes truly explicit, and makes a knowledge of the exact solution available wherever in the region we want it. Moreover, the uncoupled variables will usually, because of imposed boundary conditions, be orthogonal; this property, as we have seen in equations (4), (4'), and (12), (12'), makes for great analytical elegance and in particular allows great ease in certain checking that must be done.

§ 21. In connection with non-homogeneous equations,

$$L(u) = G, \tag{19}$$

we seek relatively simple functions $\{\Phi_i\}$ such that

(*a*) inside the region, G is conveniently expressed in the form $\mu \Sigma a_i \Phi_i$ where μ is a known function;

(*b*) the solution of $L(u) = \mu a_i \Phi_i$ is easily obtained.

It will be an added convenience if the $\{\Phi_i\}$ all satisfy certain conditions at the boundaries of the region, such as vanishing. These conditions may ordinarily be satisfied by seeking functions $\{\Phi_i\}$ such that

$$L(\Phi_i) = \mu k_i \Phi_i, \tag{20}$$

where the k_i are constants, and μ is a product of known functions, each function depending on one of the independent variables. We solve equation (20) by separation of variables, obtaining

$$\Phi_i = X_i(x) Y_i(y) Z_i(z) \ldots .;$$

we endeavour to choose μ and impose such boundary conditions on X_i, Y_i, Z_i, ... as will yield a set of relatively simple functions allowing convenient expansion of a fairly general function G. It is also sometimes desirable to insist that $k_i \neq 0$; otherwise, the solution of $L(u) = A\mu\Phi_i$ will not be $u = A\Phi_i/k_i$ but $B\Phi_i v$, where v is some auxiliary function, such as a power or a logarithm. In some cases, there will be several sets of $\{\Phi_i\}$ which we may use, the sets

differing in the degree to which they possess (or do not possess) the advantages we have mentioned above. Which set we use then becomes a matter of convenience. We might mention the case of Poisson's *differential* equation in a region of circular symmetry. Here we may use as our $\{\Phi_i\}$ either

$$(a) \qquad \Phi_{mn} = [A \cos m\theta + B \sin m\theta]\,[C\varrho^n + D\varrho^{-n}]$$

or

$$(b) \quad \Phi'_{mn} = [\mathrm{A} \cos m\theta + B \sin m\theta]\,[CJ_m(n\varrho) + DY_m(n\varrho)]$$

where m is an integer $\geqslant 0$, $n > 0$.

Powers of ϱ are of course much simpler to deal with than Bessel functions; moreover, powers of ϱ appear in the solution of the homogeneous equation and it is convenient to continue their use for the particular solution of the non-homogeneous equation. On the other hand, k_i vanishes if $m = n$ when using set (a), but k_i never vanishes if we use set (b). In addition, we may readily choose the constants C, D, and n, so that the functions Φ'_{mn} will, as functions of ϱ, have various pleasant orthogonality and boundary-vanishing properties. The restrictions on G when using set (b) seem much weaker than when using set (a). We can of course combine $\varrho^{\pm n}$ for different integral n to form Legendre polynomials, etc., allowing expansion of more general G's; but then we lose the convenient reproducing property (20).

§ 22. So far we have described methods which are ordinarily usable only for geometrically simple regions, such as rectangles, rings, spheres, etc. Such regions are really the most important ones, since their simplicity makes them occur again and again in applied mathematics—whereas a region of irregular shape may occur on only one occasion (although it may be required to obtain a solution for a considerable family of boundary distributions on that one occasion). But suppose that we do have a non-simple region. Then, of course, various techniques of mapping are available to us, of which the best known is perhaps conformal mapping applied to Laplace's equation in two dimensions. To find the transformation is itself a boundary-value problem, and we may have to carry out by hand an iterative numerical calculation to find the mapping function (see Southwell for the use of relaxation techniques in conformal mapping [1]). But such a numerical task of mapping need be

carried through only once; after that, if the transformed region and transformed equation fall into the category of problems to which our new techniques apply, we may solve rather easily for any number of different boundary distributions, the chief numerical problem being probably not the solution of the transformed equation over the transformed region, but the transformation back to the original region.

§ 23. We now describe a technique which seems to handle irregular regions and does not involve mappings. We use Laplace's equation as an example. Let us divide the boundary by $2N$ points, which may be spaced at our convenience (see fig. 3): x_α, y_α ($\alpha = 1, 2, \ldots, 2N$) with corresponding boundary values u_α.

