TR/27 September 1973

FINITE-DIFFERENCE SOLUTION OF POISSON'S EQUATION IN RECTANGLES OF ARBITRARY SHAPE

BY

J. BARKLEY ROSSER

Sponsored in part by the United States Army under Contract No.DA-31-124-AR0-D-462 and in part by the Science Research Council under grant B/RG 4121 at Brunel University.

FINITE-DIFFERENCE SOLUTION OF POISSON'S EQUATION IN RECTANGLES OF ARBITRARY SHAPE

J. BARKLEY ROSSER

1. Introduction. We consider the problem of getting an approximation of reasonably good accuracy by finite-difference methods for the function u(x,y) which satisfies Poisson's equation

(1.1) $V^2 u(x,y) = f(x,y)$

inside a rectangle R, and satisfies various boundary conditions on the boundary of R. When f(x,y) = 0, (1.1) reduces to Laplace's equation, and the problem is appreciably simpler.

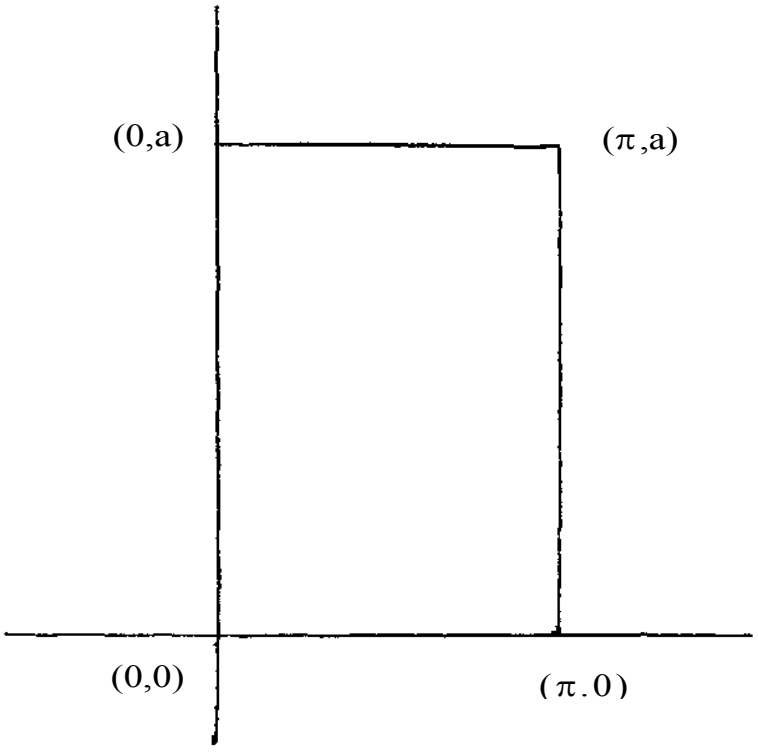
This problem has been much studied. A common approach is to cover R exactly with a mesh or grid of squares of side h, after which one can replace (1.1) by a finite-difference approximation involving values of u(x,y) at the grid points. One then tries to solve this finite-difference analogue of (1.1) to a suitable degree of accuracy. In order to employ this approach, the ratio of the sides of R must be rational; otherwise one cannot cover R exactly with a grid of squares. The conformal transformation method of Papamichael and Whiteman [3] will lead more often than not to a rectangle in which the ratio is not rational. Even when the ratio is rational, there may be difficulties. Suppose, from some engineering problem, one is confronted with a rectangle R of base six and five-eighths and height five and seven-eighths. If this

is to be covered exactly with squares, there must be $53N$ squares along the base and $47N$ squares along a vertical side, where N is a positive integer. With such a covering, many popular methods would operate at less than maximum efficiency.

Accordingly, we will propose a method of getting good accuracy for moderate labor for rectangles of arbitrary shape.

2. Formulation of the problem.

By rotation, translation, and scaling, as needed, we can take the rectangle R to be that shown in Figure 1. By rotating through another 90° and translating and scaling again, if need be, we can assure that $a \geq \pi$. If $a = \pi$, we have a square, and familiar approaches suffice. So we assume $a > \pi$.



The rectangle R

Figure 1

We consider first the case of Dirichlet boundary conditions. That is, we wish to approximate the function u(x,y) which is continuous on and inside R ,, satisfies

(2.1) $\nabla^2 u(x,y) = f(x,y)$

inside R, and on the sides of R satisfies the Dirichlet boundary conditions

(2.2) $u(0,y) = g_o(y)$ $0 < y < a$

(2.3) $u(\pi,y) = g_\pi(y)$ $0 < y < a$

(2.4) $u(x,0) = h_o(x)$ $0 < x < \pi$

(2.5) $u(x,a) = h_a(x)$ $0 < x < \pi.$

Because we seek a u(x,y) which is continuous on R , as well as inside, we are thereby assuming that $g_o(y)$ and $g_\pi(y)$ are continuous for $0 \leq y \leq a$, that $h_o(x)$ and $h_a(x)$ are continuous for $0 \leq x \leq \pi$, and that

(2.6) $g_o(0) = h_o(0),$

(2.7) $g_o(a) = h_a(0),$

(2.8) $g_\pi(0) = h_o(\pi),$

(2.9) $g_\pi(a) = h_a(\pi).$

In any reasonable problem, there will be only a finite number of discontinuities of the boundary values around the boundary of R , and these will be jump discontinuities at worst; that is, as the point of discontinuity is approached from different directions, u(x,y) will approach finite limits which are different. We shall

indicate how, if there are a finite number of jump discontinuities around the boundary of R, we can replace the problem by one in which there are fewer points of discontinuity. By repeating this reduction, we are finally brought to a problem, in which there are no points of discontinuity.

Consider first the case in which there is a jump discontinuity at a corner. We shall illustrate by treating the case in which the discontinuity is at the upper right hand corner. Very minor modifications are required to handle other corners, but the reader can work these out with little effort.

It is essential that, in "removing" the discontinuity at the upper right hand corner, we do not introduce additional discontinuities elsewhere. We shall consider in detail the case in which there is only one discontinuity, and it is at the upper right hand corner, and shall show how to "remove" this so as to produce a situation in which there are no discontinuities. It will be clear that in the more general case the procedure will "remove" the discontinuity at the upper right hand corner without introducing additional discontinuities elsewhere.

So let us assume that $g_o(y)$ and $g_\pi(y)$ are continuous for $0 \leq y \leq a$, that $h_o(x)$ and $h_a(x)$ are continuous for $0 \leq x \leq \pi$, that (2.6), (2.7), and (2.8) hold, but that (2.9) fails. We introduce the function

$$u^*(x,y) = Q \arctan \frac{\pi - x}{a - y}\, u(x,y), \quad ,$$

where

$$Q = \frac{2}{\pi} \left(g_\pi(a) - h_a(\pi)\right) .$$

Then $u^*(x,y)$ satisfies (2.1) inside R (that is,

$$\nabla^2 u^*)(x,y) = f(x, y)$$

holds inside R), and on the sides of R satisfies the boundary conditions

$$u^*(0,y) = Q \arctan \frac{\pi}{a-y} + g_o(y) \qquad 0 < y < a$$

$$u^*(\pi.y) = g_\pi(y) \qquad 0 < y < a$$

$$u^*(x,0) = Q \arctan\frac{\pi - x}{a} + h_o(x) \qquad 0 < x < \pi$$

$$u^*(x,0) = g_\pi(a) - h_a(\pi) + h_a(x) \qquad 0 < x < \pi .$$

From the boundary conditions above, we see that $u^*(x,y)$ has no discontinuities as one proceeds around the boundary of R. So, if we know how to handle such a case, we can thereby get a good approximation for $u^*(x,y)$, from which we can get a good approximation for $u(x_4y)$ by the relation

$$u(x,y) = u^*(x,y) - Q \arctan \frac{\pi - x}{a - y}.$$

It is clear that if we had started with a u(x,y) with other discontinuities around the boundary as well as at the upper right hand corner, then u^* (x,y) will have one fewer discontinuities around the boundary of R than u(x,y), and a succession of suitable reductions will bring us to the case of no discontinuities.

The case of a discontinuity other than at a corner is handled similarly. We shall illustrate by indicating how to handle

a discontinuity on the top of R. We are confining our attention to the case of jump discontinuities. So suppose that $0 < p < \pi$, that as x approaches p from the left $h_a(x)$ has a limit $h_a(p\text{-o})$, and that as x approaches p from the right $h_a(x)$ has a different limit $h_a(p+0)$. Let u(x,y) be the function for which we wish to "remove" the discontinuity at $x = p$. We introduce the function

$$u^*(x,y) - Q \arctan \frac{p-x}{a-y} \; u(x,y)$$

where

$$Q = \frac{1}{\Pi} \left(h_a(p+0) - h(p-0)\right).$$

Then $u^*(x,y)$ satisfies (2.1) inside R, and on the sides of R satisfies the boundary conditions

$$u^*(0,y) = Q \arctan \frac{\rho}{a-y} + g_o(y) \qquad 0 < y < a$$

$$u^*(\pi,y) = Q \arctan \frac{\rho-\pi}{a-y} + g_\pi(y) \qquad 0 < y < a$$

$$u^*(x,0) = Q \arctan \frac{\rho - x}{a} + h_0(x) \qquad 0 < x < \pi$$

$$u^*(x,a) = \frac{1}{2}(h_a + 0) - h_a(\rho+0)) + h_a(x) \qquad 0 < x < P$$

$$u^*(x,a) = \frac{1}{2}(h_a - 0) - h_a(\rho+0)) + h_a(x) \qquad \rho < x < \pi.$$

It is clear that $u^*(x,y)$ is continuous at (p,a), where it has the value

$$\frac{1}{2}(h_a(\rho+0) + h_a(\rho-0)).$$

Also, elsewhere around the boundary of R u^* (x,y) is continuous wherever u(x,y) is; at points of discontinuity of u(x,y) there are points of discontinuity of u^*(x,y) of exactly similar kinds.

