ON THE NUMERICAL SOLUTION OF POISSON'S EQUATION OVER A RECTANGLE

P. STEIN AND J. E. L. PECK

Introduction. We consider the equation

$$\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=f(x, y)$$

over the rectangle $0\leq x\leq a$, $0\leq y\leq b$, with given boundary values for z. Following the usual procedure (see for example Hyman [1]) we approximate the solution by solving a set of mn simultaneous equations, arising from the corresponding difference equation. If we write

$$a=(n+1)\Delta x, \quad b=(m+1)\Delta y, \quad \rho=\Delta y/\Delta x, \quad a_{i,j}=-f(j\Delta x, i\Delta y)\Delta y^2$$

and $z_{i,j}=z(j\Delta x, i\Delta y)$, the mn equations are of the form

$$(1) \qquad 2(1+\rho^2)z_{i,j}=\rho^2(z_{i,j+1}+z_{i,j-1})+z_{i+1,j}+z_{i-1,j}+a_{i,j}\,, \qquad i=1,\cdots,m, \quad j=1,\cdots,n\,.$$

A solution of this set of equations is given by Hyman [1]. In the case where the boundary values are zero, the solution takes the form $Z=C\omega D$ [1, p. 340] where C and D are matrices which depend on n and m and may be written down without any calculations, and ω is a matrix depending on m, n, ρ and the values of $f(x, y)$ at the lattice points. The matrix ω requires somewhat elaborate calculations. To obtain the solution with given boundary values, he adds to the matrix $C\omega D$ the value of u as a matrix obtained from the solution of the equation $\Delta^2u/\Delta x^2+\Delta^2u/\Delta y^2=0$ with the given boundary values. He obtains for u the matrix value $U=C\phi$ [1, p. 329], where C is the matrix mentioned above and ϕ is a matrix depending on n, m, ρ and the boundary values and requires to be recalculated for every set of boundary values.

In this paper the solutions of equations (1) are obtained, column by column, in the form $Z_j=\sum_k M_{j,k}B_k$, where the $M_{j,k}$ are matrices depending on m, n, and ρ and which require somewhat elaborate calculations, and the B_k are vectors depending on m, n, ρ, the values of $f(x, y)$ at the lattice points and the boundary values and can be written down without calculation. We may regard this solution as giving an explicit formula for the values of z at the lattice points.

The principal work in the calculation of Z_j is the calculation of the matrices $M_{j,k}$. It will be shown that it is sufficient to calculate a

Received July 3, 1953.

selection of columns of Z, as the method lends itself to a stepping off process; also that all the matrices used can be written down easily from a knowledge of their top rows. The calculation is simplest when $\rho=1$. Further, the case when $\rho=1$ or is nearly 1 is the most accurate [1, p. 332]. It will be shown that when $|\rho^2-1|<1$, Z may be obtained by successive approximations with the help of the matrices calculated for $\rho=1$. It appears to the authors that if a not very elaborate set of tables were to be prepared for selected values of j, m and n with $\rho=1$, the calculation of Z would be much simplified. Further, if such a set of tables were available, it might be of assistance in the iterative method of the solution of these simultaneous equations when the boundary is not a rectangle.

In § 1 we develop the method of solution. In §§ 2 and 3 we give methods by which the required matrices may be evaluated. Section 4 deals with the iterative process when ρ is nearly 1, and this is amplified in §§ 5 and 6.

1. We write the mn equations (1) in n sets each consisting of m equations. A typical set is

$$
\begin{aligned}
&2(1+\rho^2)z_{1,j}-z_{2,j} &&=\rho^2(z_{1,j+1}+z_{1,j-1})+a_{1,j}+z_{0,j}\\
&-z_{1,j}+2(1+\rho^2)z_{2,j}-z_{3,j} &&=\rho^2(z_{2,j+1}+z_{2,j-1})+a_{2,j}\\
&\cdot\qquad\cdot\qquad\cdot\qquad\cdot\qquad\cdot\qquad\cdot\\
&-z_{m-1,j}+2(1+\rho^2)z_{m,j} &&=\rho^2(z_{m,j+1}+z_{m,j-1})+a_{m,j}+z_{m+1,j}\,.
\end{aligned}
\tag{2}
$$

We write Z_j for the vector $(z_{1,j}, z_{2,j}, \cdots, z_{m,j})$, A_j for the vector $(a_{1,j}, a_{2,j}, \cdots, a_{m,j})$, Z_j' for the vector $(z_{0,j}, 0, 0, \cdots, 0, z_{m+1,j})$ and $M_m(\alpha)$ for the $m\times m$ matrix

$$
\begin{pmatrix}
\alpha, & -1, & 0, & \cdot\\
-1, & \alpha, & -1, & \cdot\\
0, & -1, & \alpha, & \cdot\\
\cdot & \cdot & \cdot & \cdot
\end{pmatrix}.
\tag{3}
$$

