Technical Memorandum No. 10

U.S. Joint Numerical Weather Prediction Unit

LIBRARY
FEB 17 1975
National Oceanic &
Atmospheric Administration
U.S. Dept. of Commerce

Convergence Rates of Liebmann's and Richardson's Iterative Methods when Applied to the Solution of a System of Helmholtz'-Type Equations

LIBRARY
OCT 12 1956
WEATHER BUREAU

RARE BOOK
QC
996
.T33
no. 10

Geirmundur 'Arnason
U. S. Weather Bureau

July 16, 1956

92622

National Oceanic and Atmospheric Administration

U.S. Joint Numerical Weather Prediction Unit

ERRATA NOTICE

One or more conditions of the original document may affect the quality of the image, such as:

Discolored pages
Faded or light ink
Binding intrudes into the text

This has been a co-operative project between the NOAA Central Library, National Center for Environmental Prediction and the U.S. Air Force. This project includes the imaging of the full text of each document. To view the original documents, please contact the NOAA Central Library in Silver Spring, MD at (301) 713-2607 x124 or www.reference@nodc.noaa.gov.

LASON
Imaging Contractor
12200 Kiln Court
Beltsville, MD 20704-1387
April 13, 2004

Contents

1. Introduction

In his paper, "Convergence Rates of Iterative Treatments of Partial Differential Equations", Frankel [1] has studied the convergence rates of Richardson's and Liebmann's iterative methods when applied to the numerical solution of the two-dimensional Laplace equation $\nabla^2\phi = 0$. His results apply without any modification to the Poisson equation $\nabla^2\phi + q = 0$ as well. It is shown in the third chapter of this paper that with slight modification it is also applicable to the Helmholtz equation $\nabla^2\phi + b\phi + q = 0$. Frankel's results, in particular his determination of an optimum over-relaxation coefficient have been very useful in numerical weather prediction when forecasting by means of one- and two-level models. In dealing with multilevel models, for instance of the type designed by Charney and Phillips [2], it may be desireable to solve the associated system of equations simultaneously rather than to reduce it to a number of Helmholtz equations to which Frankel's theory applies almost directly.

This paper deals with the convergence criterium and the determination of optimum overrelaxation coefficients when applying Richardson's and Liebmann's iterative methods to the numerical solution of a high order linear system of the form

$$L\,\phi_{k,i,j} + \sum_{l=1}^{m} b_{k,l}\,\phi_{l,i,j} - q_{k,i,j} = 0$$

$$(k = 1, 2, 3, \ldots . m)$$

where L is the finite difference form of the two-dimensional Laplace operator. For lack of a better notation, the linear system above is here referred to as a system of Helmholtz-type equations. The method used is essentially an extension of the one laid down in Frankel's paper, and for convenience Frankel's theory has been included as a separate chapter (No. 2) of this paper. In the third chapter his theory is extended to the special case $m = 2$ and in the fourth and last chapter to an arbitrary m.

Upon completion of this work it has come to the author's attention that much the same problem has been treated by D. Young [3].

2. Frankel's Theory for the Convergence Rates of Richardson's and Liebmann's Iterative Methods

The convergence rates of Richardson's and Liebmann's iterative methods for the solution of the two-dimensional Laplace equation $\nabla^2\phi = 0$ have been investigated by Frankel [1]. The Laplace equation is approximated by the difference equation

$$L\,\phi(x, y) \equiv \phi(x + h, y) + \phi(x - h, y) + \phi(x, y + h) + \phi(x, y - h) - 4\,\phi(x, y) = 0 \qquad (1)$$

As usual we put $x = ih$ and $y = jh$ and denote the value assigned to $\phi(ih, jh)$ at the nth stage of iteration by $\phi^n_{i,j}$. The limit to which $\phi^n_{i,j}$ approaches with increasing n is denoted by $\phi_{i,j}$ and is the solution of (1) for the given boundary conditions. The boundaries have been chosen to be rectangular and fixed values assigned to ϕ there. The coordinates i and j assume the values

$$i = 0, 1, 2, \ldots. p; \qquad j = 0, 1, 2, \ldots. q;$$

At interior points the equation

$$L\,\phi_{i,j} \equiv \phi_{i-1,j} + \phi_{i+1,j} + \phi_{i,j-1} + \phi_{i,j+1} - 4\,\phi_{i,j} = 0 \qquad (2)$$

holds and at the boundaries $\phi^n_{i,j} = b_{i,j}$ for all n.

Although Frankel deals with the Laplace equation, his theory applies to Poisson's equation

$$L\,\phi_{i,j} - q_{i,j} = 0 \qquad (3)$$

as well. Since the latter is of greater importance in numerical weather prediction, we shall refer to (3) in the following discussion; the boundary conditions are the same as those given above.

For interior points Richardson's iterative scheme is defined as follows:

$$\phi^{n+1}_{i,j} = \phi^{n}_{i,j} + \alpha(L\,\phi^{n}_{i,j} - q_{i,j}) = \phi^{n}_{i,j} + \alpha\, R^{n}_{i,j} \qquad (4)$$

where α is a real positive number, the optimum value of which is 0.25 as we will see later; $R^{n}_{i,j}$ is the residual in the point i,j after n completed scans.

The error at each stage is denoted by $\epsilon^{n}_{i,j}$,

$$\epsilon^{n}_{i,j} \equiv \phi^{n}_{i,j} - \phi_{i,j} \qquad (5)$$

where $\phi_{i,j}$ is now the solution of (3). Substitution into (4) gives the error recurrence relation

$$\epsilon^{n+1}_{i,j} = \epsilon^{n}_{i,j} + \alpha L \epsilon^{n}_{i,j} = (1 + \alpha L)\, \epsilon^{n}_{i,j} = K\, \epsilon^{n}_{i,j}; \text{ (interior points)}$$

$$= 0 \text{ (boundary points)} \qquad (6)$$

since $L\,\phi_{i,j} - q_{i,j} = 0$. The convergence rate is now studied by investigating the solution of (6). For this purpose the error field ϵ^{o} ($n = 0$) is expanded into a series

$$\epsilon^{o}_{i,j} = \sum_{r,s} A_{r,s}\, \epsilon^{r,s}_{i,j} \qquad (7)$$

where

$$\epsilon^{r,s}_{i,j} = \sin\left(\frac{r}{p}\pi i\right) \sin\left(\frac{s}{q}\pi j\right) \qquad \begin{array}{l} r = 1, 2, \ldots\ p - 1 \\ s = 1, 2, \ldots\ q - 1 \end{array} \qquad (8)$$

Application of the operator L, see eq. (2), to (8) gives

$$L\,\epsilon^{r,s}_{i,j} = \left[2 \cos \frac{r}{p}\pi + 2 \cos \frac{s}{q}\pi - 4\right]\epsilon^{r,s}_{i,j}$$

$$= -4\left[\sin^2\left(\frac{r}{2p}\pi\right) + \sin^2\left(\frac{s}{2p}\pi\right)\right] \epsilon^{r,s}_{i,j} \qquad (9)$$

The quantities

$$L_{r,s} = -4 \left[\sin^2\left(\frac{r}{2p}\pi\right) + \sin^2\left(\frac{s}{2q}\pi\right)\right] \qquad \begin{array}{l} r = 1, 2, \ldots\ p - 1 \\ s = 1, 2, \ldots\ q - 1 \end{array} \tag{10}$$

are the eigenvalues of the operator L. The smallest and the largest magnitude for these eigenvalues are obtained for $r = s = 1$ and $r = p - 1$, $s = q - 1$ and are denoted by L_o and L_m, respectively,

$$L_o = -4 \left[\sin^2 \frac{\pi}{2p} + \sin^2 \frac{\pi}{2q}\right] \cong -\pi^2(p^{-2} + q^{-2}) \tag{11}$$

$$L_m = -4 \left[\sin^2 \frac{p-1}{2p}\pi + \sin^2 \frac{q-1}{2q}\pi\right] = -(8 + L_o) \cong -8 + \pi^2(p^{-2} + q^{-2})$$

