A Two-Step Scheme for the Advection Equation with Minimized Dissipation and Dispersion Errors

LAWRENCE L. TAKACS

M/A-COM Sigma Data, Inc., Goddard Laboratory for Atmospheres, NASA/Goddard Space Flight Center, Greenbelt, MD 20771

(Manuscript received 11 February 1984, in final form 19 February 1985)

ABSTRACT

A two-step advection scheme of the Lax-Wendroff type is derived which has accuracy and phase characteristics similar to that of a third-order scheme. The scheme is exactly third-order accurate in time and space for uniform flow. The new scheme is compared with other currently used methods, and is shown to simulate well the advection of localized disturbances with steep gradients. The scheme is derived for constant flow and generalized to two-dimensional nonuniform flow.

1. Introduction

One of the most common ways of measuring the relative merit of a given numerical scheme for advection is to consider the scheme's dissipation and dispersion properties. Following the work of Crowley (1968), Molenkamp (1968), Anderson and Fattahi (1974) and Mahlman and Sinclair (1977) on the comparison of numerical schemes for the advection equation, it has been shown that high-order schemes tend to shift their dissipation and dispersion errors toward the high wavenumbers of a given solution. In choosing a scheme of a given order, however, an objective criterion must be devised which will determine the scheme which minimizes some measure of error while at the same time remain within the bounds of computational efficiency and storage requirement determined by each user.

Within the last few years, many schemes have been devised which effectively transport passive scalar quantities with impressively small errors. Examples of familiar methods include the flux-corrected transport (FCT) method developed by Boris and Book (1973, 1976), Book *et al.* (1975), the pseudo-spectral method used by Orszag (1971), the cubic spline interpolation technique as reported by Price and MacPherson (1973), and the quasi-Lagrangian cubic spline approach demonstrated by Purnell (1976). Notwithstanding the success of these methods, many modellers still employ the simpler second-order Lax-Wendroff conservative method or even the first-order upstream scheme due to computational ease and cost. In this regard, much work has been devoted to improving the standard first-order upstream and the second-order Lax-Wendroff schemes with respect to their stability, amplitude and phase characteristics. Smolarkiewicz (1982) replaced the common time-splitting technique for advecting the Lax-Wendroff scheme in two or more directions with the addition of cross-space terms and off-axis information in the first spatial derivatives to improve its stability characteristics in the case of strong deformation flow. The inherent phase lag associated with the Lax-Wendroff scheme, however, remained intact. Alternatively, Smolarkiewicz (1983) used a predictor–corrector sequence applied to the upstream scheme such that the corrector reversed the effect of the implicit diffusion in the upstream predictor. Best results were obtained when the time-splitting technique was used in conjunction with a twice repeated corrector step. Fromm (1968) linearly combined a forward and backward timestep of the Lax-Wendroff scheme which effectively "zeroed-out" the phase lag in each step. Gadd (1978) introduced a free parameter and four grid-points for use in the spatial first derivative of the Lax-Wendroff scheme which eliminated much of the inherent phase lag and reduced the order of magnitude of the dissipation error. However, the free parameter was chosen in such a way that conservation of the advected quantity was not guaranteed.

In this paper, a scheme of the Lax-Wendroff type for uniform flow has been derived which gives an optimal balance between the dissipation and dispersion errors, i.e., a scheme which dissipates most strongly those waves whose phase speed is most poorly represented. The schemes under consideration are such that the order of accuracy in space and time are equal. As mentioned in Anderson and Fattahi (1974), phase truncation errors are related to odd-ordered schemes while amplitude truncation errors are related to even-ordered schemes. This relationship will be developed for the generalized version of the scheme presented, and then examined in detail for a particular second-order scheme using four horizontal

grid-points for uniform flow. We will restrict ourselves to a two-step second-order formulation for simplicity and computational considerations. Higher-order methods in both time and space for nonuniform flow, such as the third-order methods presented by Burstein and Mirin (1970) and Rusanov (1970), are generally implemented using an iterative predictor–corrector–corrector sequence which is generally more costly than simpler two-step processes. In addition, as Petschek and Libersky (1975) and others have shown, even the standard second-order time-splitting Lax-Wendroff scheme may become unstable in the case of strong deformational flow. Therefore, in this paper an attempt is made to construct a stable two-step scheme which at least simulates some of the improved phase characteristics associated with third-order schemes for uniform flow. The generalized version of the scheme and its order of accuracy requirements are introduced in Section 2. Section 3 examines the order of magnitude of the amplitude and phase errors for the generalized scheme, while Section 4 presents the actual scheme under consideration. Sections 5 and 6 show a stability and error analysis for the particular scheme, while Section 7 shows one-dimensional results using constant flow. Sections 8 and 9 generalize the scheme to two dimensions, and Section 10 shows two-dimensional results. Section 11 concludes with a summary.

2. Generalized scheme

For simplicity, the formulation of the scheme will be based on the one-dimensional advection equation given by

$$\frac{\partial q}{\partial t} + c\frac{\partial q}{\partial x} = 0. \tag{2.1}$$

Here q represents any arbitrary quantity being advected, while c is the constant velocity of the flow field in the x-direction.

In general we can write a finite-difference version of (2.1) as

$$q_j^{n+1} = \sum_{j'} a_{j'} q_{j+j'}^n, \quad j' = 0, \pm 1, \pm 2, \cdots. \tag{2.2}$$

Here the subscripts j and $j + j'$ denote the location in space, while the superscripts n and $n + 1$ denote the time index. The weighted coefficients $a_{j'}$ are presently arbitrary.

For a given order of accuracy, certain restrictions are placed on the coefficients. By using a Taylor series substitution, (see Appendix A), it can be shown that first-order accuracy requires

$$\sum_{j'} a_{j'} = 1,$$

$$\sum_{j'} j' a_{j'} = -\mu, \tag{2.3}$$

where μ is the Courant-Friedrichs-Lewy (CFL) stability parameter given by

$$\mu = \frac{c\Delta t}{\Delta x}. \tag{2.4}$$

Second-order accuracy requires in addition to (2.3) the requirement

$$\sum_{j'} j'^2 a_{j'} = \mu^2. \tag{2.5}$$

In general, for mth-order accuracy there are $m + 1$ requirements beginning with (2.3) and ending with

$$\sum_{j'} j'^m a_{j'} = (-\mu)^m. \tag{2.6}$$

When these requirements are satisfied, the scheme will be mth-order accurate both in time and in space.

3. Error analysis

By choosing a solution of (2.2) in the form

$$q_j^n = \mathrm{Re}\{\hat{q}^n e^{ik j\Delta x}\}, \tag{3.1}$$

where $i = \sqrt{-1}$ and k is the horizontal wavenumber, substitution of (3.1) into (2.2) yields

$$\hat{q}^{n+1} = \lambda_D \hat{q}^n, \tag{3.2}$$

where λ_D is an amplification factor associated with the difference solution and is given by

$$\lambda_D = \sum_{j'} a_{j'} e^{ij'\theta}. \tag{3.3}$$

Here we have defined the nondimensional wavenumber θ as

$$\theta = k\Delta x. \tag{3.4}$$

It is easily shown that the true continuous solution of (2.1) also has an amplification factor which is given by

$$\lambda_T = e^{-i\mu\theta}. \tag{3.5}$$

We see that the magnitude of the true amplification factor is 1. Thus, each harmonic wave component of the true solution is simply advected with a phase shift of $-\mu\theta$ for each time step, while the amplitude of each wave component remains unchanged.

Using (3.3) and (3.5), we define the relative error ϵ associated with each time step of the difference scheme by the relation

$$\lambda_D = (1 + \epsilon)\lambda_T, \tag{3.6}$$

where

$$\epsilon = \epsilon_R + i\epsilon_I, \tag{3.7}$$

and

$$\epsilon_R = \sum_{j'} a_{j'} \cos(j' + \mu)\theta - 1, \tag{3.8}$$

$$\epsilon_I = \sum_{j'} a_{j'} \sin(j' + \mu)\theta. \tag{3.9}$$

By writing λ_D and λ_T as

$$\lambda_D = |\lambda_D| e^{i\varphi_D}, \tag{3.10}$$

$$\lambda_T = |\lambda_T| e^{i\varphi_T}, \tag{3.11}$$

it can be shown from (3.6) that

$$|\lambda_D|^2 = |\lambda_T|^2 + \epsilon_{|\lambda|}, \tag{3.12}$$

$$\varphi_D = \varphi_T + \epsilon_\varphi, \tag{3.13}$$

where

$$\epsilon_{|\lambda|} = \epsilon_R{}^2 + \epsilon_I{}^2 + 2\epsilon_R, \tag{3.14}$$

$$\epsilon_\varphi = \tan^{-1}\left(\frac{\epsilon_I}{\epsilon_R + 1}\right). \tag{3.15}$$

Here $\epsilon_{|\lambda|}$ is the error associated with the square of the modulus of the finite-difference amplification factor while ϵ_φ is the error associated with the phase. In obtaining (3.12), the relation $|\lambda_T| = 1$ was used.