The equations (2) then take the form

$$M_m(2+2\rho^2)Z_j=\rho^2(Z_{j+1}+Z_{j-1})+A_j+Z_j'\,, \qquad j=1,\cdots,n\,,$$

or

$$
\begin{aligned}
&\rho^{-2}M_m(2+2\rho^2)Z_1-Z_2 &&=\rho^{-2}(A_1+Z_1')+Z_0 &&=B_1\\
&-Z_1+\rho^{-2}M_m(2+2\rho^2)Z_2-Z_3 &&=\rho^{-2}(A_2+Z_2') &&=B_2\\
&\cdot\qquad\cdot\qquad\cdot\qquad\cdot\qquad\cdot\qquad\cdot\\
&-Z_{n-1}+\rho^{-2}M_m(2+2\rho^2)Z_n &&=\rho^{-2}(A_n+Z_n')+Z_{n+1} &&=B_n\,.
\end{aligned}
\tag{4}
$$

These equations can be solved by iteration. See for example Todd [3].

The class of all ordered sets of m real numbers is a vector space over the ring of polynomials in the matrix $M_m(2+2\rho^2)$. Interpreting equations (4) in this way, we may obtain their solution from Cramer's rule in the form

$$\mathscr{D} Z_j = \sum_{k=1}^{n} \mathscr{M}_{j,k} B_k \tag{5}$$

where $\mathscr{D}$ is the determinant of the matrix of matrix coefficients on the left of (4) and the $\mathscr{M}_{j,k}$ are cofactors of $\mathscr{D}$. One may readily prove that $\mathscr{M}_{j,k} = \mathscr{M}_{k,j}$ and that when $j \leq k$

$$\mathscr{M}_{j,k} = D_{j-1}(\rho^{-2} M_m(2+2\rho^2)) D_{n-k}(\rho^{-2} M_m(2+2\rho^2)) \tag{6}$$

where D_n is the polynomial defined by the nth order determinant

$$D_n(x) = \begin{vmatrix} x, & -1, & 0, & \cdot \\ -1, & x, & -1, & \cdot \\ 0, & -1, & x, & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix}$$

and $D_0(x) = 1$.

One may regard (5) as expressing z in terms of the given values for $f(x, y)$ and the boundary values. In particular when $j=1$, we have from (5) and (6)

$$Z_1 = D_n^{-1}(\rho^{-2} M_m(2+2\rho^2)) \sum_{k=1}^{n} D_{n-k}(\rho^{-2} M_m(2+2\rho^2)) B_k . \tag{7}$$

As was pointed out by Hyman [1, p. 331] it is unnecessary to calculate the remaining values of z by the use of (5). It is sufficient to use (7). Knowing Z_0 and Z_1 we may "step off" using (4) to determine Z_2 and then use it again to get Z_3 from Z_2 and Z_1.

2. In this section we obtain some properties of the polynomial D_n and of the matrices $D_n(\beta M_m(\alpha))$.

THEOREM 1.

$$D_n(x) = x D_{n-1}(x) - D_{n-2}(x) \tag{8}$$

$$D_n(x) = \sum_{r=0}^{[n/2]} (-1)^r \binom{n-r}{r} x^{n-2r} \tag{9}$$

$$D_n(x) = \frac{a^{n+1} - b^{n+1}}{2^n (a-b)}, \qquad a = x + \sqrt{x^2-4}, \qquad b = x - \sqrt{x^2-4} \tag{10}$$

(11) $$D_n(x)=2^{-n}\sum_{r=0}^{[n/2]}\binom{n+1}{2r+1}x^{n-2r}y^{2r}\,, \qquad y=\sqrt{x^2-4}$$

(12) $$D_n(x)=\frac{\sinh (n+1)\phi}{\sinh \phi}\,, \qquad x=2\cosh\phi$$

(13) $$D_n(x)=\frac{\sin (n+1)\theta}{\sin\theta}\,, \qquad x=2\cos\theta$$

(14) $$D_n(x)=\prod_{r=1}^{n}\left(x-2\cos\frac{r\pi}{n+1}\right).$$

Formulae (8), (13) and (14) are known ([2] and [4]). Formula (8) follows immediately from the definition and (9) may be proved by induction using (8). Formula (10) also follows from (8) by induction. Formula (11) comes from (10) on writing $a=x+y$, $b=x-y$. The equation $x=2\cosh\phi$ means that $a=2e^{\phi}$, $b=2e^{-\phi}$ whence (10) gives (12). Formula (13) is proved similarly. By (13) the roots of the equation $D_n(x)=0$ are $2\cos(r\pi/(n+1))$, $(r=1,\cdots,n)$ giving (14).