Referring to eq. (6) we must require that for all the components of the error field the corresponding $K_{r,s}$

$$K_{r,s} = 1 + \alpha L_{r,s} \tag{12}$$

have a magnitude less than unity, since otherwise Richardson's iterative scheme would not generally converge. The ultimate convergence rate is determined by the maximum magnitude of $K_{r,s}$ which we shall denote by K^*. All $K_{r,s}$ lie in the range

$$K_m \equiv 1 + \alpha L_m \leq K_{r,s} \leq 1 + \alpha L_o \equiv K_o \tag{13}$$

the equality signs holding for $r = s = 1$ and $r = p - 1$, $s = q - 1$. K^* is equal to $|K_o|$ or $|K_m|$ depending on which of these two has the greater magnitude. As α increases from zero K_o drops slowly and K_m rapidly from unity and the latter changes its sign when α reaches the value $\frac{1}{8 + L_o}$. For $\alpha = \frac{2}{8 + L_o}$ K_m takes on the value -1. The smallest K^* occurs when

$$K_o + K_m = 0$$

i.e., when K_o and K_m have the same magnitude but different signs. Hence, for the most rapid convergence we choose

$$\alpha_o = -\frac{2}{L_o + L_m} = \frac{1}{4} \tag{14}$$

For this optimum α-value the shortest and the longest error waves decay at the same rate while other error waves decay more rapidly. For K^* we have the following relation:

$$K^* = 1 + \frac{1}{4} L_o \cong 1 - \frac{\pi^2}{4} (p^{-2} + q^{-2}) \tag{15}$$

For a large grid K^* may be quite close to unity which means slow convergence.
For $p = 19$, $q = 29$ $\quad L_o = - 4 \cdot 10^{-2}$ $\quad$ and $K^* = 0.990$,
for $p = 30$, $q = 34$ $\quad L_o = - 2 \cdot 10^{-2}$ $\quad$ and $K^* = 0.995$

The Liebmann iteration process deviates from Richardson's in that the corrected ϕ-value is used in subsequent operations in that iteration step. We assume that the grid is scanned in the same direction along successive rows, starting at the interior point $i = 1$, $j = 1$, then the iterative scheme is defined as follows:

$$\phi^{n+1}_{i,j} = \phi^n_{i,j} + \alpha \left(L \phi^{n,\, n+1}_{i,j} - q_{i,j}\right) = \phi^n_{i,j} + \alpha R^{n,\, n+1}_{i,j} \tag{16}$$

where

$$L \phi^{n,\, n+1}_{i,j} \equiv \phi^{n+1}_{i-1,j} + \phi^n_{i+1,j} + \phi^{n+1}_{i,j-1} + \phi^n_{i,j+1} - 4 \phi^n_{i,j} \tag{17}$$

The iteration equation for the error

$$\epsilon^n_{i,j} \equiv \phi^n_{i,j} - \phi_{i,j}$$

is obtained by substitution into (16) and utilizing (17),

$$\begin{aligned} \epsilon^{n+1}_{i,j} &= \epsilon^n_{i,j} + \alpha \left[\epsilon^{n+1}_{i-1,j} + \epsilon^n_{i+1,j} + \epsilon^{n+1}_{i,j-1} + \epsilon^n_{i,j+1} - 4 \epsilon^n_{i,j}\right]; \quad \text{(interior points)} \\ &= 0 \quad \text{(boundary points)} \end{aligned} \tag{18}$$

and may be written as

$$\epsilon^{n+1} = K(\alpha)\, \epsilon^n \tag{19}$$

where $K(\alpha)$ is a linear operator, depending on the parameter α, to be determined. We seek a solution of (18) of the form

$$\epsilon^{r,s}_{i,j} = A^i \sin \left(\frac{r}{p} \pi i\right) B^j \sin \left(\frac{s}{q} \pi j\right) \tag{20}$$

and substitute into (18), which, when connected with (19), yields:

$$\begin{aligned}
&(K - 1 + 4\alpha) \cdot A^i \cdot B^j \sin \left(\frac{r}{p} \pi i\right) \cdot \sin \left(\frac{s}{q} \pi j\right) \\
&= \alpha B^j \sin \left(\frac{s}{q} \pi j\right) \left[KA^{i-1} \sin \frac{r}{p} \pi (i - 1) + A^{i+1} \sin \frac{r}{p} \pi (i + 1)\right] \\
&+ \alpha A^i \sin \left(\frac{r}{p} \pi i\right) \left[KB^{j-1} \sin \frac{s}{q} \pi (j - 1) + B^{j+1} \sin \frac{s}{q} \pi (j + 1)\right] \\
&= \alpha A^i \cdot B^j \sin \frac{s}{q} \pi j \left[(K \cdot A^{-1} + A^{+1}) \sin \left(\frac{r}{p} \pi i\right) \cos \frac{r}{p} \pi\right) \\
&\qquad - (K \cdot A^{-1} - A^{+1}) \cos \left(\frac{r}{p} \pi i\right) \sin \frac{r}{p} \pi\Big] \\
&+ \alpha A^i B^j \sin \left(\frac{r}{p} \pi i\right) \left[(K \cdot B^{-1} + B^{+1}) \sin \left(\frac{s}{q} \pi j\right) \cos \frac{s}{q} \pi\right. \\
&- \left.(KB^{-1} - B^{+1}) \cos \frac{s}{q} \pi j \sin \frac{s}{q} \pi\right]
\end{aligned} \tag{21}$$

The appearance of the $\cos \left(\frac{r}{p} \pi i\right)$ and $\cos \left(\frac{s}{q} \pi j\right)$ terms are prevented by requiring

$$A^2 = B^2 = K \tag{22}$$

whereby equation (21) reduces to

$$A^2 - 1 + 4\alpha = 2\alpha A \left[\cos \frac{r}{p} \pi + \cos \frac{s}{q} \pi\right]$$

or

$$A^2 - 2\alpha t A + 4\alpha - 1 = 0; \qquad t = \cos \frac{r}{p} \pi + \cos \frac{s}{q} \pi \tag{23}$$

If we take $\alpha = \frac{1}{4}$ the roots of (23) are

$$A = \begin{cases} 0 \\ t/2 \end{cases}$$

The larger root corresponds to

$$K = A^2 = \frac{1}{4} t^2 = \frac{1}{4} [\cos \frac{r}{p} \pi + \cos \frac{s}{q} \pi]^2 \tag{24}$$

The greatest value for K occurs when $r = s = 1$ and $r = p - 1$, $s = q - 1$, and is

$$K^* = \frac{1}{4} [\cos \frac{\pi}{p} + \cos \frac{\pi}{q}]^2 \cong [1 - \frac{\pi^2}{4} (p^{-2} + q^{-2})]^2 \tag{25}$$

Comparison with (15) shows that in each iteration cycle the most resistant errors are reduced as much as in two cycles in Richardson's procedure.