If u(x,y) has a large number of discontinuities around the boundary of R , the first impression is that it will be a lot of bother to "remove" them all. However, one is certainly working with a computer, and it is possible to shuffle off quite a bit of the bother onto the computer. Thus, suppose that there are five discontinuities along the top at $\rho_1, \rho_2, \rho_3, \rho_4,$ and ρ_5. One can write a subroutine to handle a discontinuity at p, and then successively input the five parameters $\rho_1, \rho_2, \rho_3, \rho_4$ and ρ_5.

One may prefer to retain the discontinuities, and to use one of the better known methods of solution, allotting time for enough additional computation to reduce the errors arising from the discontinuities to an acceptable size. To get some idea of the errors arising from discontinuities in the boundary conditions, the reader may consult Rosser [5] .

To use the methods of the present paper with any success, there must be no discontinuities in the boundary conditions as one proceeds around R . Indeed, it is probably desirable to arrange that the boundary conditions have continuous derivatives, except at the corners; even there the derivatives should have finite limits as the corner is approached. The errors arising from discontinuities in the derivatives of the boundary conditions are also discussed in Rosser [5] .

If there are discontinuities of the derivatives around the boundary, and if after consulting Rosser [5] it is deemed advisable to "remove" these discontinuities, this can be done as follows,

Consider first the steps outlined above taken to "remove" the discontinuities of the boundary conditions themselves. Observe that these will not introduce any discontinuities into the derivatives, except perhaps at the corners, and jumps of the derivatives at the corners do not cause any errors in the use of finite-difference methods. So we may assume that the boundary conditions themselves are continuous (perhaps because of "removal" of all discontinuities), and it is discontinuities of the derivatives that we wish to "remove", one by one. The method we shall use should have, and does have the desirable attribute that it will not reintroduce discontinuities into the boundary conditions themselves.

We may take the discontinuity . to be at the point (ρ, a), where $0 < \rho < \pi$. We assume that as x approaches ρ from the left $h_a'(x)$ has a limit $h_a'(\rho - 0)$, and that as x approaches ρ from the right $h_a'(x)$ has a different limit. $h_a'(\rho + 0)$ Let u(x,y) be the function which is continuous around the boundary of R but whose derivative of the boundary conditions around R has the indicated discontinuity at (ρ, a). We introduce the function

$$u^*(x,y) = Q\, s(x,y) + u(x,y),$$

where

$$s(x,y) = (x-\rho) \arctan\frac{\rho - x}{a - y} + \frac{1}{2}(a-y)\log\{(a-y)^2 + (\rho-x)^2\}$$

and

$$Q = \frac{1}{\pi}\,(h_a'(\rho+0) - h_a'(\rho-0))\ .$$

Then $u^*(x,y)$ satisfies (2.1) inside R, and on the sides of R satisfies the boundary conditions

$$u^*(0.,y) = Q\,s(0,s) + g_o(y) \qquad 0 < y < a$$

$$u^*(\pi,y) = Q\,s(\pi,y) + g\pi(y) \qquad 0 < y < a$$

$$u^*(x,0) = Q\,s(x,0) + h_o(x) \qquad 0 < y < \pi$$

$$u^*(x,a) = \frac{1}{2}(x-\rho)\,(h_a'(\rho+0) - h_a'(\rho-0)) + h_a(x) \qquad 0 < y < \rho$$

$$u^*(x,a) = \frac{1}{2}(x-\rho)\,(h_a'(\rho-0) - h_a'(\rho+0)) + h_a(x) \qquad \rho < x < \pi$$

It is clear that $u^*(x,y)$ is continuous around the boundary of R. Also,

$$\frac{d}{dx}\,u^*(x,a)$$

is continuous at $x=\rho$, where it has the value

$$\frac{1}{2}\,(h_a'(\rho+0) + h_a'(\rho-0))$$

Elsewhere around the boundary of R, except at the corners, the behavior of the derivative of the boundary conditions, as far as continuity is concerned, is the same for $u^*(x,y)$ as for $u(x,y)$.

By repetitions of a reduction such as indicated, we can reduce to the case in which the boundary conditions are continuous around the boundary, and their derivatives are continuous except at the corners. That is, we can assure that

$g_0(y), g_0'(y), g_\pi(y)$, and $g_\pi'(y)$ are continuous for $0 = y = a$,

$h_0(x), h_0'(x), h_\pi(x)$, and $h_\pi'(x)$ are continuous for $0 \leq x \leq \pi$,

and (2.6), (2.7), (2.8), and (2.9) hold.

Should the boundary conditions or their derivatives have more complicated singularities than jump discontinuities, one may be able to "remove" them by suitable choices of $u^*(x,y)$. We leave these difficultues to the ingenuity of the reader.

3. Finite—difference approximations.

There are finite-difference approximations of various orders. The higher order methods of solution, involving the higher order approximations, can be used only when the function f(x,y) which appears in (2.1) has suitable high order smoothness; that is when it is continuous and has continuous derivatives of suitable orders. Thus the reader must exercise discrimination in choosing which order method to use. When they can be used, the high order methods permit the use of coarse meshes. This can greatly reduce the labor of computation.

Case A. Methods of order 2. Take M and P positive integers, and set

$$h = \frac{\pi}{M}, \tag{3.1}$$

$$k = \frac{a}{p}. \tag{3.2}$$

Clearly, one can fill up the rectangle R exactly with MP small rectangles, each of base h and height k . The corners of the small rectangles, (mh,nk), are called grid points. We write,

(3.3) $u_{m,n} = u(mh,nk)$.

Let us define a 5-point difference operator by

$$(3.4) \qquad \Delta_5 u_{m,n} = \frac{k}{n}(u_{m+1,n} + u_{m-1,n}) + \frac{h}{k}(u_{m,n+1} + u_{m,n-1}) - 2\frac{h^2+k^2}{hk} u_{m,n}.$$

Assuming that u(x,y) has enough partial derivatives to justify doing so, one can expand u(x,y) in a double Taylor series about the point (mh,nk) and verify that

(3.5) $\Delta_5 u_{m,n} \cong hk \nabla^2 u(mh,nk)$

to within terms of order $h^4 + k^4$. So, to get approximate values for u(x,y) at the grid points, we approximate (2.1) by

(3.6) $\Delta_5 \bar{u}_{m,n} = hk\ f(mh,nk)$,

and undertake to solve the resulting set of difference equations. It is known that there is a unique set of values $\bar{u}_{m,n}$ which satisfy (3.6) together with the boundary conditions

$\bar{u}_{0,n} = g_o(nk)$ $\qquad 0 < n < P$

$\bar{u}_{M,n} = g\pi(nk)$ $\qquad 0 < n < P$

$\bar{u}_{m,0} = h_o(mn)$ $\qquad 0 < m < M$

$\bar{u}_{m,P} = h_o(mn)$ $\qquad 0 < m < M$.

If f(x,y) is continuous, and the boundary conditions are Continuous, then the $\overline{u}_{m,n}$ will differ from the $u_{m,n}$ by amounts of the order of $h^2 + k^2$. Thus the method of approximating $u_{m,n}$ by solving (3.6) for the $\overline{u}_{m,n}$ is known as a method of order 2.

In case the boundary conditions are discontinuous, the approximation will be only of order 1 near the points of discontinuity, or even poorer. See the discussion in Rosser [5]. So, depending on the order of accuracy required, it may be desirable to "remove" discontinuities of the boundary conditions, as explained in the previous section.

If it is f(x,y) which is discontinuous, one may be able to "remove" the discontinuity by choosing a suitable $u^*(x,y)$. The possibilities are so varied that it seems hopeless to attempt a catalogue of procedures. If the reader is both lucky and ingenious, he may be able to cope with the situation. Otherwise, he must be satisfied with an approximation of low order.

Given that there is a unique set of $\overline{u}_{m,n}$ which satisfy (3.6) and the boundary conditions, it can be a formidable task to get good numerical approximations for these $\overline{u}_{m,n}$, particularly when M and P are large. The SOR and ADI methods which are familiar for the case h =k can be generalized to this case with reasonable success.. The extremely fast direct method of Hockney [1] is written with the possibility that $h \neq k$, and so is immediately available. Nonetheless, if quite good accuracy is desired, so that a fine mesh is required, the labor can be considerable. So, if f(x,y) and the boundary conditions are

sufficiently smooth, it may be desirable to use a higher order method, for which a coarser mesh will suffice.

