An Implicit Compact Fourth-Order Algorithm for Solving the Shallow-Water Equations in Conservation-Law Form

I. M. NAVON AND H. A. RIPHAGEN

National Research Institute for Mathematical Sciences of the CSIR, Pretoria, South Africa

(Manuscript received 23 December 1978, in final form 4 May 1979)

ABSTRACT

An alternating-direction implicit finite-difference scheme is developed for solving the nonlinear shallow-water equations in conservation-law form.

The algorithm is second-order time accurate, while fourth-order compact differencing is implemented in a spatially factored form. The application of the higher order compact Padé differencing scheme requires only the solution of either block-tridiagonal or cyclic block-tridiagonal coefficient matrices, and thus permits the use of economical block-tridiagonal algorithms. The integral invariants of the shallow-water equations, i.e., mass, total energy and enstrophy, are well conserved during the numerical integration, ensuring that a realistic nonlinear structure is obtained.

Largely in an experimental way, two methods are investigated for determining stable approximations for the extraneous boundary conditions required by the fourth-order method. In both methods, third-order uncentered differences at the boundaries are utilized, and both preserve the overall fourth-order convergence rate of the more accurate interior approximation.

A fourth-order dissipative term was added to the equations to overcome the increased aliasing due to the fourth-order method. Alternatively, Wallington and Shapiro low-pass filters were applied.

The numerical integration of the shallow-water equations is performed in a channel corresponding to a middle-latitude band. A linearized version of this method is shown to be unconditionally stable.

1. Introduction

We consider the application of implicit fourth-order compact finite-difference schemes for solving numerically the nonlinear shallow-water equations in conservation-law form:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{P}(\mathbf{U})}{\partial x} + \frac{\partial \mathbf{Q}(\mathbf{U})}{\partial y} - f\mathbf{R}(\mathbf{U}) = 0, \qquad (1)$$

where **U** is an unknown p-component vector and **P**, **Q** and **R** are given vector-valued functions of the components of **U**.

We confine attention to the case of two spatial dimensions. Recently, motivated by a suggestion of Kreiss (Orszag and Israeli, 1974), several higher order finite-difference schemes with similar properties have been developed by a number of investigators.

In all these methods both the functions and their derivatives are considered unknown at several grid points or the function values are collocated at several grid points (usually three) instead of at just one.

Historically these methods may be traced back to Numerov (1924, 1927) and Fox and Goodwin (1949), who used them in the numerical integration of ordinary differential equations. These methods have been termed Hermitian finite differences by Collatz (1960), Adam (1975) and Peters (1974); Padé differencing approximations by Beam and Warming (1976) and Lomax (1976); high-order compact differencing by Hirsch (1975), Ciment and Leventhal (1975), Ciment *et al.* (1978) and Wirtz *et al.* (1977); fourth-order Mehrstellen by Krause *et al.* (1976) and Kreiss and Oliger (1973); and fourth-order finite-difference method by Orszag and Israeli (1974). Rubin and Khosla (1976) and Rubin and Graves (1975) have shown that the results obtained by compact, fourth-order differencing approximations can be recovered and improved by appropriate fourth-order spline-on-spline methods.

Recently, following Peters (1974) and Swartz (1974), Ciment and Leventhal (1978) and then Ciment *et al.* (1978), a more general spatial approximation method was proposed, called the operator-compact implicit method. In this method, instead of setting up spatial approximations for individual derivative terms, one poses the difference approximation in terms of the spatial operator.

The connection between the compact implicit methods and splines was also pointed out by Swartz and Wendroff (1974) (see Appendix A). Usually fourth-order methods require five-grid-point finite-

0027-0644/79/091107-21$09.25

difference expressions and cause the bandwidth of the system of equations which are to be solved, to increase and become pentadiagonal. In the compact implicit fourth-order algorithm used in the present investigation, the coefficient matrix remains tridiagonal. Also the number of fictitious boundary points is reduced.

In Section 2 of this paper, we present the system of nonlinear shallow-water equations in conservation-law form, which leads to system (1). In Section 3 we obtain the second-order time-accurate, factored algorithm for system (1), and then the implicit fourth-order form of the algorithm. It is shown that, as opposed to the Eulerian gas dynamic equations with a polytropic equation of state, for the shallow-water equations the nonlinear functions **P(U)** and **Q(U)** are not homogeneous functions of the components of **U**.

In Section 4 details are given of the computational procedure for solving the fourth-order compact implicit shallow-water equations. Section 5 is entirely devoted to the question of extraneous boundary conditions (i.e., boundary conditions required by the fourth-order difference equations but not by the differential equations). Two sets of boundary conditions are compared experimentally, while use is made of theoretical results due to Kreiss (1970), Kreiss and Oliger (1972, 1973), Oliger (1974), Gustafsson *et al.* (1972), Osher (1973, 1974) and Skollermo (1975a, b).

Owing to the larger aliasing error inherent in the fourth-order schemes (see Grammeltvedt, 1969; Orszag, 1971), we found it necessary to add a fourth-order dissipative term to control the increase of small-scale energy. Details of the dissipative algorithm are given in Section 6. Alternative low-pass filters due to Wallington (1962), Shuman (1955) and Shapiro (1970) are also discussed in that section.

In Section 7 a linearized stability analysis of the algorithm is provided.

The numerical results of test calculations are given in Section 8. The conservation of integral invariants of the shallow-water equations is discussed, stressing the importance of enstrophy conservation if the nonlinear structure of the equations is to be correctly modeled. The accuracy of the fourth-order compact implicit shallow-water equations scheme is then tested by using the same scheme with double-mesh resolution in each horizontal direction and by comparing it with a highly accurate nonlinear ADI method due to Gustafsson (1971a).

It is expected that if this algorithm were tested in a real numerical forecasting model, it would enable low-pressure and high-pressure systems to be located more accurately (see Navon and Alperson, 1978).

2. The shallow-water equations

We consider the shallow-water equations, that is, the primitive equations for an incompressible, inviscid fluid with a free surface confined to a channel corresponding to a middle-latitude band. The north and south boundaries are rigid walls, while the flow is assumed to be periodic in the east–west direction.

The beta plane approximation is made.

The basic nonlinear shallow-water equations in Eulerian form are

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} - fv + g\frac{\partial h}{\partial x} = 0 \\ &\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + fu + g\frac{\partial h}{\partial y} = 0 \\ &\frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} + h\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0 \end{aligned}\right\} \tag{2}$$

for a rectangular domain, $0 \leq x \leq L$, $0 \leq y \leq D$, $t \geq 0$. Variables are defined as follows:

- x, y east–west and north–south coordinates, respectively
- t time
- u, v velocity components in the x and y directions, respectively $[u = u(x,y,t),\ v = v(x,y,t)]$
- h depth of the fluid
- g acceleration of gravity, constant
- f Coriolis force $[=\hat{f} + \beta(y - D/2), \hat{f}, \beta \text{ constant}]$

Following Houghton *et al.* (1966), one can write (2) in conservation-law form (i.e., divergence form) as

$$\left.\begin{aligned} &\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}(hu^2) + \frac{\partial}{\partial y}(huv) + gh\frac{\partial h}{\partial x} - fvh = 0 \\ &\frac{\partial}{\partial t}(hv) + \frac{\partial}{\partial x}(huv) + \frac{\partial}{\partial y}(hv^2) + gh\frac{\partial h}{\partial y} + fuh = 0 \\ &\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) + \frac{\partial}{\partial y}(hv) = 0 \end{aligned}\right\} \tag{3}$$

or in matrix form as

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y} - f\mathbf{R} = 0, \tag{4}$$

where **U**, **P**, **Q** and **R** are the column matrices

$$\mathbf{U}=\begin{bmatrix}\tilde m\\ \tilde n\\ h\end{bmatrix},\quad \mathbf{P}=\begin{bmatrix}\dfrac{\tilde m^2}{h}+\tfrac12 gh^2\\ \dfrac{\tilde m\tilde n}{h}\\ \tilde m\end{bmatrix},\quad \mathbf{Q}=\begin{bmatrix}\dfrac{\tilde m\tilde n}{h}\\ \dfrac{\tilde n^2}{h}+\tfrac12 gh^2\\ \tilde n\end{bmatrix},\quad \mathbf{R}=\begin{bmatrix}\tilde n\\ -\tilde m\\ 0\end{bmatrix},\tag{5}$$

in which $\tilde m = hu$, $\tilde n = hv$. We assume periodic solutions in the x-direction, i.e.,

$$\mathbf{U}(x,y,t) = \mathbf{U}(x + L, y, t).\tag{6}$$

Then, with the boundary conditions

$$v(x,0,t) = v(x,D,t) = 0\tag{7}$$

and the initial condition

$$\mathbf{U}(x,y,0) = \boldsymbol{\psi}(x,y),\tag{8}$$

the total energy

$$E = \frac{1}{2}\int_0^L\int_0^D (u^2 + v^2 + gh)h\,dxdy\tag{9}$$

is independent of time.

Also independent of time are the average values of the height of the free surface

$$\bar h = \frac{\int_0^L\int_0^D h\,dxdy}{\int_0^L\int_0^D dxdy}\tag{10}$$

and the enstrophy

$$Z = \iint\left(\frac{q^2}{h}\right)dxdy,\tag{11}$$

where

$$q = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} + f.\tag{11a}$$

3. The basic algorithm

a. Time-differencing and linearization

Denoting by a superscript n the time level $n\Delta t$, where Δt is the time increment, we start by using a trapezoidal time-differencing scheme (Beam and Warming, 1976; Briley and McDonald, 1977):

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \frac{\Delta t}{2}\left[\left(\frac{\partial \mathbf{U}}{\partial t}\right)^n + \left(\frac{\partial \mathbf{U}}{\partial t}\right)^{n+1}\right] + O(\Delta t^3).\tag{12}$$

If the scheme (12) is applied to (4), one obtains

$$\mathbf{U}^{n+1} = \mathbf{U}^n - \frac{\Delta t}{2}\left[\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y} - f\mathbf{R}\right)^n + \left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y} - f\mathbf{R}\right)^{n+1}\right] + O(\Delta t^3).\tag{13}$$

As

$$\mathbf{P}^{n+1} = \mathbf{P}(\mathbf{U}^{n+1}) \quad\text{and}\quad \mathbf{Q}^{n+1} = \mathbf{Q}(\mathbf{U}^{n+1})\tag{14}$$

are nonlinear functions of $\mathbf{U}^{n+1}$, a linearization procedure (see Steger, 1978; Beam and Warming, 1976) involving a local Taylor expansion about $\mathbf{U}^n$, is employed to overcome the nonlinearity of the problem:

$$\left.\begin{aligned}\mathbf{P}^{n+1} &= \mathbf{P}^n + \mathbf{A}^n(\mathbf{U}^{n+1} - \mathbf{U}^n) + O(\Delta t^2)\\ \mathbf{Q}^{n+1} &= \mathbf{Q}^n + \mathbf{B}^n(\mathbf{U}^{n+1} - \mathbf{U}^n) + O(\Delta t^2)\end{aligned}\right\},\tag{15}$$

where the matrices

$$\mathbf{A} = \frac{\partial \mathbf{P}}{\partial \mathbf{U}},\quad \mathbf{B} = \frac{\partial \mathbf{Q}}{\partial \mathbf{U}}\tag{16}$$

are Jacobian matrices with elements

$$\left(\frac{\partial \mathbf{P}}{\partial \mathbf{U}}\right)_{qr} = \frac{\partial \mathbf{P}_q}{\partial \mathbf{U}_r}\quad\text{and}\quad \left(\frac{\partial \mathbf{Q}}{\partial \mathbf{U}}\right)_{qr} = \frac{\partial \mathbf{Q}_q}{\partial \mathbf{U}_r}.\tag{17}$$

Substituting (15) into (13), a linear system for $\mathbf{U}^{n+1}$ is obtained:

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\right]\right\}\mathbf{U}^{n+1} - \frac{\Delta t}{2}f\mathbf{R}^{n+1} = \left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\right]\right\}\mathbf{U}^n - \Delta t\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y}\right)^n + \frac{\Delta t}{2}f\mathbf{R}^n.\tag{18}$$

In Eq. (18) and throughout the paper the notation

$$\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\right]\mathbf{U}^{n+1}\tag{19}$$

is used to denote

$$\frac{\partial}{\partial x}(\mathbf{A}^n\mathbf{U}^{n+1}) + \frac{\partial}{\partial y}(\mathbf{B}^n\mathbf{U}^{n+1}).\tag{20}$$

and **I** is the unit matrix.