Since the error is wavenumber dependent, an insight into the order of magnitude of the error is gained by expanding the error in a Taylor series about $\theta = 0$. Doing so we find that Eqns. (3.8) and (3.9) may be expressed as

$$\epsilon_R = \left(\sum_{j'} a_{j'} - 1\right) - \frac{\Delta\theta^2}{2!}\sum_{j'}(j' + \mu)^2 a_{j'} + \frac{\Delta\theta^4}{4!}\sum_{j'}(j' + \mu)^4 a_{j'} - \cdots, \tag{3.16}$$

$$\epsilon_I = \Delta\theta \sum_{j'}(j' + \mu)a_{j'} - \frac{\Delta\theta^3}{3!}\sum_{j'}(j' + \mu)^3 a_{j'} + \cdots. \tag{3.17}$$

It can be seen that each coefficient of $\Delta\theta^m$ is a binomial expansion of $(j' + \mu)^m$. It is easily shown that for a scheme of mth order accuracy, $\sum_{j'}(j' + \mu)^m a_{j'} = 0$. Thus, by collectively satisfying the order of accuracy requirements defined in Section 2, the order of magnitude of the relative error may be systematically reduced. Furthermore, we can see that for a scheme of a given order m,

$$O(\epsilon_{|\lambda|}) = O(\epsilon_R), \tag{3.18}$$

$$O(\epsilon_\varphi) = O(\epsilon_I), \tag{3.19}$$

since

$$\tan^{-1}\left(\frac{\epsilon_I}{\epsilon_R + 1}\right) \approx \frac{\epsilon_I}{\epsilon_R + 1} = \epsilon_I(1 - \epsilon_R + \epsilon_R{}^2 - \cdots). \tag{3.20}$$

Using (3.18) and (3.19), we may rewrite (3.12) and (3.13) as

$$|\lambda_D|^2 = |\lambda_T|^2 + O(\epsilon_R), \tag{3.21}$$

$$\varphi_D = \varphi_T + O(\epsilon_I). \tag{3.22}$$

Equation (3.21) shows that the order of the error associated with the magnitude of the finite-difference amplification factor is only a function of the real component of the relative error. As can be seen from (3.16), only even-ordered schemes will reduce the order of magnitude of the real component of the relative error. Similarly, we see from (3.22) and (3.17) that only odd-ordered schemes will reduce the order of magnitude of the error associated with the phase change of the finite-difference scheme.

Figures 1 and 2 show the amplitude and phase errors defined by (3.14) and (3.15) for 1st-, 2nd-, 3rd-, and 4th-order accurate schemes found by satisfying the order of accuracy requirements presented in Section 2. The 1st-order scheme is the simple forward in time upstream scheme using two grid-points ($j' = 0, -1$). The 2nd- and 4th-order schemes are centered schemes using three and five grid-points respectively ($j' = 1, 0, -1$ and $j' = 2, 1, 0, -1, -2$). These schemes are the 2nd- and 4th-order schemes as presented in Crowley (1968), of which the 2nd-order scheme is the standard Lax-Wendroff scheme described by Leith (1965). The 3rd-order scheme uses four grid-points ($j' = 1, 0, -1, -2$).

From Fig. 1 we clearly see the reduction in the amplitude error between the 1st- and 2nd-order schemes and also between the 3rd- and 4th-order schemes although to a lesser extent. By systematically going to the next highest even-ordered scheme, the amplitude error is confined to smaller and smaller scales. It is also interesting to note that the odd-ordered upstream-type schemes have errors which are symmetric about $\mu = 0.5$.

In Fig. 2 we see the opposite effect taking place, i.e., a reduction in the phase error between the 2nd- and 3rd-order schemes. In fact, since the even-ordered schemes damp less, the phase error may appear worse due to the poor phase properties of the surviving scales. Finally, we again see the symmetric behavior associated with the odd-ordered schemes.

4. Formulation of proposed scheme

From the preceding sections it was found that using three grid-points centered about q_j^n, second-order accuracy uniquely defined the second-order Lax-Wendroff scheme for uniform flow as described by Leith (1965). We know that second-order schemes have fourth-order amplification errors and third-order phase errors as defined by (3.21) and (3.22). By including one more grid-point, third-order accuracy may be achieved thus reducing the order of magnitude of the phase error.

Achieving a stable third-order accurate scheme in time and space is trivial for the case of uniform flow. However, the same is not so easy when the advecting

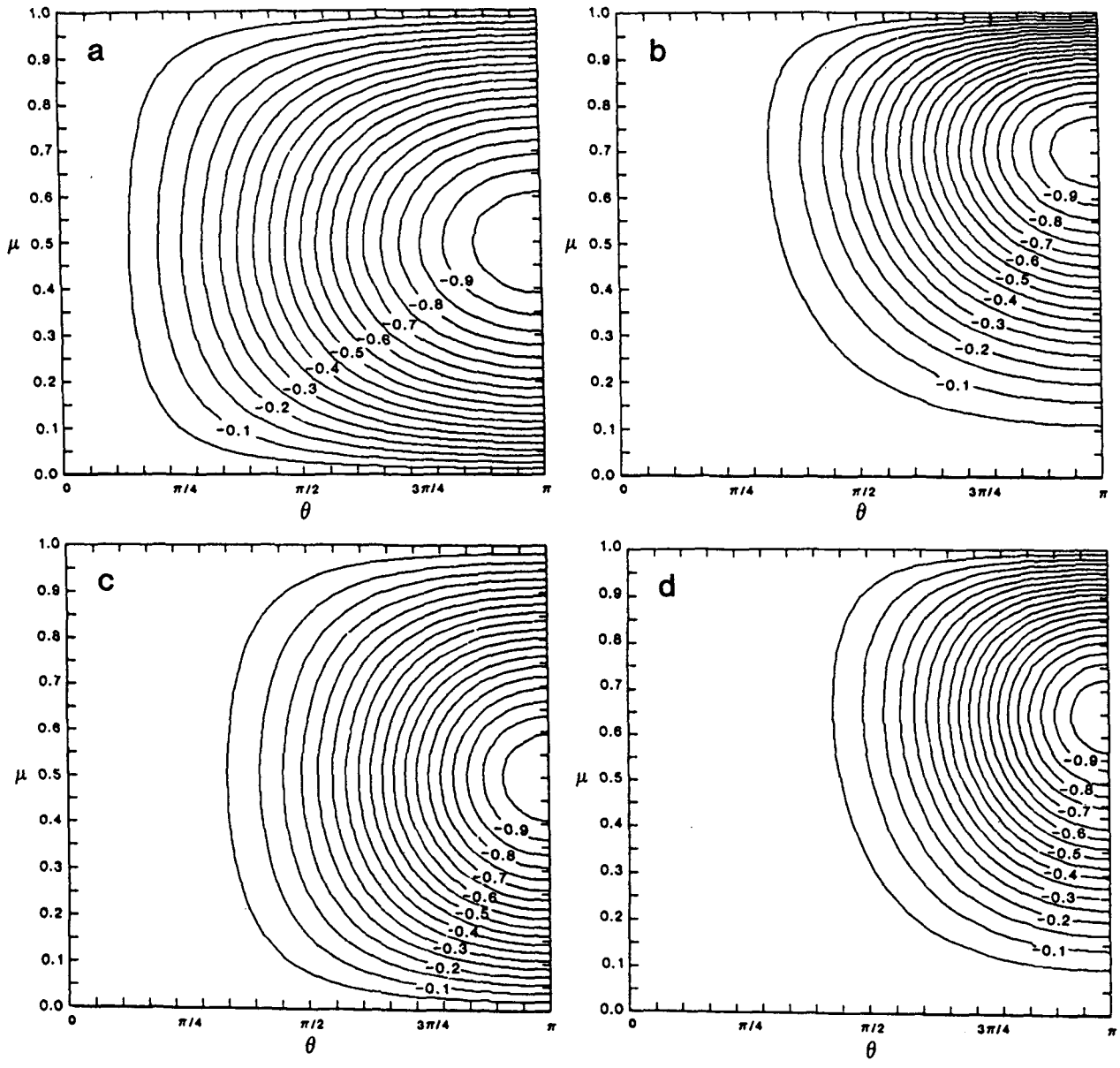


FIG. 1. Amplitude error for (a) 1st-order scheme, (b) 2nd-order scheme, (c) 3rd-order scheme and (d) 4th-order scheme.

flow varies in space. As shown by Petschek and Libersky (1975) and Smolarkiewicz (1982), even the standard second-order time-splitting Lax-Wendroff scheme may become unstable in the case of strong deformational flow. In addition, higher-order schemes in both time and space, such as the third-order methods presented by Burstein and Mirin (1970) and Rusanov (1970), are generally implemented using an iterative predictor–corrector–corrector sequence, where the higher-order accuracy comes about from repeated evaluations of the space-differencing operator. These three-step procedures are generally more costly than simpler two-step procedures. Also, as pointed out by Prof. Fedor Mesinger (personal communication, 1984), given the number of grid-points one is considering to use within the advection scheme, it does not necessarily follow that satisfying the maximum number of the order of accuracy requirements should give the best result. Therefore, rather than satisfying third-order accuracy analytically, we will use four grid-points in a stable two-step second-order scheme for uniform flow and allow the extra grid-point added to act as a free parameter to minimize specific measurements of dissipation and dispersion error. Once the second-order scheme is found for uniform flow, it will be generalized for nonuniform