There may be other cases in which use of a higher order method is desirable. Suppose one wishes to solve the biharmonic equation

(3.7) $\nabla^4 u(x,y) = g(x,y)$

inside R, being given the values of u(x,y) on the boundary (for example, by (2.2) through (2.5)), the values of

$$\frac{\partial^2}{\partial x^2} u(x,y)$$

along the vertical sides, and the values of

$$\frac{\partial^2}{\partial y^2} u(x,y)$$

along the top and bottom. By (2.2), we have along the left side

$$\frac{\partial^2}{\partial y^2} u(x,y) = g_0''(y).$$

Adding these to the given values of

$$\frac{\partial^2}{\partial x^2} u(x,y)$$

along the left side gives the values of

$$\nabla^2 u(x,y)$$

along the left side. That is, if we define f(x,y) by

(3.8) $$f(x,y) = \nabla^2 u(x,y),$$

then we have determined f(x,y) along the left side. In a similar way, we can determine f(x,y) on all the other sides. By (3.8)

and (3.7), we have

(3.9) $\nabla^2 f(x,y) = g(x,y)$.

So f(x,y) can be determined by solving a Poisson equation with Dirichlet boundary conditions. As soon as f(x,y) is known, one can determine u(x,y) by solving (3.8), which is the same as (2.1).

This is all very well in theory, but when one attempts to get approximations for u(x,y) on a computer certain complications arise. First of all, we must get approximations for f(x,y) from (3.9). We get these approximations at the grid points. These approximations must be stored, for use in solving (3.8) approximately by means of (3.6). If we have a fine mesh, there will be a large number of grid points, and the storage of the approximations for f(x,y) may exceed the capacity of the high speed memory.

A more serious point is that if we have solved (3.9) to order 2 and wish to get a solution of (3.7) to order 2, we must have a solution of

$$\nabla^4 u(x,y) = \nabla^2 f(x,y)$$

to order 2. If we are to infer this from (3.8), we must have a solution of (3.8) to order 4. The alternative is to use an extremely fine mesh, involving serious computer storage problems, to get a solution of (38) with high accuracy, which is then a solution of (3.7) only to modest, and somewhat indeterminate, accuracy.

Thus there certainly are cases in which one would wish to consider the use of higher order methods. We turn to some of these.

Case B. Methods of order 4. It seems generally believed that if one wishes to get a method of order greater than 2, one must take h = k. Indeed, this assertion is made on p.260 of Fox [6]. This may account in good part for the wide-spread acceptance of the belief.

However, the belief is unjustified. In Wirz [7] use is made of a method of order 4 for which $h \neq k$.

To derive such a method, we proceed as follows. Define

$$A = \frac{12h^2k^2}{h^2 + k^2} \tag{3.10}$$

$$b = \frac{10k^2 2h^2}{h^2 + k^2} \tag{3.11}$$

$$c = \frac{10h^2 - 2k^2}{h^2 + k^2} \; . \tag{3.12}$$

Then

$$(b + 2)h^2 = A = (c + 2)k^2 \tag{3.13}$$

$$b + c = 8. \tag{3.14}$$

Let us define a 9 - Point difference operator by

$$\Delta_9 u_{m,n} = b(u_{m+1,n} + u_{m-1,n}) + c(u_{m,n+1} u_{m,n-1}) + (u_{m+1,n+1} + u_{m-1,n+1} + u_{m+1,n-1} + u_{m-1,n-1}) - 20u_{m,n} \; . \tag{3.15}$$

Again assuming that we can expand in a double Taylor series, we get

$$(3.16)\quad \Delta_9 u_{m,n} \cong A\nabla^2 u(mh,nk) + \frac{h^2 A}{12} u_{xxxx}(mh,nk)$$

$$+ h^2 k^2 u_{xxyy}(mh,nk) + \frac{k^2 A}{12} u_{yyyy}(mh,nk)$$

to within terms of order $h^6 + k^6$.

By (2.1) we have

$$u_{xxxx}(mh,nk) + u_{xxyy}(mh,nk) = f_{xx}(mh,nk)$$

$$u_{xxyy}(mh,nk) + u_{yyyy}(mh,nk) = f_{yy}(mh,nk) .$$

If we multiply the first by $h^2 A/12$ and the second by $k^2 A/12$ and add, we see by (3.10) that we can write (3.16) as

$$(3.17)\quad \Delta_9 u_{m,n} \cong A f(mh,nk) + \frac{h^2 A}{12} f_{xx}(mh,nk) + \frac{k^2 A}{12} f_{yy}(mh,nk) .$$

So, to get approximate values for u(x,y) at the grid points, we approximate (2.1) by

$$(3.18)\quad \Delta_9 \bar{u}_{m,n} = A f(mh,nk) + \frac{h^2 A}{12} f_{xx}(mh,nk) + \frac{k^2 A}{12} f_{yy}(mh,nk) ,$$

and undertake to solve the resulting set of difference equations.

These difference equations are more complicated than (3.6). On the other hand, this method is of order 4, and one can use a much coarser mesh, with many fewer grid points, to obtain the same accuracy. Whether this counterbalances the additional complication will depend on the degree of accuracy required.

If we are dealing with Laplace's equation, in which $f(x,y) \equiv 0$, then the right sides of both (3.6) and (3.18) are zero. However, even in other cases the right side of (3.18) may not "be much worse than the right side of (3.6). For example, suppose that f(x,y) is

(3.19) $$\sqrt{(x+1)^2 (y+1)^2} .$$

Then

(3-20) $$f_{xx}(x,y) = \frac{(y+1)^2 f(x,y)}{((x+1)^2 + (y+1)^2)^2}$$

(3.21) $$f_{yy}(x,y) = \frac{(x+1)^2 f(x,y)}{((x+1)^2 + (y+1)^2)^2}$$

So

$$A f(x,y) + \frac{h^2 A}{12} f_{xx}(x,y) + \frac{k^2 A}{12} f_{yy}(x,y)$$

$$= \frac{A}{12} \left\{ 12 + \frac{h^2 (y+1)^2 + k^2 (x+1)^2}{((x+1)^2 + (y+1)^2)^2} \right\} f(x,y) .$$

On a computer, this would be only about twice as time consuming to calculate as the term

$$hk\ f(x,y)$$

which appears on the right side of (3.6). Also, recall that we have a coarser mesh, with fewer than one fourth as many grid points, perhaps one ninth as many, or even one sixteenth. So having to do twice as much calculating for each right side is rather a minor matter.

Actually, one can put the right side of (3.18) in a more convenient form. Recall that

$$h^2 f_{xx} \quad (mh,nk) \cong f((m+1)h,nk) + f((m-l)h,nk) - 2f(mh,nk) ,$$

$$k^2 f_{yy} \quad (mh,nk) \cong f(mh,(n+1)k) + f(mh,(n-l)k) - 2f(mh,nk) .$$

So, in place of (3.18), we write

$$(3.12)\quad \Delta_9 \overset{=}{u}_{m,n} = \frac{h^2k^2}{h^2+k^2}\left\{8f(m1\ nk) + f((m+1)h,nk) + f(m-1)h,nk)\right.$$

$$\left. + f(mh,(n+1)k) + f(mh,(n-1)k)\right\} .$$

Use of (3.6) would require the calculation of f(mh,nk) at all interior grid points. For (3.22) one would have also to calculate f(mh,nk) at grid points on the boundary, a rather small additional labor.

Use of (3.22) is a method of order 4, in that the $\overset{=}{u}_{m,n}$ will differ from the $u_{m,n}$ by amounts of the order of h^4+k^4 .

For the case h= k formula (3.22) is given in Collatz [8] as one of the stencils in Table VI on p. 542.

Case C.Methods of order 6. It is well known that if one can take h = k, then the 9 -point difference operator will lead to a method of order 6. There is a myth of some currency to the effect that this can be done only if f(x,y) ≡ 0. We will elucidate the true circumstances.

Assuming that we can take h = k, we can show by (2,1) that

$$(3.23)\quad \Delta 9\, u_{m,n} \cong 6h^2 f(mh\ \ nh) + \frac{h^4}{2}\nabla^2 f(mh,nh) + \frac{h^6}{60}\nabla^4 f(mh,nh)$$

$$+ \frac{h^6}{30} f_{xxyy}(mh,nh)\ ,$$

to within terms of order h^8. One may compare this with equation (20.57) on p. 194 of Forsythe and Wasow [2] ; however, in their equation the term

$$\frac{1}{2} h^2 \Delta^2 u(P)$$

should be replaced by

$$\frac{1}{2} h^2 \Delta^2 u(P)$$

as can be verified by looking up the reference which they cite for the derivation of their equation (20.57).

Thus we take

$$(3.24)\quad \Delta_9 \bar{u}_{m,n} = 6h^2 f(mh,nh) + \frac{h^4}{2}\nabla^2 f(mh,nh) + \frac{h^6}{60}\nabla^4 f(mh,nh)$$

$$+ \frac{h^6}{30} f_{xxyy}(mh,nh)$$

as a finite difference approximation for (2.1). If we solve for the $\bar{u}_{m,n}$, we have a method of order 6, in that the $\bar{u}_{m,n}$ and the $u_{m,n}$ will differ by amounts of the order h^6 . Thus we can use quite a coarse mesh, with yet fewer grid points, indeed considerably fewer. If we are dealing with Laplace's equation, then the right side of (3.24) will be zero. However, as we saw with (3.18), the matter is not hopeless even if f(x,y) is not identically zero.