b. The ADI factorization

As it stands, Eq. (18) seems to suggest that a large number of operations are required to solve the implicit equations. Clearly, if one could factor the space-difference operators into separate spatial variables, instead of having to solve a formidable matrix inversion problem, one would have only to solve block-tridiagonal systems, using efficient solution algorithms. This significant improvement in efficiency for multidimensional implicit methods is achieved by using the alternating-direction implicit (ADI) algorithm (see Douglas and Gunn, 1964). We first note that in (18) the term $f\mathbf{R}$ can be written

$$f\mathbf{R} = f\begin{pmatrix} \tilde{n} \\ -\tilde{m} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & f & 0 \\ f & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \tilde{m} \\ \tilde{n} \\ h \end{pmatrix} = \mathbf{CU}. \quad (21)$$

Therefore one can write (18) as

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\partial}{\partial y}(\mathbf{B}^n\cdot) - \mathbf{C}\right]\right\}\mathbf{U}^{n+1} = \left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \mathbf{C}\right]\right\}\mathbf{U}^n - \Delta t\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y}\right)^n + O(\Delta t^3). \quad (22)$$

The form of (22) suggests that we establish a factorizable term within the braces by adding the following third-order perturbation terms:

$$\left.\begin{aligned}
&(\mathrm{I}) \quad \frac{\Delta t^3}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\frac{(\mathbf{U}^{n+1} - \mathbf{U}^n)}{\Delta t} = \frac{\Delta t^3}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\frac{\partial}{\partial t}\mathbf{U}^n + O(\Delta t^4) \\
&(\mathrm{II}) \quad \frac{\Delta t^3}{4}\mathbf{C}^{(1)}\mathbf{C}^{(2)}\frac{(\mathbf{U}^{n+1} - \mathbf{U}^n)}{\Delta t} = \frac{\Delta t^3}{4}\mathbf{C}^{(1)}\mathbf{C}^{(2)}\frac{\partial}{\partial t}\mathbf{U}^n + O(\Delta t^4) \\
&(\mathrm{III}) \quad \frac{\Delta t^3}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\mathbf{C}^{(2)}\frac{(\mathbf{U}^{n+1} + \mathbf{U}^n)}{\Delta t} \\
&(\mathrm{IV}) \quad \frac{\Delta t^3}{4}\frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\mathbf{C}^{(1)}\frac{(\mathbf{U}^{n+1} + \mathbf{U}^n)}{\Delta t}
\end{aligned}\right\}, \quad (23)$$

where

$$\mathbf{C}^{(1)} = \begin{bmatrix} 0 & 0 & 0 \\ -f & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{C}^{(2)} = \begin{bmatrix} 0 & f & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (24)$$

$$\mathbf{C}^{(1)} + \mathbf{C}^{(2)} = \mathbf{C}. \quad (25)$$

A scale analysis shows that the last three perturbation terms (II)–(IV) are of the order 10^{-8}, 10^{-4} and 10^{-4}, respectively, compared with the magnitude of the first perturbation term. (Typical magnitudes are $h = 2000$ m, $u = 30$ m s^{-1}, $v = 5$ m s^{-1} and $f = 10^{-4}$ s^{-1}).

The factored scheme can then be written as (see Appendix C)

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) - \mathbf{C}^{(1)}\right]\right\}\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) - \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^{n+1} = \left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \mathbf{C}^{(1)}\right]\right\}\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^n - \Delta t\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y}\right)^n. \quad (26)$$

A three-level centered-time scheme is also tested, constructed from the trapezoidal time-differencing algorithm written over three time levels ($n - 1, n, n + 1$) (see Steger, 1978):

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) - \mathbf{C}^{(1)}\right]\right\}\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) - \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^{n+1} = \left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \mathbf{C}^{(1)}\right]\right\}\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^{n-1} - \Delta t\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y}\right)^n. \quad (27)$$

It is worthwhile mentioning at this point that owing to the presence of the term $\frac{1}{2}gh^2$ in the column matrices **P** and **Q**, these are not homogeneous functions of degree 1 in the variables U_i. Consequently, the simplification obtained by Beam and Warming (1976) and Steger (1978) for the Euler equations of gas dynamics is not applicable here. (See also footnote in Beam and Warming, 1976, p. 95.) This can be shown by calculating the products **AU** and **BU** and comparing the results with the column matrices **P** and **Q**, respectively.

The Jacobian matrices **A** and **B** are given by

$$\mathbf{A} = \begin{bmatrix} 2u & 0 & -u^2 + gh \\ v & u & -uv \\ 1 & 0 & 0 \end{bmatrix}, \tag{28}$$

$$\mathbf{B} = \begin{bmatrix} v & u & -uv \\ 0 & 2v & -v^2 + gh \\ 0 & 1 & 0 \end{bmatrix}, \tag{29}$$

$$\mathbf{BU} = \begin{bmatrix} huv \\ hv^2 + gh^2 \\ hv \end{bmatrix} = \begin{bmatrix} \dfrac{\tilde{n}\tilde{m}}{h} \\ \dfrac{\tilde{n}^2}{h} + gh^2 \\ \tilde{n} \end{bmatrix} = \mathbf{Q} + \begin{bmatrix} 0 \\ \tfrac{1}{2}gh^2 \\ 0 \end{bmatrix}, \tag{30}$$

$$\mathbf{AU} = \begin{bmatrix} hu^2 + gh^2 \\ huv \\ hu \end{bmatrix} = \begin{bmatrix} \tilde{m}^2 + gh^2 \\ \dfrac{\tilde{m}\tilde{n}}{h} \\ \tilde{m} \end{bmatrix} = \mathbf{P} + \begin{bmatrix} \tfrac{1}{2}gh^2 \\ 0 \\ 0 \end{bmatrix}. \tag{31}$$

A computationally convenient form of (26), which emphasizes the spatial splitting, is

$$\bar{\mathbf{U}}^{n+1} = \left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^n, \tag{32a}$$

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) - \mathbf{C}^{(1)}\right]\right\}\bar{\bar{\mathbf{U}}}^{n+1} = \left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \mathbf{C}^{(1)}\right]\right\}\bar{\mathbf{U}}^{n+1} - \Delta t\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y}\right)^n, \tag{32b}$$

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) - \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^{n+1} = \bar{\bar{\mathbf{U}}}^{n+1}. \tag{32c}$$

For the three-level centered-time scheme, we obtain

$$\bar{\mathbf{U}}^{n+1} = \left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^{n-1}, \tag{33a}$$

with the other two intermediary stages being identical to the corresponding (32b) and (32c).

An alternative way of using (30)–(31) is to write the term $-\Delta t[(\partial\mathbf{P}/\partial x) + (\partial\mathbf{Q}/\partial y)]^n$ in the form

$$-\Delta t\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y}\right)^n = -\Delta t\left[\frac{\partial}{\partial x}(\mathbf{AU})^n + \frac{\partial}{\partial y}(\mathbf{BU})^n\right] - \Delta t\left(\frac{\partial}{\partial x}\begin{bmatrix}\tfrac{1}{2}gh^2\\0\\0\end{bmatrix} - \frac{\partial}{\partial y}\begin{bmatrix}0\\ \tfrac{1}{2}gh^2\\0\end{bmatrix}\right)$$

$$= -\Delta t\left[\frac{\partial}{\partial x}(\mathbf{AU})^n + \frac{\partial}{\partial y}(\mathbf{BU})^n\right] + \Delta t\left(\frac{\partial}{\partial x}\begin{bmatrix}\tfrac{1}{2}gh^2\\0\\0\end{bmatrix} + \frac{\partial}{\partial y}\begin{bmatrix}0\\ \tfrac{1}{2}gh^2\\0\end{bmatrix}\right). \tag{34}$$

Then Eq. (26) can be written as

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) - \mathbf{C}^{(1)}\right]\right\}\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) - \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^{n+1} = \left\{\mathbf{I} - \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \mathbf{C}^{(1)}\right]\right\}$$

$$\times\left\{\mathbf{I} - \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^n + \Delta t\left(\frac{\partial}{\partial x}\begin{bmatrix}\tfrac{1}{2}gh^2\\0\\0\end{bmatrix} + \frac{\partial}{\partial y}\begin{bmatrix}0\\ \tfrac{1}{2}gh^2\\0\end{bmatrix}\right), \tag{35}$$

and the algorithm analogous to (32a)–(32c) can then be written

$$\bar{\mathbf{U}}^{n+1} = \left\{\mathbf{I} - \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n \cdot) + \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^n, \tag{36a}$$

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n \cdot) - \mathbf{C}^{(1)}\right]\right\}\bar{\bar{\mathbf{U}}}^{n+1} = \left\{\mathbf{I} - \frac{\Delta t}{2}\left[\frac{\partial}{\partial x}(\mathbf{A}^n \cdot) + \mathbf{C}^{(1)}\right]\right\}\bar{\mathbf{U}}^{n+1} + \Delta t\left(\frac{\partial}{\partial x}\begin{bmatrix}\tfrac{1}{2}gh^2\\ 0\\ 0\end{bmatrix} + \frac{\partial}{\partial y}\begin{bmatrix}0\\ \tfrac{1}{2}gh^2\\ 0\end{bmatrix}\right), \tag{36b}$$

$$\left\{\mathbf{I} + \frac{\Delta t}{2}\left[\frac{\partial}{\partial y}(\mathbf{B}^n \cdot) - \mathbf{C}^{(2)}\right]\right\}\mathbf{U}^{n+1} = \bar{\bar{\mathbf{U}}}^{n+1}. \tag{36c}$$

The same procedure is applied to the three-level centered-time ADI scheme [Eq. (27)]. Here we assumed

$$\mathbf{P} = \mathbf{A}\bar{\mathbf{U}}^{n+1} - \begin{bmatrix}\tfrac{1}{2}gh^2\\ 0\\ 0\end{bmatrix}^n.$$

c. Compact fourth-order spatial differencing

For the approximation of the first spatial derivative, the fourth-order compact spatial differencing takes the form

$$\left(\frac{\partial u}{\partial x}\right)_i = \left[\frac{D_{0x}}{(1 + \Delta x^2 D_{+x}D_{-x}/6)}\right]u_i + O(\Delta x^4), \tag{37}$$

and involves only the grid points $i+1$, i, $i-1$ ($x_i = i\Delta x$),

where

$$\left.\begin{aligned} D_{0x}u_i &= (u_{i+1} - u_{i-1})/2\Delta x\\ D_{+x}u_i &= (u_{i+1} - u_i)/\Delta x\\ D_{-x}u_i &= (u_i - u_{i-1})/\Delta x \end{aligned}\right\}. \tag{38}$$

Eq. (37) is equivalent to

$$\frac{1}{6}\left[\left(\frac{\partial u}{\partial x}\right)_{i+1} + 4\left(\frac{\partial u}{\partial x}\right)_i + \left(\frac{\partial u}{\partial x}\right)_{i-1}\right] = D_{0x}u_i. \tag{39}$$

Thus $(\partial u/\partial x)_i$, $i = 1 \ldots, N_x$, can be determined from u_i by solving a system of linear equations whose coefficient matrix is tridiagonal and of the form

$$\mathbf{J} = \frac{1}{6}\begin{bmatrix} 4 & 1 & & & 0\\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & & 1\\ 0 & & & 1 & 4 \end{bmatrix}_{(N_x \times N_x)}. \tag{40}$$

Ciment and Leventhal (1975) introduced the more convenient notation

$$\left(\frac{\partial u}{\partial x}\right)_i = \left[\frac{D_{0x}}{(1 + \delta_x^2/6)}\right]u_i = Q_x^{-1}D_{0x}u_i + O(\Delta x^4) \tag{41}$$

for Eq. (37), where

$$\left.\begin{aligned} Q_x u_i &= (1 + \delta_x^2/6)u_i = \tfrac{1}{6}(u_{i+1} + 4u_i + u_{i-1})\\ \delta_x^2 u_i &= u_{i+1} - 2u_i + u_{i-1} \end{aligned}\right\}. \tag{42}$$

Application of compact fourth-order differencing to the first space derivatives in the ADI shallow-water algorithm (32a)–(32c) yields

$$\bar{U}_{ij}^{n+1} = \left\{I + \frac{\Delta t}{2}[Q_y^{-1}D_{0y}(B_{ij}^n \cdot) + C_{ij}^{(2)}]\right\}U_{ij}^n, \tag{43a}$$

$$\left\{\mathbf{I} + \frac{\Delta t}{2}[Q_x^{-1}D_{0x}(A_{ij}^n \cdot) - C_{ij}^{(1)}]\right\}\bar{\bar{U}}_{ij}^{n+1} = \left\{I + \frac{\Delta t}{2}[Q_x^{-1}D_{0x}(A_{ij}^n \cdot) + C_{ij}^{(1)}]\right\}\bar{U}_{ij}^{n+1} - \Delta t(Q_x^{-1}D_{0x}P_{ij}^n + Q_y^{-1}D_{0y}Q_{ij}^n), \tag{43b}$$

$$\left\{I + \frac{\Delta t}{2}[Q_y^{-1}D_{0y}(B_{ij}^n \cdot) - C_{ij}^{(2)}]\right\}U_{ij}^{n+1} = \bar{\bar{U}}_{ij}^{n+1}, \tag{43c}$$

$$i = 1, \ldots, N_x,$$
$$j = 1, \ldots, N_y,$$

where $N_x\Delta_x = L$, $N_y\Delta_y = D$.