As will be shown below, $\alpha = \frac{1}{4}$ is not the optimum value for Liebmann's method. Solution of (23) gives

$$A = \alpha t \pm \sqrt{\alpha^2 t^2 - 4\alpha + 1} = \begin{cases} A_1 = \alpha t + \sqrt{\alpha^2 t^2 - 4\alpha + 1} \\ A_2 = \alpha t - \sqrt{\alpha^2 t^2 - 4\alpha + 1} \end{cases} \tag{26}$$

For $\alpha = 0$ $A = \pm 1$; increasing α causes A_1 to decrease and A_2 to increase until both become equal to αt when α reaches the value given by the smaller of the two roots of

$$\alpha^2 t^2 - 4\alpha + 1 = 0 \tag{27}$$

This α is the optimum value corresponding to that particular t. The ultimate rate of convergence is determined by the largest and smallest t-values both of which have the magnitude t_o given by

$$t_o = \cos \frac{\pi}{p} + \cos \frac{\pi}{q} = 2 + \frac{L_o}{2} \cong 2 \left(1 - \frac{\pi^2}{4} (p^{-2} + q^{-2})\right) \tag{28}$$

Equation (27) may be written as follows:

$$(2\alpha - 1)^2 - (4 - t^2)\, \alpha^2 = 0$$

We define $\cos \theta = \frac{1}{2} (\cos \frac{r}{p} \pi + \cos \frac{s}{q} \pi) = \frac{t}{2}$

then $4 - t^2 = 4 \sin^2\theta$

and

$$(2\alpha - 1 + 2\alpha \sin \theta)\ (2\alpha - 1 - 2\alpha \sin \theta) = 0$$

The smaller root is

$$\alpha = \frac{1}{2\ (1 + \sin \theta)} \tag{29}$$

where $\sin \theta$ is always positive and can range from unity (when $r = \frac{p}{2}$ and $s = \frac{q}{2}$) to $\sin \theta_o$

where

$$\cos \theta_o = \frac{1}{2} \left(\cos \frac{\pi}{p} + \cos \frac{\pi}{q}\right)$$

The optimum α-value, α_o, is

$$\alpha_o = \frac{1}{2\ (1 + \sin \theta_o)} \cong \frac{1}{2} \left(1 - \sqrt{\frac{\pi^2}{2} (p^{-2} + q^{-2})}\ \right) \tag{30}$$

When α exceeds the value given by (29), the roots of (26) become complex and both have the magnitude $4\alpha - 1$; thus

$$K^* = 4\alpha_o - 1 \cong 1 - 2 \sqrt{\frac{\pi^2}{2} (p^{-2} + q^{-2})} = 1 - \sqrt{2} \cdot \pi\ (p^{-2} + q^{-2})^{\frac{1}{2}} \tag{31}$$

For large p and q K^* as determined from (31) is a considerable improvement upon K^* as determined from (25). Attention is drawn to the fact that α cannot exceed the value 0.5 since the magnitude of K would in that case exceed unity for all t's.

For the two grids $p = 19$, $q = 29$ and $p = 30$, $q = 34$ α_o is 0.43 and 0.45 respectively. Corresponding $K^* = (\alpha_o t_o)^2 = 4\ \alpha_o - 1$ is 0.73 and 0.80 respectively.

Fig. 1 shows the two roots A_1 and A_2 of (26) when $t = t_o$ and $L_o = -2.10^{-2}$ $(p = 30,\ q = 34)$.

3. Extension of Frankel's Theory to a System of Two Helmholtz'-type Equations

Frankel's theory is easily extended to the Helmholtz equation which in difference form is

$$L\phi_{i,j} + b_{i,j}\,\phi_{i,j} - q_{i,j} = (L + b_{i,j})\,\phi_{i,j} - q_{i,j} = 0 \tag{3a}$$

Instead of dealing with the eigenvalues of the operator L as defined by (2) we now have to deal with the eigenvalues of the operator M defined as

$$M = L + b \tag{2a}$$

Because of the simple relationship between these two operators we can immediately write the eigenvalues of M which have the smallest and the largest magnitude as

$$M_o = 2[\cos\frac{\pi}{p} + \cos\frac{\pi}{q} - 2] + b = L_o + b \tag{11a}$$

$$M_m = 2[\cos\frac{p-1}{p}\pi + \cos\frac{q-1}{q}\pi - 2] + b = L_m + b = -(8 + L_o) + b$$

It follows from the study in the previous chapter that in order to ensure convergence of Richardson's iterative scheme one must require that

$$L_o + b < 0 \qquad \text{i.e.} \qquad b < \pi^2\,(p^{-2} + q^{-2}) \tag{32}$$

Similarily we find the optimum value for α to be

$$\alpha_o = -\frac{2}{M_o + M_m} = \frac{1}{4 - b} \tag{14a}$$

Thus for b negative the optimum α is less than 0.25 and approaches zero when $b \to -\infty$. For K* we obtain the following relation

$$K^* = 1 + \frac{L_o + b}{4 - b} = 1 + \frac{L_o}{4} + \frac{4 + L_o}{4\,(4 - b)}\,b \tag{15a}$$

For negative b the convergence is more rapid than in the case of Poisson's equation.

When the Liebmann method is used, equation (23) is replaced by

$$A^2 - 2\alpha t A + (4 - b)\,\alpha - 1 = 0 \tag{23a}$$

For $\alpha = \frac{1}{4 - b}$ (23) has the roots

$$A = \begin{cases} 0 \\ \frac{2t}{4 - b} \end{cases}$$

The larger root is always less than one in magnitude if (32) is fulfilled. When $t = t_o$ we get

$$K^* = A^2 = \frac{4\, t_o^{\,2}}{(4 - b)^2} \cong \left(\frac{4 + L_o}{4 - b}\right)^2$$

The optimum α-value corresponding to a particular t is the smaller one of the two roots of

$$\alpha^2 t^2 - (4 - b)\,\alpha + 1 = 0 \qquad (27a)$$

Analogous with the solution of (27) we find that the solution of (27a) is

$$\alpha = \frac{1}{2\,(1 - \frac{b}{4})(1 + \sin\theta)} \; ; \quad \sin\theta > 0 \qquad (29a)$$

where

$$\cos\theta = \frac{1}{2\,(1 - \frac{b}{4})} \quad \left(\cos\frac{r\pi}{p} + \cos\frac{s\pi}{q}\right)$$

The requirement $|\cos\theta| \leq 0$ is fulfilled if condition (32) is satisfied. As before we set $t = t_o$ in order to get the α-value which ultimately gives the most rapid rate of convergence, thus

$$\alpha_o = \frac{1}{2\,(1 - \frac{b}{4})(1 + \sin\theta_o)} \qquad (30a)$$

For large p and q and small b α_o is approximately

$$\alpha_o = \frac{1}{2}\,\left(1 - \sqrt{\frac{\pi^2}{2}\,(p^{-2} + q^{-2}) + \frac{b}{2}}\,\right) \qquad (30b)$$

and

$$K^* = 4\,\alpha_o - 1 = 1 - \pi\cdot\sqrt{2}\,(p^{-2} + q^{-2} - b)^{\frac{1}{2}} \tag{31a}$$

It can be shown that the inequality (32) must hold for the Liebmann method to converge. As in the case of Richardson's method, negative b improves the rate of convergence as compared with $b = 0$ (Poisson's equation).

