A Convergence Analysis of a Numerical Method for Solving the Balance Equation

S. J. BIJLSMA AND R. J. HOOGENDOORN

Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

(Manuscript received 8 May 1981, in final form 10 May 1982)

ABSTRACT

In this paper the convergence of an iterative method for solving the nonlinear balance equation is analyzed. It is shown that this iterative method, originally proposed by Miyakoda and Shuman, is convergent if a sufficiently accurate initial approximation is used and if the successive iterates satisfy the ellipticity condition. Otherwise the method may be divergent. Experimental results are presented.

1. Introduction

The balance equation is obtained by first applying the two-dimensional divergence operator to the primitive equations of horizontal motion. If only the divergence-free part of the velocity is retained, the resulting equation may be written, in tangent plane coordinates, as

$$f\nabla^2\psi + 2(\psi_{xx}\psi_{yy} - \psi_{xy}^2) + \nabla f \cdot \nabla\psi - \nabla^2\Phi = 0, \quad (1.1)$$

where ψ is the stream function, f the Coriolis parameter, Φ the geopotential and $\nabla^2\psi$ the vorticity. The condition that equation (1.1) be elliptic is (see Courant and Hilbert, 1962)

$$(f + 2\psi_{xx})(f + 2\psi_{yy}) - 4\psi_{xy}^2 > 0. \quad (1.2)$$

In combination with equation (1.1) condition (1.2) can be written in the form

$$\nabla^2\Phi + \frac{f^2}{2} - \nabla f \cdot \nabla\psi > 0. \quad (1.3)$$

Condition (1.2) implies that there are two distinct solutions of (1.1) for given boundary values of ψ, one for which the quantities $f + 2\psi_{xx}$ and $f + 2\psi_{yy}$ are both positive and the other for which these quantities are negative. Since it is generally observed that the absolute vorticity

$$f + \nabla^2\psi$$

is positive, we impose apart from (1.2) or (1.3) the conditions

$$\left.\begin{array}{l} f + 2\psi_{xx} > 0, \\ f + 2\psi_{yy} > 0 \end{array}\right\}. \quad (1.4)$$

Using the identity

$$4\psi_{xx}\psi_{yy} = (\nabla^2\psi)^2 - (\psi_{xx} - \psi_{yy})^2,$$

Eq. (1.1) reads

$$\tfrac{1}{2}[(\nabla^2\psi)^2 - (\psi_{xx} - \psi_{yy})^2 - 4\psi_{xy}^2] + f\nabla^2\psi + \nabla f \cdot \nabla\psi - \nabla^2\Phi = 0 \quad (1.5)$$

or with (1.4)

$$\nabla^2\psi = -f + (2\nabla^2\Phi + f^2 + A^2 + B^2 - 2\nabla f \cdot \nabla\psi)^{1/2}, \quad (1.6)$$

where

$$A = \psi_{xx} - \psi_{yy}, \quad B = 2\psi_{xy}.$$

In the following we shall analyze the convergence behaviour of a commonly used iterative method (see for instance Shuman (1957a) or Miyakoda (1956)) for solving equation (1.6). 'Arnason (1958) showed that a similar method to solve equation (1.1) might be divergent if the relative vorticity takes large positive values.

Based on 'Arnason's work, Paegle and Tomlinson (1975) introduced a modification of the method which is convergent in the case of large positive vorticity. For experiments the reader is referred to page 534 of their article. Applying the same modification to the method of Shuman (1957a), Paegle and Tomlinson found that the modified method also converged where the original method diverged. No vorticity criterion was apparent and no convergence analysis was attempted. We shall try to determine the convergence conditions of the Shuman and modified Shuman methods.