The same procedure is applied to the three-level centered-time scheme [Eq. (27)].

4. Computational procedure

To evaluate (43a) we write it in the form

$$\bar{\mathbf{U}}^{n+1} = \left[I + \frac{\Delta t}{2}\mathbf{C}^{(2)}\right]\mathbf{U}^n + \frac{\Delta t}{2}Q_y^{-1}D_{0y}(B^nU^n). \tag{44}$$

For this one-dimensional problem we first solve a block-tridiagonal system

$$Q_y W^n = D_{0y}(B^nU^n) \tag{45}$$

to obtain

$$W^n = Q_y^{-1}D_{0y}(B^nU^n). \quad (46)$$

In the block-tridiagonal system the individual blocks are (3 × 3). We then evaluate

$$\bar{\mathbf{U}}^{n+1} = \left[I + \frac{\Delta t}{2}\mathbf{C}^{(2)}\right]\mathbf{U}^n + \frac{\Delta t}{2}W^n \quad (47)$$

and with the definition

$$\alpha_2 = \Delta t/2, \quad (48)$$

we obtain

$$\begin{bmatrix}\bar{U}_1^{n+1}\\ \bar{U}_2^{n+1}\\ \bar{U}_3^{n+1}\end{bmatrix}_{ij} = \begin{bmatrix}h^nu^n + \alpha_2 fh^nv^n + \alpha_2 W_1^n\\ h^nv^n + \alpha_2 W_2^n\\ h^n + \alpha_2 W_3^n\end{bmatrix}_{ij}. \quad (49)$$

For Eq. (43b) we start by evaluating the right-hand term

$$-\Delta t\frac{\partial \mathbf{Q}^n}{\partial y} = -\Delta t Q_y^{-1}D_{0y}\mathbf{Q}^n. \quad (50)$$

We define

$$\mathbf{Y} = Q_y^{-1}D_{0y}\mathbf{Q}^n \quad (51)$$

and solve the block-tridiagonal system

$$Q_y\mathbf{Y} = D_{0y}\mathbf{Q}^n. \quad (52)$$

We can then write the right-hand side of (43b) as

$$\bar{\mathbf{V}}^{n+1} = \left[I + \frac{\Delta t}{2}\mathbf{C}^{(1)}\right]\bar{\mathbf{U}}^{n+1} - \Delta t\mathbf{Y} + \frac{\Delta t}{2}(Q_x^{-1}D_{0x}(\mathbf{A}^n\bar{\mathbf{U}}^{n+1} - 2\mathbf{P}^n)). \quad (53)$$

Multiplying (43b) from the left by the operator Q_x, we then obtain

$$\left[Q_x + \frac{\Delta t}{2}(D_{0x}(\mathbf{A}^n\cdot) - Q_x\mathbf{C}^{(1)})\right]\bar{\mathbf{U}}^{n+1} = Q_x\bar{\mathbf{V}}^{n+1} = \left[Q_x + \frac{\Delta t}{2}Q_x\mathbf{C}^{(1)}\right]\bar{\mathbf{U}}^{n+1} - \frac{\Delta t}{2}Q_x\mathbf{Y} + \frac{\Delta t}{2}[D_{0x}(\mathbf{A}^n\bar{\mathbf{U}}^{n+1} - 2\mathbf{P}^n)]. \quad (54)$$

Here, owing to the cyclic boundary conditions in the x-direction, cyclic block-tridiagonal systems have to be solved for each $j = 1, \ldots, N_y$.

Efficient algorithms for solving cyclic tridiagonal systems were proposed by Temperton (1975), Navon (1977) and Hindmarsh (1977), among others, and were generalized to block-cyclic tridiagonal matrices by Navon (1977).

For a given j, the cyclic block-tridiagonal matrix resulting from the discretization of (54) has the form

$$\mathbf{R} = \begin{bmatrix} E_1 & F_1 & & & D_1\\ D_2 & \ddots & \ddots & 0 & \\ & \ddots & \ddots & \ddots & \\ & 0 & \ddots & \ddots & F_{N_x-1}\\ F_{N_x} & & & D_{N_x} & E_{N_x}\end{bmatrix} \quad (55)$$

with

$$D_{ij} = \begin{bmatrix} 1 - 2\alpha u & 0 & -\alpha(-u^2 + gh)\\ -\alpha v + \dfrac{\Delta t f}{2} & 1 - \alpha u & \alpha uv\\ -\alpha & 0 & 1\end{bmatrix}_{ij}^{\overline{(n+1)}}, \quad (56)$$

$$E_{ij} = \begin{bmatrix} 4 & 0 & 0\\ 2\Delta t f & 4 & 0\\ 0 & 0 & 4\end{bmatrix}_{ij}^{\overline{(n+1)}}, \quad (57)$$

$$F_{ij} = \begin{bmatrix} 1 + 2\alpha u & 0 & \alpha(-u^2 + gh)\\ \alpha v + \dfrac{\Delta t f}{2} & 1 + \alpha u & -\alpha uv\\ \alpha & 0 & 1\end{bmatrix}_{ij}^{\overline{(n+1)}}, \quad (58)$$

where $\alpha = 6\Delta t/4\Delta x$.

Having obtained $\bar{\mathbf{U}}^{n+1}$ we finally multiply (43c) from the left by the operator Q_y to obtain

$$Q_y\mathbf{U}^{n+1} + \frac{\Delta t}{2}D_{0y}(\mathbf{B}^n\mathbf{U}^{n+1}) - \frac{\Delta t}{2}Q_y(\mathbf{C}^{(2)}\mathbf{U}^{n+1}) = Q_y\bar{\mathbf{U}}^{n+1}. \quad (59)$$

A block-tridiagonal matrix with (3×3) individual blocks and of dimension N_y has to be inverted for each $i = 1, \ldots, N_x$ at each time step.

For given i and j the (3×3) blocks have the following entries

$$D_{ij} = \begin{bmatrix} 1 - \alpha v & -\alpha u - \dfrac{\Delta t f}{2} & \alpha u v \\ 0 & 1 - 2\alpha v & -\alpha(-v^2 + gh) \\ 0 & -\alpha & 1 \end{bmatrix}_{ij}, \tag{60}$$

$$E_{ij} = \begin{bmatrix} 4 & -2\Delta t f & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}_{ij}, \tag{61}$$

$$F_{ij} = \begin{bmatrix} 1 + \alpha v & \alpha u - \Delta t f & -\alpha u v \\ 0 & 1 + 2\alpha v & \alpha(-v^2 + gh) \\ 0 & \alpha & 1 \end{bmatrix}_{ij}. \tag{62}$$

The inverse of a (3×3) matrix was explicitly calculated to increase the efficiency of the program.

5. Boundary conditions for the fourth-order compact ADI algorithm

It appears to be no simple matter to deal with extraneous boundary conditions when fourth-order accurate approximation of the hyperbolic mixed initial-boundary value problems is attempted.

After the theoretical advances made by Gustafsson *et al.* (1972) and Oliger and Sundström (1976), numerous investigators such as Elvius and Sundström (1973), Adam (1977), Gary (1978), Oliger (1974), Osher (1973, 1974), Skollermo (1975a, 1975b), Chu and Sereny (1974) and Gottlieb and Turkel (1978) have recently addressed the problem.

Oliger (1974) proposed an $O(h^4)$ approximation with an $O(h^3)$ extrapolation at the boundaries using time averaging to stabilize the boundary approximation. This method has been further elaborated by Gary (1978). Adam (1977) proposed a third-order accurate boundary condition which can be applied to fourth-order accurate compact finite differencing and preserves the tridiagonal character of the algorithm. This method has since been successfully employed by Peyret (1978) for a compact implicit finite-difference solution of the stationary Navier-Stokes equations.

In both approaches a convergence result is used due to Gustafsson (1971b, 1975), according to which, provided the scheme is stable with regard to boundary conditions, it is possible to use at the boundaries approximations one order lower in accuracy and yet retain the convergence of the more accurate interior approximation. Elvius and Sundström (1973) succeeded in demonstrating stability of a set of boundary conditions for a second-order space approximation of the shallow-water equations. However, no proof of stability has yet been derived for the boundary conditions of a fourth-order space approximation to the nonlinear shallow-water equations, and one has to rely on experimental evidence.

$O(h^3)$ approximations for the extraneous boundary conditions are often stabilized by adding one-dimensional dissipative operators acting in the direction of the coordinate tangential to the boundary (Kreiss and Oliger, 1973).

In this work we decided to use the Adam (1977) and Oliger (1974) boundary conditions.

In Eqs. (43a) and (43b) we make use of the periodicity in the x direction.

In the y direction we first use the Adam (1977) boundary conditions. For instance, for

$$\frac{\partial}{\partial y}(\mathbf{B}^n\mathbf{U}^n)_{ij} = Q_y^{-1}D_{0y}(\mathbf{B}^n\mathbf{U}^n)_{ij} = \bar{W}_{ij}^{n+1}, \tag{63}$$

we write

$$Q_y \bar{W}_{ij}^{n+1} = D_{0y}(\mathbf{B}^n\mathbf{U}^n)_{ij} \tag{64}$$

and, at $j = 1$,

$$\bar{W}_{i,1}^{n+1} + 2\bar{W}_{i,2}^{n+1} = \frac{1}{2\Delta y}(-5(\mathbf{B}^n\mathbf{U}^n)_{i,1} + 4(\mathbf{B}^n\mathbf{U}^n)_{i,2} + (\mathbf{B}^n\mathbf{U}^n)_{i,3}) + O(h^3),$$

$$2\bar{W}_{i,2}^{n+1} + \bar{W}_{i,3}^{n+1} = \frac{1}{2\Delta y}((\mathbf{B}^n\mathbf{U}^n)_{i,1} - 4(\mathbf{B}^n\mathbf{U}^n)_{i,2} + 5(\mathbf{B}^n\mathbf{U}^n)_{i,3}) + O(h^3), \tag{65}$$

and the analogue at $j = N_y$, i.e.,

$$\bar{W}^{n+1}_{i,N_y} + 2\bar{W}^{n+1}_{i,N_y-1} = \frac{1}{2\Delta y}(5(\mathbf{B}^n\mathbf{U}^n)_{i,N_y} - 4(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-1} - (\mathbf{B}^n\mathbf{U}^n)_{i,N_y-2}),$$

$$\bar{W}^{n+1}_{i,N_y-2} + W^n_{i,N_y-1} = \frac{1}{2\Delta y}(-5(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-2} + 4(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-1} + (\mathbf{B}^n\mathbf{U}^n)_{i,N_y}) + O(h^3). \quad (66)$$

For (43c), however, both the value of the derivative and that of the unknown $\mathbf{U}^{n+1}$ are required at the y boundaries. Here, we use an inward–backward extrapolation formula for the unknown U_b^{n+1}, due to Gustafsson *et al.* (1972) and which has been shown to be stable by Elvius and Sundström (1973)

$$(U)_b^{n+1} = 2(U)_{b-1}^n - U_{b-2}^{n-1}, \quad (67)$$

where b denotes the boundary grid point.

The Oliger (1974) method is also adequate for an $O(h^4)$ approximation with $O(h^3)$ extrapolation at the boundaries. For Eq. (63), we obtain

$$\bar{W}^{n+1}_{i,0} = \frac{1}{6\Delta y}[-11/2((\mathbf{B}^n\bar{\mathbf{U}}^{n+1})_{i,0} + (\mathbf{B}^n\bar{\mathbf{U}}^{n-1})_{i,0}) + 18(\mathbf{B}^n\mathbf{U}^n)_{i,1} - 9(\mathbf{B}^n\mathbf{U}^n)_{i,2} + 2(\mathbf{B}^n\mathbf{U}^n)_{i,3}] \quad \text{for} \quad j = 0$$

$$\bar{W}^{n+1}_{i,1} = \frac{1}{6\Delta y}[-2(\mathbf{B}^n\mathbf{U}^n)_{i,0} - 3/2((\mathbf{B}^n\bar{\mathbf{U}}^{n+1})_{i,1} + (\mathbf{B}^n\mathbf{U}^{n-1})_{i,1}) + 6(\mathbf{B}^n\mathbf{U}^n)_{i,2} - (\mathbf{B}^n\mathbf{U}^n)_{i,3}] \quad \text{for} \quad j = 1$$

$$\bar{W}^{n+1}_{i,N_y-1} = \frac{1}{6\Delta y}[(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-3} - 6(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-2} + 3/2((\mathbf{B}^n\mathbf{U}^{n+1})_{i,N_y-1} + (\mathbf{B}^n\mathbf{U}^{n-1})_{i,N_y-1} + 2(\mathbf{B}^n\mathbf{U}^n)_{i,N_y}] \quad (68)$$

$$\text{for} \quad j = N_y - 1,$$

$$\bar{W}^{n+1}_{i,N_y} = \frac{1}{6\Delta y}[-2(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-3} + 9(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-2} - 18(\mathbf{B}^n\mathbf{U}^n)_{i,N_y-1} + 11/2((\mathbf{B}^n\bar{\mathbf{U}}^{n+1}_{i,N_y}) + (\mathbf{B}^n\mathbf{U}^{n-1})_{i,N_y})]$$

$$\text{for} \quad j = N_y.$$

In this work we experimented with both the boundary conditions given by Eqs. (65)–(66) and those given by (68).