We shall now deal with the convergence rate of two Helmholtz-type equations

$$\begin{aligned} L\,\phi_{1,i,j} + b_{11}\,\phi_{1,i,j} + b_{12}\,\phi_{2,i,j} - q_1 &= 0 \\ L\,\phi_{2,i,j} + b_{21}\,\phi_{1,i,j} + b_{22}\,\phi_{2,i,j} - q_2 &= 0 \end{aligned} \tag{3b}$$

where the b's are constants or given functions of i and j and ϕ is known on the boundaries. As in the simple two dimensional case we will use an iterative scheme of the form

$$\phi^{n+1}_{k,i,j} = \phi^{n}_{k,i,j} + \alpha_k\,[L\phi^{n,n+1}_{k,i,j} + b_{k1}\,\phi^{n,n+1}_{1,i,j} + b_{k2}\,\phi^{n}_{2,i,j} - q_k];\; k = 1,\,2 \tag{16a}$$

where the double superscript $n,\, n+1$ refers to Liebmann's method and where it is assumed that the first level $(k = 1)$ is scanned first; the operator L is defined by (17). Equation (16a) applies to Richardson's scheme if $n,\, n+1$ is replaced by n and L defined by Eq. (2) (except for the equality sign in (2)).

We define the error $\epsilon^n_{k,i,j}$ as

$$\epsilon^n_{k,i,j} \equiv \phi^n_{k,i,j} - \phi_{k,i,j} \qquad k = 1,\,2 \tag{5a}$$

where $\phi_{k,i,j}$ is the solution of (3b). Substitution for ϕ^n and ϕ^{n+1} by means of (5a) into (16a) gives the error recurrence relation

$$\epsilon^{n+1}_{k,i,j} = \epsilon^{n}_{k,i,j} + \alpha_k\,\{L\,\epsilon^{n,n+1}_{k,i,j} + b_{k1}\,\phi^{n,n+1}_{1,i,j} + b_{k2}\,\phi^{n}_{2,i,j}\}; \qquad k = 1,\,2 \tag{6a}$$

at interior points and zero at boundary points.

We will now define an operator $K_k\ (\alpha_k,\ b_{k1}, b_{k2})$ such that

$$\epsilon^{n+1}_{k,i,j} = K_k\ \epsilon^{n}_{k,i,j}; \qquad k = 1,\ 2 \qquad (19a)$$

and substitute for ϵ^{n+1} from (19a) into (6a) which then yields

$$\begin{aligned} K_1\epsilon_{1,i,j} &= \epsilon_{1,i,j} + \alpha_1[K_1\epsilon_{1,i-1,j} + \epsilon_{1,i,j+1} + \epsilon_{1,i+1,j} \\ &\quad + K_1\epsilon_{1,i,j-1} - (4 - b_{11})\ \epsilon_{1,i,j} + b_{12}\ \epsilon_{2,i,j}] \\ K_2\epsilon_{2,i,j} &= \epsilon_{2,i,j} + \alpha_2\ [K_2\epsilon_{2,i-1,j} + \epsilon_{2,i,j+1} + \epsilon_{2,i+1,j} \\ &\quad + K_2\epsilon_{2,i,j-1} - (4 - b_{22})\ \epsilon_{2,i,j} + b_{21}\ K_2\epsilon_{1,i,j}] \end{aligned} \qquad (33)$$

where the superscript n has been omitted. We seek a solution to (33) of the form

$$\epsilon^{r,s}_{k,i,j} = A^{i}_{k}\ B^{j}_{k}\cdot C_k\ \sin\ (\tfrac{r}{p}\ \pi\ i)\ \sin\ (\tfrac{s}{q}\ \pi\ j); \qquad k = 1,2 \qquad (20a)$$

and substitute in (33) which yields

$$\begin{aligned} &[K_1 - 1 + (4 - b_{11})\alpha_1]\ A^{i}_{1}\cdot B^{j}_{1}\cdot C_1\ \sin\ (\tfrac{r}{p}\ \pi\ i)\cdot \sin\ (\tfrac{s}{q}\ \pi\ j) \\ &= \alpha_1 A^{i}_{1}\cdot B^{j}_{1}\cdot C_1\cdot(\sin \tfrac{r}{p}\ \pi\ i)\cdot \sin\ (\tfrac{s}{q}\ \pi\ j)[(K_1\cdot A^{-1}_{1} + A_1)\ \cos \tfrac{r}{p}\ \pi\ + \\ &\quad (K_1 A^{-1}_{1} - A_1)\ \frac{\cos\ (\frac{r}{p}\ \pi\ i)\cdot \sin \frac{r}{p}\ \pi}{\sin\ (\frac{r}{p}\ \pi\ i)} + (K_1 B^{-1}_{1} + B_1)\ \cos \tfrac{s}{q}\ \pi\ + \\ &\quad (K_1 B^{-1}_{1} - B_1)\ \frac{\cos(\frac{s}{q}\ \pi\ j)\cdot \sin \frac{s}{q}\ \pi}{\sin\ (\frac{s}{q}\ \pi\ j)} + b_{12}\cdot \frac{C_2}{C_1}] \end{aligned}$$

for k = 1 and a similar equation for k = 2.

We now require

$$K_k = A^2_{\ k} = B^2_{\ k} \qquad (22a)$$

which when substituted in the equation above yields

$$K_1 - 1 + (4 - b_{11})\ \alpha_1 = 2\ \alpha_1 A_1 t + \alpha_1 b_{12} \frac{C_2}{C_1}$$

and for $k = 2$ gives

$$K_2 - 1 + (4 - b_{22})\ \alpha_2 = 2\ \alpha_2 A_2 t + \alpha_2 b_{21}\ K_2 \frac{C_1}{C_2}$$

where $t = \cos\frac{r\pi}{p} + \cos\frac{s\pi}{q}$ $\qquad r = 1, 2, \ldots p - 1; \quad s = 1, 2, \ldots q - 1$

Utilizing (22a) we may write

$$\begin{aligned} &A^2{}_1 - 2\ \alpha_1 t A_1 + (4 - b_{11} - b_{12} \frac{C_2}{C_1})\alpha_1 - 1 = 0 \\ &(1 - \alpha_2 b_{21} \frac{C_1}{C_2})\ A^2{}_2 - 2\ \alpha_2 t A_2 + (4 - b_{22})\ \alpha_2 - 1 = 0 \end{aligned} \qquad (23b)$$

If Richardson's iterative method is used (23b) is replaced by

$$\begin{aligned} &K_1 - 2\ \alpha_1 t + (4 - b_{11} - b_{12} \cdot \frac{C_2}{C_1})\alpha_1 - 1 = 0 \\ &K_2 - 2\ \alpha_2 t + (4 - b_{22} - b_{21} \cdot \frac{C_1}{C_2})\alpha_2 - 1 = 0 \end{aligned} \qquad (23c)$$

When the equations (23b) and (23c) are compared with (23a) the main difference is the occurrence of the unknown amplitude ratio $\frac{C_1}{C_2}$ which initially can be quite arbitrary. For this reason there is generally no means for keeping the magnitudes of K_1 and K_2 both less than unity in the first iteration steps unless both b_{12} and b_{21} are equal to zero. In this case, however, there is no feed-back from one level to the other and (23b) and (23c) are reduced to the types already treated. Before dealing with (23c) in general, we will consider the two special cases $b_{12} = 0$, $b_{21} \neq 0$ and, $b_{21} = 0$, $b_{12} \neq 0$ i.e., a one-sided influence from level 1 to level 2 or vice versa. One of the equations (3b) is in each of these cases reduced to the Helmholz equation dealt with in the beginning of this chapter and will converge if condition (32) is fulfilled. Both iterative schemes initially and for a certain number of scans may show divergence when applied to the other equations. Take the case that $b_{12} = 0$, then after a sufficient number of scans C_1 is for all practical purposes equal to zero and the term $\alpha_2 b_{21} \frac{C_1}{C_2}$ in (23b) as well, which reduces the equation containing this term to a type dealt with earlier. Thus, convergence is ensured when (32) is satisfied where, however, b is to be identified with b_{11} and b_{22} respectively. The same conclusion is reached when $b_{21} = 0$, $b_{12} \neq 0$.