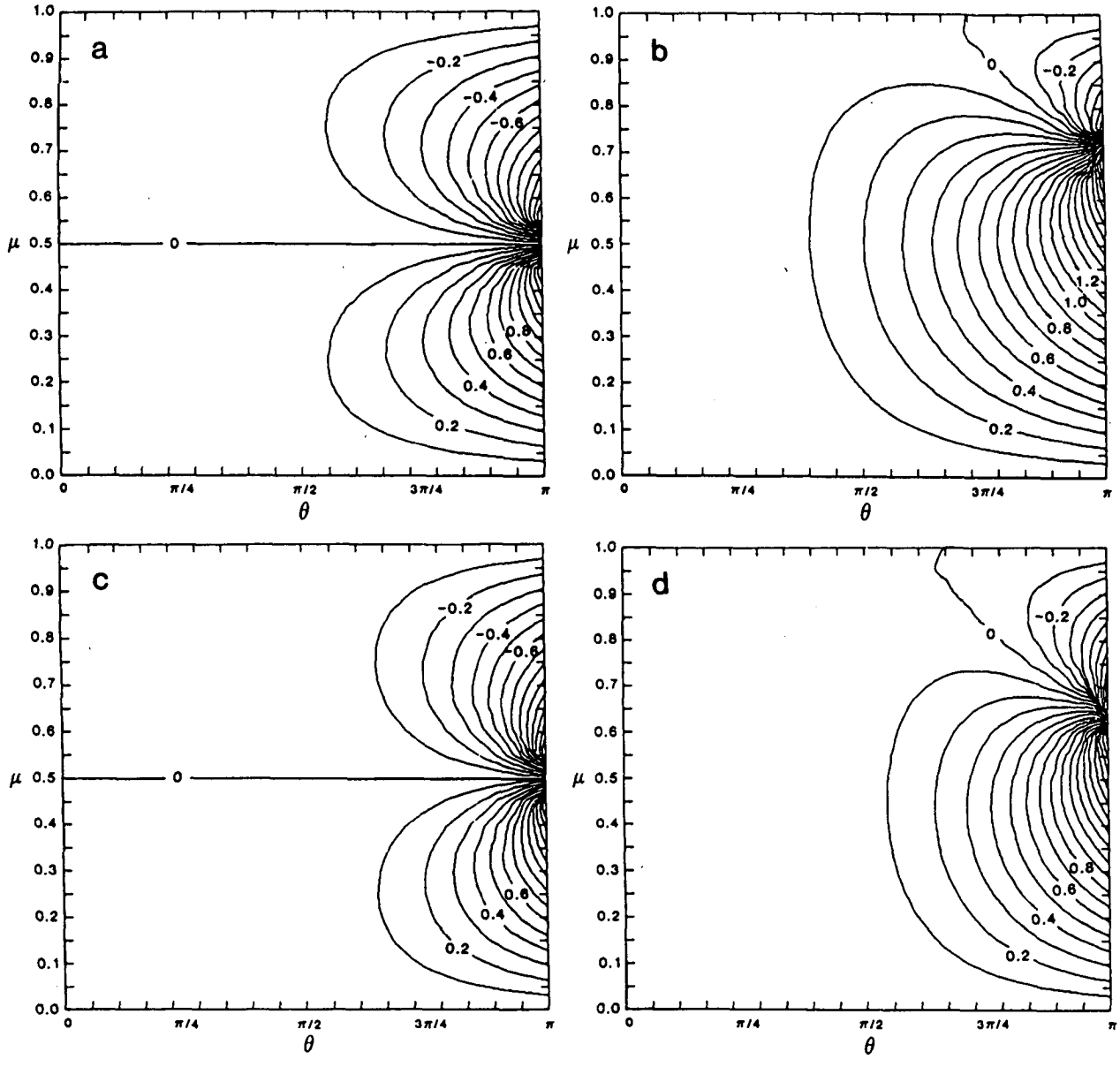


FIG. 2. Phase error for (a) 1st-order scheme, (b) 2nd-order scheme, (c) 3rd-order scheme and (d) 4th-order scheme.

flow with special attention given to its stability characteristics. The following, then, is a summary of the steps taken in the formulation of the second-order scheme.

1. Use one more grid-point than the minimum needed for the desired accuracy. In this case, four grid-points are used ($j' = 1, 0, -1, -2$) for a second-order scheme.
2. Let the coefficient of the extra grid-point be a free parameter.
3. Using the free parameter, obtain a family of stable schemes.
4. From this family of stable schemes, find the scheme which minimizes some measure of error.

The form of the second-order scheme under consideration is given by

$$q_j^{n+1} = a_1 q_{j+1}^n + a_0 q_j^n + a_{-1} q_{j-1}^n + a_{-2} q_{j-2}^n, \quad (4.1)$$

where the coefficients $a_{j'}$ are found by requiring second-order accuracy and have the form

$$a_1 = \mu(\mu - 1)/2 - a_{-2}$$

$$a_0 = 1 - \mu^2 + 3a_{-2}$$

$$a_{-1} = \mu(\mu + 1)/2 - 3a_{-2}$$

$$a_{-2} = \alpha\mu(\mu - 1). \quad (4.2)$$

For the present analysis, we still assume that the flow is positive and constant.

The form of the free parameter a_{-2} was chosen so that for $\mu = 0$ or 1, the scheme is exact. Substituting (4.2) into (4.1) we have

$$q_j^{n+1} = q_j^n - \frac{\mu}{2}(q_{j+1}^n - q_{j-1}^n) + \frac{\mu^2}{2}(q_{j+1}^n - 2q_j^n + q_{j-1}^n)$$

$$- \alpha\mu(\mu - 1)(q_{j+1}^n - 3q_j^n + 3q_{j-1}^n - q_{j-2}^n). \quad (4.3)$$

We see that for $\alpha = 0$, the scheme reduces to the second-order Lax-Wendroff scheme. For nonzero α, an extra term is added corresponding to the third derivative of q_j^n in space. If we allow the free parameter α to be a function of the wind speed μ, then choosing $\alpha = (1 + \mu)/6$ yields the unique third-order accurate scheme in time and space using four grid-points.

5. Stability analysis

In order to obtain a family of stable schemes associated with the free parameter, restrictions were placed on the first nonzero derivative of the expansion described by (3.16) so that the slope of the magnitude of the amplification factor was negative (i.e., damping). Thus

$$\left.\frac{\partial^4 \epsilon_R}{\partial\theta^4}\right|_{\theta=0} \leq 0. \quad (5.1)$$

In addition, it was also required that the scheme be stable for the shortest waves, i.e.,

$$|\lambda_D|^2/_{\theta=\pi} \leq 1. \quad (5.2)$$

Together, (5.1) and (5.2) produce bounds on the free parameter α given by

$$0 \leq \alpha \leq \frac{1}{2}, \quad (5.3)$$

for

$$0 \leq \mu \leq 1. \quad (5.4)$$

It is useful at this point to examine the stability of the proposed scheme by means of the energy method. Squaring (4.3) and summing over all space assuming cyclic boundary conditions we find after rearranging terms

$$\sum_j (q_j^{n+1^2} - q_j^{n^2}) = -\mu(1 - \mu)\left[\frac{\mu(1 + \mu)}{4} + \alpha(1 - 2\mu)\right]$$

$$\times \sum_j (q_{j+1}^n - 2q_j^n + q_{j-1}^n)^2 - \alpha\frac{\mu^2(1 - \mu)^2}{2}(1 - 2\alpha)$$

$$\times \sum_j (q_{j+1}^n - 3q_j^n + 3q_{j-1}^n - q_{j-2}^n)^2. \quad (5.5)$$

We see that for $\mu \leq 1$ and $\alpha \leq \frac{1}{2}$, the scheme is absolutely stable and damping. When $\alpha = 0$ the scheme reduces to the dissipative Lax-Wendroff scheme. For nonzero α the scheme's dissipative properties may actually increase depending on the value of μ. However, it will be shown in Section 7 that the dissipation error is an order of magnitude smaller than the phase or dispersion error associated with the scheme, and that the reduction in the dispersion error for nonzero α more than compensates the possible increase in dissipation error.

6. Dissipation and dispersion error

In order to obtain the optimum value for α, experiments similar to those of Crowley (1968), Molenkamp (1968), and others were performed in which a cone-shaped disturbance with a base-width of 10 grid-points was advected over a distance of 70 grid-points, (see Fig. 3). Three measurements of error were taken representing the total error, the dissipation error, and the dispersion error. The total error was defined to be the mean square error for the experiment, given by

$$E_{TOT} = \frac{1}{M}\sum_j (q_T - q_D)^2, \quad (6.1)$$

where q_T is the true solution, q_D is the finite-difference solution, and M is the total number of grid-points. By expressing (6.1) as the variance of $(q_T - q_D)$ plus the difference of the means squared,

$$\frac{1}{M}\sum_j (q_T - q_D)^2 = \sigma^2(q_T - q_D) + (\bar{q}_T - \bar{q}_D)^2, \quad (6.2)$$