That caution must be observed when experimenting computationally with extraneous boundary conditions, has been proved by Osher (1974) for the case of linearized shallow-water equations. Taking a set of seemingly reasonable boundary conditions, he showed that this could lead to either non-uniqueness or non-existence of solutions. A thorough study of how the boundary conditions affect the stability and accuracy of implicit methods for hyperbolic equations has been initiated by Skollermo (1975a, 1975b).

6. Fourth-order dissipation

Owing to the larger aliasing error introduced by the fourth-order accurate scheme (Grammeltvedt, 1969; Orzag, 1971), we found it necessary to add a fourth-order dissipative term of the form

$$\frac{-\epsilon}{8}\Delta x^4 D_{+x}{}^2 D_{-x}{}^2 U_i = \frac{-\epsilon}{8}\Delta x^4\left(\frac{\partial^4 \mathbf{U}}{\partial x^4}\right)_i = \frac{-\epsilon}{8}(U_{i+2} - 4U_{i+1} + 6U_i - 4U_{i-1} + U_{i-2}). \quad (69)$$

For the simple hyperbolic system

$$\frac{\partial \mathbf{U}}{\partial t} = \mathbf{A}\frac{\partial \mathbf{U}}{\partial x} \quad (70)$$

the eigenvalues of the amplification matrix, if a leapfrog time discretization of the following form is used:

$$\mathbf{U}(x, t + \Delta t) = \left(I - \frac{\epsilon \Delta x^4}{16} D_{+x}{}^2 D_{-x}{}^2\right)\mathbf{U}(x - \Delta t) + 2\Delta t A D_{0x}\mathbf{U}(x,t), \quad (71)$$

are $|K_j| = 1 - \epsilon \sin^4\xi/2$ for $\epsilon < 1$ and $\Delta t/\Delta x \leq 1 - \epsilon$, where

$$\xi = W\Delta x = \frac{2\pi}{J}\Delta x, \quad (72)$$

where J is the wavelength.

The result illustrates the fact that waves shorter than four times the grid size are the ones most affected by the dissipative term. On the other hand these very waves are the ones subject to aliasing. The fourth-order dissipative term was appended to (43a)–(43c) as follows (see Beam and Warming, 1976):

$$\bar{U}_{ij}^{n+1} = \left[I + \frac{\Delta t}{2}(Q_y^{-1}D_{0y}(B_{ij}^n\cdot) + C_{ij}^{(2)})\right]U_{ij}^n - \frac{\epsilon_y}{16}D_{+y}{}^2D_{-y}{}^2\bar{U}_{ij}^n, \quad (73a)$$

$$\left[I + \frac{\Delta t}{2}(Q_x^{-1}D_{0x}(A_{ij}^n\cdot) - C_{ij}^{(1)})\right]\bar{\bar{U}}_{ij}^{n+1} = [I + \Delta t(Q_x^{-1}D_{0x}(A_{ij}^n\cdot) + C_{ij}^{(1)})]\bar{U}_{ij}^{n+1}$$

$$- \Delta t(Q_x^{-1}D_{0x}P_{ij}^n + Q_y^{-1}D_{0y}Q_{ij}^n) - \frac{\epsilon_x}{8}D_{+x}{}^2D_{-x}{}^2\bar{U}_{ij}^{n+1} \quad (73b)$$

$$\left[I + \frac{\Delta t}{2}(Q_y^{-1}D_{0y}(B_{ij}^n\cdot) - C_{ij}^{(2)})\right]U_{ij}^{n+1} = \bar{\bar{U}}_{ij}^{n+1} - \frac{\epsilon_y}{16}D_{+y}{}^2D_{-y}{}^2\bar{\bar{U}}_{ij}^{n+1}. \quad (73c)$$

In (73a) $\bar{U}_{ij}^n$ on the right-hand side comes from the previous computational sequence where the numerical solution is advanced from time level $n - 1$ to time level n. A value of $\epsilon_x = \epsilon_y = 0.5$ was employed throughout.

Filtering techniques

Instead of a higher order dissipation term to control aliasing, one can also use a very selective low-pass filter.

Three filtering methods were tested. The first is the Wallington (1962) filter which is an extension of the Shuman (1955, 1957) filter. It consists in the periodic successive application of the following two-point operators:

$$\bar{U}_i = 4.28U_i - 2.16(U_{i+1} + U_{i-1}) + 0.52(U_{i+2} + U_{i-2}),$$
$$\bar{U}_i = 0.375\bar{U}_i + 0.25(\bar{U}_{i+1} + \bar{U}_{i-1}) + 0.0625(\bar{U}_{i+2} + \bar{U}_{i-2}). \quad (74)$$

This filter completely eliminates waves with wavelengths less than $3\Delta x$.

A second filtering method tested consists of a periodic application of a high-order (16th) Shapiro (1970) filter. The filter is of the form

$$\bar{\mathbf{U}} = \{1 - (F_x{}^2)^8\}\{1 - (F_y{}^2)^8\}\mathbf{U},$$
$$F_x{}^2(U_{ij}) = (U_{i+1,j} - 2U_{ij} + U_{i-1,j})/4 \quad (75)$$

and has a response

$$F_x{}^2[\exp ikx] = -\sin^2(k\Delta x)\exp ik\Delta x. \quad (76)$$

This filter eliminates waves shorter than $4\Delta x$ and even after hundreds of applications has only a negligible damping effect on waves longer than $4\Delta x$. However, its main shortcoming is that because of the number of grid points employed, it is effective only in the center of the domain.

A third, very selective, low-pass filter due to Paul Long (Mahrer and Pielke, 1978) was also tested. This filter is

$$(1 - \delta)\bar{U}_{i+1} + 2(1 - \delta)\bar{U}_i + (1 - \delta)\bar{U}_{i-1} = U_{i+1} + 2U_i + U_{i-1}, \quad (77)$$

where $\bar{U}_i$ is the filtered field.

This filter completely eliminates the $2\Delta x$ waves with each application, while its smoothing effect on other wavelengths is a function of δ. Its response function is such that

$$\bar{\mathbf{U}} = \mathbf{U}\frac{1}{1 + \delta\tan^2(\lambda\Delta x/2)}, \quad (78)$$

where $\lambda = 2\pi/J$ is the wavenumber and J the wavelength. For $\delta \leq 0.1$ there is little damping of waves larger than $6\Delta x$.

7. Numerical results

a. The test problem

We decided to use two different initial conditions employed by Grammeltvedt (1969), both describing a westerly jet flow with north–south perturbations

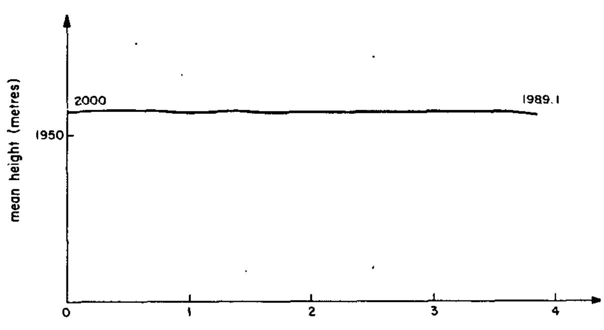

FIG. 1. Time evolution of the mean height for initial condition (I).

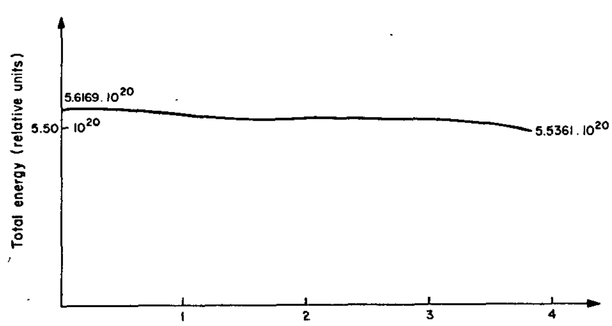

FIG. 2. As in Fig. 1 except for total energy.

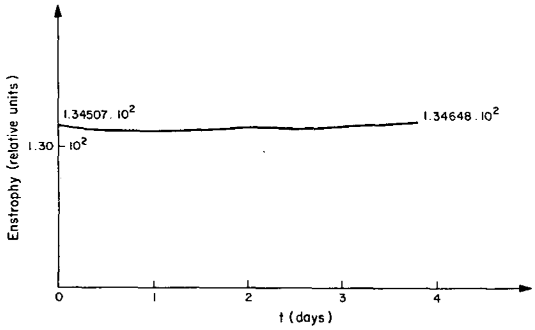

FIG. 3. As in Fig. 1 except for enstrophy.

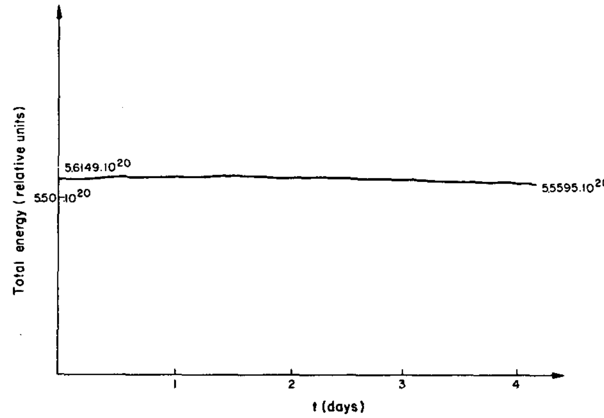

FIG. 5. As in Fig. 4 except for total energy.

of different wavelengths and amplitudes along the zonal axis of the jet. Initial condition I was also employed by Gustafsson (1971a) and Navon (1978) amongst others, and this provides a basis for comparison.

Initial condition II was employed by Gerrity *et al.* (1972) with a fourth-order finite-difference scheme and by Cullen (1975, 1977) for linear spline Galerkin finite-element schemes.

The initial height fields are

$$\text{(I)}\quad h(x,y) = H_0 + H_1 \tanh\frac{(9(D/2-y))}{2D} + H_2\,\text{sech}^2\frac{(9(D/2-y))}{D}\sin\left(\frac{2\pi x}{L}\right) \tag{79}$$

$$\text{(II)}\quad h(x,y) = H_0 + H_1 \tanh\frac{(9(D/2-y))}{2D} + H_2\,\text{sech}^2\frac{(9(D/2-y))}{D}\left[0.7\sin\frac{(2\pi x)}{L} + 0.6\sin\frac{(6\pi x)}{L}\right]. \tag{80}$$

The initial velocity field components u and v are derived from the initial height field using the geostrophic approximation

$$u = \left(\frac{-g}{f}\right)\frac{\partial h}{\partial y}, \quad v = \left(\frac{g}{f}\right)\frac{\partial h}{\partial x}. \tag{81}$$

Initial condition I initially has energy only in wavenumber 1 in the x direction, whereas initial condition II initially contains energy in wavenumbers 1 and 3 in the x direction.

The dimensions of the rectangular domain were

$$L = 4400 \text{ km}, \quad D = 6000 \text{ km} \tag{82}$$

and the following constant values were adopted

$$H_0 = 2000 \text{ m}, \quad H_1 = +220 \text{ m}, \quad H_2 = 133 \text{ m}.$$
$$g = 10 \text{ m s}^{-2}, \quad \hat{f} = 10^{-4} \text{ s}^{-1},$$
$$\beta = 1.5\ 10^{-11} \text{ m}^{-1}\text{ s}^{-1}, \tag{83}$$

where

$$f = \hat{f} + \beta(y - D/2). \tag{84}$$

The fourth-order compact scheme was run with the spatial resolution

$$\Delta x = \Delta y = 200 \text{ km} \tag{85}$$

and time steps of $\Delta t = 900$ s or $\Delta t = 600$ s.

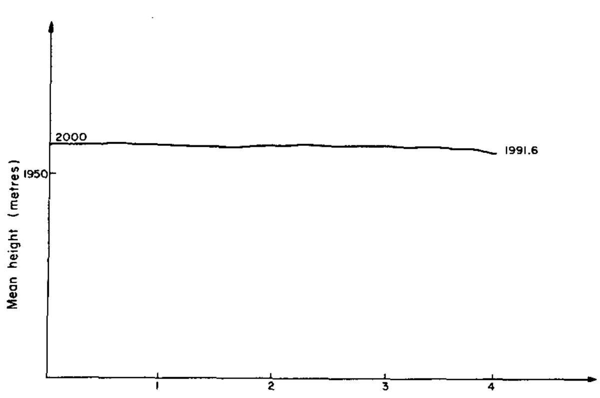

FIG. 4. Time evolution of the mean height for initial condition (II).