2. The method of solution

We consider a sequence of functions ψ^i ($i = 0, 1, 2, \cdots$), satisfying conditions (1.2) and (1.4), determined by

$$\nabla^2\psi^n = -f + (2\nabla^2\Phi + f^2 + A_{n-1}^2 + B_{n-1}^2 - 2\nabla f \cdot \nabla\psi^{n-1})^{1/2}, \quad (2.1)$$

with $A_{n-1} = \psi_{xx}^{n-1} - \psi_{yy}^{n-1}$, $B_{n-1} = 2\psi_{xy}^{n-1}$, which satisfy the same Dirichlet boundary condition as ψ from equation (1.6). If the difference $\epsilon^n = \psi^n - \psi^{n-1}$ converges towards zero, the sequence of approximate solutions from (2.1) will converge towards the solution ψ of equation (1.6). From (2.1) we have

$$\nabla^2\epsilon^{n+1} = F(\psi^n) - F(\psi^{n-1}),$$

where

$$F(\psi^n) = (2\nabla^2\Phi + f^2 + A_n^2 + B_n^2 - 2\nabla f\cdot\nabla\psi^n)^{1/2}. \quad (2.2)$$

Writing

$$F(\psi^n) - F(\psi^{n-1}) = \frac{F^2(\psi^n) - F^2(\psi^{n-1})}{F(\psi^n) + F(\psi^{n-1})},$$

we find, neglecting the term $2\nabla f\cdot\nabla\epsilon^n$,

$$\nabla^2\epsilon^{n+1} = \frac{A_n^2 - A_{n-1}^2 + B_n^2 - B_{n-1}^2}{2f + \nabla^2(\psi^{n+1} + \psi^n)}, \quad (2.3)$$

where

$$\left.\begin{aligned} A_n^2 - A_{n-1}^2 &= ((\psi^n + \psi^{n-1})_{xx} \\ &\quad - (\psi^n + \psi^{n-1})_{yy})(\epsilon_{xx}^n - \epsilon_{yy}^n), \\ B_n^2 - B_{n-1}^2 &= 4(\psi^n + \psi^{n-1})_{xy}\epsilon_{xy}^n, \\ \nabla^2(\psi^{n+1} + \psi^n) &= \nabla^2(\psi^n + \psi^{n-1}) + \nabla^2(\epsilon^{n+1} + \epsilon^n) \end{aligned}\right\}.$$

With $\bar\psi = \psi^n + \psi^{n-1}$, equation (2.3) becomes

$$\nabla^2\epsilon^{n+1} = \frac{(\bar\psi_{xx} - \bar\psi_{yy})(\epsilon_{xx}^n - \epsilon_{yy}^n) + 4\bar\psi_{xy}\epsilon_{xy}^n}{2f + \nabla^2\bar\psi + \nabla^2(\epsilon^{n+1} + \epsilon^n)}. \quad (2.4)$$

We note that $\epsilon^i = 0$ ($i = 1, 2, \cdots$) on the boundary. For simplicity we consider equation (2.4) on a square ($0 \leq x, y \leq 2\pi$) treating the functions $\bar\psi_{xx}$, $\bar\psi_{yy}$ and $\bar\psi_{xy}$ as constants. Further, we linearize equation (2.4) by neglecting the terms $\nabla^2\epsilon^{n+1}$ and $\nabla^2\epsilon^n$ in the rhs, so that we have instead of (2.4),

$$\nabla^2\epsilon^{n+1} = \frac{(\bar\psi_{xx} - \bar\psi_{yy})(\epsilon_{xx}^n - \epsilon_{yy}^n) + 4\bar\psi_{xy}\epsilon_{xy}^n}{2f + \nabla^2\bar\psi}. \quad (2.5)$$

In view of the foregoing assumptions we may expand ϵ^n and ϵ^{n+1} in the double Fourier series

$$\left.\begin{aligned} \epsilon^n &= \sum_{k,m=-\infty}^{\infty} A_{k,m}^n e^{i(kx+my)}, \\ \epsilon^{n+1} &= \sum_{k,m=-\infty}^{\infty} A_{k,m}^{n+1} e^{i(kx+my)} \end{aligned}\right\}. \quad (2.6)$$

Substitution of (2.6) into (2.5) gives the following relation between $A_{k,m}^{n+1}$ and $A_{k,m}^n$ ($k, m \neq 0$),

$$A_{k,m}^{n+1} = \frac{(\bar\psi_{xx} - \bar\psi_{yy})(k^2 - m^2) + 4km\bar\psi_{xy}}{(k^2 + m^2)(2f + \nabla^2\bar\psi)} A_{k,m}^n. \quad (2.7)$$

From (1.4) we have

$$f + \bar\psi_{xx} = \epsilon_1 > 0, \quad f + \bar\psi_{yy} = \epsilon_2 > 0. \quad (2.8)$$