COROLLARY. *If M is a square matrix and I is the corresponding identity matrix:—*

(15) $$D_n(M)=MD_{n-1}(M)-D_{n-2}(M)$$

(16) $$D_n(M)=\sum_{r=0}^{[n/2]}(-1)^r\binom{n-r}{r}M^{n-2r}$$

(17) $$D_n(M)=2^{-n}\sum_{r=0}^{[n/2]}\binom{n+1}{2r+1}M^{n-2r}(M^2-4I)^r$$

(18) $$D_n(M)=\prod_{r=1}^{n}\left(M-2\cos\frac{r\pi}{n+1}I\right).$$

THEOREM 2. *If P is any polynomial, then $P(M_m(\alpha))$ is an $m\times m$ matrix which is symmetric about both diagonals.*

If two matrices which commute are symmetric about both diagonals, then so is their sum, product and any scalar multiple. This theorem therefore proves that the matrices $D_n(\beta M_m(\alpha))$ are symmetric about both diagonals.

THEOREM 3. *Let P be any polynomial and let $a_{i,j}$ be the elements of $P(M_m(\alpha))$. Then if we interpret $a_{i,j}=0$ whenever i or j is <1 or $>m$, we have, for $1\leq i\leq m$, $1\leq j\leq m$ and $i+j\leq m+2$,*

(19) $$a_{i,j}=a_{i-1,j-1}+a_{1,i+j-1}\,,$$

and for $1 \leq i \leq j \leq m$ *and* $i+j \leq m+1$,

$$a_{i,j}=a_{1,-i+j+1}+a_{1,-i+j+3}+\cdots+a_{1,i+j-1}\,. \tag{20}$$

Theorems 2 and 3 enable us to write down all the elements of $P(M_m(\alpha))$ from a knowledge of the elements in the first row.

Proof of Theorem 3. We observe that (19) is invariant under addition and scalar multiplication. We observe also that in the case $i=1$, (19) reduces to a triviality. If $j=1$ it becomes $a_{i,1}=a_{1,i}$ which is true by symmetry about the main diagonal, and if $i+j=m+2$, it becomes $a_{i,j}=a_{i-1,j-1}$, which is true because of symmetry about the other diagonal (Theorem 2). Formula (19) will therefore be established if we can show that it is true for $M_m^r(\alpha)$, where r is a nonnegative integer, and when $2 \leq i \leq m$, $2 \leq j \leq m$ and $i+j \leq m+1$.

By inspection it is true when $r=0$, 1. Let $a_{i,j}^r$ denote the i, jth element of $M_m^r(\alpha)$. Then $a_{i,j}^r=-a_{i,j-1}^{r-1}+\alpha a_{i,j}^{r-1}-a_{i,j+1}^{r-1}$ for $1 \leq i \leq m$, $1 \leq j \leq m$. If we assume that it is true for $r-1$ we have for $2 \leq i \leq m$, $2 \leq j \leq m$, $i+j \leq m+1$, that

$$\begin{aligned}a_{i,j}^r=&-a_{i-1,j-2}^{r-1}+\alpha a_{i-1,j-1}^{r-1}-a_{i-1,j}^{r-1}\\&-a_{1,i+j-2}^{r-1}+\alpha a_{1,i+j-1}^{r-1}-a_{1,i+j}^{r-1}=a_{i-1,j-1}^{r}+a_{1,i+j-1}^{r}\end{aligned}$$

which completes the proof of (19).

Formula (20) follows from a repeated use of (19).

THEOREM 4. *If we denote the* i, j*th element of* $D_n(\beta M_m(\alpha))$ *by* $a_{i,j}^{m,n}$ *then*

$$a_{1,j}^{k,n}=a_{1,j}^{m,n}, \qquad 1 \leq j \leq k\,, \qquad n \leq k \leq m\,. \tag{21}$$