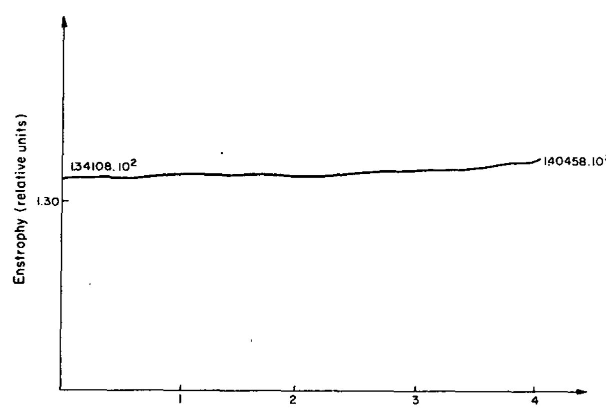

FIG. 6. As in Fig. 4 except for enstrophy.

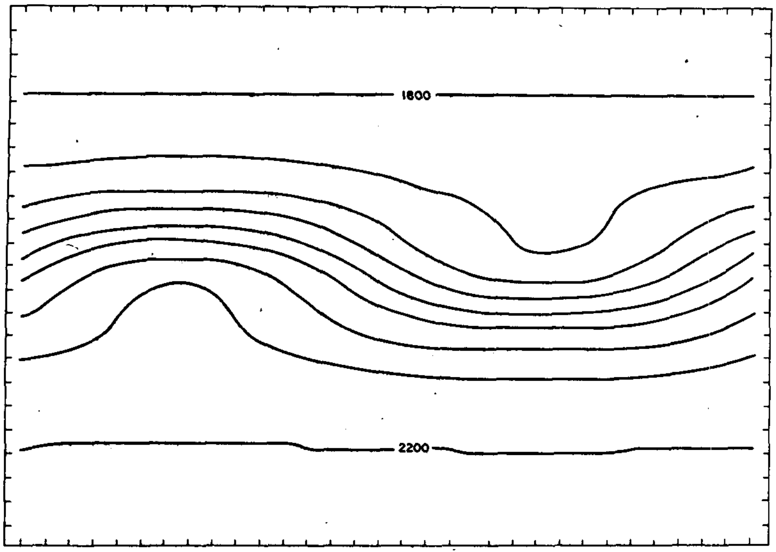

FIG. 7. The initial distribution of the height field depicted by isopleths drawn at 50 m intervals [initial condition (I)]. The channel walls are shown as solid horizontal lines, the cyclic boundaries as vertical lines. The domain is covered by 31 grid points in horizontal (east-west) and 23 points in the vertical (north-south).

b. *Discussion of the numerical results of the simulation*

It is well known (Haltiner and Williams, 1973; Gerrity *et al.*, 1972) that if linear computational stability is to be maintained, the use of the fourth-order space approximation requires a time-step smaller by about 30%. Using standard stability criteria for the linearized two-dimensional shallow-water equations, one gets

$$\frac{\Delta t\sqrt{2}}{\Delta x}\,[\bar{u} + (gH)^{1/2}] \leq 1, \tag{86}$$

where $\bar{u}$ and H are mean velocity and height, respectively. Using $\bar{u} = 40$ m s^{-1}, $H = 2000$ m, $g = 10$ m s^{-2}, $\Delta x = 200$ km, we obtain

$$(\Delta t)_2 = 750 \text{ s}, \tag{87}$$

where the subscript 2 stands for a second-order

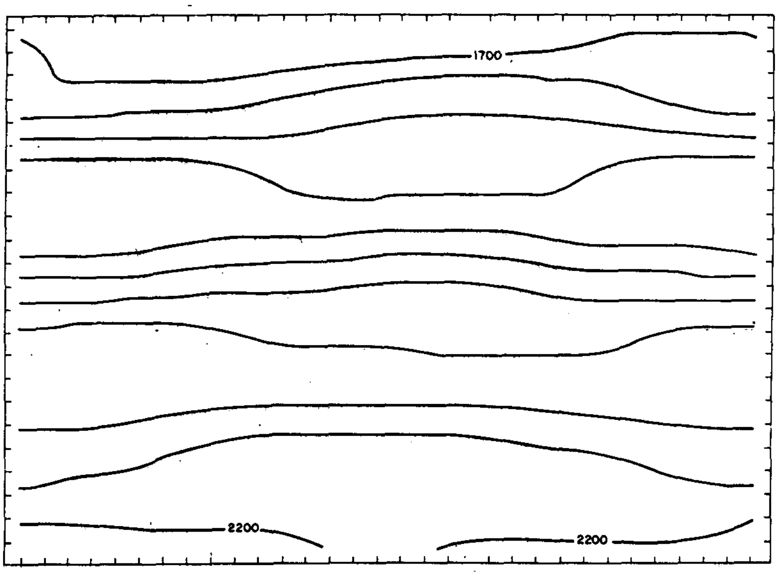

FIG. 8. The 24 h forecast of height field by the compact fourth-order scheme isoplethed at intervals of 50 m [initial condition (I)].

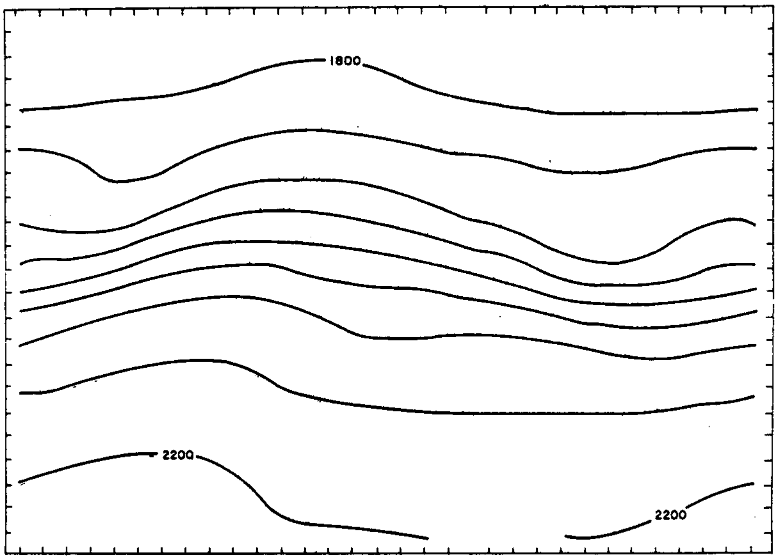

FIG. 9. The 48 h forecast of height field by the compact fourth-order scheme at intervals of 50 m [initial condition (I)].

space approximation. Following Gerrity *et al.* (1972),

$$(\Delta t)_4 = 0.735(\Delta t)_2 \approx 550 \text{ s}. \quad (88)$$

In the numerical experiments carried out with the implicit compact fourth-order ADI algorithm [Eqs. (43a)–(43c)] a time step was used of either 900 or 600 s and a spatial resolution of $\Delta x = \Delta y = 200$ km. It was found experimentally that the best results in respect of accuracy were obtained by a periodical application of the Wallington filter.

Acceptable results were also obtained by employing fourth-order dissipation. The Shapiro filter proved to be inefficient in removing short-wave noise near the boundaries.

The Long filter (using $\delta = 0.1$) was also inefficient near the boundaries.

In all the numerical experiments to be discussed, the Wallington filter was applied periodically (every three time steps). Application of the Oliger (1974), $O(h^3)$ boundary conditions to the implicit compact fourth-order ADI algorithm for the shallow-water

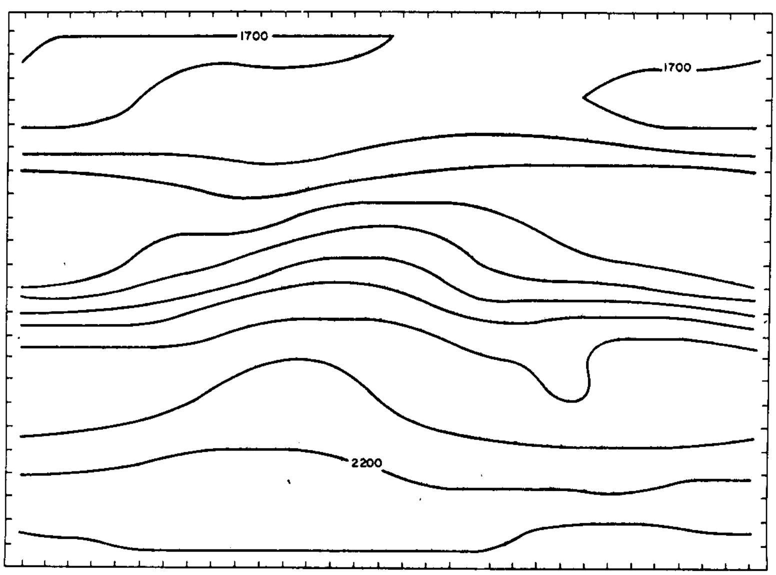

FIG. 10. The 72 h forecast of height field by compact fourth-order scheme isoplethed at intervals of 50 m [initial condition (I)].

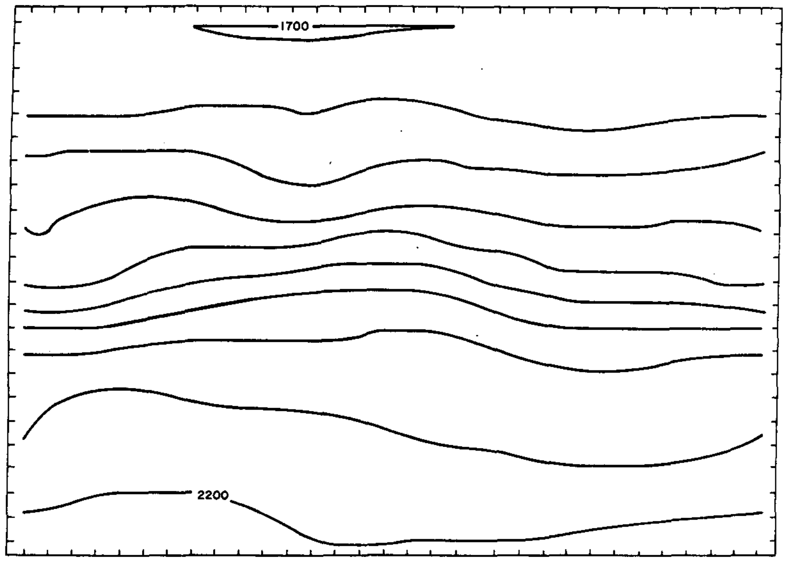

FIG. 11. The 72 h forecast of height field by compact fourth-order scheme isoplethed at intervals of 50 m [initial condition (I)].

equations [Eqs. (43a)–(43c)] gave rise to instability near the boundaries after two days of simulation. On the other hand the Adam (1977) $O(h^3)$ boundary conditions [Eqs. (65)–(66)] maintained stability during the entire period of simulation (four days) and these boundary conditions were then adopted for all the numerical experiments.

The adequacy of the Adam (1977) boundary conditions was also established experimentally by Peyret (1978).

Energy and enstrophy conservation, allowing correct nonlinear transfers among explicit scales, have emerged as fundamental concepts in finite-differencing for the shallow-water equations (Arakawa, 1966; Sadourny, 1975; Fairweather and Navon, 1977).

The time evolution of the three well-known invariants for the shallow-water equations: mean height, total energy and enstrophy, was calculated at each time step of the numerical integration. An almost perfect conservation of total energy, enstrophy and mean height (proportional to mass) was obtained, as is evident from Figs. 1–3, which show total energy, enstrophy and mean height, respectively, as functions of time, for initial condition I.

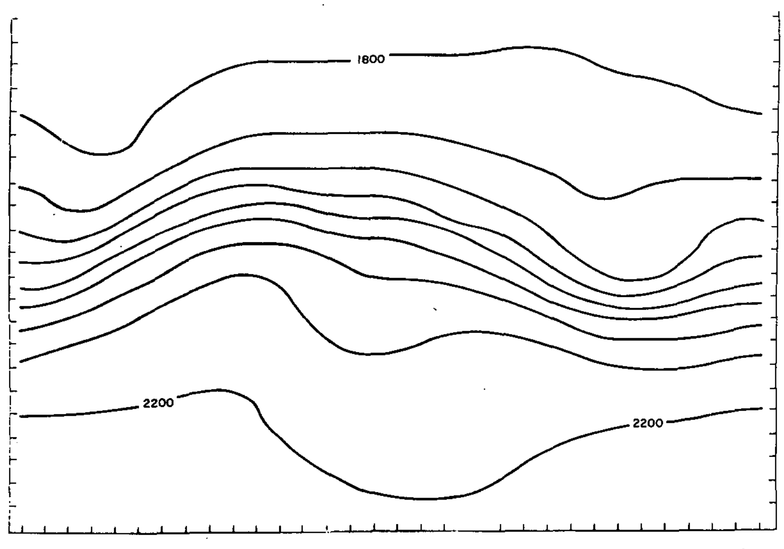

FIG. 12. The 48 h forecast of height field by Gustafsson's QNEX1 (M = 6) nonlinear ADI scheme isoplethed at intervals of 50 m [initial condition (I)].