If f(x,y) happens to be a reasonable function in closed form, such as (3.19)s then the right side of (3.24 is not too burdensome on a computer; this is especially true since we now have quite a coarse mesh, so that one needs to compute the right side of (3.24) at relatively few grid points.

There is merit in deriving something more analogous to (3.22), though we cannot do quite as well this time. We note that

$$(3.25)\quad f(x+h,y+h) + f(x,y+h) + f(x-h,y+h) + f(x+h,y) - 8\, f(x,y) + f(x-h,y) + f(x+h,y-h) + f(x,y-h) + f(x-h,y-h) \cong 3h^2 \nabla^2 f(x,y) + \frac{h^4}{4} \nabla^4 f(x,y) + \frac{h^4}{2} f_{xxyy}(x,y)$$

to vithin terms of order h^6. So we can get the following analogue for (3.22)

$$(3.26)\quad \Delta_9 u_{m,n} = \frac{h^2}{15}\Big\{ 82\, f(mh,nh) + f((m+1)h,(n+1)h) + f(mh,(n+1)h) + f((m-1)h,(n+1)h) + f((m+1)h,nh) + f((m-1)h,nh) + f((m+1)h,(n-1)h) + f(mh,(n-1)h) + f((m-1)h,(n-1)h)\Big\} + \frac{3h^4}{10} \nabla^2 f(mh,nh)\,.$$

As with (3.24, use of (3.26) gives a method of order 6. If f(x,y) is something like (3.19), the right side of (3.26) would be much simpler to compute than the right side of (3.24). Indeed, if f(x,y) is (3.19), then

$$\nabla^2 f(x,y) = \frac{1}{f(x,y)}\,,$$

so that computation of the right side of (3.26) would involve little more computer time than the computation of the right side of (3.6).

The formula (3.26) is given in Col atz [8] as one of the stencils in Table VI on p. 543.

As an indication of the advantage of (3.26), suppose we are trying to solve (3.7). If we wish a solution of (3-7) to order 4, we must get a solution of (3.9) to order 4 and then a solution of (3.8) to order 6. So we use (3.22) to get an approximate solution for (39). Accordingly, we can use a coarse enough mesh that storage of our approximations for f(mh,nh) is not a problem. We then undertake to get an approximate solution for (3.8) by use of (3.26). As we have stored approximations for f(mh,nh), we can fill in all of the right side of (3.26) except the term

$$\frac{3h^4}{10}\nabla^2 f(mh,nh).$$

However, by (3.9) this is given (to quite high order) by

$$\frac{3h^4}{10} g(mh,nh).$$

So we can proceed to an approximate solution of (3.8) to order 6. By (3.9), this is a solution of (3.7) to order 4.

It could be the case that we have f(x,y) given in tabular form, with no ready means to approximate

$$\frac{3h^4}{10}\nabla^2 f(mh,nh).$$

So we seek a formula to approximate the right side of (3.23) exclusively in terms of values of f(x,y).

We observe that while we would wish to restrict the left side of (3.23) to involve values of $u_{m,n}$ at only the familiar

9 grid points, there is no need for a similar restriction on the right side. We note further that

$$h^2 f_{xx}(mh,nh) \cong -\frac{1}{12} f((m+2)h,nh) + \frac{4}{3} f((m+1)h,nh) - \frac{5}{2} f(mh,nh)$$
$$+ \frac{4}{3} f((m-2)h,nh) - \frac{1}{12} f((m-2)h,nh)$$

to terms of order h^6. By using this, and the corresponding relation for f_{yy} (mh,nh), in (3.26), we see that we can take

$$(3.27)\quad \Delta_9 u_{m,n} \equiv \frac{h^2}{120}\Big\{476f(mh,nh) + 56(f((m+1)h,nh) + f((m-1)h,nh)$$
$$+ f(mh,(n+1)h) + f(mh,(n-1)n)) + 8(f((m+1)h,(n+1)h)$$
$$+ f((m-1)h,(n+1)h) + f((m+1)h,(n-1)h)$$
$$+ f((m-1)h,(n-1)h)) - 3(f((m+2)h,nh) + f((m-2)h,nh)$$
$$+ f(mh,(n+2)h) + f(mh,(n-2)h))\Big\}$$

as a finite-difference approximation for (2.1) to get a method of order 6.

Quite clearly, the right side of (3.27) is not the only possible choice. There may be others which, for some reason or other, are more advantageous. As alternatives are easily found, we leave it to the reader to choose which one pleases him best.

We turn to a more troublesome matter. Clearly (3.27) cannot be used at grid points which are only a distance h from the boundary (unless values of f(x,y) outside the boundary are available). For such grid points we will have to use a "stencil"

on the right side of (3.27) which is not centered on the point (mh,nh). One can use the off center difference approximation

$$h^2 f_{xx}(x,y) \cong \frac{5}{6} f(x-h,y) - \frac{5}{4} f(x,y) - \frac{1}{3} f(x+h,y) + \frac{7}{6} f(x+2h,y)$$
$$- \frac{1}{2} f(x+3h,y) + \frac{1}{12} f(x+4h,y),$$

which is correct to within terms of order h^6, to modify (3.26). However, this uses six consecutive grid points in the x-direction. One wonders if one one could avoid having to go so far in the x-direction if one would use additional points above and below those indicated. When one has to deal with the grid point nearest a corner, one must use the off center difference approximation in both the x-direction and the y-direction. It would seem that by using additional points, one should not have to go so far from the principal grid point.

This is not the case. We shall give an analysis which shows this, and illuminates the question of more general difference approximations in two dimensions.

We start by seeking a positive integer S and coefficients $a_{m,n}$ ($-1 \le m \le S$, $-1 \le n \le S$) such that

$$(3.28) \quad \sum_{m=-1}^{s} \sum_{n=-1}^{s} a_{m,n} f(x+mh, y+nh)$$
$$\cong 6 f(x,y) + \frac{h^2}{2} \nabla^2 f(x,y) + \frac{h^4}{60} \nabla^4 f(x,y)$$
$$+ \frac{h^4}{30} f_{xxyy}(x,y)$$

to within terms of order h^6.

Because the right side is invariant under interchange of

x and y, if we find a set of $a_{m,n}$ which satisfy (3.28)-so would $\bar{a}_{m\,n}$, where we define

$$\bar{a}_{m,n} = a_{n,m} .$$

Then, equally, we could replace $a_{m,n}$ in (3.28) by

$$\frac{1}{2}\left(a_{m,n} + \bar{a}_{m,n} \right).$$

So we might as well assume from the start that

(3.29) $$a_{m,n} = a_{n,m} .$$

Assuming that we can expand f(x,y) in a double Taylor series about the point (x,y) we have

(3.30) $$\sum_{m=-1}^{s} \sum_{n=-1}^{s} a_{m,n}\, f(x+mh,\, y+nh)$$

$$\cong \sum_{r=0}^{5} h^{r} \sum_{s=0}^{r} \frac{K_{rs}}{s!\,(r-s)!}\, D_x^{r-s}\, D_y^{s}\, f(x,\, y)$$

to within terms of order h^6, where

(3.31) $$K_{rs} = \sum_{m=-1}^{s} \sum_{n=-1}^{s} m^{r-s}\, n^{s}\, a_{m,n} .$$

If we write

$$A_n = \sum_{m=-1}^{s} a_{m,n} \quad , \tag{3.32}$$

then because of (3.29) we have

$$K_{r0} = K_{rr} = \sum_{m=-1}^{s} n^r A_n \quad . \tag{3.33}$$

By (3.30), If we are to satisfy (3.28) we must have

$$K_{00} = 6 \tag{3.34}$$

$$K_{10} = K_{01} = 0 \tag{3.35}$$

$$K_{20} = K_{02} = \tag{3.36}$$

$$K_{30} = K_{03} = \circ \tag{3.37}$$

$$K_{40} = k_{04} = \frac{2}{5} \tag{3.38}$$

$$K_{50} = K_{05} = 0 \tag{3.39}$$

By (3.33), this is a set of six simultaneous linear equations for the A_n . They have no solution unless $S \geq 4$. So we take $S = 4$, and find the solution