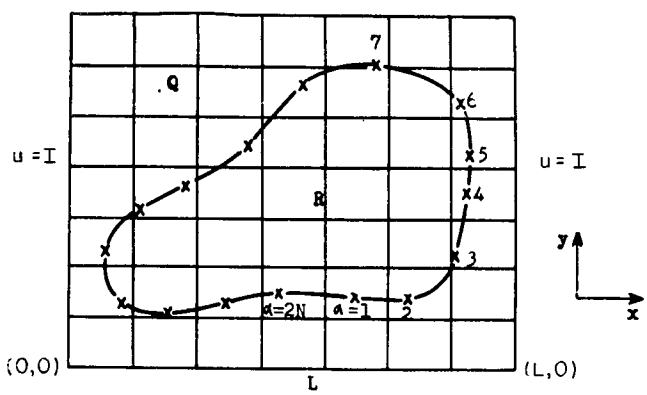

Fig. 3. Computation over an irregular domain.

Next we enclose the whole figure in a rectangle. We divide the rectangle by N evenly spaced verticals and as many evenly spaced horizontals as are convenient (it is desirable but not essential that $\Delta x = \Delta y$ approximately).

The $2N$ boundary points need not be points of the lattice. Now (cf. formula (3)), the solution of the difference equation (1) for this lattice can be written as

$$u(x, y) = I + \sum_{n=1}^{N} \sin \frac{n\pi x}{L} [P_n \lambda_{1n}^{\Delta y} + Q_n \lambda_{2n}^{\Delta y}]; \qquad (21)$$

I may be taken as zero or as the average of the $\{u_\alpha\}$. The $2N$ constants $\{P_n, Q_n\}$ are determined by imposing the boundary conditions $u = u_\alpha - I$ at x_α, y_α (note that λ_{1n} and λ_{2n} are real and positive —

equation (2.3) — and therefore $\lambda_{1n}^{y/\Delta y}$ and $\lambda_{2n}^{y/\Delta y}$ have real meanings for non-integral values of $y/\Delta y$). (21) is a "sine formula"; the condition $u = I$ on $x = 0$ and $x = L$ in general induces new singularities outside R of the differential solution u^*. Alternatively, we may use a "sine-cosine formula" and leave unspecified the values of u on $x = 0, L$. It should be noted that the presence of singularities causes loss of significant figures in numerical work; for example, we should avoid evaluating formula (21) near singularities.

§ 24. Thus we have to invert a $(2N \times 2N)$ matrix by some means after which (21) is an explicit solution of Laplace's difference equation over the irregular region, taking on the prescribed boundary values; starting from (21), we apply our usual stepping-ahead methods *). We make the following remarks:

(*a*) The matrix to be inverted is independent of the values $\{u_\alpha\}$ assigned at the $2N$ boundary points, and can be inverted once and for all for any particular region (provided the same boundary points are used). In changing from one boundary distribution to another for this region, we need only multiply the $(2N \times 2N)$ inverse matrix by the new vector $(u'_\alpha - I')$ to obtain the new coefficients $\{P'_n, Q'_n\}$ (primes denote new values).

(*b*) Suppose the number of interior and boundary points to be about M^2 and $4M$, respectively. The usual methods (relaxation, etc.) seek to invert for each boundary distribution a specialized $(M^2 \times M^2)$ matrix, while here the matrix to be inverted once is merely $(4M \times 4M)$. If the region is a perfect rectangle, we have seen that the matrix problem can be reduced much further still, to the trivial inversion of $2M$ (2×2) matrices. If the region, while not completely elementary, still possesses a good deal of symmetry, the typical matrix to be inverted may be somewhere in order between 2 and $4M$.

(*c*) We have illustrated the method by enclosing the irregular region R in a rectangle. This is usually most convenient. However, if R is nearly circular, or nearly elliptic, we may find it desirable to use one of the other "standard" lattices over which we can solve. The chief difficulty would seem to be in finding and evaluating the

*) To solve Poisson's equation over R, we first get a particular solution over the rectangle by an obvious adaptation of our previous technique — then solve Laplace's equation over R with modified boundary-values.

explicit solution (ES) at the "starting lines"; but the stepping-ahead which we conceive to be the major part of the numerical work, always consists of simple arithmetic operations. It should be noted that the rectangle, circle, parallelopiped, etc., which we put around the region R to be solved, is merely a "scaffolding", a device for setting up a convenient (ES) over R.

(*d*) There is no question that (21) gives us a valid solution of the difference problem; that is, (21) certainly satisfies the equation and takes on the assigned boundary values. There is some question of convergence, however; that is, what happens as the mesh dimensions go to zero? For example, there might be an infinite singularity at the point Q; as the mesh dimensions go to zero, (21) must go to infinity, at least near Q.