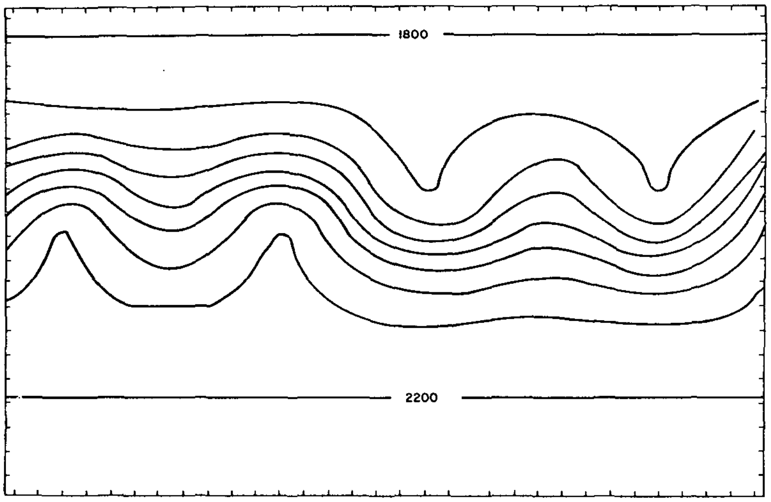

FIG. 13. The initial distribution of the height field by isopleths drawn at 50 m intervals [initial condition (II)].

Figs. 4–6 show total energy, enstrophy and mean height respectively as functions of time, for initial condition II.

Adam (1977) O(h^3) boundary conditions were applied to equations 43a–43b, while Elvius-Sundström (1973) boundary conditions [Eq. (67)] were used for the variable U^{n+1} in (43c). The initial height field for initial condition I is shown in Fig. 7, the contours having been drawn at 50 m intervals. The height field after 1, 2, 3 and 4 days of simulation with a time step of Δt = 900 s is shown in Figs. 8–11, respectively.

The three-level time-centered scheme [Eqs. (27) and (33)] was also experimented with. Owing to the appearance of a computational mode as well as a typical odd-even time-step separation, we employed a Robert time filter; however, the results obtained with this scheme are still unsatisfactory.

This is due to the fact that the uncoupling between time levels makes the three-time level scheme more sensitive to the linearization procedure (Steger, 1978). We found its accuracy to be somewhat lower and it also necessitated higher dissipative terms. It is anticipated, however, that when this scheme is used without linearization (as done by Gustafsson, 1971a) it could yield more accurate results.

Visual comparison shows agreement between the results of Gustafsson (1971a) after two days (Fig. 12) and our results on using Eqs. (43a)–(43c), i.e., the compact fourth-order ADI algorithm, in both the positions and amplitudes of the main troughs and ridges.

Another set of experiments was conducted this time using the initial height field for initial condition II (Fig. 13).

We then compared our results with those of Gerrity *et al.* (1972) after two days, also with the results obtained by Cullen (1975).

Table 1 gives the extreme amplitude values of the height field in each trough and ridge at the midpoint of the channel after two days, while Table 2 gives the corresponding positions as a fraction of the distance along the channel of the corresponding extreme values of troughs and ridges.

Figs. 14 and 15 show the height field after one and two days of integration, using initial condition II.

The results show that the compact fourth-order ADI results as far as the amplitudes and detailed positions of the troughs and ridges are concerned match the Gerrity results with a Δx = 100 km spatial resolution. It is also evident that the compact fourth-order scheme has translated the systems of troughs and ridges faster than the corresponding scheme of Gerrity *et al.* (1972) with Δx = 200 km.

c. Accuracy tests

In order to provide a basis for accuracy comparison in the absence of an analytic solution to the full nonlinear shallow-water equations, a refer-

TABLE 1. Amplitudes (after two days) in decameters.

Finite difference (Δx = 200 km) (Gerrity *et al.*, 1972)	209	202	209	190	200	189
Finite difference (Δx = 100 km) (Gerrity *et al.*, 1972)	208	204	206	192	197	189
Compact fourth-order ADI method (Δx = 200 km)	208	204	207	193	198	189
Finite-element (400 km) using the two-stage Galerkin method (Cullen, 1975)	210	204	205	193	197	186

TABLE 2. Phases after two days.

Finite difference (Δx = 200 km) Gerrity *et al.* (1972)	0.198	0.341	0.494	0.703	0.852	1.000
Finite difference (Δx = 100 km) Gerrity *et al.* (1972)	0.235	0.399	0.499	0.730	0.857	1.000
Compact fourth-order ADI method (Δx = 200 km)	0.225	0.373	0.497	0.716	0.854	1.000
Finite element (Δx = 400 km) using the two-stage Galerkin method (Cullen 1975)	0.225	0.419	0.475	0.668	0.775	0.968

ence solution was obtained by integrating the system by a fourth-order compact ADI method with double resolution in both horizontal space dimensions.

Then, following Gustafssòn (1971), we define a Hilbert space H, by considering all vector functions satisfying

$$U_{ij} = U_{i+N_x,j}$$

and

$$V_{i,0} = V_{i,N_y} = 0. \quad (89)$$

The inner product of two vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and the norm are then defined by

$$(\alpha,\beta) = \Delta x \Delta y \sum_{i=1}^{N_x} \{ \sum_{j=1}^{N_{y-1}} \alpha_{ij}^T \beta_{ij} + \tfrac{1}{2}(\alpha_{i,0}^T \beta_{i,0} + \alpha_{i,N_y}^T \beta_{i,N_y})\}$$

$$\|\boldsymbol{\alpha}\|^2 = (\alpha,\alpha). \quad (90)$$

The relative error between the approximate and the true solutions of the scheme, represented by U_{sw4s} and U_{sw4D}, respectively, is

$$\text{relative error} = \frac{\|\epsilon_{sw4}\|}{\|U_{sw4D}\|} \quad \begin{array}{l} t = 1 \text{ day} \\ \Delta t = 900 \text{ s}, \end{array} \quad (91)$$

where

$$\epsilon_{sw4} = U_{sw4s} - U_{sw4D} . \quad (92)$$

The error is summarized in Table 3.

The results are somewhat disappointing, as theoretically the truncation error is $\Delta x^4/180 + \ldots$ (Orszag and Israeli, 1974), but one has to take into account both the influence of the boundary conditions and the increased aliasing characteristic of the fourth-order finite-difference schemes.

The relative error was also calculated by comparing the fourth-order compact implicit scheme results with those obtained with Gustafsson's (1971) most accurate nonlinear scheme QN3, which was assumed to be the reference true solution. The results are given in Table 4.

8. Summary

It was shown that an implicit compact fourth-order ADI solution to the shallow-water equations in conservation-law form is feasible. The trade-off between efficiency and programming effort inherent in the use of Padé finite differences (Boyd, 1978) proved to be not very significant.

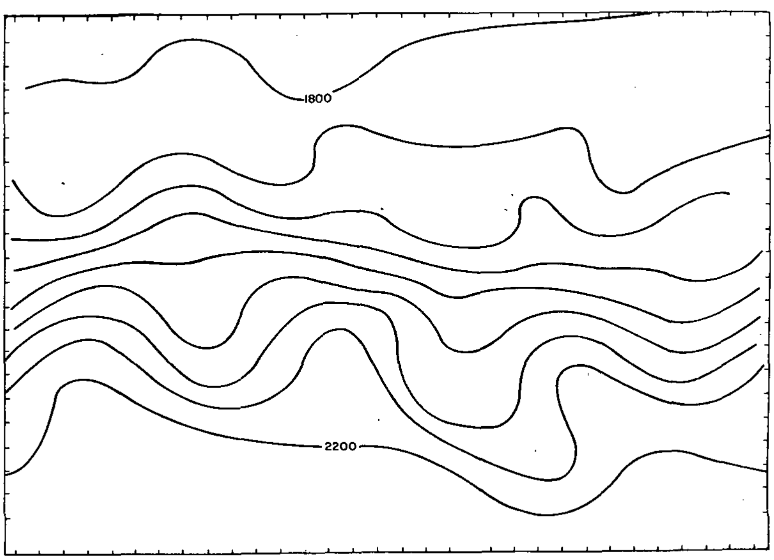

FIG. 14. The 24 h forecast of height field by compact fourth-order scheme isoplethed at intervals of 50 m [initial condition (II)].

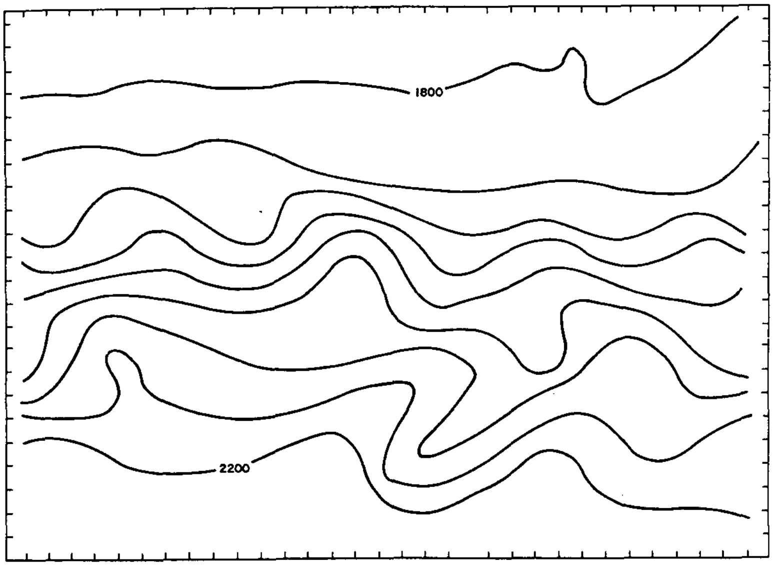

FIG. 15. The 48 h forecast of height field by compact fourth-order scheme, isoplethed at intervals of 50 m [initial condition (II)].

Quadratic invariants of the shallow-water equations were almost perfectly conserved for the period of integration. The application of numerically well-posed boundary conditions and the control of increased aliasing inherent in fourth-order schemes proved to be the main computational issues.

Extensive comparisons with the results of other investigators showed the results to be comparable to those obtained by a second-order scheme with a double resolution. Some implications from this study are the following:

1) The fourth-order compact implicit method should be tested in a numerical operational weather forecasting model and compared with the usual five-point fourth-order method (Kalnay-Rivas, 1977).

2) The operator-compact implicit method suggested by Ciment *et al.* (1978) should be adapted to our problem and tested for efficiency and accuracy.

In general, the compact fourth-order implicit technique appears to offer a fruitful alternative to classical fourth-order methods.

TABLE 3. $\|\epsilon_{sw4}\|/\|U_{sw4D}\|$.

Method	Δt = 900 s
Compact fourth-order algorithm Eqs. (43a)–(43c) no smoothing	8.4 10^{-4}
Same, smoothing with Wallington filter at every time-step	1.8 10^{-3}
Same smoothing with Wallington filter periodically (i.e., every three time-steps)	6.1 10^{-4}

It also links the finite-difference method with the spline method (see Appendix A) and with the finite-element method (Cullen, 1977) and in a sense generalizes the latter.

The fourth-order compact algorithm offers a computationally efficient alternative to the finite element approach because the linear systems to be solved are tridiagonal, as opposed to the more complex coefficient matrices usually generated by the finite-element method. Morton (1977) points out that for regular linear elements, the coefficients of the mass-matrix in the finite-element method correspond to an operator $(1 + \delta x^2/6)$ acting on $\dot{U}_j$. The approximation

$$\left(1 + \frac{\delta x^2}{6}\right)^{-1} = \left(1 - \frac{\delta x^2}{6}\right) + O(h^4)$$

characteristic of fourth-order compact schemes is equivalent to a "half-lumped" mass matrix for the finite-element method. Recently, Navon (1979) used a generalized mixed-mass (GMM) finite-element scheme for solving the shallow-water equations. The GMM mass scheme uses a convex combination of lumped and consistent mass matrices for the finite-element method.

TABLE 4. $\|\epsilon_{SW4}\|/\|U_{QN3}\|$ t = 1 day. $\epsilon_{SW4} = U_{SW4S} - U_{QN3}$.