$$A_{-1} = \frac{9}{20}$$

$$A_0 = \frac{209}{40}$$

$$A_1 = \frac{1}{10}$$

$$A_2 = \frac{7}{20}$$

$$A3 = -\frac{3}{20}$$

$$A_4 = \frac{1}{40} .$$

Analogously, if we write

$$(3.40) \qquad B_n = \sum_{m=-1}^{s} m\, a_{m,n} ,$$

then to satisfy (3.28) we must have

$$(3.41) \qquad \sum_{n=-1}^{s} n^r B_n = 0 \qquad (0 \leq r \leq 4).$$

This set of 5 equations has the solution

$$B_{-1} = \frac{1}{5} B_0$$

$$B_1 = -2B_0$$

$$B_2 = 2 B_0$$

$$B_3 = - B_0$$

$$B_4 = \frac{1}{5} B_0 .$$

We write

$$(3.42) \qquad C_n = \sum_{m=-1}^{s} m^2 a_{m,n} \quad .$$

Then, to satisfy (3.28) we would have

$$(3.43) \qquad C_{-1} + C_0 + C_1 + C_2 + C_3 + C_4 = 1$$

$$(3.44) \qquad -C_{-1} + C_1 + 2C_2 + 3C_3 + 4C_k = 0$$

$$(3.45) \qquad C_{-1} + C_1 + 4C_2 + 9C_3 + 16C_4 = \frac{4}{15}$$

$$(3.46) \qquad -C_{-1} + C_1 + 8C_2 + 27C_3 + 64C_4 = 0.$$

These have the solution

$$C_{-1} = \frac{19}{60} - \frac{1}{4} C_0 + \frac{1}{4} C_4$$

$$C_1 = \frac{37}{30} - \frac{3}{2} C_0 - \frac{5}{2} C_4$$

$$C_2 = \frac{11}{15} + C_0 + 5 C_4$$

$$C_3 = \frac{11}{60} - \frac{1}{4} C_0 + \frac{15}{4} C_4 \quad .$$

It can be seen that satisfaction of the equations (3.34-(3.39), (3.41), and (3.43) - (3.46) is sufficient as well as necessary for the satisfaction of (3.28). This we have accomplished, and with three parameters, B_0, C_0 , and C_4 , at our disposal.

There is still the question how to choose the $a_{m,n}$ to give the desired values of the A_n , B_n , and C_n . If we specify values of the A_n , B_n and C_n , then (3.32), (3.40), and (3.42) are a set of 18 simultaneous linear equations for the $a_{m,n}$. In view of (3.29), there are 21 distinct $a_{m,n}$, so that we have more unknowns than equations. Since we can produce one solution by using the off center difference approximation in both the x—direction and the y—direction, it follows that there is a whole family of solutions. We will show how to find others of these solutions.

We choose arbitrarily

$$a_{-1,4} = a_{4,-1} = 0$$

$$a_{0,4} = a_{4,0} = 0$$

$$a_{1,4} = a_{4,1} = 0$$

$$a_{2,4} = a_{4,2} = 0$$

$$a_{3,4} = a_{4,3} = 0 \quad .$$

As $A_4 = 1/40$, we conclude by (3.32) that

$$a_{4,4} = \frac{1}{40} \quad .$$

Then by (3.40)

$$B_4 = \frac{1}{10} \ ,$$

so that

$$B_{-1} = -\frac{1}{10}$$
$$B_0 = \frac{1}{2}$$
$$B_1 = -1$$
$$B_2 = 1$$
$$B_3 = -\frac{1}{2} .$$

Also, by (3.42)

$$C_4 = \frac{2}{5} .$$

We arbitrarily choose

$$C_0 = \frac{1}{3} ,$$

and infer

$$C_{-1} = \frac{1}{3}$$
$$C_1 = -\frac{4}{15}$$
$$C_2 = \frac{8}{5}$$
$$C_3 = -\frac{7}{5} .$$

Using the values of A_3 , B_3 , and C_3 that we have determined, (3.32), (3.40), and (3.42) are three linear equations involving $a_{m,3}$ $(-1 \leq m \leq 4)$. We have already determined $a_{4,3}$. We arbitrarily set

$$a_{3,3} = 0$$
$$a_{2,3} = a_{3,2} = 0 .$$

We then have three equations in three unknowns. Solving gives

$$a_{-1,3} = a_{3,-1} = -\frac{9}{20}$$
$$a_{-0,3} = a_{3,0} = \frac{5}{4}$$
$$a_{-1,3} = a_{3,-1} = -\frac{19}{20} \quad .$$

Using the values of A_2 ,B_2,and C_2 that we have determined, (3.32), (3.40), and (342) are three linear equations for the six quantities $a_{m.2}$ (-1 $\leq$ m $\leq$ 4*)*. Two of the six quantities have been already determined. We arbitrarily set

$$a_{2,2} = 0$$

and solve for the other three, getting

$$a_{-1,2} = a_{2,-1} = \frac{3}{10}$$
$$a_{-0,2} = a_{2,0} = -\frac{5}{4}$$
$$a_{-1,2} = a_{2,1} = \frac{13}{10} \quad .$$

We consider the quantities a_{m-1} (-1 $\leq$ m $\leq$ 4*)*. Three have been determined already. The values of A_1 , B_1 and C_1. give three equations for the others, and we determine

$$a_{-1,1} = a_{1,-1} = \frac{23}{12}$$
$$a_{-0,1} = a_{1,-0} = -\frac{10}{3}$$
$$a_{1,1} = \frac{7}{6} \quad .$$

We now know four of the six quantities $a_{m,0}$ (-1 $\leq$ m $\leq$ 4).

So we have three equations for the two unknowns $a_{-1,0}$ and $a_{0,0}$.

Fortunately, one of the equations is dependent on the other two.

This had been anticipated from an analysis of the matrix of the coefficients. So we determine

$$a_{-1,0} = a_{0,-1} - \frac{31}{12}$$

$$a_{0,0} = \frac{1337}{120}$$

Finally, we have three equations for $a_{-1,1.}$. Again there is dependency, and we get

$$a_{-1,-1} \quad \frac{91}{15} \ .$$

It appears that, in general, one can choose the $a_{m,4}$ at will, subject to the condition that their sum A_4 be 1/40. Use of the off center difference approximation in both the x-direction and the y-direction amounts to choosing

$$a_{-1,4} = a_{1,4} = a_{2,4} = a_{3,4} = a_{4,4} = 0$$

$$a_{0,4} = \frac{1}{40} \ .$$

Having chosen the $a_{m,4}$, we will have determined B_4 and C_4 . This then fixes the B_n , but leaves C_0 at our disposal. With an arbitrary choice of the $a_{m,4,}$ one will likely not have the dependencies which we exploited above. However, by proceeding first to the $a_{m.3}$, then to the $a_{m.2.}$ and so on, it is a modest labor to express all the $a_{m,n}$ as linear combinations of C_0 , $a_{3,3}$, $a_{2,2}$, and $a_{2,2}$. In the process, one will likely derive certain relationships between these quantities, so that at the end one will probably have all the other $a_{m,n}$ expressed in terms

of C_0. It seems not worthwhile to explore the details any further at this point.

We noted earlier that if the principal grid point is at a distance h from the left edge, but further than that from the top or bottom, then one can use the off center difference approximation in the x-direction. We use the methods given above to see if one can get a "stencil" which does not extend as far as six grid points in the x-direction. It turns out that one cannot, but we will present the analysis anyhow, since it shows how to generate all possible "stencils".

Without causing confusion, we can use the same letters as before, but with slightly altered denotations.

So for our $a_{m,n}$ we will now have $-2 \leq m \leq 2$, $-1 \leq n \leq S$.

In place of (3.29), we will have

(3.4) $$a_{m,m} = a_{-m,n}.$$

All summations on m should be from -2 to +2. Specifically, this change should be made in (3.30), (3.31), (3.32), (3.40), and (3.42). In (3.33) the term K_{ro} should be omitted.
As before, we see that we must have $S > 4$. Taking $S = 4$, we get the same values of A_n as before.

By (3.47) and (3.40), we have $B_n = 0$ for all n. Thus (3.4l) is trivially satisfied.

We get the same determination as before for the C_n.

Finally, we write

$$(3.48)\quad D_n = \sum_{m=-2}^{2} m^4 a_{m,n} .$$

We must have

$$(3.49)\quad D_{-1} + D_0 + D_1 + D_2 + D_3 + D_4 = \frac{2}{5}$$

$$(3\text{-}50)\quad -D_{-1} + D_1 + 2D_2 + 3D_3 + 4D_4 = 0.$$

Given the A_n , B_n , and C_n , there is no question how to determine the $a_{m,n}$. We have immediately

$$a_{-2,n} = a_{2,n} = \frac{D_n - C_n}{24} ,$$

$$a_{-1,n} = a_{1,n} = \frac{4C_n - D_n}{6} ,$$

$$a_{0,n} = A_n - 2a_{1,n} - 2a_{2,n} .$$

Thus we can easily determine sets of $a_{m,n}$

In view of these developments, we see that it is entirely practical to get an approximation of order 6 for the solution of Poisson's equation, provided one can take h = k. Also there are cases where it is desirable to get an approximation of order 6. So it would be quite useful to have the possibility of taking h = k. We shall show how this can be done even if a/π is irrational.

4. <u>The case h = k.</u> As before, we take

$$(4.1)\qquad h = \frac{\pi}{M} ,$$

where M is a positive integer. We take N to be the integer part of aM/π. ; in symbols

(4.2) $$N = \left[\frac{aM}{\pi} \right] .$$

Then

(4.3) $$Nh \leq a,$$

(4.4) $$(N + 1)\, h > a.$$

If

(4.5) $$Nh = a,$$

then we can fill up the rectangle R exactly with MN squares of side h. That is, we have immediately available the case h = k, and the 9-point difference approximation leads to the well known method of order 6 which we discussed in the previous section. So we are interested only in the case

$$Nh < a.$$

We could assume this, but it is not required for the analysis which follows. If we should have (4.5) holding, then some of the steps of the subsequent analysis would be quite trivial but not incorrect in any way.