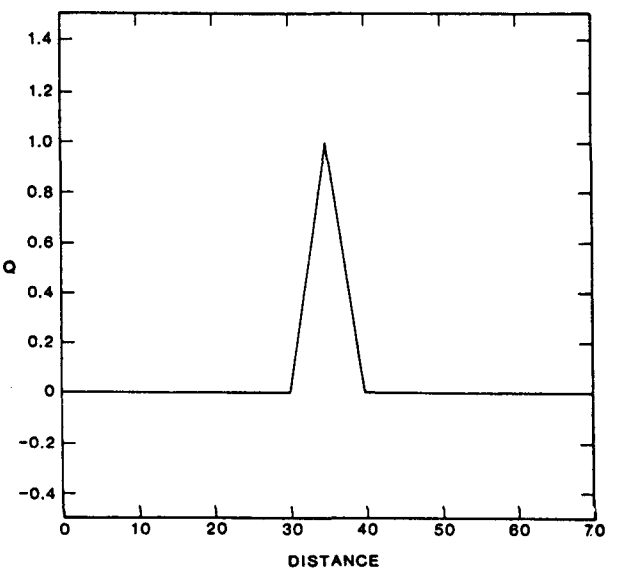
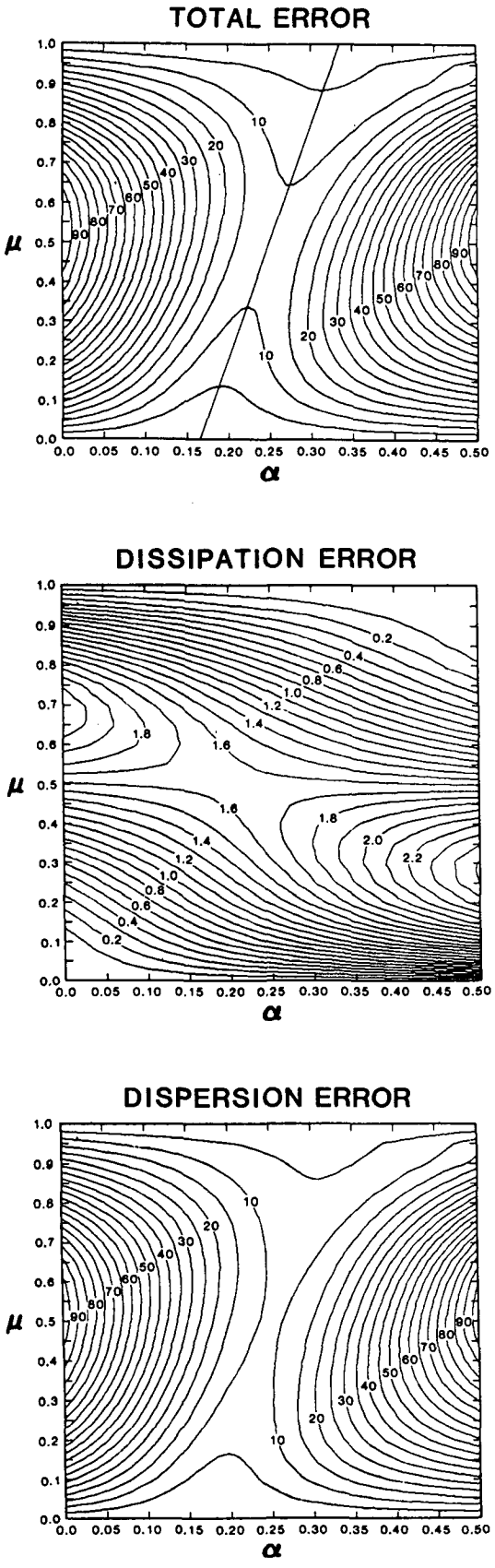


FIG. 3. Initial condition used for testing numerical schemes for advection in one dimension.



it can be shown that

$$E_{TOT} = \sigma^2(q_T) + \sigma^2(q_D) - 2\,\mathrm{cov}(q_T, q_D) + (\bar{q}_T - \bar{q}_D)^2, \quad (6.3)$$

or

$$E_{TOT} = \sigma^2(q_T) + \sigma^2(q_D) - 2\rho\sigma(q_T)\sigma(q_D) + (\bar{q}_T - \bar{q}_D)^2, \quad (6.4)$$

where ρ is the correlation coefficient between q_T and q_D. By rearranging terms we see that

$$E_{TOT} = [\sigma(q_T) - \sigma(q_D)]^2 + (\bar{q}_T - \bar{q}_D)^2 + 2(1 - \rho)\sigma(q_T)\sigma(q_D). \quad (6.5)$$

Now, if q_T and q_D are exactly correlated, i.e. $\rho = 1$, then the only error that can occur is that due to dissipation. Thus we define the dissipation error as

$$E_{DISS} = [\sigma(q_T) - \sigma(q_D)]^2 + (\bar{q}_T - \bar{q}_D)^2. \quad (6.6)$$

For $\rho \neq 1$, the additional error introduced is due to dispersion. Thus, the dispersion error is defined by

$$E_{DISP} = 2(1 - \rho)\sigma(q_T)\sigma(q_D). \quad (6.7)$$

By varying the free parameter α for different values of the advective speed, (i.e., the Courant number), an examination of the three error types may be made and, thus, an optimal scheme chosen.

7. One-dimensional results

Figure 4 shows the results for the three error measurements previously discussed. Each error measurement is scaled by the maximum mean square error and by the number of iterations required to advect the cone the 70 grid-point distance. Thus, the error measurements reflect the error per iteration as a function of μ and α.

From Fig. 4 we see that a minimization of the total error occurs within the possible range of α, with the optimal α in fact lying along the third-order accurate coefficient $\alpha = (1 + \mu)/6$. As μ approaches 0 or 1, the error tends toward zero as expected. It is interesting to note here that choosing $\alpha = 0.25$ yields the one-dimensional scheme derived by Fromm (1968) in which a forward and backward timestep of the second-order Lax-Wendroff scheme was linearly combined to "zero-out" the phase lag in each step. We see here that 0.25 is the linearly averaged value of α along its third-order accurate line.

The next plot in Fig. 4 shows that the dissipation error is affected more strongly by the advective speed

FIG. 4. Total error, dissipation error and dispersion error as a function of the free parameter α and Courant number μ after advecting initial condition one complete translation of the 70 grid-point domain. The relation $\alpha = (1 + \mu)/6$ is also shown.

rather than the free parameter α. Also, the dissipation error is an order of magnitude less than the total error. Thus we see that the extra grid-point added to the second-order scheme has very little effect on reducing dissipation error. This is in agreement with the results from Section 3 when we were considering the errors within the wavenumber domain. Finally, from the last plot in Fig. 4 we see that the error associated with dispersion is most responsible for the observed total error. The order of magnitude of the dispersion error is the same as that for the total error, and is minimized by the same optimal α. From these three error measurements we see that choosing α to be equal to the third-order accurate coefficient is the most effective way of minimizing the total error, with the major influence coming from a reduction in the dispersion error.

Figures 5–8 show the results for advecting two initial profiles through two translations of the 70 grid-point domain using cyclic boundary conditions. In each figure a comparison is shown for the four schemes discussed in Section 3, thus including the derived scheme with the optimal α (i.e., the 3rd-order scheme). The 4th-order scheme for uniform flow is shown for comparative purposes only, since a complete generalization to two-dimensional nonuniform flow with proper stability characteristics is beyond the scope of this paper. Figures 5 and 6 show the results for advecting the cone with $\mu = 0.2$ and $\mu = 0.7$, while Figs. 7 and 8 show the same for the step

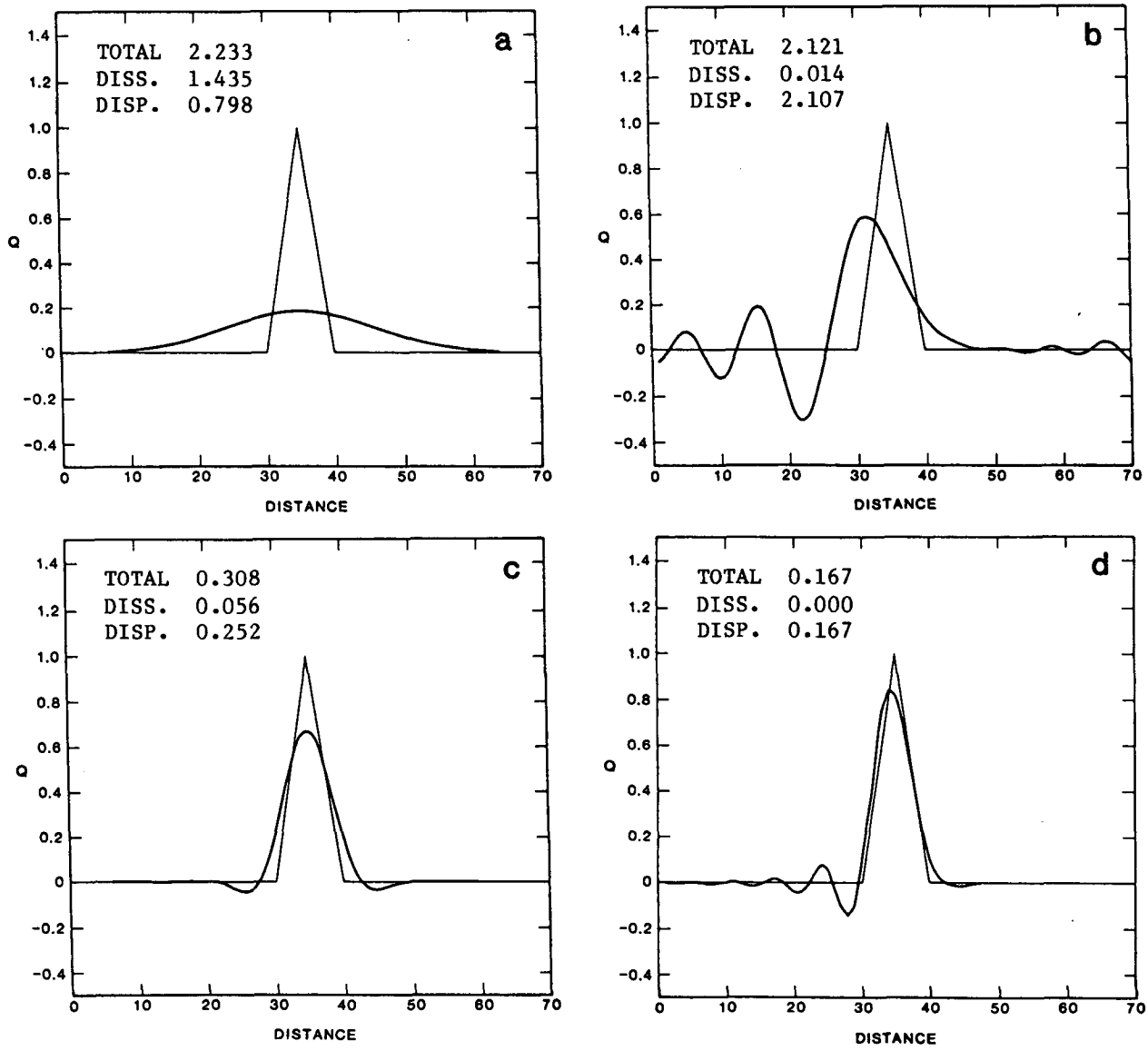


FIG. 5. Numerical solutions for (a) 1st-order scheme, (b) 2nd-order scheme, (c) 3rd-order scheme, and (d) 4th-order scheme, after two complete translations using $\mu = 0.2$