As a further step towards generalization we shall deal with the special case ~~the~~ that initially $C_2 = \pm C_1$. The application of previous results to Eqs. (23c) immediately renders the following criteria as sufficient for the magnitudes of K_1 and K_2 to both be less than unity after the first iteration step.

$$L_o + b_{11} + |b_{12}| < 0$$
$$L_o + b_{22} + |b_{21}| < 0 \qquad (32a)$$

The optimum α, which in general is different for the two levels, is

$$\alpha_{o,1} = \frac{1}{4 - b_{11} - |b_{12}|} \; ; \quad \alpha_{o,2} = \frac{1}{4 - b_{22} - |b_{21}|} \; ; \qquad (14b)$$

Corresponding expressions for K* follow from (23c). When α_2 takes on the value $\frac{1}{4 - b_{22}}$, α_1 the value $\alpha_{o,1}$ given above and these or substituted in (23b) it is easily shown that the magnitudes of both A_1 and A_2 are less than unity if conditions (32a) are fulfilled.

We now solve (23b) with respect to A_1 and A_2 ($C_2 = \pm C_1$) and obtain:

$$A_1 = \alpha_1 t \pm \sqrt{\alpha^2{}_1 t^2 - (4 - b_{11} \mp b_{12})\,\alpha_1 + 1}$$
$$A_2 = \frac{\alpha_2 t}{\kappa} \pm \sqrt{\frac{\alpha^2{}_2 t^2}{\kappa^2} - \frac{4 - b_{22}}{\kappa}\,\alpha_2 + \frac{1}{\kappa}} \qquad (26a)$$

where

$$\kappa = 1 \mp \alpha_2 b_{21} \qquad (34)$$

In order to investigate under which conditions $|A_1|$ and $|A_2|$ are less than unity we differentiate (23b) with respect to α and obtain the following expressions for $\frac{dA_1}{d\alpha_1}$ and $\frac{dA_2}{d\alpha_2}$:

$$\frac{dA_1}{d\alpha_1} = \frac{tA_1 - 2 + \frac{1}{2}(b_{11} \pm b_{12})}{A_1 - \alpha_1 t}$$

$$\frac{dA_2}{d\alpha_2} = \frac{tA_2 - 2 + \frac{1}{2}(b_{22} \pm b_{21}\, A^2{}_2)}{\kappa\, A_2 - \alpha_2 t} \tag{35}$$

For $\alpha_1 = \alpha_2 = 0$ (26a) is reduced to $A_1 = \pm 1$ and $A_2 = \pm 1$ and since the denominators of the expressions for $\frac{dA_1}{d\alpha_1}$ and $\frac{dA_2}{d\alpha_2}$ are positive for the two larger roots of (26a) it is easily verified that these roots will decrease with increasing α if the conditions (32a) are fulfilled. It is furthermore clear that the derivatives will stay negative for increasing α. Similarily it can be shown that the smaller roots of (26a) will increase with increasing α when (32a) is satisfied.

The optimum values for α_1 and α_2 are the smaller roots of each of the equations

$$\alpha^2{}_1 t^2{}_o - (4 - b_{11} \mp b_{12})\,\alpha_1 + 1 = 0$$

$$\alpha^2{}_2 t^2{}_o - (4 - b_{22})\,\kappa\,\alpha_2 + \kappa = 0 \tag{27b}$$

where κ is given by (34). The solution of (27b) renders the following expressions for $\alpha_{o,1}$ and $\alpha_{o,2}$ (for details see Appendix),

$$\alpha_{o,1} = \frac{1}{2\,(1 - \frac{1}{4}(b_{11} + |b_{12}|)(1 + \sin\theta_{o,1})}$$

$$\alpha_{o,2} = \frac{1}{2\,(1 - \frac{1}{4}(b_{22} + |b_{21}|)(1 + \sin\theta_{o,2})} \tag{30c}$$

where

$$\cos \theta_{o,1} = \frac{1}{2\left(1 - \frac{1}{4}\left(b_{11} + |b_{12}|\right)\right)} \left(\cos \frac{\pi}{p} + \cos \frac{\pi}{q}\right)$$

$$\cos \theta_{o,2} = \frac{1}{2\left(1 - \frac{1}{4}\left(b_{22} + |b_{21}|\right)\right)} \left(\cos \frac{\pi}{p} + \cos \frac{\pi}{q}\right)$$

If the inequalities (32a) are fulfilled $\cos \theta_{o,1}$, $\cos \theta_{o,2}$ as well as K^*_1 and K^*_2 are all less than unity. Approximate values for $\alpha_{o,1}$, $\alpha_{o,2}$, K^*_1, and K^*_2 for large p and q and small b_{11}, b_{12}, b_{21}, and b_{22} are obtained from (30b) and (31a) where b is to be replaced by $b_{11} + |b_{12}|$ and $b_{22} + |b_{21}|$ respectively.

So far it has been shown that if initially the ratio $\frac{C_2}{C_1} = \pm 1$ for all components of the error field, the K^* for the first scan will be less than unity for both levels provided (32a) holds true. After the first scan $\frac{C_2}{C_1}$ in equations (23b) is replaced by $\frac{K_2}{K_1} \cdot \frac{C_2}{C_1} = \frac{C^{(1)}_2}{C^{(1)}_1}$. For one or more components of the error field this new ratio or its inverse may, however, exceed unity at one of the levels and we will therefore have to consider the general case that C_2 and C_1 are unequal in magnitude either initially or after a certain number of scans. It is sufficient to deal with the largest wave component corresponding to the eigenvalue L_o. Let us assume that $|C_2| > |C_1|$ initially and write the conditions (32a) in the more convenient form

$$L_o + b_{11} + M_1 \, |b_{12}| = 0$$

$$L_o + b_{22} + M_2 \, |b_{21}| = 0 \qquad (32b)$$

where M_1 and M_2 are positive numbers both larger than unity. If the ratios $\left|\frac{C_2}{C_1}\right|$ and $\left|\frac{C_1}{C_2}\right|$ are initially less than M_1 and M_2 respectively and remain so throughout the relaxation procedure, this is analogous to the case $C_2 = \pm 1$ already dealt with and its convergence is ensured when (32a) is satisfied. The remaining cases to be considered are therefore those where $\left|\frac{C_2}{C_1}\right| \geq M_1$

either initially or after a certain number of scans. Let us first take the case where $|C_2/C_1|$ at the outset exceeds M_1. It then follows from the case $C_2 = \pm C_1$ discussed above that after the first scan $K^*_1 > 1$ and $K^*_2 < 1$, i.e., the smaller error C_1 is magnified and the larger one C_2 is diminished. Consequently the terms $b_{12} \cdot C_2/C_1$ and $b_{21} \cdot C_1/C_2$ ~~and~~ in (23b) will become respectively smaller and larger which, in turn, causes K^*_1 and K^*_2 of the second scan to respectively decrease and increase. Thus, the smaller error will increase and the larger one decrease at a decreasing rate until the condition $|C_2/C_1| < M_1$ is fulfilled after which both errors are decreased until the required solution has been reached. With increasing number of scans both errors will either become equal and then decrease essentially at the same rate towards the tolerance limit, or if one is initially below this limit the larger error may surpass this limit before equality is reached. It therefore follows that the convergence criteria (32a) obtained for the case $C_2/C_1 = \pm 1$ also ensures convergence for the more general case $|C_2/C_1| > 1$.

The case that $|C_2/C_1|$ is initially less than M_1 after a certain number of scans is obviously in essence the same as the one just treated. Finally the case $|C_1/C_2| > M_2$ can be treated in a similar way as the case $|C_2/C_1| > M_1$ and leads to the same results.

