condition that there was stability when $b'' < 0$, and instability when $b'' > 0$. The need for a second condition was suggested by the author[2] in 1952; in [1], the case of a vapor-filled bubble was treated. The present note supplies such a condition for the general case that b changes by a large ratio.

References

[1] G. Birkhoff, *Note on Taylor instability*, Quart. Appl. Math. 12, 306-9 (1954)

ERROR BOUNDS FOR A NUMERICAL SOLUTION OF A RECURRING LINEAR SYSTEM*

By A. DE LA GARZA (*Carbide and Carbon Chemicals Co., Oak Ridge, Tenn.*)

1. Introduction. Suppose that we are given a bounded region $R, f \geq 0$, a function ϕ_B on the boundary B of R, and point functions f and g in R, and that we are required to determine ϕ so that

$$(P): \quad \partial^2\phi/\partial x^2 + \partial^2\phi/\partial y^2 = f\phi + g \text{ in } R, \qquad \phi = \phi_B \text{ on } B.$$

In order to calculate ϕ approximately, we first construct a square grid in a rectangle S that has sides parallel to the coordinate axis and is large enough to contain R in its interior. Then we approximate (P) by a system of linear algebraic difference equations, arriving at a nonhomogeneous linear algebraic system

$$A\phi = \eta, \tag{1}$$

where ϕ is a vector having a component associated with each of the N grid points inside R, η is a known N-vector, and A is a known $N \times N$ matrix. Finally, we obtain an approximate solution ϕ_0 of (1), committing an error

$$\epsilon = A^{-1}(\eta - A\phi_0).$$

A direct estimate of ϵ is given in Eq. (4) below. Our principal object is to show that ϵ can be estimated far more conveniently, though less exactly, from a properly chosen system L of linear algebraic equations set up on the $M \geq N$ grid points interior to the rectangle S. Though motivated and illustrated by problem (P), L can be constructed as soon as A (with the properties listed below) is known. This method applies generally to any system (1) with those properties.

2. Estimate for ϵ. We first list properties of $A = (a_{ij})$:

$$a_{ii} < 0, \quad i = 1(1)N; \qquad a_{ij} \geq 0, \quad i \neq j, \quad i, j = 1(1)N;$$

$$\sum_{j=1}^{N} a_{ij} \leq 0, \qquad i = 1(1)N,$$

the inequality holding for at least one value of i;

*Received January 27, 1955. The work described in this paper was done under AEC Contract W-7405-eng-26.

the matrix A cannot be transformed by the same permutation of rows and columns to the form

$$\begin{pmatrix} P & U \\ 0 & Q \end{pmatrix},$$

where P and Q are square matrices, and 0 consists of zeros. It follows from Theorem II in [1] that $| A | \neq 0$.

We now develop a series expression for ϵ. Let m be the diagonal matrix with diagonal entries $m_i = - a_{ii}$, $i = 1(1)N$. It may be shown by induction that

$$A^{-1} = -\sum_{p=0}^{n-1} D^p m^{-1} + D^n A^{-1}, \tag{2}$$

where $D = I + m^{-1}A$. Let λ_q, $q = 1(1)N$, be the characteristic roots of D. From Theorem II in [1], it is seen that for $| \lambda | \geq 1$, the determinant $| D - \lambda I | \neq 0$; therefore, $| \lambda_q | < 1$ for $q = 1(1)N$. This implies that $D^n \to 0$ as $n \to \infty$. Hence, if ϕ_0 is an approximate solution to (1), the error in ϕ_0 being $\epsilon = (e_i)$, we have from (1) and the limit in (2) that

$$\epsilon = -\sum_{p=0}^{\infty} D^p m^{-1} \rho, \tag{3}$$

where the residual vector $\rho = (r_i)$ is $\eta - A\phi_0$.

