A positive finite-difference advection scheme applied on locally refined grids

W.H. Hundsdorfer, B. Koren, M. van Loon, J.G. Verwer

Department of Numerical Mathematics

Report NM-R9309 April 1993

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications. SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

P.O. Box 4079, 1009 AB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

A Positive Finite-Difference Advection Scheme Applied on Locally Refined Grids

W. Hundsdorfer, B. Koren, M. van Loon and J.G. Verwer

CWI

P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Abstract

This paper examines a class of explicit finite-difference advection schemes derived along the method of lines. An important problem field is large-scale atmospheric transport. The paper therefore focuses on the demand of positivity and on the use of locally refined grids. For the spatial discretization, attention is confined to directionally-split schemes in conservation form using 5 points per direction. The 4-th order central scheme and the family of κ-schemes comprising the 2-nd order central, the 2-nd order upwind and the 3-rd order upwind biased discretization are studied. Positivity is enforced through flux limiting. It is concluded that the limited 3-rd order upwind discretization is the best candidate from the four examined. For the time integration attention is confined to a number of explicit Runge-Kutta methods of orders 2 up to 4. With regard to the demand of positivity, these integration methods turn out to behave almost equally and no best method could be identified. For solutions with steep gradients, a local uniform grid refinement technique is proposed as an effective tool to speed up computations.

1991 Mathematics Subject Classification: Primary: 65M20. Secondary: 65M50.

1991 CR Categories: G.1.1, G.1.8

Keywords & Phrases: numerical mathematics, method of lines, local grid refinement, transport equations, advection schemes, air-pollution modeling

Note: This report will be submitted for publication elsewhere.

Note: The research reported belongs to the project EUSMOG which is carried out in cooperation with the Air Laboratory of the RIVM - The Dutch National Institute of Public Health and Environmental Protection. The RIVM is acknowledged for financial support.

1. Introduction

The subject of this paper is the numerical solution of the partial differential equation for linear advection of a scalar quantity ρ in a non-divergent velocity field $\underline{u}$, given by

$$\rho_t + \nabla \cdot (\underline{u}\rho) = 0. \tag{1.1}$$

Linear advection is an important (classical) problem in computational fluid dynamics and has been the subject of numerous investigations. The central theme is how to approximate the advection term $\nabla \cdot (\underline{u}\rho)$, such that the resulting errors in both phase and amplitude are minimized and the computational cost is still affordable. An important application we have in mind concerns atmospheric transport of chemical species. Then ρ represents a concentration

Report NM-R9309
ISSN 0169-0388
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

or density and $\underline{u}$ a wind field (in atmospheric transport modeling it is a standard rule to suppose $\underline{u}$ divergence free; see, e.g., Rood [10]). In addition to the usual accuracy and efficiency requirements, here the main consideration is that the transported concentrations must remain positive, because in actual applications also chemical reactions are modeled for which positivity is a prerequisite for avoiding non-physical chemical instabilities. We emphasize that the demand of positivity is important and that it severely restricts the choice of method, as it is essentially equivalent to the demand of avoiding numerical under- and overshoots in regions of strong variation.

The research objective of this paper is twofold. Our first objective is to examine a class of positive, finite-difference advection schemes which we consider promising for, a.o., atmospheric transport applications, and to select from this class the best possible candidate. We hereby follow the method-of-lines approach which means that the spatial discretization and temporal integration are considered separately. Our second objective is to combine this candidate scheme with a local uniform grid refinement technique (in two space dimensions). Local grid refinement is an attractive tool to save computational effort, as it offers the possibility to use fine grids in regions of strong variation (steep fronts, sharp peaks) and coarse grids elsewhere.