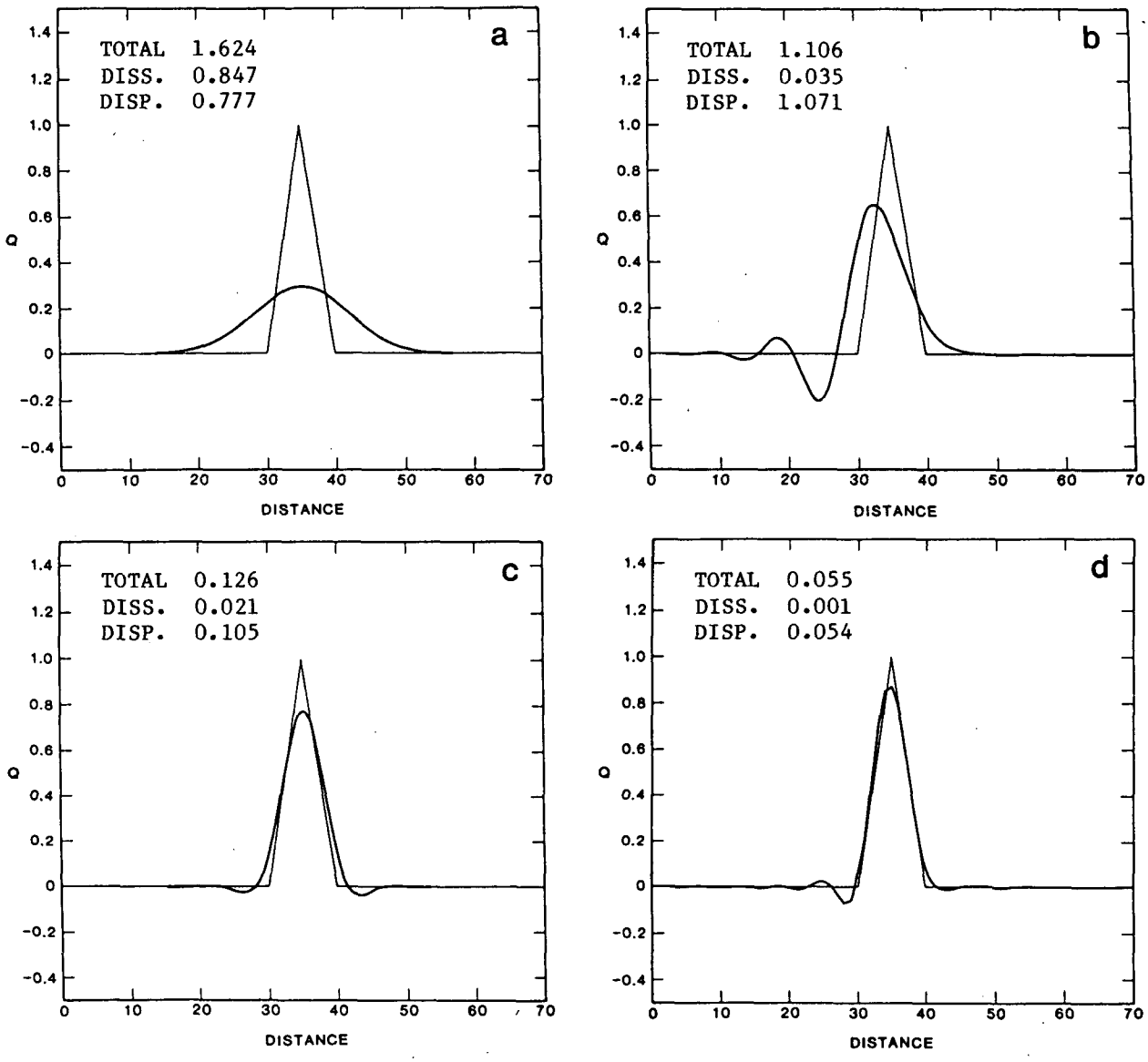


FIG. 6. As in Fig. 5 but for $\mu = 0.7$.

initial profile. In each plot the total error, dissipation error, and dispersion error as defined by (6.1), (6.6) and (6.7) are also presented scaled by the total number of iterations and by the number of grid-points used.

In all figures we see the drastic decrease in dissipation error obtained by going from the 1st-order scheme (a) to the 2nd-order scheme (b). The dispersion error, however, generally increases due to the poor phase properties of the nondissipated wave components. The dispersion error associated with the 3rd-order scheme (c) is seen to decrease significantly, thus greatly improving the numerical solutions over that of schemes (a) and (b). The 4th-order scheme (d) is seen to generally improve the numerical solution by reducing the dissipation error, although the solutions may contain more wake-like oscillations due to the nondissipated small scales.

8. Two-dimensional version

There is no unique generalization to two dimensions for the scheme presented in Section 4 for nonuniform velocity fields. However, two-dimensional schemes may be constructed by the application of two one-dimensional operators in succession. The stability of such a scheme is reduced to the stability of each separate step. In this paper, therefore, two-dimensional schemes are derived by applying two passes of a one-dimensional operator. The one-dimensional operator was constructed by imposing three constraints on the

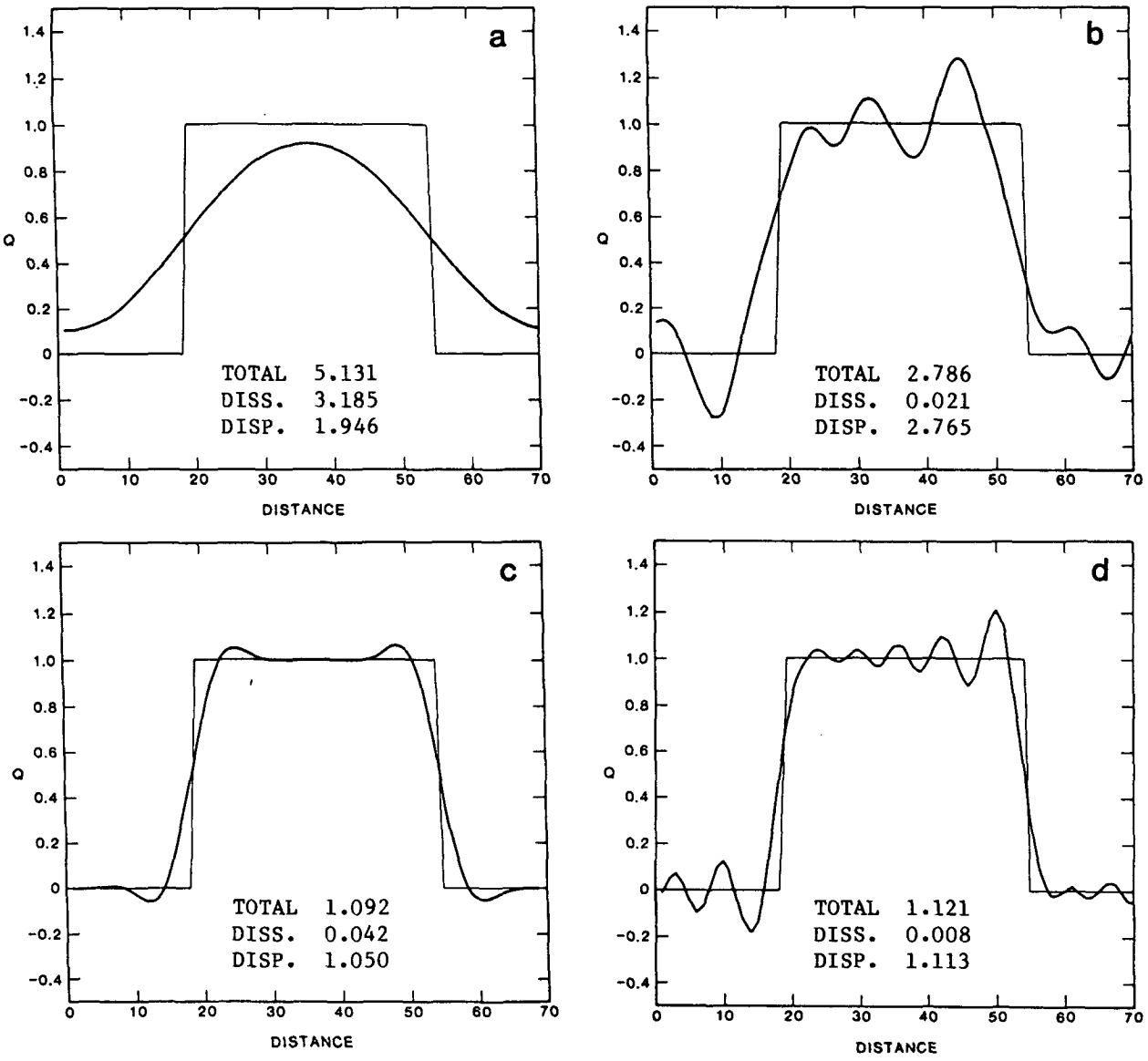


FIG. 7. As in Fig. 5 but for step profile.

time-continuous version (for simplicity) of the scheme given by (4.3).

First, the requirement of conservation of mass was imposed yielding a flux formulation of the scheme,

$$\frac{\partial}{\partial t}\sum_j q_j = 0. \tag{8.1}$$

Second, it was required that when the flow field is constant, the time-continuous flux form reduce to the time-continuous advective form given by

$$\frac{\partial q_j}{\partial t} = \frac{-c}{2\Delta x}(q_{j+1} - q_{j-1}) + \frac{\alpha c}{\Delta x}(q_{j+1} - 3q_j + 3q_{j-1} - q_{j-2}). \tag{8.2}$$

In Eq. (8.2), the notation for the free parameter α is kept generalized. In the time-continuous version, α is simply equal to $\frac{1}{6}$. For the time-discrete version, however, α must become a function of space in keeping with the requirement given by (8.1) since it carried information about the wind field.