We proceed to give a direct estimate for ϵ. Use is made of the abmatrix; see [2]. Let $B = (b_{pq})$ be a matrix. The matrix $\alpha(B) = (| b_{pq} |)$ is called the abmatrix of B. If $B = (b_{pq})$ and $C = (c_{pq})$, $p = 1(1)M$, $q = 1(1)N$, $\alpha(B) \geq \alpha(C)$ means $| b_{pq} | \geq | c_{pq} |$ for all p, q. We see from these definitions that $\alpha(A + B) \leq \alpha(A) + \alpha(B)$ and $\alpha(AB) \leq \alpha(A)\alpha(B)$, provided the indicated operations are permissible. Finally, when $b_{pq} \geq 0$ for all p, q, $B = \alpha(B)$; and the prefix α will not be used. With this notation, since $D \geq 0$, $m_i > 0$, $m_i^{-1} \leq m_L^{-1}$, and $\alpha(\rho) \leq \psi_N r_U$, where $m_L = \min(m_i)$, $r_U = \max | r_i |$, and $\psi_N = (1, 1, \cdots 1)'$ in N dimensions, we have from (3) that

$$\alpha(\epsilon) \leq (r_U/m_L) \sum_{p=0}^{\infty} D^p \psi_N . \tag{4}$$

Let us now consider an $M \times M$ matrix $E \geq 0$, $M \geq N$, which contains an $N \times N$ submatrix $E_R \geq D$ and which satisfies $E^p \to 0$ as $p \to \infty$. Let $\psi_M = (1, 1 \cdots)'$ in M dimensions. We see that

$$\sum_{p=0}^{\infty} (E^p \psi_M)^* \geq \sum_{p=0}^{\infty} D^p \psi_N , \tag{5}$$

where $(E^p \psi_M)^*$ is the column vector formed by the N elements of $E^p \psi_M$ corresponding to the N rows of E which contain the submatrix E_R. Since $E^p \to 0$ as $p \to \infty$, $(I - E)$ is non-singular. We may verify that

$$\sum_{p=0}^{\infty} E^p = (I - E)^{-1}. \tag{6}$$

Hence,

$$\tau = \sum_{p=0}^{\infty} E^p \psi_M \tag{7}$$

is the solution of the system

$$(I - E)\tau = \psi_M . \tag{8}$$

Therefore, if $\tau^* = \sum_{p=0}^{\infty} (E^p\psi_M)^*$, we see from (4), (5), and (6) that

$$\alpha(\epsilon) \leq \tau^* r_U/m_L , \tag{9}$$

which is an estimate of the error ϵ in the approximate solution ϕ_0 of (1). In applying this bound, we choose E in such a way that (8) has a readily obtainable solution.

3. Application to a Poisson-type equation. To illustrate use of the methods in part 2, we return to the problem (P) in part 1 and its approximation by finite differences. Let the grid spacing in S be $\Delta x = \Delta y = \delta$. Assign one of the numbers $q = 1(1)N$ as an index to each of the grid points interior to R. The first order system approximating (P) then is:

$$-m_q\phi_q + T_q(\phi) = g_q\delta^2 - L_q(B), \qquad q = 1(1)N, \tag{10}$$

where:

$$m_q \geq 4;$$

$\mathrm{T}_q(\phi)$ is the sum of values of ϕ on non-boundary points adjacent to the qth point; $L_q(B)$ is a linear form of boundary values required only where the qth point is adjacent to a boundary. An example will make this situation clear. With first order linear approximation throughout, the equation associated with ϕ_2 in Figure 1 is

$$\phi_1 - [4 + f_2\delta^2 + c/(1 - c)]\phi_2 + \phi_3 + \phi_4 = g_2\delta^2 - \phi_B/(1 - c).$$