In order that ϵ^n converges towards zero, the inequality

$$\left|\frac{(\bar\psi_{xx} - \bar\psi_{yy})(k^2 - m^2) + 4km\bar\psi_{xy}}{(k^2 + m^2)(2f + \nabla^2\bar\psi)}\right| < 1 \quad (2.9)$$

must hold for every $k, m \neq 0$. Using (2.8), the amplification factor in (2.7) may be written as

$$\frac{k^2\epsilon_1 + m^2\epsilon_2 - k^2\epsilon_2 - m^2\epsilon_1 + 4km\bar\psi_{xy}}{k^2\epsilon_1 + m^2\epsilon_2 + k^2\epsilon_2 + m^2\epsilon_1}. \quad (2.10)$$

If the successive approximate solutions ψ^i ($i = 0, 1, 2, \cdots$) are required to satisfy the ellipticity condition (1.2), it follows that

$$|\bar\psi_{xy}| < (\epsilon_1\epsilon_2)^{1/2}. \quad (2.11)$$

Writing

$$\left.\begin{aligned} k^2\epsilon_1 + m^2\epsilon_2 &= (|k|\epsilon_1^{1/2} - |m|\epsilon_2^{1/2})^2 + 2|km|(\epsilon_1\epsilon_2)^{1/2}, \\ k^2\epsilon_2 + m^2\epsilon_1 &= (|k|\epsilon_2^{1/2} - |m|\epsilon_1^{1/2})^2 + 2|km|(\epsilon_1\epsilon_2)^{1/2} \end{aligned}\right\},$$

the amplification factor (2.10) reads

$$\frac{(|k|\epsilon_1^{1/2} - |m|\epsilon_2^{1/2})^2 - (|k|\epsilon_2^{1/2} - |m|\epsilon_1^{1/2})^2 + 4km\bar\psi_{xy}}{(|k|\epsilon_1^{1/2} - |m|\epsilon_2^{1/2})^2 + (|k|\epsilon_2^{1/2} - |m|\epsilon_1^{1/2})^2 + 4|km|(\epsilon_1\epsilon_2)^{1/2}},$$

so that, in view of (2.11), condition (2.9) holds. To summarize, the functions ψ^i ($i = 0, 1, 2, \cdots$) which satisfy (2.1) will converge towards the solution ψ of (1.6), if a sufficiently accurate initial approximation ψ^0 of this solution is available in order to justify the linearization of (2.4), provided all the approximate functions satisfy conditions (1.2) and (1.4).

If the approximate solutions are such that an oscillation develops in the solution of the numerical procedure (2.1) resulting in a very slow convergence, it may be advantageous to solve, instead of (2.1), the pair of equations

$$\nabla^2\phi^{n+1} = -f + F(\psi^n),$$

$$\psi^{n+1} = \phi^{n+1} + \omega(\psi^n - \phi^{n+1}), \quad 0 < \omega < 1,$$

where $F(\psi^n)$ is given by (2.2). Elimination of ϕ^{n+1} gives

$$\nabla^2\psi^{n+1} = \omega\nabla^2\psi^n + (1 - \omega)(-f + F(\psi^n)). \quad (2.12)$$

If we write $\omega = \alpha/(1 + \alpha)$, equation (2.12) reads

$$(1 + \alpha)\nabla^2\psi^{n+1} = \alpha\nabla^2\psi^n - f + F(\psi^n), \quad (2.13)$$

which is the form of the modified iterative method of Shuman, used by Paegle and Tomlinson (1975, Eq. (11), p. 531). Analogous to the treatment of (2.1) we can analyze the convergence of the iterative method (2.12). From this analysis it appears that also this method diverges if the original method (2.1) diverges. In the experiments described in Section 4 we shall show that the convergence of the iterative method (2.1) is dominated by the ellipticity condition.