Thus, we may conclude that the condition (32a) is a sufficient condition for the convergence of both Richardson's and Liebmann's iterative methods when applied to the system (3b). In general this condition is also necessary.

4. Extension to a System of an Arbitrary Number of Helmholtz'-type Equations

The theory of the preceding chapter is readily extended to an arbitrary number m of Helmholtz'-type equations (m-levels). This we may do by first replacing the system (3b) by

$$
\begin{aligned}
L\, \phi_{k,i,j} + b_{k1}\, \phi_{1,i,j} + b_{k2}\, \phi_{2,i,j} + \cdots b_{km}\, \phi_{m,i,j} - q_{k,i,j} &= \\
M_k\, \phi_{l,i,j} - q_{k,i,j} = 0 \qquad\qquad k,l &= 1, 2 \,.\!.\!.\!.\, m
\end{aligned}
\tag{3c}
$$

The iterative scheme is of the form

$$\phi^{n+1}_{k,i,j} = \phi^{n}_{k,i,j} + \alpha_k \left[M_k \, \phi^{n,n+1}_{l,i,j} - q_{k,i,j}\right];$$

$$k, l = 1, 2, \cdots m \tag{16b}$$

where the linear operator M_k is defined by (3c). As in the case $m = 2$ we will assume that the first level ($k = 1$) is scanned first, then the level $k = 2$, etc. successively, ending with the mth level after which we again scan the first level and repeat the whole procedure as often as necessary. The double superscript $n, n + 1$ refers as before to Liebmann's method and is replaced by the single superscript n when Richardson's method is used. In the latter case $M_k \, \phi^{n}_{l,i,j}$ is unambiguously defined by (3c); this is not the case for $M\phi^{n,n+1}_{l,i,j}$ where the distribution of superscripts are determined by the particular sequence in which the m levels are scanned. With the sequence described above $M\phi^{n,n+1}_{l,i,j}$ where the distribution of superscripts are determined by the particular sequence in which the n levels are scanned. With the sequence described above $M\phi^{n,n+1}_{l,i,j}$ is defined as follows:

$$M_k \phi^{n,n+1}_{l,i,j} \equiv L\phi^{n,n+1}_{k,i,j} + \sum_{l=1}^{k-1} b_{k,l} \phi^{n+1}_{l,i,j} + \sum_{l=k+1}^{m} b_{k,l} \, \phi^{n}_{l,i,j}$$

$$+ b_{kk} \phi^{n}_{k,i,j} \tag{2b}$$

where

$$\sum_{l=1}^{k-1} b_{k,l} \, \phi^{n+1}_{l,i,j} = 0 \text{ for } k = 1$$

As before the error $\epsilon^{n}_{k,i,j}$ is defined as

$$\epsilon^{n}_{k,i,j} \equiv \phi^{n}_{k,i,j} - \phi_{k,i,j}; \quad k = 1, 2, \cdots m \tag{5b}$$

where $\phi_{k,i,j}$ is the solution of (3c) with proper boundary conditions; as in the previous cases we assume $\phi_{k,i,j}$ to be known in the boundary points.

Substitution for ϕ^n and ϕ^{n+1} in (16b) by means of (5b) results in the recurrence relation

$$\epsilon^{n+1}_{k,i,j} = \epsilon^{n}_{k,i,j} + \alpha_k \cdot M_k \epsilon^{n,n+1}_{l,i,j}; \qquad k, l = 1, 2 \cdots m \quad (6c)$$

where $M_k\ \epsilon^{n,n+1}$ is defined by (2b). We now define an operator $K_k\ (\alpha_k, b_{k,l})$ such that

$$\begin{aligned} \epsilon^{n+1}_{k,i,j} &= K_k\ \epsilon^{n}_{k,i,j} \quad \text{(interior points)} \\ &= 0 \qquad\qquad\quad \text{(boundary points)} \end{aligned} \qquad (19b)$$

and substitute for ϵ^{n+1} from (19b) into (6c) which then yields

$$\begin{aligned} K_k \epsilon_{k,i,j} = \epsilon_{k,i,j} + \alpha_k\ [&K_k \epsilon_{k,i-1,j} + \epsilon_{k,i,j+1} + \epsilon_{k,i+1,j} \\ &+ K_k\ \epsilon_{k,i,j-1} - (4 - b_{kk})\ \epsilon_{k,i,j} + \sum_{l=1}^{k-1} K_k b_{k,l}\ \epsilon_{l,i,j} \\ &+ \sum_{l=k+1}^{m} b_{k,l}\ \epsilon_{l,i,j}] \end{aligned} \qquad (33a)$$

where the superscript n has been omitted. We seek a trigonometric solution of the form (20a) to (33a), utilizing (22a) to arrive at the following system of equations

$$K_k - 1 + (4 - b_{kk})\ \alpha_k = 2\alpha_k \cdot tA_k + \alpha_k\ [\sum_{l=1}^{k-1} K_k b_{k,l} \frac{C_l}{C_k} + \sum_{l=k+1}^{m} b_{k,l} \frac{C_l}{C_k}]$$

where

$$t = \cos \frac{r\pi}{p} + \cos \frac{s\pi}{q}; \qquad r = 1, 2, \cdots p - 1;\ s = 1, 2, \cdots q - 1$$

Since $K_k = A^2_k$ we may write

$$(1 - \alpha_k \sum_{l=1}^{k-1} b_{k,l} \frac{C_l}{C_k}) A^2_k - 2\alpha_k tA_k + (4 - b_{kk} - \sum_{l=k+1}^{m} b_{k,l} \frac{C_l}{C_k}) \alpha_k - 1 = 0$$

$$k = 1, 2, \cdots m \qquad (23d)$$

If Richardson's iterative method is used (23d) is replaced by

$$K_k - 2\alpha_k t + (4 - \sum_{l=1}^{m} b_{k,l} \frac{C_l}{C_k}) \alpha_k - 1 = 0 \qquad (23e)$$

As in the special case m = 2 we will first assume that initially $|C_l| = |C_k|$, whereas the C_k's for different k's may differ, and subsequently treat the general case that $|C_l| \neq |C_k|$. Furthermore, since the ultimate convergence rate is determined by the t-values corresponding to r = s = 1 and r = p - 1, s = q - 1 we will confine ourselves to these two t's; in Liebmann's method they are both equal to t_o defined by (28). For these two t-values (23e) becomes

$$K_{o,k} = 1 + \alpha_k [L_o + b_{kk} + \sum_{l=1,\ l\neq k}^{m} \delta_k \cdot b_{k,l}] \qquad r = s = 1$$

$$K_{m,k} = 1 + \alpha_k [-(8 + L_o) + b_{kk} + \sum_{l=1,\ l\neq k}^{m} \delta_k\, b_{k,l}] \qquad r = p - 1;\ s = q - 1 \qquad (36)$$

where the subscripts O and m have the meaning given to them in chapter one (see Eq. (11)), and $\delta = \pm 1$. It follows from the first Eq. (36) that after the first iteration step $|K_{o,k}| < 1$ only if

$$L_o + b_{kk} + \sum_{l=1,\ l\neq k}^{m} |b_{k,l}| < 0 \qquad (32c)$$

Moreover, the optimum value for α_k is

$$\alpha_{o,k} = [4 - b_{kk} - \sum_{l=1,\ l\neq k}^{m} |b_{k,l}|]^{-1} \qquad (14c)$$

for which $|K_{o,k}| = |K_{m,k}| = K^*$

We now return to Liebmann's iterative method by solving (23d) with respect to A_k,

$$A_k = \frac{1}{\kappa_k} [\alpha_k t \pm \sqrt{\alpha^2_k t^2 - (4 - b_{kk} - \sum_{l=k+1}^{m} \delta \cdot b_{k,l}) \kappa_k \alpha_k + \kappa_k}\ ] \qquad (26b)$$

where

$$\kappa_k = 1 - \alpha_k \sum_{l=1}^{k-1} \delta\, b_{k,l}; \quad \kappa_1 = 1, \quad \delta = \pm 1$$