We begin by defining

(4.6) $$b - Nh$$

(4.7) $$c = a - b = a - Nh.$$

We take R_b to be the rectangle with corners (0,0), (0,b), $(\pi,0)$, and (π,b), and take R_c to he the rectangle with corners (0,c), (0,a), (π,C) and (π,a).

We choose h_b (x) to be a function such that

$$h_b(0) = g_0(b)$$

$$h_b(\pi) = g_\pi(b) \ .$$

The better we can choose h_b (x) to approximate u(x,b), the more we can curtail certain computations later. With the limited information available at this stage, we content ourselves with taking

$$h_b(x) = h_a(x) + (1-\frac{x}{\pi})\ (g_0(b) - h_a(0)) + \frac{x}{\pi}(g_\pi(b) - h_a(\pi)) \ .$$

We take u_b (x,y) to be the function which is continuous on and inside R_b satisfies (2.1) inside R_b and on the sides of R_b satisfies the boundary conditions

(4.8) $u_b(0,y) = g_0(y)$ $\qquad 0 \leq y \leq b$

(4.9) $u_b(\pi,y) = g_\pi(y)$ $\qquad 0 \leq y \leq b$

(4.10) $u_b(x,0) = h_0(x)$ $\qquad 0 <\_ x <\_ \pi$

(4.11) $u_b(x,b) = h_b(x)$ $\qquad 0 \leq x \leq \pi$.

We take u_c (x,y) to be the function which is continuous on and inside R_c , satisfies (2,1) inside R_c. and on the sides of R_c satisfies the boundary conditions

(4.12) $u_c(0,y) = g_0(y) \qquad c \leq y \leq a$

(4.13) $u_c(\pi,y) = g_\pi(y) \qquad c \leq y \leq a$

(4.14) $u_c(x,c) = u_b(x,c) \qquad 0 \leq x \leq \pi$

(4.15) $u_c(x,a) = h_a(x) \qquad 0 \leq x \leq \pi$.

We recall that we are assuming that u(x,y) has continuous boundary conditions around the rectangle R. If it had not been so, we would have replaced u(x,y) by some u**(x,y) for which it is so, as explained in Section 2. By our definition of h_b (x), we see that $u_b(x,y)$ has continuous boundary conditions around the rectangle R_b . Then it follows by (4.14) that the same holds for u_c (x,y) relative to the rectangle R_c. This is why in (4.8) through (4.15) we can use $\leq$ rather than $<$.

By (4.1) and (4.6) we can fill up the rectangle R_b exactly with MN squares of side h. Thus we can use the familiar 9-point difference approximation to get accurate approximations for u_b (x,y) inside R_b at the grid points (mh,nh). From these, we can get accurate approximations for $u_b(mh,c)$. By (4.14) these are part of the boundary values for u_c (x,y). Thus it is necessary to determine them to order h^6 . By the principle of the maximum, it is also sufficient. For a given m, the point (mh,c) is on a vertical grid line. Thus one can determine u_b (mh,c) to order h^6 by using a high order interpolation formula in one dimension

on the values at the six grid points $(mh,0)$, (mh,h), $(mh,2h)$, $(mh,3h)$, $(mh,4h)$, and $(mh,5h)$. Although the spline function interpolation method of Papamichael and Whiteman [4] is simpler and more convenient, it likely does not give the needed accuracy.

By (4.14), this gives us good approximations to $u_c(x,c)$ at $x = h, 2h, \ldots, (M-1)h$. By (4.1) and (4.7) we can fill up the rectangle R_c exactly with MN squares of side h. Thus we can use the familiar 9-point difference approximation to get accurate approximations for $u_c(x,y)$ inside R_c at the grid points $(mh,c+nh)$. Then we can get accurate approximations for $u_c(mh,b)$ by the method mentioned earlier.

We define R_{bc} to be the rectangle which is the intersection of the rectangles R_b and R_c. In R_{bc}, the function $u_c(x,y) - u_b(x,y)$ is harmonic. Also, it is zero along the bottom and along the two vertical sides. So on and inside R_{bc} we have

$$(4.16)\quad u_c(x,y) - u_b(x,y) = \sum_{r=1}^{\infty} a_r \frac{\sinh r(y-c)}{\sinh r(b-c)} \sin r x$$

where

$$(4.17)\quad a_r = \frac{2}{\pi} \int_0^{\pi} \{u_c(x,b) - u_b(x,b)\} \sin r x \, dx \quad .$$

Clearly the $|a_r|$ are bounded by

$$(4.18) \qquad 2 \max_{0 \le x \le \pi} |u_c(x,b) - u_b(x,b)|$$

We recall (see (4.11)) that

$$u_b(x,b) = h_b(x) \ .$$

Presumably $u_c(x,b)$ is fairly close to $u(x,b)$. If also we were lucky enough to choose $h_b(x)$ fairly close to $u(x,b)$, then by (4.18) the a_r will be fairly small. This will save computational effort later.

On and inside R define

$$(4.19) \qquad v(x,y) = \sum_{r=1}^{\infty} a_r \, b_r \, \frac{\sinh r(a-y)}{\sinh r a} \sin r x \ ,$$

where

$$(4.20) \qquad b_r = \frac{\sinh r c}{\sinh r (b-c)} \ ,$$

On and inside R_b define

(4.21) $u(x,y) = u_b(x.y) + v(x,y)$

$$+ \sum_{r=1}^{\infty} a_r \frac{\sinh r(y-c)}{\sinh r(b-c)} \sin r x \quad .$$

We see that u(x,y) is continuous on and inside the rectangle R_b, satisfies (2.1) inside r_b, and on three sides satisfies the boundary conditions (4.8), (4.9), and (4.10). By (4.16), we see that, on and inside R_{bc} we have

(4.22) $u(x,y) = u_c(x,y) + v(x,y).$

We use (4.22) to define u(x»y) for the rest of the rectangle R_c. Then u(x,y) is continuous on and inside the rectangle Rc, satisfies (2.1) inside R_c, and on three sides satisfies the boundary conditions (4.12), (4.13), and (4.15).

Thus we see that u(x,y) is exactly the function u(x,y) that we were seeking to obtain.

We have obtained accurate approximations for $u_b(x,y)$ and $u_c(x,y)$ at various grid points. If M is of reasonable size, then c is small, since $0 \leq c < h$ by (4.7), (4.3), and (4.4). As a is greater than π, and b=a-c by (4.7), we see that the series on the right of (4.19) is rapidly convergent for $0 \leq y \leq a$. Also, the series appearing on the right of (4.21) is rapidly convergent for small y, certainly for $0 < y < h$. If in addition the a_r are all

quite small (see (4.18)), then very few terms of the series are needed to get high accuracy. So, using the known approximations for $u_b(mh,nh)$ with the spline function method of Papamichael and Whiteman [4], we can get approximate values for u(x,y) for small y by (4.21). For all other values of y, we can use the known approximations for u_c (mn, c + nh) with the spline function interpolation method of Papamichael and Whiteman [4] to get approximate values for u(x,y) by (4.22).

The calculation of the a presents no problem. Not more than four or five will be required; fewer if the a_r are all small. Observe that the values of $u_b(x,b)$ are given by (4.11). Also, we had got accurate approximations for u_c (mh, b). So we can use a numerical quadrature formula to calculate the a_r by (4.17).

CAUTION. If r is not fairly small compared to N, then there will be fairly few abscissa points in each cycle of sin r x in (4.17); in such case the usual quadrature formulas are not trustworthy. One can get twice, or four times, or eight times, as many abscissa points by interpolating to get approximations for u_c (x,b) at the additional abscissa points (recall that u_b (x,b) is given by (4.11)). For this interpolation one can either use a high order one dimensional interpolation formula on the values u_c (0,b), u_c (h,b), u_c (2h,b), . . . , or one can use the spline function interpolation method of Papamichael

and Whiteman [4]. However, before using the latter, one should check if it gives sufficiently high accuracy. We need high accuracy for only the first one or two of the a_r . Perhaps by the time we need to interpolate additional points for use with the quadrature formula, the needed accuracy will be low enough that one can use splines. In any case, *one* should increase the number of abscissa points, as needed, to the point where one can use a quadrature formula with assurance.

One should keep in mind the very rapid convergence of the series appearing on the right of (4.19) and (4.21). Thus, while we probably require full accuracy for a_1 , we need less for a_2, still less for a_3, down to hardly any for a_5, and probably none at all for a_6, a_7 , ... However, attention should be given to be sure there are enough abscissa points to give such accuracy as is needed. Also, by a little foresight in the choice of N, one can arrange that, after increasing the number of abscissa points if needed, one can use a high order quadrature formula, like Bode's Rule, for example.

5. Possible curtailment of the calculation.

In case f(x,y) is not aero, we have had to calculate the right side of (3.26), or something of the sort, at one set of grid points to calculate u_b (mh,nh), and at a different set of grid points to calculate u_c (mh, c + nh). It would be desirable if we could eliminate the second set of calculations of the

right side of (3.26). In some cases we can.