For the spatial discretization we confine ourselves to directionally-split stencils using 5 points per direction. We consider this a good starting point since a 5-point stencil is computationally attractive for the following reasons. First, a 5-point stencil is still relatively compact which is an advantage for implementing inflow and outflow boundaries. This is an important practical consideration when local uniform grid refinement is used, since local grid refinement introduces internal grid interfaces which have to be treated numerically as if they were domain boundaries. Second, a 5-point stencil allows orders of consistency up to 4 and comprises a number of potentially interesting spatial discretizations, viz. the 2-nd order central, the 2-nd order upwind, the 3-rd order upwind biased, and the 4-th order central discretization. In our investigation all four discretizations show up. We provide them with a flux-limiting procedure to enforce positivity. Our examination of positivity specifically involves a comparison between a variant of the well-known 3-rd order upwind ($\kappa = \frac{1}{3}$) discretization of Van Leer [7] (see also [4]) and the 4-th order central discretization, both limited in the same way. The derivation of the specific limiting procedure we use goes along the lines of Sweby's analysis [13]. This comparison leads to the conclusion that, for the case of pure advection, the limited 3-rd order discretization is to be preferred above the limited 4-th order one, although their performances are very much alike. However, in our tests the limited 3-rd order upwind discretization has turned out to perform notably better than the two limited 2-nd order discretizations.

For the time integration we confine ourselves to a number of explicit Runge-Kutta methods of orders of consistency two up to four. These methods are often used in the method of lines approach for solving hyperbolic partial differential equations. However, given a positive semi-discretization, stability of the time integration is in general not sufficient for maintaining positivity for the fully discrete solution. As a rule, the step size must satisfy an additional constraint which forces the admissible range of step size values to be smaller. Therefore, our focus is again on the positivity property, but now for the fully discrete solution where the limited 3-rd order upwind discretization is used for the spatial discretization. We have examined results from the linear theory of Bolley & Crouzeix [1] and the nonlinear theory of

Shu & Osher [11, 12] and Kraaijevanger [6]. Our tests indicate that for the current application both theories are of limited practical value. With regard to the demand of positivity, the integration methods tested turn out to behave almost equally and no best method could be identified.

For the local grid refinement we have borrowed the technique discussed by Trompert & Verwer [14, 15, 16]. In this technique the refinement takes place in a nested manner, from coarse to fine, in a way that coarse grids are used where the solution varies slowly and finer-and-finer subgrids are used in regions of rapid variation. These fine subgrids are adapted in time so as to move with regions of rapid variation. The aim of local refinement is to reduce the total number of grid points in space and time, so that a substantial reduction of CPU time and memory can be realized. The local refinement approach is therefore of great interest for atmospheric transport problems, as these problems are truly large scale (see McRae, Goodin & Seinfeld [9], Zlatev & Wasniewsky [19]).

Section 2 of this paper is devoted to the spatial discretization. Here we focus on a simple positivity analysis based on the idea of flux limiting and select the best spatial 5-point discretization from the ones discussed, which is the limited 3-rd order upwind discretization. 1D numerical test examples are used to illustrate our findings. Positivity of the time integration is discussed in Section 3. In Section 4 we formulate the 3-rd order upwind discretization for the 2D case and provide it with a simple boundary scheme for inflow and outflow boundaries. In Section 5 we present 2D numerical test examples for this discretization when combined with the classical 4-th order RK method for time integration. This combination is subsequently applied on locally refined grids in Section 6. In Section 7 we summarize our main conclusions.

2. The spatial discretization

The schemes are built from their one-space dimensional forms. Therefore, for most of the discussion it suffices to consider the constant coefficient 1D problem

$$\rho_t + f_x = 0, \quad f = u\rho, \quad u > 0, \tag{2.1}$$

which we spatially approximate, on the uniformly distributed grid points x_i, by the semi-discrete conservation form

$$\frac{d}{dt}w_i + \frac{F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}}{h} = 0. \tag{2.2}$$