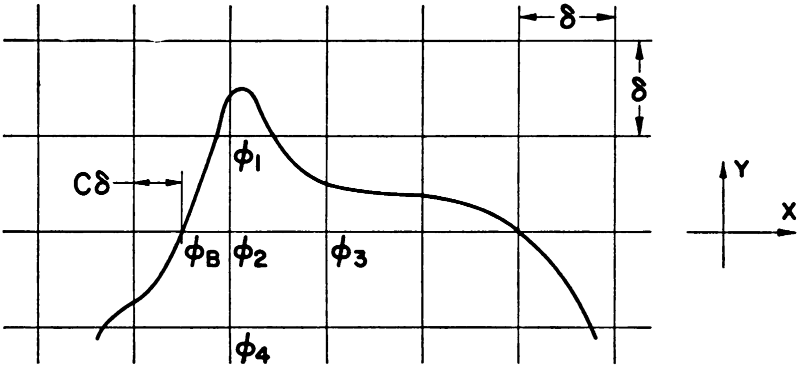

FIG. 1

We see that the qth row of the matrix D associated with (10) has elements m_q^{-1} corresponding to the unit coefficients of $T_q(\phi)$, all other elements of D being zero. Let the sides of the rectangle S be of length $H\delta$ and $I\delta$. Consider the system (11) corresponding to the $(H - 1) \times (I - 1)$ grid points interior to S:

$$\begin{gathered} t_{h,i} - m_L^{-1}(t_{h,i+1} + t_{h+1,i} + t_{h,i-1} + t_{h-1,i}) = 1, \qquad i = 0(1)I, \qquad h = 0(1)H, \\ t_{h,0} = t_{h,I} = t_{0,i} = t_{H,i} = 0, \end{gathered} \tag{11}$$

where $m_L = \min(m_q)$. Write the system (11) in the matrix form (8), the equations corresponding to the grid points interior to R being written in the same order as in (10). We see that the resulting matrix E then contains an $N \times N$ submatrix $E_R \geq 0$ with entries m_L^{-1} in the same positions as the entries m_q^{-1} of the matrix D associated with (10). Hence, $E_R \geq D$. From Theorem II in [1], we see that the characteristic roots of E are less than one in absolute value; hence, $E^p \to 0$ as $p \to \infty$. It follows that the solution of (11) may be used in (9) to compute an error bound for an approximate solution of (1). We may verify that the solution of (11) is

$$t_{h,i} = (HI)^{-1} \sum_{r=1, s=1}^{r=H-1, s=I-1} C_{rs} \sin hr\pi/H \sin is\pi/I, \tag{12}$$

where

$$C_{rs} = \{[1 - (-1)^r][1 - (-1)^s] \sin r\pi/H \sin s\pi/I\} \cdot \{(1 - \cos r\pi/H)(1 - \cos s\pi/I)[1 - 2m_L^{-1}(\cos r\pi/H + \cos s\pi/I)]\}^{-1}.$$

For specifying the largest absolute residual which may be tolerated in an iterated solution, a bound for the maximum error $e_U = \max | e_i |$ is useful. From (9),

$$e_U \leq t_U r_U / m_L , \tag{13}$$

where $t_U = \max(t_{h,i})$. For simplicity, let both H and I be even. We may verify that t_U occurs at $h = H/2$, $i = I/2$, and that

$$t_U = -4(HI)^{-1} \sum_{r=1, s=1, \text{odd}}^{r=H-1, s=I-1} (-1)^{(r+s)/2}[1 - 2m_L^{-1}(\cos r\pi/H + \cos s\pi/I)]^{-1} \cdot \cot r\pi/2H \cot s\pi/2I, \tag{14}$$

where the summation is on odd values of r and s. For illustration, t_U in (14) with $m_L = 4$ is tabulated in Table 1 for H, $I = 8$ (2) 20. Since t_U in (14) increases as m_L decreases, the tabulated values may be used in (13) to state error bounds for an approximate solution of the linear system (10).

TABLE 1

Factor t_U for use in (13)

H, I = 8(2) 20

H/I	8	10	12	14	16	18	20
8	18.6						
10	22.7	29.2					
12	25.6	34.5	42.2				
14	27.6	38.5	48.6	57.5			
16	29.0	41.5	53.8	65.1	75.2		
18	30.0	43.8	57.9	71.4	84.0	95.2	
20	30.6	45.4	61.1	76.6	91.4	105	118

Acknowledgements. Mr. E. B. Carter of the Numerical Analysis Department at K-25 computed Table 1. The referee's comments simplified and generalized the original results.

References

1. O. Taussky, *A recurring theorem on determinants*, Am. Math. Monthly **56**, 672-676 (1949)
2. A. de la Garza, *Error bounds on approximate solutions to systems of linear algebraic equations*, **MTAC 7**, 81-84 (1953)