Finally, a stability criterion for the time-continuous case when the flow field is nondivergent was imposed, given by

$$\frac{\partial}{\partial t}\sum_j \frac{1}{2} q_j^2 \leq 0. \tag{8.3}$$

Using these requirements, it can be shown that the flux form of (8.2) using a staggered grid is given by

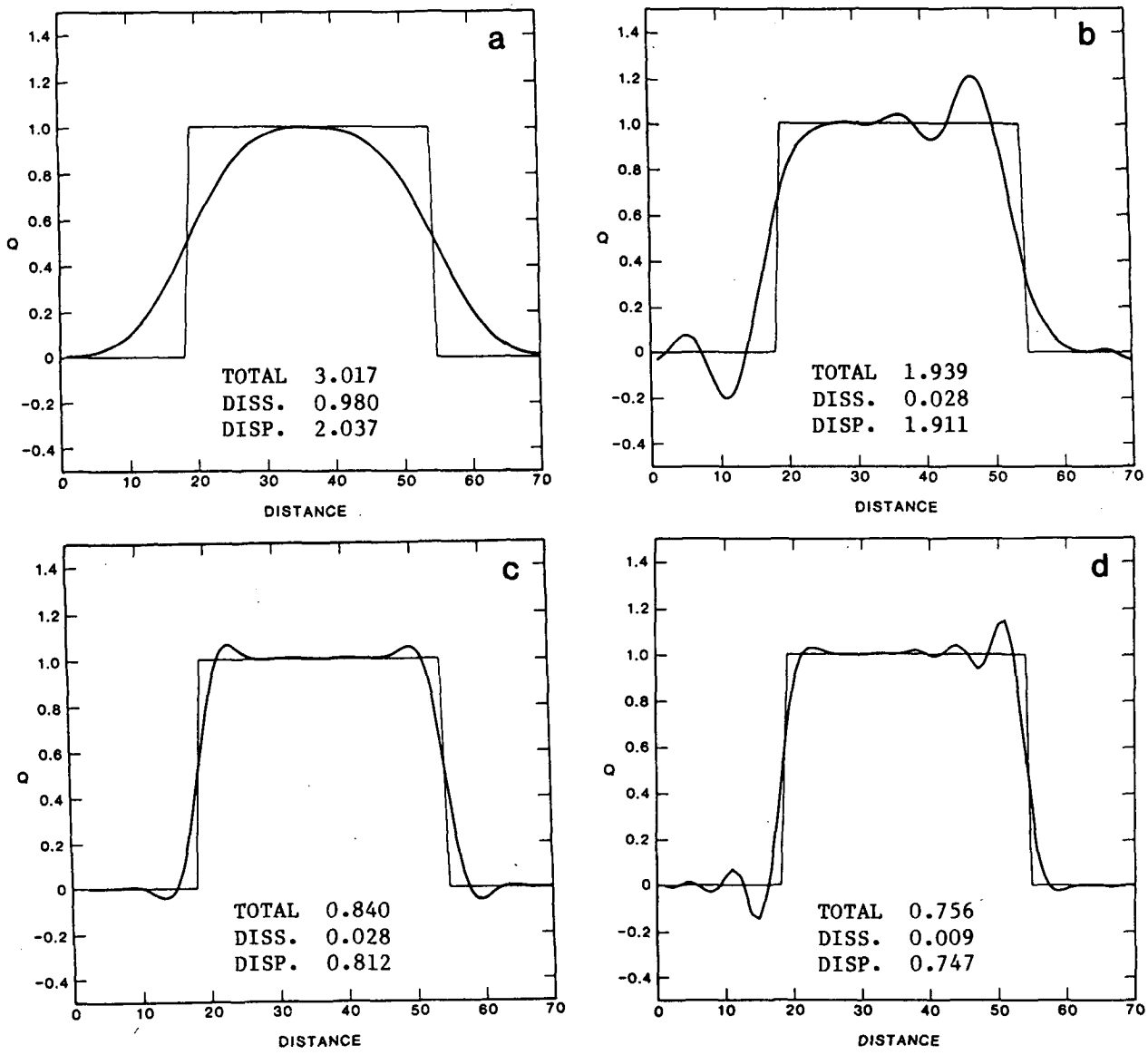


FIG. 8. As in Fig. 5 but for $\mu = 0.7$.

$$\frac{\partial q_j}{\partial t} = -\frac{1}{2\Delta x}[P_{j+1/2} - P_{j-1/2}] + \frac{\alpha}{\Delta x}[Q_{j+1/2} - Q_{j-1/2}], \tag{8.4}$$

where

$$P_{j+1/2} = c_{j+1/2}(q_{j+1} + q_j)$$

$$Q_{j+1/2} = c_{j+1/2}(q_{j+1} - q_j) - \hat{c}_{j+1/2}\hat{c}_{j-1/2}(q_j - q_{j-1})$$

$$\hat{c}_{j+1/2} = \sqrt{c_{j+1/2}}.$$

A detailed analysis of the stability characteristics of (8.4) is given in Appendix B.

For the time-discrete case ($\mu > 0$) and changing wind speeds, a predictor-corrector sequence can be devised which will reduce to Eq. (4.3) when the flow field is constant. Substitution will show that Eqs. (8.5) and (8.6) satisfy this requirement.

PREDICTOR:

$$q_j^* = q_j^n - [F_{j+1/2} - F_{j-1/2}],$$
$$F_{j+1/2} = \mu_{j+1/2} q_j^n. \tag{8.5}$$

CORRECTOR:

$$q_j^{n+1} = q_j^n - \frac{1}{2}[P_{j+1/2} - P_{j-1/2}] + [\alpha_{j+1/2}Q_{j+1/2} - \alpha_{j-1/2}Q_{j-1/2}] \tag{8.6}$$

where

$$P_{j+1/2} = \mu_{j+1/2}(q^*_{j+1} + q^n_j)$$

$$Q_{j+1/2} = \mu_{j+1/2}(q^*_{j+1} - q^n_j) - \hat{\mu}_{j+1/2}\hat{\mu}_{j-1/2}(q^*_j - q^n_{j-1})$$

$$\alpha_{j+1/2} = (1 + \mu_{j+1/2})/6$$

$$\hat{\mu}_{j+1/2} = \sqrt{\mu_{j+1/2}}.$$

Here we see that the free parameter α is now a function of space.

9. Nonpositive flow fields

In all discussions thus far, the flow field, although varying in space, was assumed to be positive. In general, however, advection by a nonuniform two-dimensional flow requires additional information to ensure the inclusion of the proper grid-points for upstream-type schemes. The two-dimensional scheme of Section 8 generalized to allow for nonpositive flow is given by

PREDICTOR:

$$q^*_j = q^n_j - [F_{j+1/2} - F_{j-1/2}]$$

$$F_{j+1/2} = \mu^+_{j+1/2}q^n_j + \mu^-_{j+1/2}q^n_{j+1} \qquad (9.1)$$

where

$$\mu^+_{j+1/2} = \left(\frac{\mu + |\mu|}{2}\right)_{j+1/2}$$

$$\mu^-_{j+1/2} = \left(\frac{\mu - |\mu|}{2}\right)_{j+1/2}.$$

CORRECTOR:

$$q^{n+1}_j = q^n_j - \frac{1}{2}[P_{j+1/2} - P_{j-1/2}] + [\alpha_{j+1/2}Q_{j+1/2} - \alpha_{j-1/2}Q_{j-1/2}] \qquad (9.2)$$

where

$$P_{j+1/2} = \mu^+_{j+1/2}(q^*_{j+1} + q^n_j) + \mu^-_{j+1/2}(q^*_j + q^n_{j+1})$$

$$Q_{j+1/2} = [\mu^+_{j+1/2}(q^*_{j+1} - q^n_j) - \hat{\mu}^+_{j+1/2}\hat{\mu}^+_{j-1/2}(q^*_j - q^n_{j-1})] - [\mu^-_{j+1/2}(q^n_{j+1} - q^*_j) + \hat{\mu}^-_{j+1/2}\hat{\mu}^-_{j+3/2}(q^n_{j+2} - q^*_{j+1})]$$

$$\alpha_{j+1/2} = \left(\frac{1 + |\mu|}{6}\right)_{j+1/2}$$

$$\hat{\mu}^\pm_{j+1/2} = (|\mu^\pm_{j+1/2}|)^{1/2}.$$

10. Two-dimensional results

As the first two-dimensional test, a cone-shaped disturbance was advected by solid-body rotation in a 100 × 100 grid-point domain. A streamfunction defined by

$$\psi_{i,j} = \frac{\Omega}{2}[(x_i - 50)^2 + (y_j - 50)^2] \qquad (10.1)$$

was used to determine the velocity field. The constant Ω was equal to $1/80$, thus the maximum Courant number was equal to 0.625 at the domain boundary.

Figure 9 shows the initial condition for this experiment and the numerical solutions after one and two complete revolutions using the proposed scheme and the standard Lax-Wendroff scheme with the time-splitting technique for advection in two dimensions. We again see the improved phase characteristics of the new scheme when compared to the standard Lax-Wendroff scheme. Also included in these plots are the total, dissipation and dispersion errors as previously defined.