Hence, $w_i(t)$ is a continuous time approximation to $\rho(x_i, t)$ at $x_i = ih$. We will interpret $w_i(t)$ as a point value in the finite-difference sense and we suppose a cell-vertex centered grid. $F_{i+\frac{1}{2}}$ is a numerical flux expression that will determine the actual semi-discretization. $F_{i+\frac{1}{2}}$ depends on neighboring values $f_j = uw_j$, such that it represents a consistent approximation to the true, analytical flux value at the cell center $x_{i+\frac{1}{2}} = (x_{i+1} + x_i)/2$. Throughout Section 2 we suppose u constant. Note that in 1D non-divergence implies a spatially constant velocity. Also note that the constant coefficient 1D formulations are extended in a straightforward manner to the 2D (and 3D) case, where the velocity can be both space and time dependent (cf. Section 4).

2.1. The 5-point discretizations

Numerous semi-discrete schemes can be brought in the conservation form (2.2). In this paper, we confine ourselves to discretizations on 5-point stencils for reasons outlined in the introduction. These are the 2-nd order central, the 2-nd order upwind, the 3-rd order upwind biased, and the 4-th order central discretization. Note that to obtain a higher order discretization that fits in (2.2), a wider stencil would be necessary. The first three discretizations mentioned above all belong to the κ-family that has been introduced by Van Leer for application to the nonlinear Euler equations (see [7] and the references therein). The numerical flux for the κ-scheme reads

$$F_{i+\frac{1}{2}} = f_i + \frac{1-\kappa}{4}(f_i - f_{i-1}) + \frac{1+\kappa}{4}(f_{i+1} - f_i), \tag{2.3}$$

where the values $\kappa = 1, -1$ and $\frac{1}{3}$ correspond with the 2-nd order central, the 2-nd order upwind, and the 3-rd order upwind biased discretization, respectively. Hence, for our purpose, this κ-formulation is very convenient. Note that for κ=1 a 3-point stencil suffices. However, for κ=1 the limited form needs a 5-point stencil too. In a similar way we can write the numerical flux for the 4-th order central scheme,

$$F_{i+\frac{1}{2}} = f_i + \frac{1}{12}(f_i - f_{i-1}) + \frac{1}{2}(f_{i+1} - f_i) - \frac{1}{12}(f_{i+2} - f_{i+1}). \tag{2.4}$$

Because later in the paper the 3-rd order scheme will play an important role, at this stage it is appropriate to recall its close resemblance with the 4-th order central one. Both fit in the form

$$\frac{d}{dt}w_i + \frac{f_{i-2} - 8f_{i-1} + 8f_{i+1} - f_{i+2}}{12h} = -\gamma\frac{1}{12}h^3\frac{f_{i-2} - 4f_{i-1} + 6f_i - 4f_{i+1} + f_{i+2}}{h^4}, \tag{2.5}$$

where γ=0 (central) or γ=1 (upwind). The right-hand side is the standard, central approximation to the fourth derivative, i.e.,

$$\frac{f_{i-2} - 4f_{i-1} + 6f_i - 4f_{i+1} + f_{i+2}}{h^4} = \frac{\partial^4 f}{\partial x^4}(x_i) + \frac{1}{60}h^2\frac{\partial^6 f}{\partial x^6}(x_i) + O(h^4), \tag{2.6}$$

Hence the upwind scheme is completely identical to the central one if the latter is applied to

$$\rho_t + f_x = -\frac{1}{12}h^3\frac{\partial^4 f}{\partial x^4} \tag{2.7}$$

and (2.6) is used for the 4-th derivative term. This term introduces dissipation due to the minus sign. So, from the central difference point of view, this term plays the role of artificial diffusion, entirely similar as in the case of the familiar 2-nd order central and 1-st order upwind scheme. The effect of the artificial diffusion term, i.e., the precise difference between central and upwind, is nicely illustrated from simple Fourier analysis. Introduce the trial function $w_i(t) = \rho_\omega(t)e^{\sigma\omega x_i}$, $\sigma = \sqrt{-1}$. We find