In order to investigate under which condition $|A_k|$ is less than unity we differentiate (23d) with respect to α_k and arrive at the following expression for $\frac{dA_k}{d\alpha_k}$

$$\frac{dA_k}{d\alpha_k} = \frac{tA_k - 2 + \frac{1}{2}\left(b_{kk} + A^2{}_k \sum_{l=1}^{k-1} \delta\, b_{k,l} + \sum_{l=k+1}^{m} \delta\, b_{k,l}\right)}{\kappa_k A_k - \alpha_k t} \tag{35a}$$

Similar reasoning as used for the case m = 2 leads to the conclusion that both roots of (26b) are less than unity in magnitude when the condition (32c) is fulfilled. The two roots become equal when α_k assumes the optimum value $\alpha_{o,k}$ which is the smaller root of the equation

$$\alpha^2{}_k t^2 - \left(4 - b_{kk} - \sum_{l=k+1}^{m} \delta\, b_{k,l}\right) \kappa_k \alpha_k + \kappa_k = 0 \tag{27c}$$

When we substitute for κ_k from (34a), (27c) may be written in the form

$$\left[t^2 + \left(4 - b_{kk} - \sum_{l=k+1}^{m} \delta\, b_{k,l}\right) \sum_{l=1}^{k-1} \delta\, b_{k,l}\right] \alpha^2{}_k -$$
$$\left[4 - b_{kk} - \sum_{l=k+1}^{m} \delta\, b_{k,l} + \sum_{l=1}^{k-1} \delta\, b_{k,l}\right] \alpha_k + 1 = 0 \tag{27d}$$

The solution of this equation renders the following expression for $\alpha_{o,k}$ (optimum value of α_k),

$$\alpha_{o,k} = [x\ (1 + \sin \theta_o)]^{-1} \tag{30d}$$

where

$$x = 2\left[1 - \frac{1}{4}\left(b_{kk} + \sum_{l=1,\ l \neq k}^{m} |b_{k,l}|\right)\right] \tag{37}$$

and

$$\cos\,\theta_o = \frac{1}{x}\,[\cos\frac{\pi}{p} + \cos\frac{\pi}{q}] \tag{38}$$

The details of the derivation of (30d) are found in the appendix.

As for the general case that $|{}^{C}l| \neq |{}^{C}k|$, a reasoning analogous to the one applied to the case m = 2 leads to the general conclusion that provided (32c) is satisfied, the initially largest errors $({}^{C}l)_{max}$ will decay during the relaxation procedure until after a certain number of scans it becomes equal to, or possibly less, than one or more of the other errors. The new maximal error will go through a similar procedure until all ${}^{C}l$'s have been diminished below the tolerance limit. Similar to the conclusion reached in the special case, we find that the condition (32c) is sufficient to ensure convergence of both Richardson's and Liebmann's iterative methods when applied to the system (3c). In general (32c) is also necessary.

Appendix

The first of the two equations (27b) may be written as follows:

$$\alpha^2 t^2_o - (4 - b_{11} \mp b_{12})\,\alpha + 1 = (x\,\alpha - 1)^2 - z^2\alpha^2 =$$
$$(x\,\alpha - 1 + z\,\alpha)(x\,\alpha - 1 - z\,\alpha) = 0 \tag{1}$$

where the unknown x and z are to be determined. It is readily verified that

$$x = 2\,[1 - \tfrac{1}{4}\,(b_{11} \pm b_{12})]; \qquad z^2 = x^2 - t^2_o \tag{2}$$

We now define

$$\cos\,\theta_o = \frac{t_o}{x}\,, \quad \text{thus } t^2_o = x^2\cos^2\theta_o \qquad \text{and,}$$

$$z^2 = x^2\,\sin^2\,\theta_o$$

Eq. (1) may now be written as

$$(x\ (1 + \sin\ \theta_o)\ \alpha - 1)(x\ (1 - \sin\ \theta_o)\ \alpha - 1) = 0 \tag{3}$$

The smaller root for α is therefore

$$\alpha = [x\ (1 + \sin\ \theta_o]^{-1} \tag{4}$$

where x is given by (2)

Since α necessarily must be real, one must ascertain that $\cos\ \theta_o \leq 1$; now

$$\cos\ \theta_o = \frac{t_o}{x} = \frac{2 + \frac{1}{2} L_o}{2\ [1 - \frac{1}{4}\ (b_{11} \pm b_{12})]} = \frac{1 + \frac{1}{4} L_o}{1 - \frac{1}{4}\ (b_{11} \pm b_{12})} \leq$$

$$\frac{1 + \frac{1}{4} L_o}{1 - \frac{1}{4}\ (b_{11} + |b_{12}|)}$$

Therefore, $\cos\ \theta_o < 1$ if the first inequality of (32a) is fulfilled; moreover, we replace $\pm\ b_{12}$ by $+\ |b_{12}|$ in the expressions for x and α.

The second of the two equations (27b) is now treated similarly to the first one as follows

$$\alpha^2 t_o - (4 - b_{22})\ \kappa\ \alpha + \kappa = 0$$

where $\kappa = 1 \mp \alpha\ b_{21}$

We substitute for κ and obtain after rearranging terms

$$[t^2_o \pm b_{21}\ (4 - b_{22})]\ \alpha^2 - (4 - b_{22} \pm b_{21})\ \alpha + 1 = (x\ \alpha - 1) - z^2\alpha^2 = 0 \tag{1a}$$

The unknown x and z are given by the equations

$$x = 2\left[1 - \frac{1}{4}(b_{22} \mp b_{21})\right]; \qquad z^2 = x^2 \mp b_{21}(4 - b_{22}) - t^2_o$$

$$x^2 \mp b_{21}(4 - b_{22}) = 4\left[1 - \frac{1}{2}(b_{22} \mp b_{21}) + \frac{1}{16}(b_{22} \mp b_{21})^2\right]$$

$$\mp 4\, b_{21} \pm b_{21} \cdot b_{22} \tag{2a}$$

$$= 4\left[1 - \frac{1}{2}(b_{22} \pm b_{21}) + \frac{1}{16}(b_{22} \pm b_{21})^2\right] = 2^2\left[1 - \frac{1}{4}(b_{22} \pm b_{21})\right]^2 = x^2_1$$