Over what region, then, does (21) converge? To what values does it converge? Does this region contain R? These are interesting questions about which we shall say only a little here. The discussion in references [4] [5] [6] [7] shows that a solution of equation (1) which takes on the proper values at the boundary points of a general region R *must converge* to the solution of Laplace's differential equation analytic throughout R and taking on the prescribed boundary distribution. We believe that this general convergence theorem (which holds for more general problems than Dirichlet's) covers our treatment (summed up in equation (21)) of an irregular region R *inside* which a non-singular solution is sought. Outside R, the region of convergence of (21) will depend on the position of singularities, generally present, of the differential solution u^*.

It should be remarked that the *rate* of convergence at any point P (inside or outside R) is influenced by the nearness of P to these singularities.

§ 25. Note that our method in effect yields, not the Green's function for the *differential* equation and domain, but an arbitrarily good approximation thereof — and the result is in a form well-suited to computation. Elsewhere we shall discuss the interesting integral equation analogue of (21), obtained by passing to the limit of infinitely many boundary points. As far as we know, this integral equation approach to the solution of linear boundary-value problems is new.

§ 26. In this paper we have usually spoken in terms of second order elliptic equations in which the values of the unknown function are given along the closed boundary of the region. But considering the remarks we have just made concerning extensions of the technique, it is clear that higher order equations (involving, for example, the biharmonic operator) and other boundary conditions (involving the unknown function and/or several of its external derivatives) introduce no difficulties in principle. We reserve for a later paper explicit discussion of these matters, as well as eigenvalue problems.

§ 27. We conclude with two remarks. First, our use of stepping-ahead techniques allows a unified approach to the numerical solution of mixed problems (where the governing differential equation is sometimes elliptic, sometimes hyperbolic or parabolic in the region of integration). A good example of a mixed problem is computation of the transonic flow of a gas; in the numerical solution of this problem it is usual at the present time to use iterative methods in the subsonic (elliptic) regime and stepping-ahead methods in the supersonic (hyperbolic) regime. Our second remark consists of the observation that, in solving a difference problem by our methods, we have in effect inverted a high-order matrix. This is the reverse of the usual "direct" procedure for the non-iterative solution of a difference equation, which consists of the systematic inversion of the high-order matrix. We thus have discovered, as a by-product of our research, a new method for inverting high-order matrices of simple (but non-trivial) structure. Indeed the only essential attribute of this structure seems to be that there should be zeros everywhere except on certain diagonals (if we order the points and the equations in the most obvious way); otherwise, the matrix elements may be fairly general.

Received 21st November, 1951

REFERENCES

1) Southwell, R. V., Relaxation Methods in Theoretical Physics; University Press, Oxford, 1946.
2) Shortley, G., and R. Weller, J. appl. Phys. **9** (1938) 334.
3) Grinter, L. E., ed., Numerical Methods of Analysis in Engineering; see especially chapters 3, 4, 5, 8, 9, 10; chapter 10 by T. J. Higgins contains a very extensive bibliography; Macmillan, 1949.

4) Courant, R., K. Friedrichs, and H. Lewy, Math. Ann. **100** (1928) 32.
5) Wiener, N., and H. B. Phillips, J. Math. Phys. **2** (1923) 105.
6) Petrowsky, I. G., Uspekhi Matem., Nauk. **8** (1941) 161.
7) Rosenbloom, P. C., The difference equation method for solving the Dirichlet problem, to appear. We learned of Rosenbloom's work, which is parallel to ours at one point, after our analysis was completed. His paper has an extensive list of references.
8) Moskovitz, D., Quart. J. appl. Math. **2** (1944) 148.
9) Richardson, L. F., Phil. Trans. **210A** (1910) 307.
10) O'Brien, G., M. Hyman, and S. Kaplan, J. Math. Phys. **29** (1951) 223.
11) Allen, D. N. de G., and S. C. R. Dennis, Quart. J. Mech. appl. Math. **4** (1951) 199.
12) Scarborough, J. B., Numerical Mathematical Analysis; Johns Hopkins Press, 1950.
13) Milne, W. E., Numerical Calculus; Princeton University Press, 1949.
14) Willers, F., Practical Analysis; Dover, 1948.
15) Runge, C., Z. angew. Math. Phys. **56** (1908) 225. Runge's use of stepping-ahead methods has been revived and emphasized by E. Stiefel (Zürich). D. M. Young and K. S. Kunz (Harvard University) have prepared but not published tables of weighting coefficients (i.e., contributions of boundary points to values on "first interior line") for squares with 4, 6, 8 intervals on a side.