$$w_i(t) = \rho_\omega(0)e^{-\frac{1}{3}\gamma\mu(\cos\xi - 1)^2 t}e^{\sigma\omega(x - u(\xi)t)}, \tag{2.8}$$

where $\mu = u/h$, $\xi = \omega h$ and $u(\xi)$ is the numerical phase velocity given by

$$u(\xi) = \frac{\sin\xi(4 - \cos\xi)}{3\xi}u. \tag{2.9}$$

We see that both schemes generate the same dispersion errors because they have the same phase velocity. The only difference is the spurious dissipation term in the upwind case. This dissipation is largest for the shorter wavelengths, where also the dispersion error is maximal. Hence one can argue that the upwind scheme just damps away the short wavelength errors of the central scheme, in a manner prescribed by (2.8).

One might also argue that this is an advantage when solving pure advection problems, since no finite-difference method can resolve arbitrarily short wavelengths without excessive dispersion errors. Unfortunately, in spite of this spurious damping, the 3-rd order upwind scheme still suffers from under- and overshoot and lack of positivity in regions of truly strong variation. In fact, in this respect there appears to be little difference between all four schemes considered here. In applications merely smaller wiggles for the 3-rd order upwind scheme are observed, when compared with the other three ones. However, with regard to positivity all four fail.

2.2. Positive semi-discretizations

Scheme (2.2) is called positive (or non-negative), if for any non-negative initial solution $\{w_i(t_0)\}$ ($w_i(t_0) \geq 0 \quad \forall i$) the evolving solution $\{w_i(t)\}$ remains non-negative for all $t \geq t_0$. Obviously, a scheme is positive, if and only if for all i and all $t \geq t_0$,

$$w_i(t) = 0, \quad w_j(t) \geq 0, \quad \forall j \neq i \quad \Longrightarrow \quad \frac{d}{dt} w_i(t) \geq 0. \tag{2.10}$$

If we check the above four schemes for this condition, their lack of positivity immediately follows. Lack of positivity is of course intimately related to undershoot, but also to overshoot. This can be concluded from the following observation. Let α, β be arbitrary real constants and consider the linear transformation $w_i(t) = \alpha v_i(t) + \beta$. Suppose that $F_{i+\frac{1}{2}}$ satisfies the linear invariance property

$$F_{i+\frac{1}{2}}(\{w_i(t)\}) = \alpha F_{i+\frac{1}{2}}(\{v_i(t)\}) + \beta u. \tag{2.11}$$

Then $v_i(t)$ is also a solution of scheme (2.2), so that undershoot is equivalent to overshoot, simply because the graph of the solution of (2.2) can be folded around and shifted upward and downward in an arbitrary way, according to the transformation $w_i(t) = \alpha v_i(t) + \beta$. Note that with a divergence-free velocity field, the advection problem (1.1) itself is also linear invariant. Therefore, a positive scheme that satisfies (2.11) exhibits no under and overshoot.

To achieve positivity we apply flux limiting. Consider the general flux expression

$$F_{i+\frac{1}{2}} = f_i + \frac{1}{2}\phi_{i+\frac{1}{2}}(f_i - f_{i-1}) \tag{2.12}$$

with limiter ϕ. This limiter is supposed to be a nonlinear function of neighboring fluxes that defines a high order accurate scheme in smooth monotone regions of the solution, where no wiggles will arise, whereas in regions of sharp gradients the limiter must prevent wiggles and thus enforce monotonicity and positivity. This means that ϕ is to work as an intelligent nonlinear switch between a high order scheme and a low order, positive one. Note that for ϕ=0 the 1-st order upwind scheme is recovered, which is positive. Following Koren [4], we have adopted the limiting procedure that has been proposed by Sweby [13]. However,

other limiting procedures exist that can be followed too (see, e.g, Hirsch [3], LeVeque [8] and Zalesak [18]).