From (15) we have

$$a_{1,j}^{m,n}=-\beta a_{1,j-1}^{m,n-1}+\alpha\beta a_{1,j}^{m,n-1}-\beta a_{1,j+1}^{m,n-1}-a_{1,j}^{m,n-2} \tag{22}$$

where $a_{1,0}^{m,n}=a_{1,m+1}^{m,n-1}=0$ and $1 \leq j \leq m$. From (22) we have by induction on n that $a_{1,j}^{m,n}=0$ if $j \geq n+2$, which means that $a_{1,k+1}^{m,n-1}=0$. Since we must write $a_{1,0}^{m,n-1}=a_{1,0}^{k,n-1}=a_{1,k+1}^{k,n-1}=0$, this allows the following induction on n

$$\begin{aligned}a_{1,j}^{m,n}&=-\beta a_{1,j-1}^{m,n-1}+\alpha\beta a_{1,j}^{m,n-1}-\beta a_{1,j+1}^{m,n-1}-a_{1,j}^{m,n-2}\\&=-\beta a_{1,j-1}^{k,n-1}+\alpha\beta a_{1,j}^{k,n-1}-\beta a_{1,j+1}^{k,n-1}-a_{1,j}^{k,n-2}=a_{1,j}^{k,n}\,.\end{aligned}$$

However the theorem is true for $n=1$ and $n=2$.

Formula (21) shows that the top row of the matrix $D_n(\beta M_m(\alpha))$ is essentially the same for all useful values of n, while (22) gives a recursive method of computing this top row. Theorem 3 and the remark after Theorem 2 show how the remainder of the matrix can be filled in from the top row. Thus the computation of the matrices $D_n(\beta M_m(\alpha))$

for $n \leq m$ is simplified, and the matrices lend themselves to easy tabulation.

3. **In this section** we give some results which are useful in the calculation of $D_n^{-1}(\beta M_m(\alpha)) = [D_n(\beta M_m(\alpha))]^{-1}$.

Since the inverse of any matrix is a polynomial in that matrix, we have $D_n^{-1}(\beta M_m(\alpha))$ a polynomial in $D_n(\beta M_m(\alpha))$ and therefore a polynomial in $M_m(\alpha)$. Theorems 2 and 3 therefore apply. It is thus sufficient to compute only its first row. From the first row we may obtain its elements $a_{i,j}$ for $1 \leq i \leq j \leq m$ and $i+j \leq m+1$ by (19) or (20) and then the other elements can be filled in by symmetry.

THEOREM 5. *If the element in the ith row and jth column of* $M_m^{-1}(\alpha)$ *is* $a_{i,j}$ *and if* $\alpha = 2\cosh\phi$, *then*

$$a_{i,j} = \frac{\sinh i\phi \sinh (m+1-j)\phi}{\sinh\phi \sinh (m+1)\phi}, \qquad i \leq j. \tag{23}$$

For proof we have first that

$$a_{i,j} = \frac{D_{i-1}(\alpha) D_{m-j}(\alpha)}{D_m(\alpha)}, \qquad i \leq j. \tag{24}$$

The result then follows from (12).

THEOREM 6. *If*

$$\alpha_r = \alpha - \frac{2}{\beta} \cos \frac{r\pi}{n+1},$$

then

$$D_n^{-1}(\beta M_m(\alpha)) = \beta^{-n} \prod_{r=1}^{n} M_m^{-1}(\alpha_r). \tag{25}$$

From (18) we have

$$\begin{aligned} D_n(\beta M_m(\alpha)) &= \prod_{r=1}^{n} \left(\beta M_m(\alpha) - 2\cos \frac{r\pi}{n+1} I\right) \\ &= \beta^n \prod_{r=1}^{n} M_m\left(\alpha - \frac{2}{\beta} \cos \frac{r\pi}{n+1}\right) \\ &= \beta^n \prod_{r=1}^{n} M_m(\alpha_r). \end{aligned} \tag{26}$$

The result follows immediately.

A result which may be easier from the computational point of view is to express $D_n^{-1}(\beta M_m(\alpha))$ as a sum of matrices. This is done in the following theorem.

THEOREM 7. *If*

$$\alpha_r = \alpha - \frac{2}{\beta}\cos\frac{r\pi}{n+1},$$

then

$$D_n^{-1}(\beta M_m(\alpha)) = \frac{2}{\beta(n+1)}\sum_{r=1}^{n}(-1)^{r+1}\sin^2\frac{r\pi}{n+1}\, M_m^{-1}(\alpha_r). \tag{27}$$

From (26) we have that

$$D_n(\beta M_m(\alpha)) = \prod_{r=1}^{n}\beta M_m(\alpha_r).$$

Therefore

$$D_n^{-1}(\beta M_m(\alpha)) = \sum_{r=1}^{n} c_r\{\beta M_m(\alpha_r)\}^{-1}$$

where the c_r's are suitably chosen scalars.