Method	Δt = 900 s
Compact fourth-order algorithm (Wallington smoothing at every time step)	8.1 10^{-4}
Same algorithm with Wallington smoothing every three time steps	6.7 10^{-4}

The numerical results obtained with the GMM mass scheme are very similar to those obtained with the fourth-order compact implicit algorithm and tend to confirm Morton's (1977) remark.

Acknowledgment. The authors are thankful to the reviewer for his helpful comments.

APPENDIX A

Generalized Compact Implicit Difference Approximations

Swartz and Wendroff (1974) pointed out that the compact implicit-difference approximation to $\partial/\partial x$ can be generalized to implicit schemes with truncation $O(h^{4l})$ and bandwidth $2l + 1$. If we consider S as a difference operator and apply it to the function $l^{2\pi i\omega x}$, evaluated at mesh points for integral ω, then

$$(Se^{2\pi i\omega x})_k = \sum_{j=0}^{J_1-1} S_{kj}e^{2\pi i\omega j\Delta x} = \Delta x \sum_{j=0}^{J_1-1} S_{j-k}e^{2\pi i\omega j\Delta x}$$

$$= \Delta x e^{2\pi i\omega k\Delta x} \sum_{j=0}^{J_1-1} S_j e^{ij\theta} \quad \text{with} \quad \theta = 2\pi\omega\Delta x. \quad \text{(A1)}$$

Defining the symbol of the difference operator S to be

$$a(\theta) = \sum_{j=0}^{J_1-1} S_j e^{ij\theta}, \quad \text{(A2)}$$

the symbols of the high-order compact implicit centered-difference operators are given by

$$b_l(\theta)/a_l(\theta) = i \sin\theta g_l[\sin^2(\theta/2)], \quad \text{(A3)}$$

where g_l is the following truncation (Wall, 1948, pp. 345, 380) of a continued-fraction approximation of arc $\sin(\tau)/[\tau(1 - \tau^2)^{1/2}]$

$$g_l(\tau) = \frac{1}{1} - \frac{1.2.\tau^2}{3} - \frac{1.2\tau^2}{5} \cdots - \frac{(2j-1)2j\tau^2}{(4j-1)}$$

$$- \frac{(2j-1)2j\tau^2}{(4j+1)} \cdots - \frac{(2l-1)2l\tau^2}{4l-1}. \quad \text{(A4)}$$

Each scheme can be reconstituted from its symbol by replacing $i \sin\theta$ by D_{0x} and $\sin^2(\theta/2)$ in g_l by $-\Delta x^2 D_{+x}D_{-x}/4$. The scheme with $l = 1$ coincides with the $O(\Delta x^4)$ fourth-order compact first derivative and also with the $O(\Delta x^4)$ piecewise linear spline scheme, with

$$A_1 = (1,4,1)/6 \quad \text{and} \quad B_1 = (-1, 0, 1)/2\Delta x \quad \text{(A5)}$$

(see also Swartz and Wendroff, 1974).

The phase error per period of the compact implicit schemes is

$$\epsilon_{il}(N,l) = 2\pi\{1 - b_l(\theta)/[i\theta a_l(\theta)]\} \quad \text{(A6)}$$

$$\theta = 2\pi\omega\Delta x = 2\pi/N, \quad \text{(A7)}$$

where N is the number of space intervals per wavelength.

APPENDIX B

Linearized Stability Analysis

A system of linearized-perturbation shallow-water equations (Kurihara 1965, Navon 1978) is given by

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} + \bar{u}\frac{\partial u}{\partial x} - fv + \frac{\partial \Phi}{\partial x} \\ &\frac{\partial v}{\partial t} + \bar{u}\frac{\partial v}{\partial x} + fu \\ &\frac{\partial \Phi}{\partial t} + \bar{u}\frac{\partial \Phi}{\partial x} - f\bar{u}v + \bar{\Phi}\frac{\partial u}{\partial x} \end{aligned}\right\} = 0, \quad \text{(B1)}$$

where $\bar{u}$ is a basic constant zonal wind and $\bar{\Phi}$ is the mean geopotential. This system can be written as

$$\frac{\partial \mathbf{w}}{\partial t} + \mathbf{A}\frac{\partial \mathbf{w}}{\partial x} + \tilde{\mathbf{C}}\mathbf{w} = 0, \quad \text{(B2)}$$

where $\mathbf{w} = (u,v,\Phi)^T$,

$$\mathbf{A} = \begin{bmatrix} \bar{u} & 0 & 1 \\ 0 & \bar{u} & 0 \\ \bar{\Phi} & 0 & \bar{u} \end{bmatrix}, \quad \tilde{\mathbf{C}} = \begin{bmatrix} 0 & -f & 0 \\ f & 0 & 0 \\ 0 & -f\bar{u} & 0 \end{bmatrix}. \quad \text{(B3)}$$

As $\mathbf{A}$ is a constant matrix, the system can also be written as

$$\frac{\partial \mathbf{w}}{\partial t} + \partial\frac{(\mathbf{A}\mathbf{w})}{\partial x} + \tilde{\mathbf{C}}\mathbf{w} = 0. \quad \text{(B4)}$$

Applying to it our algorithm, one obtains

$$\left[\mathbf{I} + \frac{\Delta t}{2}\left(\frac{\partial}{\partial x}(\mathbf{A}\cdot) + \tilde{\mathbf{C}}\right)\right]\mathbf{w}^{n+1}$$

$$= \left[\mathbf{I} - \frac{\Delta t}{2}\left(\frac{\partial}{\partial x}(\mathbf{A}\cdot) + \tilde{\mathbf{C}}\right)\right]\mathbf{w}^{n}. \quad \text{(B5)}$$

Introducing compact fourth-order differencing and multiplying throughout by Q_x, we obtain

$$\left[Q_x + \frac{\Delta t}{2}(D_{0x}(\mathbf{A}\cdot) + Q_x\tilde{\mathbf{C}})\right]\mathbf{w}^{n+1}$$

$$= \left[Q_x - \frac{\Delta t}{2}(D_{0x}(\mathbf{A}\cdot) + Q_x\tilde{\mathbf{C}})\right]\mathbf{w}^{n}. \quad \text{(B6)}$$

Substituting Fourier terms of the form $u_j^l = u^l \times \exp(i\lambda j\Delta x)$, $l = n, n+1$, where λ is the wavenumber and $x_j = j\Delta x$ and similar expressions are written for v_j^l and Φ_j^l, we obtain

$$(4 + 2\cos\lambda\Delta x)\mathbf{w}^{n+1} + \frac{\Delta t}{2} 6i \frac{\sin\lambda\Delta x}{\Delta x} A\mathbf{w}^{n+1} + \frac{\Delta t}{2}(4 + 2\cos\lambda\Delta x)\tilde{\mathbf{C}}\mathbf{w}^{n+1}$$
$$= (4 + 2\cos\lambda\Delta x)\mathbf{w}^{n} - \frac{\Delta t}{2} 6i \frac{\sin\lambda\Delta x}{\Delta x} \mathbf{A}\mathbf{w}^{n} - \frac{\Delta t}{2}(4 + 2\cos\lambda\Delta x)\tilde{\mathbf{C}}\mathbf{w}^{n}. \quad \text{(B7)}$$

Dividing throughout by the factor $4 + 2\cos\lambda\Delta x$, we obtain

$$\left[\mathbf{I} + \frac{\Delta t}{2\Delta x}\frac{3i\sin\lambda\Delta x}{2 + \cos\lambda\Delta x}\mathbf{A} + \frac{\Delta t}{2}\tilde{\mathbf{C}}\right]\mathbf{w}^{n+1} = \left[\mathbf{I} - \frac{\Delta t}{2\Delta x}\frac{3i\sin\lambda\Delta x}{2 + \cos\lambda\Delta x}\mathbf{A} - \frac{\Delta t}{2}\tilde{\mathbf{C}}\right]\mathbf{w}^{n}, \quad \text{(B8)}$$

so that the amplification matrix **G** is given by

$$\mathbf{G} = (\mathbf{I} + \mathbf{T})^{-1}(\mathbf{I} - \mathbf{T}), \quad \text{(B9)}$$

where

$$\mathbf{T} = \gamma\alpha\mathbf{A} + \frac{\Delta t}{2}\tilde{\mathbf{C}}, \quad \text{(B10)}$$

$$\gamma = \frac{\Delta t}{\Delta x}, \quad \alpha = \frac{3i\sin\lambda\Delta x}{4 + 2\cos\lambda\Delta x}. \quad \text{(B11)}$$

More explicitly,

$$\mathbf{T} = \begin{bmatrix} \gamma\alpha\bar{u} & \frac{-\Delta t}{2}f & \gamma\alpha \\ \frac{\Delta t}{2}f & \gamma\alpha\bar{u} & 0 \\ \gamma\alpha\bar{\Phi} & \frac{-\Delta t}{2}f\bar{u} & \gamma\alpha\bar{u} \end{bmatrix}. \quad \text{(B12)}$$

Observing that $(\mathbf{I} + \mathbf{T})^{-1}$ exists, if κ_i is an eigenvalue of **T** the corresponding eigenvalue g_i of **G** is (Henrici, 1974)

$$g_i = \frac{1 - \kappa_i}{1 + \kappa_i}. \quad \text{(B13)}$$

Therefore, we have to calculate the eigenvalues κ_i of **T** given by

$$(\gamma\alpha\bar{u} - \kappa_i)^3 - \frac{\Delta t^2}{4}f^2\bar{u}\gamma\alpha - (\gamma\alpha)^2\bar{\Phi}(\gamma\alpha\bar{u} - \kappa_i) + \frac{\Delta t^2}{4}f^2\bar{u}\gamma\alpha = (\gamma\alpha\bar{u} - \kappa_i)^3 - (\gamma\alpha)^2\bar{\Phi}(\gamma\alpha\bar{u} - \kappa_i) = 0. \quad \text{(B14)}$$

The cubic equation (92) has three distinct roots, all complex. One of them is

$$\kappa_1 = \gamma\alpha\bar{u} = \frac{\Delta t}{\Delta x}\frac{3i\sin\lambda\Delta x}{4 + 2\cos\lambda\Delta x}\bar{u} \quad \text{(B15)}$$

and the other two are given by the solution of the quadratic equation

$$(\gamma\alpha\bar{u} - \kappa_i)^2 - (\gamma\alpha)^2\bar{\Phi} = 0,$$
$$\kappa_i^2 - 2\gamma\alpha\bar{u}\kappa_i + (\gamma\alpha)^2\bar{u}^2 - (\gamma\alpha)^2\bar{\Phi} = 0, \quad \text{(B16)}$$
$$\kappa_{1,2} = \tfrac{1}{2}(2\gamma\alpha\bar{u} \pm (4(\gamma\alpha)^2\bar{u}^2 - 4\gamma^2\alpha^2(\bar{u}^2 - \bar{\Phi}))^{1/2} = \gamma\alpha\bar{u} \pm \gamma\alpha\bar{\Phi}^{1/2}. \quad \text{(B17)}$$

We can now make use of a well-known theorem stating that if $\text{Re}\kappa_i \leq 0$ and κ_i are distinct, then $|g_i| = 1 - \kappa_i/1 + \kappa_i$ are inside the unit circle, or on it. In our case $\text{Re}\kappa_i = 0$, $i = 1, 2, 3$ and it follows that all the eigenvalues g_i, $i = 1, 2, 3$ of **G** are on the unit circle. Hence we have unconditional stability for the linear case.