3. The computational method

The transformation of equation (1.1), written in tangent plane coordinates, onto the south polar stereographic projection (the projection plane passes through the circle of 60°N latitude) is given by

$$m^2 f\nabla^2\psi + 2m^4(\psi_{xx}\psi_{yy} - \psi_{xy}^2) + m^2\nabla f\cdot\nabla\psi - m^2\nabla^2\Phi = 0, \quad (3.1)$$

where $m(\varphi)$ is the map factor

$$m(\varphi) = \frac{(1 + \sin\pi/3)}{(1 + \sin\varphi)},$$

and φ is the latitude. Analogous to (3.1), Eq. (1.6) reads

$$m^2\nabla^2\psi = -f + [f^2 + m^4(A^2 + B^2) - 2m^2\nabla f\cdot\nabla\psi + 2m^2\nabla^2\Phi]^{1/2}. \quad (3.2)$$

If Γ denotes the boundary of the region where (3.2) applies, then $\psi(x, y)$ is to satisfy the Dirichlet boundary condition

$$\psi(x, y) = \Phi(x, y)/f, \quad (x, y) \in \Gamma.$$

We now impose a uniform square grid (x_i, y_j) on this region, with mesh side d. The boundary consists of rectangular segments along horizontal and vertical mesh lines, so that an interior point is surrounded by eight mesh points, each of which is an interior or a boundary mesh point. Using the notation $\psi(i, j) = \psi(x_i, y_j)$, $f(i, j) = f(x_i, y_j)$, we can replace partial derivatives by the following usual finite difference approximations at an interior point (i, j):

$$\left.\begin{aligned}
&\psi_{xx}: \frac{(\psi(i-1, j) + \psi(i+1, j) - 2\psi(i, j))}{d^2},\\
&\psi_{yy}: \frac{(\psi(i, j-1) + \psi(i, j+1) - 2\psi(i, j))}{d^2},\\
&\psi_{xy}: \frac{(\psi(i+1, j+1) + \psi(i-1, j-1) - \psi(i+1, j-1) - \psi(i-1, j+1))}{4d^2},\\
&2\nabla f\cdot\nabla\psi: [(f(i+1, j) - f(i-1, j))(\psi(i+1, j) - \psi(i-1, j))\\
&\qquad + (f(i, j+1) - f(i, j-1))(\psi(i, j+1) - \psi(i, j-1))]/2d^2 = \frac{\gamma}{d^2}
\end{aligned}\right\}$$

so that

$$\nabla^2\psi: \frac{(\psi(i+1, j) + \psi(i-1, j) + \psi(i, j-1) + \psi(i, j+1) - 4\psi(i, j))}{d^2} = \frac{D\psi(i, j)}{d^2},$$

$$A^2 + B^2: [(\psi(i-1, j) + \psi(i+1, j) - \psi(i, j-1) - \psi(i, j+1))^2 + \tfrac{1}{4}(\psi(i+1, j+1) + \psi(i-1, j-1) - \psi(i+1, j-1) - \psi(i-1, j+1))^2]/d^4 = \frac{(\alpha^2 + \beta^2)}{d^4},$$

and Eq. (3.2) gives the system of finite-difference equations

$$m^2 D\psi/d^2 = -f + \left[f^2 + \frac{m^4(\alpha^2 + \beta^2)}{d^4} - \frac{m^2\gamma}{d^2} + \frac{2m^2 D\Phi}{d^2}\right]^{1/2}, \quad (3.3)$$

where we dropped the indices i, j. Eq. (3.3) is solved by the iteration process

$$D\psi^{n+1} = -\frac{fd^2}{m^2} + \left[\left(\frac{fd^2}{m^2}\right)^2 + \alpha_n{}^2 + \beta_n{}^2 - \frac{\gamma_n d^2}{m^2} + \frac{2D\Phi d^2}{m^2}\right]^{1/2}, \quad (3.4)$$

where the quantities α_n, β_n and γ_n are calculated using the nth iterate ψ^n. As an initial guess we take $\psi^0 = \Phi/\bar{f}$, where $\bar{f}$ is the average Coriolis parameter. If we choose a fixed ordering of the interior mesh points, then (3.4) is equivalent to a system of linear equations with the grid function values $\psi^{n+1}(i, j)$ at the ordered mesh points as unknowns.