If $f(x) = \Pi_{r=1}^{n}(x-\gamma_r)$, $\gamma_r \neq \gamma_s$ when $r \neq s$, then

$$f(x)^{-1} = \sum_{r=1}^{n} f'(\gamma_r)^{-1}(x-\gamma_r)^{-1}.$$

To obtain the values of the scalars c_r, we put $f = D_n$. From (13) we have

$$D_n'(2\cos\theta) = \frac{(n+1)\cos(n+1)\theta\sin\theta - \sin(n+1)\theta\cos\theta}{\sin^2\theta}\left(\frac{-1}{2\sin\theta}\right).$$

This gives

$$D_n'\left(2\cos\frac{r\pi}{n+1}\right) = \frac{(-1)^{r+1}(n+1)}{2\sin^2\frac{r\pi}{n+1}},$$

and therefore

$$c_r = (-1)^{r+1}2(n+1)^{-1}\sin^2\frac{r\pi}{n+1}.$$

With the help of (23) we can obtain a more explicit result.

COROLLARY. *If* $a_{i,j}$ *is the* i, j*th element of* $D_n^{-1}(\beta M_m(\alpha))$ *and*

$$2\cosh\phi_r = \alpha - \frac{2}{\beta}\cos\frac{r\pi}{n+1},$$

then $a_{i,j} = a_{j,i}$ *for all* i *and* j, *and for* $i \leq j$

$$(28)\quad a_{i,j}=[\beta(n+1)]^{-1}\sum_{r=1}^{n}(-1)^{r+1}\left(1-\cos\frac{2r\pi}{n+1}\right)\frac{\sinh i\phi_r \sinh(m+1-j)\phi_r}{\sinh\phi_r \sinh(m+1)\phi_r}\ .$$

In the case $i=1$, this reduces to

$$(29)\qquad a_{1,j}=[\beta(n+1)]^{-1}\sum_{r=1}^{n}(-1)^{r+1}\left(1-\cos\frac{2r\pi}{n+1}\right)\frac{\sinh(m+1-j)\phi_r}{\sinh(m+1)\phi_r}\ .$$

4. In the formulae of the two previous sections, if $\rho=\Delta y/\Delta x=1$ we have $\alpha=4$, $\beta=1$ and there is some simplification in the resulting calculations. It is pointed out by Hyman [1, p. 322] that the case $\rho=1$ is the one which gives the most accurate results. Hence it is suggested that in arranging the lattice points of a rectangle, one should attempt to have ρ approximately one. We now give a method of finding a correction, when ρ is approximately one, to the solution obtained by assuming $\rho=1$. It is found that in this way we can make use of tables prepared for the case $\rho=1$.

We write $\rho^2=1+\delta$. The equations (1) then become

$$(30)\qquad (4+2\delta)z_{i,j}=(1+\delta)(z_{i,j+1}+z_{i,j-1})+(z_{i+1,j}+z_{i-1,j})+a_{i,j}\ .$$

Let

$$\Delta x_{i,j}=2x_{i,j}-x_{i,j+1}-x_{i,j-1}$$
$$\Box x_{i,j}=4x_{i,j}-x_{i+1,j}-x_{i-1,j}-x_{i,j+1}-x_{i,j-1}\ .$$

Then (30) may be written

$$(31)\qquad \Box z_{i,j}=-\delta\Delta z_{i,j}+a_{i,j}\ .$$

We suppress the first term on the right of (31) and find $u_{i,j}^{(1)}$ so that

$$(32)\qquad \Box u_{i,j}^{(1)}=a_{i,j}$$

and $u_{i,j}^{(1)}=z_{i,j}$ on the boundaries. Let Z denote the values of $z_{i,j}$ at the lattice points, with similar notation for $U^{(r)}$ and $V^{(r)}$ with $r\geq 1$. Let $Z=U^{(1)}+V^{(1)}$, then $U^{(1)}$ is an approximation to the values of Z with error $V^{(1)}$ for which an equation is obtained by subtracting (32) from (31). Thus

$$\Box v_{i,j}^{(1)}=-\delta\Delta z_{i,j}=-\delta\Delta v_{i,j}^{(1)}-\delta\Delta u_{i,j}^{(1)}$$

and $V^{(1)}$ is zero on the boundary.

We now find $U^{(2)}$ such that

$$\Box u_{i,j}^{(2)}=-\delta\Delta u_{i,j}^{(1)}$$

and $U^{(2)}$ is zero on the boundary. Writing $V^{(1)}=U^{(2)}+V^{(2)}$ we obtain, by subtraction,

$$\Box v_{i,j}^{(2)} = -\delta \Delta v_{i,j}^{(1)} .$$

Proceeding in this manner we obtain for $r \geq 1$

$$V^{(r)} = V^{(r+1)} + U^{(r+1)}$$

(33)
$$\Box u_{i,j}^{(r+1)} = -\delta \Delta u_{i,j}^{(r)}$$

(34)
$$\Box v_{i,j}^{(r+1)} = -\delta \Delta v_{i,j}^{(r)}$$

where $V^{(r)}$ and $U^{(r)}$, $r \geq 2$ are zero on the boundary. A formal solution of equations (30) is thus