As a second test for the two-dimensional version, experiments following Smolarkiewicz (1982) were performed in which the initial cone profile was subjected to strong deformational flow derived from the streamfunction.

$$\psi_{i,j} = 3.94 \sin\left(\frac{\pi x_i}{25}\right)\cos\left(\frac{\pi y_i}{25}\right). \qquad (10.2)$$

The maximum Courant number in this case was equal to 0.5. Figure 10 shows the initial condition and the numerical solutions for the proposed scheme and the standard Lax-Wendroff scheme after 3000 iterations in addition to the integrated square of the field normalized to its initial value. As in the results presented by Smolarkiewicz, the standard Lax-Wendroff scheme becomes unstable producing a growth in the integrated square of the field. Smolarkiewicz showed that even periodic filtering of the small scales could not prevent this instability since it was due to an amplification of the longer wave components. The proposed scheme, however, behaves well and produces a solution very similar to that obtained by Smolarkiewicz, although no extra smoothing or filtering was needed as was in Smolarkiewicz's study.

One final note should be mentioned in the case of the two-dimensional version of the new scheme. Although the $Q_{j+1/2}$ term in the corrector [eq. (9.2)] contains geometric means of the velocity field in order to guarantee stability in the time-continuous case, it has been experimentally found that the overall damping nature of the scheme allows a simpler expression to be used, given by

$$Q_{j+1/2} = \mu^+_{j+1/2}[(q^*_{j+1} - q^n_j) - (q^*_j - q^n_{j-1})] - \mu^-_{j+1/2}[(q^n_{j+1} - q^*_j) - (q^n_{j+2} - q^*_{j+1})]. \qquad (10.3)$$

In all experiments that were performed, using (10.3) in (9.2) produced results as good as those given by the original formulation.

11. Summary

By using four grid-points for a two-step finite-difference formulation of the advection equation, a scheme with optimal dissipation and dispersion prop-

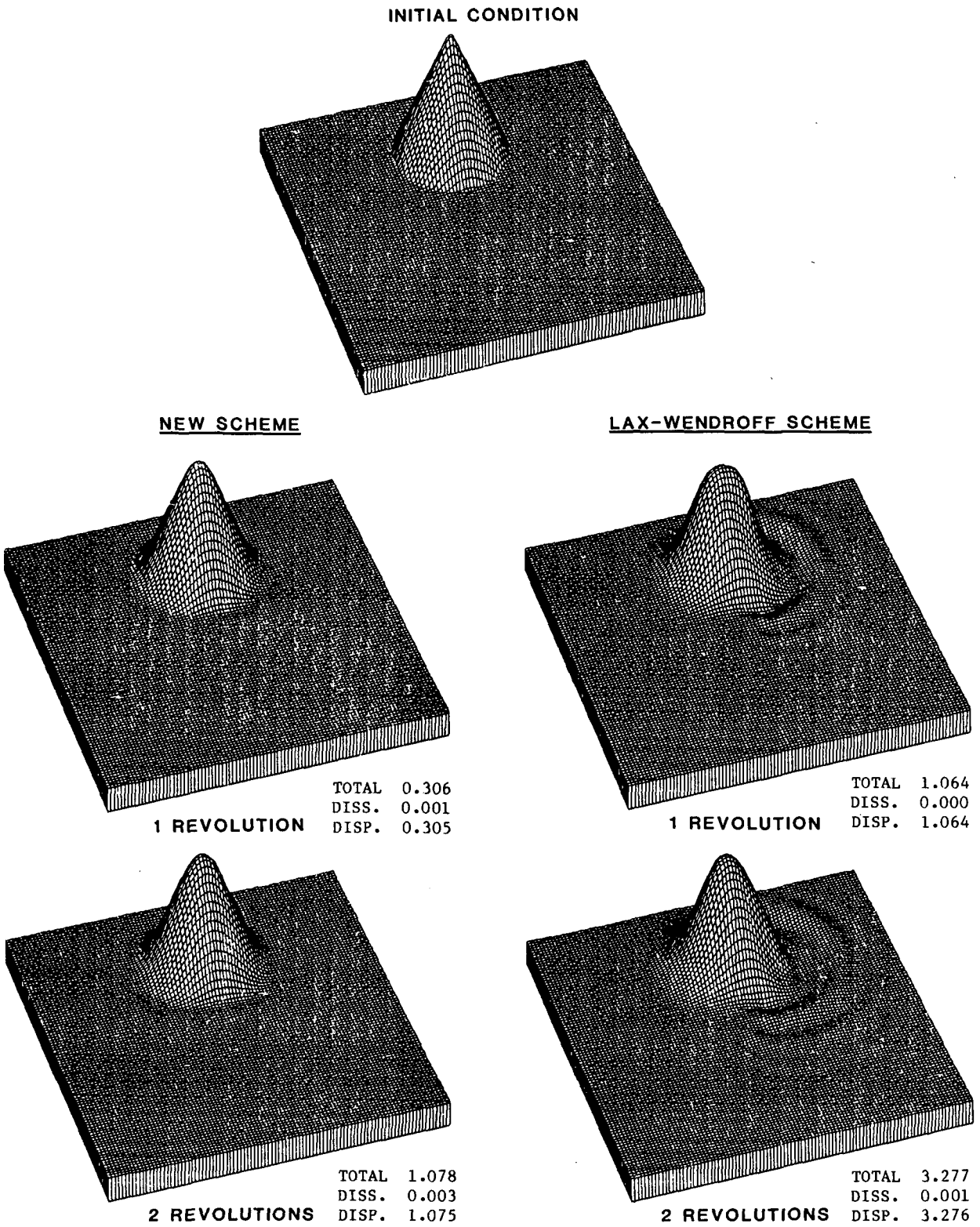


FIG. 9. Initial condition and numerical solutions for proposed scheme and standard Lax-Wendroff scheme using solid-body rotation.

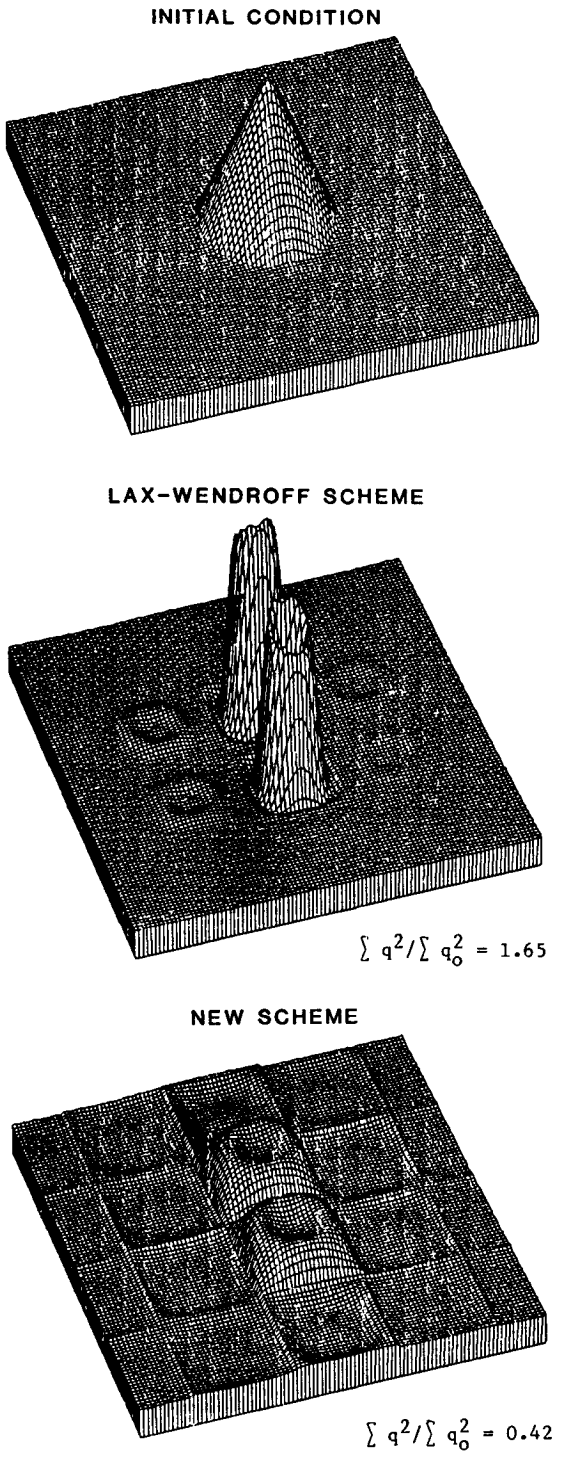


FIG. 10. As in Fig. 9 but using deformation flow field.

erties is devised which simulates the characteristics of third-order accurate schemes. For uniform flow, the scheme is exactly third-order accurate in both time and space. A generalization to two dimensions and nonuniform flow is also presented. Numerical experiments indicate that the new scheme advects well localized disturbances with steep gradients, and also has favorable stability properties in the presence of strong deformational flow.

Acknowledgments. I would like to gratefully acknowledge Professors Akio Arakawa and Max Suarez for their help and useful discussions relating to the formulation of the one- and two-dimensional versions of the numerical scheme presented. Part of this work formed the basis of graduate research while at UCLA. I would also like to thank Dr. Eugenia Kalnay of the Goddard Laboratory for Atmospheres for her many useful comments and suggestions and continuing support for this work. Thanks are also given to the anonymous reviewers of the manuscript who helped in clarifying and focusing the presentation. Finally, thanks are due Mary Ann Wells for typing the manuscript and Laura Rumburg and Brian Sherbs for drafting the figures.