Analogous to condition (1.2) of Section 1, the condition that Eq. (3.1) be elliptic is given by

$$\left(\frac{f}{m^2} + 2\psi_{yy}\right)\left(\frac{f}{m^2} + 2\psi_{xx}\right) - 4\psi_{xy}^2 > 0 \quad (3.5)$$

or, equivalently,

$$2\nabla^2\Phi + \frac{f^2}{m^2} - 2\nabla f \cdot \nabla\psi > 0. \quad (3.6)$$

Of course, conditions (3.5) and (3.6) are equivalent if ψ is a solution of (3.1). They are not equivalent for functions ψ^n which are approximate solutions of (3.1) obtained by an iterative method such as (2.1).

In the following we shall consider the finite-difference version of (3.5) as the ellipticity condition of our problem, and in the next section we shall show that the iterative method (3.4) will be convergent if the successive approximate solutions ψ^n satisfy the finite-difference version of (3.5). Otherwise, very slow convergence or even divergence may occur. In practice, however, we shall be concerned with condition (3.6). The better (3.6) is satisfied by an approximate solution ψ^n of (3.4), the better (3.5) will be satisfied. In the experiments of the next section we use condition (3.6) in three different ways:

1) In the first iteration of (3.4) the geopotential is changed at those points where $2D\Phi + f^2d^2/m^2 < 0$, so that $2D\Phi = -f^2d^2/m^2$. During further iteration, the geopotential is modified again at those points where

$$2D\Phi + \frac{f^2d^2}{m^2} - \gamma_n < 0$$

so that

$$2D\Phi = \gamma_n - \frac{f^2d^2}{m^2}.$$

This will give only small changes in Φ. For instance, at the point where the modification of Φ is largest, the change of Φ due to the latter operation is, on the average, 0.5% of the change caused by the operation in the first iteration.

2) At those points where $2D\Phi + f^2d^2/m^2 < 0$ the geopotential is modified, so that $2D\Phi = -f^2d^2/m^2$. During the iteration process (3.4) we substitute zero for the expression

$$\left(\frac{fd^2}{m^2}\right)^2 + \alpha_n{}^2 + \beta_n{}^2 - \frac{\gamma_n d^2}{m^2} + \frac{2D\Phi d^2}{m^2}$$

when it is negative.

3) We now apply a method proposed by Shuman (1957a):

"The field of $Z = 2D\Phi + f^2d^2/m^2$ is scanned with a test for negative values. When a negative value of Z is encountered, the values at the surrounding nearest four points are reduced by ¼ of the magnitude of Z at the central point and the value of Z at the central point is increased to zero. Boundary values are excepted from change."

This operation is applied 20 times. With respect to the expression

$$\left(\frac{fd^2}{m^2}\right)^2 + \alpha_n{}^2 + \beta_n{}^2 - \frac{\gamma_n d^2}{m^2} + \frac{2D\Phi d^2}{m^2}$$

we refer to 2).

When the numerical solutions start to oscillate, we may solve the pair of equations

$$D\phi^{n+1} = -\frac{fd^2}{m^2} + \left[\left(\frac{fd^2}{m^2}\right)^2 + \alpha_n{}^2 + \beta_n{}^2 - \frac{\gamma_n d^2}{m^2} + \frac{2D\Phi d^2}{m^2}\right]^{1/2},$$

$$\psi^{n+1} = \phi^{n+1} + \omega(\psi^n - \phi^{n+1}), \quad 0 < \omega < 1. \quad (3.7)$$

Iteration is stopped when a function ψ^n is obtained which makes the maximum of the absolute values of the residuals with respect to equation (3.3) less than a prescribed quantity. A residual is defined as the difference between the right and left member of (3.3). In each iteration step we must solve a Poisson equation. If this is done by an iterative method (as is the case in this paper), the required accuracy must be chosen in accordance with this stop criterion (see also Section 4).

4. Experiments

The tabulated geopotential is smoothed by means of the one-dimensional 3-element operator of Shuman [1957b, Eqs. (2) and (8)], by applying it six times successively in x and y direction.