(35)
$$Z = \sum_{r=1}^{\infty} U^{(r)} .$$

We observe that equations (32) and (33) to determine $U^{(r)}$ are the equations (1) where $\rho = 1$ and where different sets of values are successively used in place of the $a_{i,j}$. The formal solution (35) will be the solution provided $V^{(r)}$ tends to 0 as r tends to ∞. This will certainly be the case if, given any arbitrary $X^{(1)}$ we can show that the iteration

$$\Box x_{i,j}^{(r+1)} = -\delta \Delta x_{i,j}^{(r)}$$

leads to the result $X^{(r)} \to 0$ as $r \to \infty$. In the next two sections we obtain the condition on δ that this should be the case, and we obtain an estimate of the error if we take $Z = \sum_{r=1}^{s} U^{(r)}$.

5. **We proceed to** the solution of (33) (and (34)) when $r \geq 2$. The equation may be written

$$-u_{i,j+1}^{(r+1)} + 4u_{i,j}^{(r+1)} - u_{i,j-1}^{(r+1)} = u_{i+1,j}^{(r+1)} + u_{i-1,j}^{(r+1)} - \delta(-u_{i,j+1}^{(r)} + 2u_{i,j}^{(r)} - u_{i,j-1}^{(r)}) .$$

If $U_i^{(r)}$ is the vector $(u_{i,1}^{(r)}, u_{i,2}^{(r)}, \cdots, u_{i,n}^{(r)})$ then since all boundary values are zero, these n equations can be written

(36)
$$M_n(4)U_i^{(r+1)} = U_{i+1}^{(r+1)} + U_{i-1}^{(r+1)} - \delta M_n(2) U_i^{(r)} .$$

If now $\mathscr{M}$ is the $m \times m$ matrix of matrices defined by

$$\mathscr{M} = \begin{pmatrix} M_n(4), & -I, & 0, & \cdot \\ -I, & M_n(4), & -I, & \cdot \\ 0, & -I, & M_n(4), & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

$\mathscr{M}^*$ is the $m \times m$ matrix of matrices defined by

$$\mathscr{M}^* = \begin{pmatrix} M_n(2), & 0, & 0, & \cdot \\ 0, & M_n(2), & 0, & \cdot \\ 0, & 0, & M_n(2), & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

and $U^{(r)}$ is the vector $(U_1^{(r)},\cdots,U_m^{(r)})$, then since all boundary values are zero the m equations (36) can be written

$$\mathscr{M}U^{(r+1)}=-\delta\mathscr{M}^*U^{(r)}$$

and so

$$U^{(r+1)}=-\delta\mathscr{M}^{-1}\mathscr{M}^*U^{(r)}. \tag{37}$$

In the case $r=1$, we must take into account some boundary values. Thus let $Z'=(Z_1',Z_2',\cdots,Z_m')$ where Z_i' $(i=1,\cdots,m)$ is the vector $(z_{i,0},0,\cdots,0,z_{i,n+1})$. Then the solution of (33) for $r=1$ is

$$U^{(2)}=-\delta\mathscr{M}^{-1}(\mathscr{M}^*U^{(1)}-Z'). \tag{38}$$

Returning to (37) we have

$$U^{(r)}=(-\delta\mathscr{M}^{-1}\mathscr{M}^*)^{r-2}U^{(2)}.$$

Hence $U^{(r)}$ and $V^{(r)}$ tend to zero as r tends to ∞ provided that a circle of radius $|\delta|^{-1}$ and center the origin contains the spectrum of $\mathscr{M}^{-1}\mathscr{M}^*$.

The spectrum of $\mathscr{M}^{-1}\mathscr{M}^*$ is found most easily by considering the matrix $\mathscr{M}^{*-1}\mathscr{M}$. Writing $M=M_n(2)$, we have

$$\mathscr{M}^{*-1}\mathscr{M}$$

$$=\begin{pmatrix} M^{-1}, & 0, & 0, & \cdot \\ 0, & M^{-1}, & 0, & \cdot \\ 0, & 0, & M^{-1}, & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}\cdot\begin{pmatrix} M+2I, & -I, & 0, & \cdot \\ -I, & M+2I, & -I, & \cdot \\ 0, & -I, & M+2I, & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

$$=\begin{pmatrix} I+2M^{-1}, & -M^{-1}, & 0, & \cdot \\ -M^{-1}, & I+2M^{-1}, & -M^{-1}, & \cdot \\ 0, & -M^{-1}, & I+2M^{-1}, & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