APPENDIX A

Order of Accuracy Requirements for Generalized Scheme

A modified partial differential equation may be obtained to investigate the space and time truncation errors of a given scheme by expanding the generalized difference equation in a Taylor series about the point q_j^n. Beginning with the advection equation given by

$$\frac{\partial q}{\partial t} = -c\frac{\partial q}{\partial x}, \tag{A1}$$

we define the generalized difference equation by

$$q_j^{n+1} = \sum_{j'} a_{j'} q_{j+j'}^n, \quad j' = 0, \pm 1, \pm 2, \cdots. \tag{A2}$$

Expanding (2) in a Taylor series, we have

$$q_j^n + \left(\Delta t\frac{\partial q}{\partial t} + \frac{\Delta t^2}{2!}\frac{\partial^2 q}{\partial t^2} + \frac{\Delta t^3}{3!}\frac{\partial^3 q}{\partial t^3} + \cdots\right)\Big|_{j,n}$$
$$= \sum_{j'} a_{j'}\left[q_j^n + \left((j'\Delta x)\frac{\partial q}{\partial x} + \frac{(j'\Delta x)^2}{2!}\frac{\partial^2 q}{\partial x^2} + \frac{(j'\Delta x)^3}{3!}\frac{\partial^3 q}{\partial x^3} + \cdots\right)\Big|_{j,n}\right]. \tag{A3}$$

Regrouping, we obtain

$$q_j^n(1 - \sum a_{j'}) + \left(\Delta t\frac{\partial q}{\partial t} - \Delta x\sum j' a_{j'}\frac{\partial q}{\partial x}\right)$$
$$+ \frac{1}{2!}\left(\Delta t^2\frac{\partial^2 q}{\partial t^2} - \Delta x^2\sum j'^2 a_{j'}\frac{\partial^2 q}{\partial x^2}\right)$$
$$+ \frac{1}{3!}\left(\Delta t^3\frac{\partial^3 q}{\partial t^3} - \Delta x^3\sum j'^3 a_{j'}\frac{\partial^3 q}{\partial x^3}\right) + \cdots = 0 \tag{A4}$$

where all partial derivatives are evaluated at the point j, n and all summations are over all possible j'.

From the advection equation, (1), we see that

$$\frac{\partial^m q}{\partial t^m} = (-c)^m \frac{\partial^m q}{\partial x^m}, \quad m = 1, 2, 3, \cdots. \tag{A5}$$

Using (5) in (4) and dividing through by Δt gives

$$q_j^n\left(\frac{1 - \sum a_{j'}}{\Delta t}\right) + \left(\frac{\partial q}{\partial t} - \frac{\sum j' a_{j'}}{\mu} c \frac{\partial q}{\partial x}\right) + \frac{\Delta t}{2!}\left(1 - \frac{\sum j'^2 a_{j'}}{\mu^2}\right)\frac{\partial^2 q}{\partial t^2} + \frac{\Delta t^2}{3!} \times \left(1 + \frac{\sum j'^3 a_{j'}}{\mu^3}\right)\frac{\partial^3 q}{\partial t^3} + \cdots = 0, \tag{A6}$$

where μ is the Courant-Friedrichs-Lewy stability parameter defined by $\mu = c\Delta t/\Delta x$. We see that in the limit as Δx and Δt approach zero while μ remains finite, Eq. (6) is consistant with the advection equation provided

$$\sum_{j'} a_{j'} = 1$$

and

$$\sum_{j'} j' a_{j'} = -\mu. \tag{A7}$$

A scheme will have second-order accuracy in both time and space provided, in addition to (7), that

$$\sum_{j'} j'^2 a_{j'} = \mu^2. \tag{A8}$$

In general, for mth-order accuracy in both time and space, accuracy requirements beginning with (7) and ending with

$$\sum_{j'} j'^m a_{j'} = (-\mu)^m. \tag{A9}$$

must be satisfied.

APPENDIX B

Stability Analysis of Two-Dimensional Scheme in Flux Form

The stability characteristics of the time-continuous two-dimensional flux form of Eq. (8.4) may be obtained by multiplying the scheme by q_j and summing over all space. By applying two passes of the one-dimensional scheme for advecting in two dimensions, only the one-dimensional flux form need be considered. The time-continuous version of the scheme is given by

$$\frac{\partial q_j}{\partial t} = -\frac{1}{2\Delta x}[c_{j+1/2}(q_{j+1} + q_j) - c_{j-1/2}(q_j + q_{j-1})] + \frac{\alpha}{\Delta x}[c_{j+1/2}(q_{j+1} - q_j) - \hat{c}_{j+1/2}\hat{c}_{j-1/2}(q_j - q_{j-1}) - c_{j-1/2}(q_j - q_{j-1}) + \hat{c}_{j-1/2}\hat{c}_{j-3/2}(q_{j-1} - q_{j-2})], \tag{B1}$$

where $\hat{c}_{j+1/2} = (c_{j+1/2})^{1/2}$. Multiplying (1) by q_j and summing over all space, we have

$$\frac{\partial}{\partial t}\sum_j \frac{1}{2} q_j^2 = -\frac{1}{2\Delta x}\sum_j [c_{j+1/2} q_{j+1} q_j + q_j^2(c_{j+1/2} - c_{j-1/2}) - c_{j-1/2} q_j q_{j-1}] + \frac{\alpha}{\Delta x}\sum_j \left[c_{j+1/2}\left(q_{j+1}q_j - \frac{1}{2}q_j^2\right) - c_{j-1/2}\frac{1}{2}q_{j-1}^2 - c_{j-1/2}\left(\frac{1}{2}q_j^2 - q_j q_{j-1}\right) - c_{j+1/2}\frac{1}{2}q_{j+1}^2 - \hat{c}_{j+1/2}\hat{c}_{j-1/2}(q_j^2 - q_j q_{j-1} - q_{j+1}q_j + q_{j+1}q_{j-1})\right], \tag{B2}$$

where we have assumed cyclic boundary conditions. Equation (2) may be rewritten as

$$\frac{\partial}{\partial t}\sum_j \frac{1}{2} q_j^2 = -\frac{1}{2\Delta x}\sum_j q_j^2(c_{j+1/2} - c_{j-1/2}) - \frac{\alpha}{2\Delta x}\sum_j [c_{j+1/2}(q_{j+1} - q_j)^2 - 2\hat{c}_{j+1/2}\hat{c}_{j-1/2}(q_{j+1} - q_j) \times (q_j - q_{j-1}) + c_{j-1/2}(q_j - q_{j-1})^2], \tag{B3}$$

or

$$\frac{\partial}{\partial t}\sum_j \frac{1}{2} q_j^2 = -\frac{1}{2}\sum_j q_j^2\left(\frac{c_{j+1/2} - c_{j-1/2}}{\Delta x}\right) - \frac{\alpha}{2\Delta x} \times \sum_j [\hat{c}_{j+1/2}(q_{j+1} - q_j) - \hat{c}_{j-1/2}(q_j - q_{j-1})]^2. \tag{B4}$$

For two-dimensional nondivergent flow, the first summation in (4) will vanish when both dimensions are added. For one-dimensional constant flow, this term is automatically zero. The second summation in (4) is a sum of squares and is thus absolute positive for $\alpha > 0$. Therefore, the requirement for stability given by

$$\frac{\partial}{\partial t}\sum_j \frac{1}{2} q_j^2 \leq 0. \tag{B5}$$

is satisfied.

REFERENCES

Anderson, D., and B. Fattahi, 1974: A comparison of numerical solutions of the advection equation. *J. Atmos. Sci.*, **31**, 1500–1506.

Book, D. L., J. P. Boris and K. Hain, 1975: Flux-corrected transport. II: Generalizations of the method. *J. Comput. Phys.*, **18**, 248–283.

Boris, J. P., and D. L. Book, 1973: Flux-corrected transport. I: SHASTA, a fluid transport algorithm that works. *J. Comput. Phys.*, **11**, 38–69.

——, and ——, 1976: Flux-corrected transport. III: Minimal-error FCT algorithms. *J. Comput. Phys.*, **20**, 397–431.

Burstein, S. Z., and A. A. Mirin, 1970: Third order differencing methods for hyperbolic equations. *J. Comput. Phys.*, **5**, 547–571.

Crowley, W. P., 1968: Numerical advection experiments. *Mon. Wea. Rev.*, **96**, 1–11.

Fromm, J. E., 1968: A method for reducing dispersion in convective difference schemes. *J. Comput. Phys.*, **3**, 176–189.

Gadd, A. J., 1978: A numerical advection scheme with small phase speed errors. *Quart. J. Roy. Meteor. Soc.*, **104**, 583–594.

Leith, C. E., 1965: Numerical simulations of the earth's atmosphere, *Methods in Computational Physics*, Vol. 4, Academic Press, 1–28.

Mahlman, J. D., and R. W. Sinclair, 1977: Tests of various numerical algorithms applied to a simple trace constituent air transport problem, *Fate of Pollutants in the Air and Water Environments, Part I*, Vol. 8, Wiley, 295 pp.

Molenkamp, C. R., 1968: Accuracy of finite difference methods applied to the advection equation. *J. Appl. Meteor.*, **7**, 160–167.

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

Petschek, A. G., and L. D. Libersky, 1975: Stability, accuracy, and improvement of Crowley advection schemes. *Mon. Wea. Rev.*, **103**, 1104–1109.

Price, G. V., and A. K. MacPherson, 1973: A numerical weather forecasting method using cubic splines on a variable mesh. *J. Appl. Meteor.*, **12**, 1102–1113.

Purnell, D. K., 1976: Solution of the advection equation by upstream interpolation with a cubic spline. *Mon. Wea. Rev.*, **104**, 42–48.

Rusanov, V. V., 1970: On difference schemes of third order accuracy for nonlinear hyperbolic systems. *J. Comput. Phys.*, **5**, 507–516.

Smolarkiewicz, P. K., 1982: The multidimensional Crowley advection scheme. *Mon. Wea. Rev.*, **110**, 1968–1983.

——, 1983: A simple positive definite advection scheme with small implicit diffusion. *Mon. Wea. Rev.*, **111**, 479–486.