In the following $\bar{F}(\psi^n)$ stands for the square root in the right member of (3.4). Using the stop criterion

$$\max_{i,j}|D\tilde{\psi}^n + \frac{fd^2}{m^2} - \bar{F}(\tilde{\psi}^n)| < \delta \quad (4.1)$$

for the iterative method (3.4), where $\tilde{\psi}^n$ is the approximate solution of

$$D\psi^n = -\frac{fd^2}{m^2} + \bar{F}(\psi^{n-1}) \quad (4.2)$$

or

$$D\psi^n = \omega D\psi^{n-1} + (1-\omega)\left(-\frac{fd^2}{m^2} + \bar{F}(\psi^{n-1})\right).$$

When solving (3.7), this approximate solution must be computed with an accuracy so that inequality (4.1) can be satisfied at the end of the iteration process. It appears sufficient to solve the discrete Poisson equation (4.2) with an accuracy of

TABLE 1. Experimental results. For explanation of cases 1, 2, 3, see Section 3.

January 1978				
Day	Time (GMT)	Case 1	Case 2	Case 3
5	0000	16 (4)	17 (12)	32 (95)
	1200	17 (1)	17 (11)	* (74)
6	0000	19 (4)	19 (16)	25 (88)
	1200	17 (3)	18 (10)	29 (46)
7	0000	24 (1)	42 (11)	48 (65)
	1200	15 (1)	15 (11)	26 (45)
8	0000	27 (0)	28 (18)	36 (55)
	1200	17 (0)	17 (5)	33 (49)

$$\max_{i,j}|D\tilde{\psi}^n - D\psi^n| < 10^{-q} \max_{i,j}\left|D\psi^{n-1} + \frac{fd^2}{m^2} - \bar{F}(\psi^{n-1})\right|$$

with $q = 1$. In the following experiments, $\delta = 10^3$ and $d = 375.10^3$, so that

$$\max_{i,j}\left|\frac{m^2D\tilde{\psi}^n}{d^2} + f - \frac{m^2\bar{F}(\tilde{\psi}^n)}{d^2}\right| < \frac{10^{-7}}{(3.75)^2}\max_{i,j} m^2(i, j).$$

It is the intention of the experiments to show that convergence will be slower as the ellipticity is violated to a greater extent at a larger number of points. The balance equation is solved on an octagonal region covering the northern hemisphere to 15°N latitude and containing 1624 interior mesh points. The 12-h 500 mb geopotential fields used in this paper are taken from the period 5–8 January 1978. The results presented below are representative of other experiments we have done. In Table 1 the number of iterations is given, needed to satisfy the stop criterion (4.1) (the asterisk means more than 50 iterations). Between parens we give the number of points at which the ellipticity condition (3.5) was not satisfied during the last iteration. For instance, in Case 1 of 0000 GMT 8 January 1978 the ellipticity was violated to a great extent at a few points during many preceding iterations, resulting in a relatively slow convergence. It was not possible to reduce the number of iterations in Case 3 of 1200 GMT 5 January 1978 by applying the iterative method (3.7).

5. Concluding remarks

In the Introduction we mentioned that the iterative method of 'Arnason (1958) for solving (1.1) may diverge if the relative vorticity, calculated with approximate streamfunctions, takes large positive values, although these successive iterates satisfy the ellipticity condition. Motivated by this fact, Paegle and Tomlinson (1975) introduced a modification, analogous to (2.13), which made the method convergent. From their paper one might get the impression that the introduction of α in (2.13) would have a comparable effect. To demonstrate that this is not the case we showed in this paper that

1) the iterative method (2.1) is convergent if the ellipticity condition is satisfied by successive iterates and

2) the introduction of α in (2.13) is equivalent to taking weighted means between two successive iterates, which may accelerate convergence in special cases, but cannot change a divergent method (2.1) into a convergent one.

REFERENCES

'Arnason, G., 1958: A convergent method for solving the balance equation. *J. Meteor.*, **15**, 220–225.

Courant, R., and D. Hilbert, 1962: *Methods of Mathematical Physics*, Vol. II. Interscience Publishers, 322–325.

Miyakoda, K., 1956: On a method of solving the balance equation. *J. Meteor. Soc. Japan*, **34**, 364–367.

Paegle, J., and E. M. Tomlinson, 1975: Solution of the balance equation by Fourier transform and Gauss elimination. *Mon. Wea. Rev.*, **103**, 528–535.

Shuman, F. G., 1957a: Numerical methods in weather prediction. I: The balance equation. *Mon. Wea. Rev.*, **85**, 329–332.

——, 1957b: Numerical methods in weather prediction. II: Smoothing and filtering. *Mon. Wea. Rev.*, **85**, 357–361.