If $\{\mu_r,\ r=1,\cdots,n\}$ is the spectrum of M, we may use a theorem of Williamson [5, Theorem 1] to find that the spectrum of $\mathscr{M}^{*-1}\mathscr{M}$ consists of the spectra of the n $m\times m$ matrices

$$\begin{pmatrix} 1+2\mu_r^{-1}, & -\mu_r^{-1}, & 0, & \cdot \\ -\mu_r^{-1}, & 1+2\mu_r^{-1}, & -\mu_r^{-1}, & \cdot \\ 0, & -\mu_r^{-1}, & 1+2\mu_r^{-1}, & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}=\mu_r^{-1}M_m(\mu_r+2).$$

By (14) this is the set $\mu_r^{-1}\left(\mu_r+2+2\cos\dfrac{s\pi}{m+1}\right)$, $r=1,\cdots,n;\ s=1,\cdots,m$.

However by (14) also $\mu_r = 2+2\cos\dfrac{r\pi}{n+1}$, $r=1, \cdots, n$. Thus the spectrum of $\mathscr{M}^{-1}\mathscr{M}^*$ is

$$\frac{2+2\cos\dfrac{r\pi}{n+1}}{4+2\cos\dfrac{r\pi}{n+1}+2\cos\dfrac{s\pi}{m+1}} = \frac{1}{1+\dfrac{\sin^2\dfrac{s\pi}{2(m+1)}}{\sin^2\dfrac{r\pi}{2(n+1)}}}, \qquad \begin{array}{l} r=1, \cdots, n. \\ s=1, \cdots, m. \end{array}$$

This spectrum therefore lies in the open interval (0, 1). $Z=\sum_{r=1}^{\infty} U^{(r)}$ is thus a solution of (1) if

$$|\rho^2-1| = |\delta| < 1+\frac{\sin^2\dfrac{\pi}{2(m+1)}}{\sin^2\dfrac{n\pi}{2(n+1)}}$$

and certainly if $|\delta| \leq 1$.

6. **We shall now** estimate the error if we take $Z=\sum_{r=1}^{s} U^{(r)}$. We suppose that $|\delta| < 1$, and consider first the case $s \geq 2$. Since the spectrum of $\mathscr{M}^{-1}\mathscr{M}^*$ lies in the open interval $(-1, 1)$, using (37) we have

$$\begin{aligned} (39) \qquad Z-\sum_{r=1}^{s} U^{(r)} &= \sum_{r=s+1}^{\infty} U^{(r)} = \sum_{r=1}^{\infty} (-\delta\mathscr{M}^{-1}\mathscr{M}^*)^r U^{(s)} \\ &= -\delta\mathscr{M}^{-1}\mathscr{M}^*(I+\delta\mathscr{M}^{-1}\mathscr{M}^*)^{-1}U^{(s)}. \end{aligned}$$

Now the spectrum of the matrix $(I+\delta\mathscr{M}^{-1}\mathscr{M}^*)^{-1}$ is

$$\left[1+\delta\left\{1+\frac{\sin^2\dfrac{s\pi}{2(m+1)}}{\sin^2\dfrac{r\pi}{2(n+1)}}\right\}^{-1}\right]^{-1}, \qquad r=1, \cdots, n;\ s=1, \cdots, m.$$

If $\delta > 0$ this lies within a circle of radius one, while if $\delta < 0$ it lies within a circle of radius $(1+\delta)^{-1} = (1-|\delta|)^{-1}$. Therefore we obtain from (39)

$$\begin{aligned} (40) \qquad \left\| Z-\sum_{r=1}^{s} U^{(r)} \right\| &< |\delta|(1-|\delta|)^{-1}\, \|U^{(s)}\|, \qquad &\text{when} \quad \delta < 0, \\ &< |\delta|\, \|U^{(s)}\|, \qquad &\text{when} \quad \delta > 0[1]. \end{aligned}$$

[1] The norm $\|T\|$ of a matrix is $\sup \|Tx\|/\|x\|$, where $\|x\|$ is the square root of the sum of the squares of the coordinates of the vector x. If T is symmetric it is known that $\|T\| = |\lambda|$ where λ is the characteristic root of T of maximum modulus.