APPENDIX C

Factorization Technique

We will show here that the addition of the perturbation terms (I)–(IV) to Eq. (22) is equivalent to the factored equation (26). To this end we expand (26) by multiplying the terms within the brackets on both sides of the equation and we obtain

$$\left[\mathbf{I} + \frac{\Delta t}{2}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\Delta t}{2}\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) - \frac{\Delta t}{2}\mathbf{C}^{(1)} - \frac{\Delta t}{2}\mathbf{C}^{(2)} + \frac{\Delta t^2}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \frac{\Delta t^2}{4}\mathbf{C}^{(1)}\mathbf{C}^{(2)} - \frac{\Delta t^2}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\mathbf{C}^{(2)} - \frac{\Delta t^2}{4}\frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\mathbf{C}^{(1)}\right]\mathbf{U}^{n+1}$$
$$= \left[\mathbf{I} + \frac{\Delta t}{2}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\Delta t}{2}\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \frac{\Delta t}{2}\mathbf{C}^{(1)} + \frac{\Delta t}{2}\mathbf{C}^{(2)} + \frac{\Delta t^2}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \frac{\Delta t^2}{4}\mathbf{C}^{(1)}\mathbf{C}^{(2)} + \frac{\Delta t^2}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\mathbf{C}^{(2)} + \frac{\Delta t^2}{4}\frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\mathbf{C}^{(1)}\right]\mathbf{U}^{(n)} - \Delta t\left(\frac{\partial\mathbf{P}}{\partial x} + \frac{\partial\mathbf{Q}}{\partial y}\right)^n. \quad \text{(C1)}$$

Using Eq. (25) we can write (C1) as

$$\left[\mathbf{I} + \frac{\Delta t}{2}\left(\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\partial}{\partial y}(\mathbf{B}^n\cdot) - \mathbf{C}\right)\right]\mathbf{U}^{n+1}$$

$$+ \underbrace{\frac{\Delta t^3}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\frac{(\mathbf{U}^{n+1}-\mathbf{U}^n)}{\Delta t}}_{\text{(I)}} + \underbrace{\frac{\Delta t^3}{4}\mathbf{C}^{(1)}\mathbf{C}^{(2)}\frac{(\mathbf{U}^{n+1}-\mathbf{U}^n)}{\Delta t}}_{\text{(II)}}$$

$$- \underbrace{\frac{\Delta t^3}{4}\frac{\partial}{\partial x}(\mathbf{A}^n\cdot)\mathbf{C}^{(2)}\frac{(\mathbf{U}^{n+1}+\mathbf{U}^n)}{\Delta t}}_{\text{(III)}} - \underbrace{\frac{\Delta t^3}{4}\frac{\partial}{\partial y}(\mathbf{B}^n\cdot)\mathbf{C}^{(1)}\frac{(\mathbf{U}^{n+1}+\mathbf{U}^n)}{\Delta t}}_{\text{(IV)}}$$

$$= \left[\mathbf{I} + \frac{\Delta t}{2}\left(\frac{\partial}{\partial x}(\mathbf{A}^n\cdot) + \frac{\partial}{\partial y}(\mathbf{B}^n\cdot) + \mathbf{C}\right)\right]\mathbf{U}^n - \Delta t\left(\frac{\partial \mathbf{P}}{\partial x} + \frac{\partial \mathbf{Q}}{\partial y}\right)^n. \quad \text{(C2)}$$

Eq. (C2) is exactly identical to (22) with the perturbation terms (I)–(IV) added on its left-hand side.

REFERENCES

Adam, Y., 1975: A Hermitian finite-difference method for the solution of parabolic equations. *Comput. Math. Appl.*, **1**, 393–406.

——, 1977: Highly accurate compact implicit methods and boundary conditions. *J. Comput. Phys.*, **24**, 10–22.

Arakawa, A., 1966: Computational design for long term numerical integration of the equations of fluid motion—Two dimensional incompressible flow. Part 1. *J. Comput. Phys.*, **1**, 119–143.

Beam, R. M., and R. F. Warming, 1976: An implicit finite difference algorithm for hyperbolic systems in conservation-law form. *J. Comput. Phys.*, **22**, 87–116.

Boyd, J. P., 1978: Spectral and pseudospectral methods for eigenvalue and nonseparable boundary value problems. *Mon. Wea. Rev.*, **106**, 1192–1203.

Briley, W. R., and H. McDonald, 1977: Solution of the multidimensional compressible Navier-Stokes equations by a generalized implicit method. *J. Comput. Phys.*, **19**, 372–397.

Chu, C., and A. Sereny, 1974: Boundary conditions in finite-difference fluid-dynamic codes. *J. Comput. Phys.*, **15**, 476–491.

Ciment, M., and S. H. Leventhal, 1975: Higher-order compact implicit schemes for the wave equation. *Math. Comput.*, **29**, 985–994.

——, and ——, 1978: A note on the operator compact implicit method for the wave equation. *Math. Comput.*, **32**, 143–147.

——, —— and B. C. Weinberg, 1978: The operator compact implicit method for parabolic equations. *J. Comput. Phys.*, **28**, 135–166.

Collatz, L., 1960: *The Numerical Treatment of Differential Equations*. Springer-Verlag, Berlin, 568 pp.

Cullen, M. J. P., 1975: Application of the finite element method to numerical weather prediction. Ph.D. thesis, University of Reading, U.K., 90 pp.

——, 1977: The application of finite element methods to the primitive equations of fluid motion. *Finite Elements in Water Resources*. Gray, Pinder and Brebbia, Eds., Pentech Press, London (see pp. 4.231–4.246).

Douglas, J., Jr., and J. E. Gunn, 1964: A general formulation of alternating-direction methods. *Numer. Math.*, **6**, 428–453.

Elvius, T., and A. Sundström, 1973: Computationally efficient schemes and boundary conditions for a fine mesh barotropic model based on the shallow-water equations. *Tellus.*, **25**, 132–156.

Fairweather, G., and I. M. Navon, 1977: A linear alternating direction implicit (ADI) method for solving the shallow-water equations. Submitted to the *Journal of Computational Physics*.

Fox, L., and E. T. Goodwin, 1949: Some new methods for the numerical integration of ordinary differential equations. *Proc. Cambridge Phil. Soc. Math. Phys. Sci.*, **45**, 373–388.

Gary, J., 1978: On boundary conditions for hyperbolic difference schemes. *J. Comput. Phys.*, **26**, 339–351.

Gerrity, J. P., R. D. McPherson and P. D. Polger, 1972: On the efficient reduction of truncation error in numerical weather prediction models. *Mon. Wea. Rev.*, **100**, 637–643.

Gottlieb, D., and E. Turkel, 1978: Boundary conditions for multistep finite difference methods for time-dependent equations. *J. Comput. Phys.*, **26**, 181–196.

Grammeltvedt, A., 1969: A survey of finite difference schemes for the primitive equations for a barotropic fluid. *Mon. Wea. Rev.*, **97**, 384–404.

Gustafsson, B., 1971a: An ADI method for solving the shallow-water equations. *J. Comput. Phys.*, **7**, 239–253.

——, 1971b: On the convergence rate for difference approximations to mixed initial boundary value problems. Rep. 33, Dept. of Computer Sciences, University of Uppsala, Sweden.

——, 1975: The convergence rate for difference approximations to mixed initial boundary value problems. *Math. Comput.*, **29**, 396–406.

——, H. O. Kreiss and A. Sundström, 1972: Difference approximations to mixed initial boundary value problems. II. *Math. Comput.*, **26**, 649–686.

Haltiner, G. J., and R. T. Williams, 1973: Numerical weather prediction: Some problems and solutions. Naval Postgraduate School Report.

Henrici, P., 1974: *Applied and Computational Complex Analysis*, Vol. 1. Wiley, 682 pp.

Hindmarsh, A. C., 1977: Solution of block-tridiagonal systems of linear algebraic equations. Lawrence Livermore Labs., 25 pp.

Hirsch, R., 1975: High-order accurate difference solutions of fluid-mechanics problems by a compact differencing technique. *J. Comput. Phys.*, **19**, 90–109.

Houghton, D., A. Kasahara and W. Washington, 1966: Long-term integration of the barotropic equations by the Lax-Wendroff method. *Mon. Wea. Rev.*, **94**, 141–150.

Kalnay-Rivas, E., 1977: The 4th order GISS model of the global atmosphere. *Contrib. Atmos. Phys.*, **50**, 299–311.

Krause, E., H. Hirschel and W. Kordulla, 1976: Fourth-order

"Mehrstellen" integration for three-dimensional turbulent boundary layers. *Comput. Fluids*, **4**, 77–92.

Kreiss, H. O., 1970: Initial boundary value problems for hyperbolic systems. *Commun. Pure Appl. Math.*, **23**, 277–298.

——, and J. Oliger, 1972: Comparison of accurate methods for the integration of hyperbolic equations. *Tellus* **27**, 199–215.

——, and ——, 1973: Methods for the approximate solution of time-dependent problems. GARP Publ. Ser. No. 10, WMO, Geneva, 107 pp.

Kurihara, Y., 1965: On the use of implicit and iterative methods for the time-integration of the wave equation. *Mon. Wea. Rev.*, **93**, 33–45.

Lomax, H., 1976: Recent progress in numerical techniques for flow-simulation. *AIAA J.*, **14**, 512–518.

Mahrer, Y., and R. A. Pielke, 1978: A test of an upstream spline interpolation technique for the advective terms in a numerical mesoscale model. *Mon. Wea. Rev.*, **106**, 818–830.

Morton, K. W., 1977: Initial value problems by finite difference and other methods. *The State of the Art in Numerical Analysis*, D. Jacobs, Ed., Academic Press, 699–756.

Navon, I. M., 1977: Algorithms for solving scalar and block-cyclic tridiagonal systems of linear algebraic equations. CSIR Spec. Rep., WISK 265, 29 pp.

——, 1978: Application of a new partly implicit time differencing scheme for solving the shallow-water equations. *Contrib. Atmos. Phys.*, **51**, 281–305.

——, 1979: Finite-element simulation of the shallow-water equations model on a limited area domain. Accepted for publication in *Applied Mathematical Modelling*.

——, and Z. Alperson, 1978: Application of fourth-order finite differences to a baroclinic model of the atmosphere. *Arch. Meteor. Geogra. Bioklim.*, **A27**, 1–19.

Numerov, B. V., 1924: A method of extrapolation of perturbations. *Roy. Astron. Soc. Mon. Notice.*, **84**, 592–601.

——, 1927: Note on the numerical integration of $y'' = f(x,y)$. *Astron. Nachrict.*, **230**, 359–364.

Oliger, J., 1974: Fourth-order difference methods for the initial-boundary value problems for hyperbolic equations. *Math. Comput.*, **28**, 15–25.

——, and A. Sundström, 1976: Theoretical and practical aspects of some initial boundary value problems in fluid dynamics. Computer Sci. Dept. Rep., Stanford University, 61 pp.

Orzag, S. A., 1971: Numerical simulation of incompressible flows within simple boundaries: Accuracy. *J. Fluid Mech.*, **49**, Part I, 75–112.

——, and M. Israeli, 1974: Numerical simulation of viscous incompressible flow. *Annual Reviews of Fluid Mechanics*, Vol. 6, Annual Reviews, Inc., 281–319.

Osher, S., 1973: Initial boundary value problems for hyperbolic systems in regions with corners. *Trans. Amer. Math. Soc.*, **176**, 141–165.

——, 1974: Initial boundary value problems for hyperbolic systems in regions with corners II. *Trans. Amer. Math. Soc.*, **198**, 155–175.

Peters, N., 1974: Boundary-layer calculations by a Hermitian finite-difference method. *Proc. Fourth Int. Conf. Numerical Methods in Fluid Dynamics*, Boulder, University of Colorado, 313–318.

Peyret, R., 1978: A Hermitian finite-difference method for the solution of the Navier-Stokes equations. *Numerical Methods in Laminar and Turbulent Flow*, Taylor Morgan, and Brebbia, Eds., Pentech Press, London, 43–54.

Rubin, S. G., and P. K. Khosla, 1976: Higher-order numerical solutions using cubic-splines. *AIAA J.*, **14**, 851–858.

——, and R. A. Graves, 1975: Cubic-spline approximation for problems in fluid dynamics. NASA TR-R-436, 91 pp.

Sadourny, R., 1975: The dynamics of finite-difference models of the shallow-water equations. *J. Atmos. Sci.*, **32**, 680–689.

Shapiro, R., 1970: Smoothing, filtering and boundary effects. *Rev. Geophys. Space Phys.*, **8**, 359–387.

Shuman, F. G., 1955: A method of designing finite difference smoothing operators to meet specifications. U.S. Joint Numerical Weather Prediction Unit Tech. Memo., No. 7, 14 pp.

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

Skollermo, G., 1975a: How the boundary conditions affect the stability and accuracy of some implicit methods for hyperbolic equations. Dept. Comput. Sci. Rep. No. 62, Uppsala University, 33 pp.

——, 1975b: Error analysis for the mixed initial boundary value problem for hyperbolic equations. Dept. Comput. Sci. Rep. No. 63, Uppsala University, 37 pp.

Steger, J. L., 1978: Coefficient matrices for implicit finite difference solution of the inviscid fluid conservation-law equations. *Comput. Methods Appl. Mech. Eng.*, **13**, 175–188.

Swartz, B. K., 1974: The construction of finite difference analogs of some finite element schemes. *Mathematical Aspects of Finite Elements in Partial Differential Equations*, C. de Boor, Ed., Academic Press, 279–312.

——, and B. Wendroff, 1974: The relative efficiency of finite-difference and finite element methods, I: Hyperbolic problems and splines. *SIAM J. Num. Anal.*, **11**, 979–993.

Temperton, C., 1975: Algorithms for the solution of cyclic tridiagonal systems. *J. Comput. Phys.*, **19**, 317–323.

Wall, H. S., 1948: *Analytic Theory of Continued Fractions*. Van Nostrand, 433 pp.

Wallington, C. E., 1962: The use of smoothing or filtering operators in numerical forecasting. *Quart. J. Roy. Meteor. Soc.*, **88**, 473–484.

Wirtz, H. J., F. de Schutter and A. Turi, 1977: An implicit compact finite difference method to solve hyperbolic equations. *Math. Comput. Simulation*, **19**, 241–261.