Suppose that $g_0(y)$, $g_\pi(y)$, and $f(x,y)$ are given in such forms that they can be continued up to $y = (N + 1)h$. In that case, we redefine

(5.1) $b = (N + 1)h$

(5.2) $c = a - Nh = a + h - b.$

With the new b and c (the new c is the same as the old one), we define R_b. and R_c as before, we take $n_b(x)$ as before, and $U_b(x,y)$ as before. Note that now $u_b(x,y)$ is defined in a region which contains R, and which therefore contains R_c.

Instead of defining $u_c(x,y)$ as before, we will define $w(x,y)$ as the function which is continuous on and inside R_c, is harmonic inside R_c, is zero on three sides of R_c, and on the top of R_c satisfies the boundary condition

(5.3) $w(x,a) = h_a(x) - u_b(x,a) \quad 0 \le x \le \pi.$

If now we define

(5.4) $u_c(x,y) = u_b(x,y) + w(x,y)$

on and inside R_c, then $u_c(x,y)$ satisfies the same conditions as before. Also, on and inside R_c we will have

(5.5) $u_c(x,y) - u_b(x,y) = w(x,y)$

$$= \sum_{r=1}^{\infty} a_r \frac{\sinh r(y-c)}{\sinh r(a-c)} \sin r x$$

where

(5.6) $$a_r = \frac{2}{\pi} \int_0^{\pi} \left\{ h_a(x) - u_b(x,a) \right\} \sin r x \, dx .$$

As with b, the new a_r are a trifle different from the old.

We define v(x,y) as before, except that now we take

(5.7) $$b_r = \frac{\sinh r c}{\sinh r (a-c)}$$

Then, on and inside R, u(x,y) is given by (4.21) while, on and inside R_c, u(x,y) would be given by (4.22), which in view of (5.4), had better be written as

(5.8) $u(x,y) = u_b(x,y) + w(x,y) + v(x,y).$

As before, (4.21) is suitable for computation for small y, while (5.8) is suitable for computation for $c \leq y \leq a$,.

With this approach, the computational details are as follows. As before, we use the 9-point difference approximation to get approximations for $u_b(x,y)$ at the grid points (mh,nh).

Then a suitable interpolation formula will give accurate approximations for u_b (mh,a). Then we can use the 9-point difference approximation to get approximations for w(x,y) at the grid points (mh, c + nh) ; see (5.3). However, now we are determining w(x,y) to be harmonic, and so avoid having to calculate the right side of (3.26) at a new set of grid points.

As before, we must calculate approximations for a few of the a_r but the details are quite similar.

One difficulty that arises is that we will now be using (5.8) instead of (4.22) to calculate values of u(x,y). It is inconvenient that our approximate values for $u_b(x,y)$ and w(x,y) which appear on the right of (5.8), are known at <u>different</u> grid points. However, by storing a suitable subroutine for interpolation in the computer, the matter can be dealt with. In any given situation, one would have to determine if the extra interpolations are less or more effort than calculating the right side of (3.26) at an additional set of grid points.

6. <u>Tests for accuracy</u>.

One advantage of using the 9-point difference approximation when one can exactly fill up the ractangle with squares is that one can make a first calculation, for less than a quarter of the

calculating effort, with squares twice as large on a side, and then repeat with the smaller squares. Because the error is of the order of h^6, one can get an estimate of the error.

This can be done with the present procedure by choosing M divisible by 2. If N is not divisible by 2, the values of b and c which are used with the squares of side 2h will not be the same as those which are used with the squares of side h. However, this does not matter.

One dividend that will accrue from making an initial calculation with squares of side 2h is that from this calculation one can derive a very good approximation to take for $h_b(x)$. Then, for the calculation with squares of side h, the a_r will be very small, so that not more than two or three of them will be needed.

7. Neumann boundary conditions.

Suppose we have the same rectangle R, and impose on u(x,y) the same conditions as before, except that on top of the rectangle R we specify values to be taken by $u_y(x,a)$. That is we replace (2.5) by the Neumann condition

$$(7.1) \qquad u_y(x,a) - k_a(x) \qquad 0<x<\pi.$$

We postpone to the latter part of the section a discussion

of how one would handle this in the case in which a/π is rational, so that one can fill up R exactly with squares of side h. For the moment, let us assume that this can be done, and explain how to generalize to the case in which a/π is irrational.

We proceed very nearly as in Section 4. Instead of the definition given there of $h_b(x)$, we use

$$(7.2) \qquad h_b(x) = (1-\frac{x}{\pi})\, g_0(b) + \frac{x}{\pi}\, g_\pi(b) \ .$$

We take $u_b(x,y)$ as before, but for $u_c(x,y)$ we replace (4.15) by the analogue of (7.1), namely

$$(7.3) \qquad \frac{\partial}{\partial y}\, u_c(x,a) = k_a(x) \qquad 0<x<\pi \ .$$

Everything now goes the same, down to the definition of $v(x,y)$. Let us pause a moment, and think what we require of $v(x,y)$. Clearly it should be harmonic, so that $u(x,y)$, as defined in part by *(4.21)* and in part by (4.22), will satisfy (2.1) inside R. Also, we wish $v(x,y)$ to be zero on the vertical sides of R, so that there $u(x,y)$ will satisfy the proper boundary conditions. Also, on the bottom of R, we must have

$$(7.4) \qquad v(x,0) = \sum_{r=1}^{\infty} a_r \frac{\sinh r\, c}{\sinh r(b-c)} \sin r\, x \qquad 0 < x < \pi$$

so that by (4.21) u(x,y) will satisfy the right boundary conditions on the bottom of R. Finally, looking at (4.22), we see that if u(x,y) is to satisfy the right boundary conditions on the top of R, we must have

$$(7.5) \qquad v_y(x,a) = 0 \qquad 0<x<\pi.$$

All these conditions can be met by simply replacing the factor

$$\frac{\sinh r(a-y)}{\sinh r\, a}$$

in the definition of v(x,y) by

$$\frac{\cosh r(a-y)}{\cosh r\, a} \ .$$

In this case, since it is unlikely that (7.2) makes $h_b(x)$ come out very close to u(x,b), we cannot count on the a_r being particularly small, so that two or three more of them might have to be calculated. It might be better to turn the rectangle R upside down and proceed as follows.

Consider next the case in which the Neumann condition is at the bottom of R. That is, u(x,y) satisfies (2.2), (2.3), and (2.5), but (2.4) is replaced by

(7.6) $u_y(x,0) = k_o(x) \qquad 0<x<\pi.$

Again, we proceed nearly as in Section 4. We can now take $h_b(x)$ the same as in Section 4, which should lead to smaller values of the a_r, so that we can get by with calculating fewer of then. For the definition of $u_b(x,y)$, we replace (4.10) by the analogue of (7.6), namely

(7.7) $\frac{\partial}{\partial y} u_b(x,o) = k_o(x) \qquad 0 < x < \pi.$

We take $u_c(x,y)$ as in Section 4, and continue the same down to the definition of $v(x,y)$. A key requirement is that u(x,y), as defined by (4.21), shall satisfy the proper boundary conditions at the bottom of R. In Section 4, this required that

(7.8) $v(x,y) + \sum_{r=1}^{\infty} a_r \frac{\sinh r(y-c)}{\operatorname{Sinh} r(b-c)} \sin r x$

should be zero when y=0. This was accomplished by the proper

choice of the b_r . Now we must assume that the partial derivative of (7.8) with respect to y shall be zero when y=0. Again, this is accomplished by the proper choice of the b_r ; specifically we now take

$$(7.9) \qquad b_r = \frac{\sinh\ r\ a}{\sinh\ r\ (b-c)} \quad \frac{\cosh\ r\ c}{\cosh\ r\ a} .$$

All else remains the same.

Next consider the case in which there are Neumann conditions both at the top and the bottom of R. That is, u(x,y) satisfies (2.2) and (2.3), but (2.4) is replaced by (7.6) and (2.5) is replaced by (7.l). We proceed much as in Section 4. In the definition of u_b (x,y) we replace (4.10) by (7.7), and in the definition of u_c (x,y) we replace (4.15) by (7.3). We define h_b (x) by (7.2). It is then easily verified that we should replace

$$\frac{\sinh\ r(a-y)}{\sinh\ r\ a}$$

in the definition of v(x,y) by

$$\frac{\cosh\ r\,(a-y)}{\cosh\ r\ a}$$

and define

$$(7.10) \qquad b_r = \frac{\cosh r\,a}{\sinh r(b-c)} \qquad \frac{\cosh r\,c}{\sinh r\,a} .$$

One can of course have Neumann conditions on one or both of the vertical sides. Let us consider first the case in which there are Neumann conditions on both vertical sides, but Dirichlet conditions at the top and bottom. Rotation by 90° would reduce this to the case just considered. However, this is not desirable, since we would then lose the qualification that the height is greater than the base. It was this that assured the rapid convergence of the Fourier series in (4.19) and (4.21).