We now define

$$\cos\theta_o = \frac{t_o}{x_1} \quad \text{i.e.,} \qquad t^2_o = x^2_1 \cos^2\theta_o \qquad \text{and}$$

$$z^2 = x^2 \mp b_{21}(4 - b_{22}) - t^2_o = x^2_1 \sin^2 \theta^2_o$$

The smaller root for α is therefore

$$\alpha = [x + x_1 \sin\theta_o]^{-1} \tag{4a}$$

It is readily verified that α is real and $\cos\theta_o < 1$ if the second inequality of (32a) is satisfied. In selecting the optimum value for α we replace the term $\pm b_{21}$ by $+ |b_{21}|$ in the expressions for x, x_1, and α, whereby x becomes equal to x_1 and (4a) is reduced to

$$\alpha = [x(1 + \sin\theta_o)]^{-1} \tag{4b}$$

Both equations (27b) are special cases of Eq. (27d):

$$\left[t^2_o + \left(4 - b_{kk} - \sum_{l=k+1}^{m} \delta\, b_{k,l}\right) \sum_{l=1}^{k-1} \delta\, b_{k,l}\right] \alpha^2_k - \left[4 - b_{kk} - \sum_{l=k+1}^{m} \delta\, b_{k,l}\right.$$

$$\left. + \sum_{l=1}^{k-1} \delta\, b_{k,l}\right] \cdot \alpha_k + 1 \tag{27d}$$

$$= (x\alpha_k - 1) - z^2\alpha^2_k = 0$$

Analogous to the special cases we find the unknown x and z to be given by the equations

$$x = 2\,[1 - \tfrac{1}{4}\,(b_{kk} + \sum_{l=k+1}^{m} \delta\, b_{k,l} - \sum_{l=1}^{k-1} \delta\, b_{k,l})]$$

$$z^2 = x^2 - (4 - b_{kk} - \sum_{l=k+1}^{m} \delta\, b_{k,l}) \sum_{l=1}^{k-1} \delta\, b_{k,l} - t^2{}_o \tag{2b}$$

Now is

$$x^2 - (4 - b_{kk} - \sum_{l=k+1}^{m} \delta\, b_{k,l}) \sum_{l=1}^{k-1} \delta\, b_{k,l} = 4\,[1 - \tfrac{1}{2}\,(b_{kk} + \sum_{l=k+1}^{m} \delta b_{k,l} - \sum_{l=1}^{k-1} \delta b_{k,l})$$

$$+ \tfrac{1}{16}\,(b_{kk} + \sum_{l=k+1}^{m} \delta b_{k,l} - \sum_{l=1}^{k-1} \delta b_{k,l})^2] - (4 - b_{kk} - \sum_{l=k+1}^{m} \delta b_{k,l}) \sum_{l=1}^{k-1} \delta b_{k,l}$$

$$= 4\,[1 - \tfrac{1}{2}\,(b_{kk} + \sum_{l=k+1}^{m} \delta b_{k,l} + \sum_{l=1}^{k-1} \delta b_{k,l})]^2 = x^2{}_1$$

We now define

$$\cos\,\theta_o = \frac{t_o}{x_1} \text{ which renders}$$

$$z^2 = x^2{}_1 - x^2{}_1 \cos^2\,\theta_o = x^2{}_1 \sin^2\,\theta^2{}_o$$

Thus, the smaller root of (27d) is

$$\alpha = [x + x_1 \sin\,\theta_o]^{-1} \tag{4c}$$

In order to ensure ultimate convergence, which in turn requires $\cos\,\theta_o < 1$ inequality (32c) must be satisfied, and in selecting the optimum value for α_k we replace $b_{k,l}$ by $|b_{k,l}|$, dropping the symbol δ in the expressions for x, x_1, and α_k. Hereby x becomes equal to x_1 and (4c) is reduced to

$$\alpha = [x\,(1 + \sin\,\theta_o)]^{-1} \tag{4d}$$

32622

References

1. Frankel, S.P.: Convergence rates of iterative treatments of partial differential equations. Mathematical Tables and Other Aids to Computation, Vol. 4 (1950), pp. 65-75.

2. Charney, J.G. and Phillips, N.A.: Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. Journal of Meteorology, Vol. 10, No. 2, pp. 71-99.

3. Young, D.: Iterative methods for solving partial difference equations of elliptic type. Transactions of the American Mathematical Society, Vol. 76 (1954), pp. 92-111.

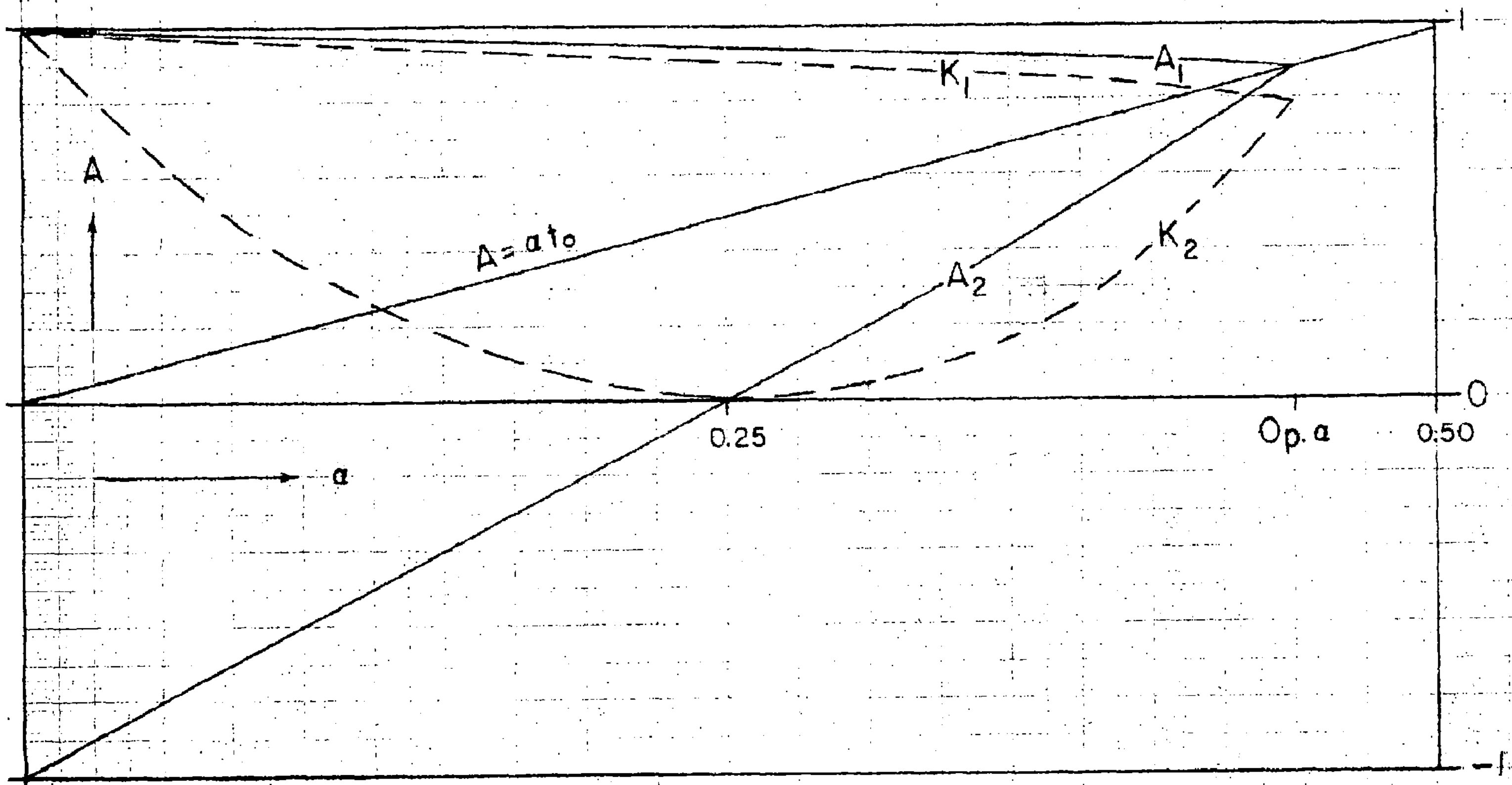

Fig. 1. A graph showing the two roots A_1 and A_2 of equation (26) as functions of the over-relaxation parameter α for $t = t_o$, $p = 30$ and $q = 34$. The graph also shows the curves $K_1 = A^2{}_1$ and $K_2 = A^2{}_2$.