Consider now the case $r=1$. By (37) and (38) we have

$$
\begin{aligned}
Z-U^{(1)} &= \sum_{r=0}^{\infty}(-\delta\mathscr{M}^{-1}\mathscr{M}^*)^r(-\delta\mathscr{M}^{-1})(\mathscr{M}^*U^{(1)}-Z') \\
&= \sum_{r=1}^{\infty}(-\delta\mathscr{M}^{-1}\mathscr{M}^*)^r(U^{(1)}-\mathscr{M}^{*-1}Z') \qquad (41)\\
&= -\delta\mathscr{M}^{-1}\mathscr{M}^*(I+\delta\mathscr{M}^{-1}\mathscr{M}^*)^{-1}(U^{(1)}-\mathscr{M}^{*-1}Z')\,.
\end{aligned}
$$

We wish now to obtain a formula corresponding to (40). We observe that

$$
\mathscr{M}^{*-1}=\begin{pmatrix} M_n^{-1}(2)\,, & 0\,, & 0\,, & \cdot \\ 0\,, & M_n^{-1}(2)\,, & 0\,, & \cdot \\ 0\,, & 0\,, & M_n^{-1}(2)\,, & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}
$$

and the i, jth element $c_{i,j}$ of $M_n^{-1}(2)$ is given by (24) and (8) as

$$
c_{i,j}=c_{j,i}=\frac{D_{i-1}(2)D_{n-j}(2)}{D_n(2)}=\frac{i(n-j+1)}{n+1}\,, \qquad i\leq j\,.
$$

By direct multiplication $\mathscr{M}^{*-1}Z'$ is thus a vector $P=(P_1,\ P_2,\ \cdots,\ P_m)$, where

$$
P_i=M_n^{-1}(2)Z_i'=(n+1)^{-1}(nz_{i,0}+z_{i,n+1},\ (n-1)z_{i,0}+2z_{i,n+1},\ \cdots,\ z_{i,0}+nz_{i,n+1})\,.
$$

Now

$$
\begin{aligned}
\|P\|^2 &= \sum_{i=1}^{m}\|P_i\|^2 \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}(n+1)^{-2}((n-j+1)z_{i,0}+jz_{i,n+1})^2 \\
&=(n+1)^{-2}\sum_{j=1}^{n}\{j^2(\|Z_0\|^2+\|Z_{n+1}\|^2)+2j(n-j+1)(Z_0,\ Z_{n+1})\}\ ^{2} \\
&= \frac{n}{6(n+1)}[(2n+1)(\|Z_0\|^2+\|Z_{n+1}\|^2)+(2n+4)(Z_0,\ Z_{n+1})] \\
&\leq \frac{n(2n+1)}{6(n+1)}\{\|Z_0\|^2+\|Z_{n+1}\|^2+2(Z_0,\ Z_{n+1})\} \\
&\leq \frac{n}{3}(\|Z_0\|+\|Z_{n+1}\|)^2\,.
\end{aligned}
$$

Thus

$$
\|\mathscr{M}^*Z'\|=\|P\|\leq\sqrt{\frac{n}{3}}(\|Z_0\|+\|Z_{n+1}\|)\,.
$$

[2] $(Z_0,\ Z_{n+1})$ is the inner product $\sum_{i=1}^{m} z_{i,0}\ z_{i,n+1}$, and $|(Z_0,\ Z_{n+1})|\leq\|Z_0\|\ \|Z_{n+1}\|$.

From (41) using the same arguments as for (40) we obtain

$$\|Z-U^{(1)}\| \leq |\delta| \left\{ \|U^{(1)}\| + \sqrt{\frac{n}{3}} (\|Z_0\| + \|Z_{n+1}\|) \right\}, \qquad \text{when } \delta > 0,$$

$$\leq |\delta| (1-|\delta|)^{-1} \left\{ \|U^{(1)}\| + \sqrt{\frac{n}{3}} (\|Z_0\| + \|Z_{n+1}\|) \right\}, \qquad \text{when } \delta < 0.$$

REFERENCES

1. M. A. Hyman, *Non-iterative numerical solutions of boundary value problems*, Appl. Sci. Res., B, **2** (1952), 325–351.
2. D. E. Rutherford, *Some continuant determinants arising in physics and chemistry*, Proc. Roy. Soc. Edinburgh, Sect. A, **62** (1947), 229–236.
3. J. Todd, *The condition of certain matrices, I*, Quart. J. Mech. Appl. Math., **2** (1949), 469–472.
4. ———, *The condition of a certain matrix*, Proc. Cambridge Philos, Soc. **46** ( ), 116–118.
5. J. Williamson, *The latent roots of a matrix of special type*, Bull. Amer. Math. Soc., **37** (1931), 585–590.

UNIVERSITY OF NATAL
DURBAN, SOUTH AFRICA

Added in proof: Some of the results of §§ 1, 2, and 3 have been found also by:
Th. J. Burgerhout, *On the numerical solution of partial differential equations of elliptic type, I*, Appl. Sci. Res. B, **4** (1954), 161–172.