So we assume that (2.4) and (2.5) hold, but that (2.2) and (2.3) are replaced by

$$(7.11) \qquad u_x(o,y) = j_o(y) \qquad 0<y<a$$

$$(7.12) \qquad u_x(\pi,y) = j_\pi(y) \qquad 0<y<a.$$

We proceed analogously to Section 4, except that we use cosines instead of sines throughout. Because it is desirable to have $u_x(x,y)$ continuous around the boundary we define

$$(7.13)\qquad h_b(x) = h_a(x) + \frac{1}{2\pi}(x-\pi)^2\left(h_a'(0) - j_0(b)\right) + \frac{x^2}{2\pi}\left(j_\pi(b) - h_a'(\pi)\right).$$

We define $u_b(x,y)$ and $u_c(x,y)$ as in Section 4, except that they now have Neumann conditions on their vertical sides. We replace (4.16) and (4.17) by

$$(7.14)\qquad u_c(x,y) - u_b(x,y) = \sum_{r=0}^{\infty} a_r \frac{\sinh r(y-c)}{\sinh r(b-c)} \cos r x$$

Where

$$(7.15)\qquad a_0 = \frac{1}{\pi}\int_0^{\pi} \left\{u_c(x,b) - u_b(x,b)\right\} dx$$

$$(7.16)\qquad a_r = 2\pi \int_0^{\pi} \left\{u_c(x,b) - u_b(x,b)\right\} \cos r x \, dx.$$

When r=0, we define

$$\frac{\sinh r(y-c)}{\sinh r(b-c)} = \frac{y-c}{b-c}.$$

Exactly analogous changes are made in (4.19) and (4.21).

If, in addition to the Neumann conditions on the vertical sides, we replace one or both of the Dirichlet conditions on the top or bottom by Neumann conditions, we can modify the

procedure just outlined quite analogously to the way in which we modified the procedure of Section 4 earlier in this section.

It will he noted that we are allowing the possibility of Neumann conditions on all four sides. In this case, the solution is not unique. However, any two solutions differ by a constant. The procedure outlined produces one of the infinity of solutions.

To handle the case of a Dirichlet condition on the left side and a Neumann condition on the right side, we replace sin r x by

$$\sin \left(r-\frac{1}{2}\right) x ,$$

with suitable related changes. To handle the case of a Dirichlet condition on the right side and a Neumann condition on the left side, we replace sin r x by

$$\text{con} \left(r-\frac{1}{2}\right) x ,$$

When we had Dirichlet conditions on all sides, we were at some pains in Section 2 to ensure that these would be continuous all around the boundary, and that they would have continuous derivatives except for jumps at the corners. This

assured that the first partial derivatives of u(x,y) would be continuous throughout. When there are Neumann conditions on some sides, attention must be paid to the continuity of the first partial derivatives around the boundary. As far as the parts of the boundary where there are Dirichlet conditions, the procedures given in Section 2 for the "removal" of discontinuities are still applicable. They do not introduce discontinuities of u(x,y) or its partial derivatives at other spots; indeed, at the points where a discontinuity of u(x,y) or one of its partial derivatives is being "removed", the attendant discontinuity in the other partial derivative is also "removed".

We turn our attention to the "removal" of jump discontinuities in the Neumann conditions. Suppose we have the condition (7.1) along the top of R, and $k_a(x)$ is discontinuous. Specifically, let $0 < p < \pi$, and let the limits $k_a(p\text{-}0)$ and $k_a(p+0)$ exist and be different. We introduce the function

$$u^*(x,y) = Q\, s(x,y) + u(x,y)$$

Where

$$s(x,y) = (y-a)\ \arctan\frac{\rho-x}{a-y} - \frac{1}{2}(\rho-x)\ \log\left\{(a-y)^2 + (\rho-x)^2\right\}$$

and

$$Q = \frac{1}{\pi} (k_a (\rho+0) - k_a (\rho-0)).$$

The other sort of discontinuity that might occur would be in case

$$g_o' (a) \neq k_a (0).$$

We introduce the function

$$u^* (x,y) = Q\, s(x,y) + u(x,y)$$

where

$$s(x,y) = (y-a) \arctan \frac{a-y}{x} + \frac{1}{2} x \log \left\{ (a-y)^2 + x^2 \right\}$$

and

$$Q = \frac{2}{\pi} (k_a (0) - g_o' (a)) .$$

We consider finally how to handle the case in which the

rectangle has a rational ratio of the sides, and we have filled it exactly with squares of side h, and wish to approximate u(x,y) at the grid points. At interior grid points, we can use (3.26), (3.27), or something of the sort. On boundaries where there are Dirichlet boundary conditions, we assign $\bar{u}_{m,n}$ the specified value. This leaves only the boundary points where there is a Neumann condition to be dealt with. Suppose, for example, that the condition (7.11) holds on the left side of R. We note that

$$(7.17)\quad hf_x(x,y) = -\frac{137}{60}f(x,y) + 5f(x+h,y) - 5f(x+2h,y) + \frac{10}{3}f(x+3h,y) - \frac{5}{4}f(x+4h,y) + \frac{1}{5}f(x+5h,y)$$

holds to within terms of order h^6. If we take x = 0 and y = nh, we get by (7.11)

$$(7.18)\quad hj_o(mh) = -\frac{137}{60}\tilde{u}_{0,m} + 5\bar{u}_{1,m} - 5\bar{u}_{2,m} + \frac{10}{3}\bar{u}_{3,m} - \frac{5}{4}\bar{u}_{4,m} + \frac{1}{5}\bar{u}_{5,m}.$$

One could use a higher order formula than (7.17), but it probably suffices. A heuristic argument for this is as follows. By the principle of the maximum, if we wish to determine

interior points to order h^6, it is sufficient to determine the boundary points to order h^6. However, if the interior points are given to order h^6, (7.18) will determine $\bar{u}_{0.m}$ to order h^6.

Use of (7.18) with (3.26) or (3.27) results in a rather messy matrix of coefficients of the $\bar{u}_{m,n}$. However, one is probably using such a coarse mesh that this matrix would be less than 100 x 100, perhaps even less than 50 x 50. If so, probably the quickest method of solution is to use the standard computer routine for solving simultaneous linear equations. If this is done, it does not matter if the matrix is messy or not.

If it happens that one is solving the Laplace equation, with $f(x,y) \equiv 0$, and has a zero normal derivative along one side, say $j_o(y) \equiv 0$, one can use the reflection principle to replace (7.18) by something which seems conceptually simpler. However, it involves three boundary grid points and three interior points, and so is probably about as much bother on a computer as (7.18), which also involves six grid points.

If one has Neumann conditions on one or more sides, and so is using (7.18), one might consider the following procedure, which would bypass the treatment in Section 4 altogether. Almost always, there is at least one side with Dirichlet conditions. By rotating and relinquishing the qualification

$a > \pi$, if need be, we can arrange to have Dirichlet conditions on top. If, in the notation of Section 4, we have $0 < c. < h$, the difficulty is that we have no good way to write down an equivalent of (3.26) for the values of u(x,y) at the row of grid points (mh , Nh), $1 \leq m \leq M-1$. As a substitute, write down (3.26) for the 9-point formula centered at (mh,a-h). It involves values of u(x,y) at ((m-1)h, a-h), ((m-1)h, a-2h), (mh, a-h), (mh, a-2h), ((m+1)h, a-h), ((m+l)h, a-2h), as well as at the boundary points ((m-l)h, a), (mh, a), and ((m+l)h, a), where u(x,y) is known. Nov, by a high order one dimensional interpolation formula, we can write each of u(rh, a-h) and u(rh, a-2h), approximately as a linear combination of u(rh, nh) for $n \leq N$; we do this for r = m-1, r = m, and r = m+1. So we get a formula involving u(rh, Nh), u(rh, (N-1)h), etc., for r = m-1, m, m+1, which we can use in place of (3.26). Probably interpolation of order seven should be used. This might make it impossible for us to use as coarse a mesh as we otherwise could. However, if we are having to deal with a messy matrix anyhow, because of the Neumann conditions, the idea might be worth considering.

REFERENCES

[1] R.W. Hockney, "The potential calculation and some applications," Methods in Computational Physics, vol. 9 (1970), pp. 135-211.

[2] George E. Forsythe and Wolfgang R. Wasow, "Finite-difference method for partial differential equations;" John Wiley and Sons, Inc., New York, 1960.

[3] N. Papamichael and John R. Whiteman, "A numerical conformal transformation method for harmonic mixed boundary value problems in polygonal domains," to appear in Zeit. für Angew. Math. Phys., vol. 24 (1973).

[4] N. Papamichael and John R. Whiteman, "Cubic spline interpolation of harmonic functions," Technical Report TR/28, Brunei University, 1973.

[5] J. Barkley Rosser, "Effect of discontinuous boundary conditions on finite-difference solutions," Technical Report TR/30, Brunei University, 1973, and MRC Technical Summary Report, # 1383, 1973.

[6] L. Fox, Editor, "Numerical solution of ordinary and partial differential equations," Pergamon Press, Oxford 1962.

[7] H.J. Wirz, "Eine Erweiterung des Verfahrens der Zwischenschritte auf allgemeinere parabolische und elliptische Differentialgleichungen," Zeitschrift für angewandte Mathematik und Mechanik, Vol. 52 (1972), pp. 329 - 336.

[8] L. Collatz, "The numerical treatment of differential equations," Springer-Verlag, Berlin, 1960.