For the flux-limited form (2.12) it is straightforward to derive a sufficient condition on the limiter ϕ to guarantee positivity of the semi-discrete scheme (2.2). For (2.12) scheme (2.2) reads

$$\frac{d}{dt}w_i + \frac{(1+\frac{1}{2}\phi_{i+\frac{1}{2}})(f_i - f_{i-1}) - \frac{1}{2}\phi_{i-\frac{1}{2}}(f_{i-1} - f_{i-2})}{h} = 0. \tag{2.13}$$

Let

$$r_{i-\frac{1}{2}} = \frac{f_i - f_{i-1}}{f_{i-1} - f_{i-2}} \tag{2.14}$$

and assume $f_i - f_{i-1} \neq 0$, i.e. $r_{i-\frac{1}{2}} \neq 0$. Then (2.13) is identical to

$$\frac{d}{dt}w_i + \frac{1}{h}\left[(1+\frac{1}{2}\phi_{i+\frac{1}{2}}) - \frac{\frac{1}{2}\phi_{i-\frac{1}{2}}}{r_{i-\frac{1}{2}}}\right](f_i - f_{i-1}) = 0. \tag{2.15}$$

Next assume $r_{i-\frac{1}{2}}$=0. Then (2.13) is also identical to (2.15) if we assume, a priori, that $\phi_{i-\frac{1}{2}}$=0 if $r_{i-\frac{1}{2}}$=0 (both formulas then yield $\frac{d}{dt}w_i = 0$, which is sensible in this case). If we now apply the positivity rule (2.10) to (2.15), then we immediately conclude that the flux (2.12) will define a positive scheme if the bracketed term is non-negative. This is true if the limiter values $\phi_{i-\frac{1}{2}}$ and $\phi_{i+\frac{1}{2}}$ satisfy the inequality

$$\frac{\phi_{i-\frac{1}{2}}}{r_{i-\frac{1}{2}}} - \phi_{i+\frac{1}{2}} \leq 2. \tag{2.16}$$

If we next replace the above a priori assumption by the stronger assumption $\phi_{i-\frac{1}{2}} = 0$ if $r_{i-\frac{1}{2}} \leq 0$, and further suppose that always $\phi_{i-\frac{1}{2}}, \phi_{i+\frac{1}{2}} \geq 0$, then (2.16) is true if $\phi_{i-\frac{1}{2}} \leq 2r_{i-\frac{1}{2}}$.

To sum up, the general numerical flux (2.12) guarantees a positive semi-discrete solution, if the limiter ϕ satisfies the constraints

$$\phi_{i-\frac{1}{2}} = 0 \quad \text{if} \quad r_{i-\frac{1}{2}} \leq 0, \quad 0 \leq \phi_{i-\frac{1}{2}},\ \phi_{i+\frac{1}{2}} \leq \delta, \quad \phi_{i-\frac{1}{2}} \leq 2r_{i-\frac{1}{2}} \tag{2.17}$$

for any constant $\delta > 0$. This constant may serve as a parameter. If we take $\delta = 2$, and in addition suppose that $\phi_{i-\frac{1}{2}}$ and $\phi_{i+\frac{1}{2}}$ can be uniquely expressed as a function value of the respective slope ratios $r_{i-\frac{1}{2}}$ and $r_{i+\frac{1}{2}}$, then (2.17) defines the TVD region given in Figure 1(a) of Sweby [13] (for the Lax-Wendroff and Beam-Warming method). For the semi-discretization alone, however, we are free to choose any $\delta > 0$ for obtaining positivity and by increasing δ we can obtain more accuracy near peaks, see Figure 3.1. On the other hand, in Section 3 we will also support the choice $\delta = 2$ and henceforth assume that $\delta = 2$, unless noted otherwise.

2.3. The flux-limited schemes

We will associate (2.12) with the original higher order flux forms (2.3) and (2.4). First we rewrite (2.3) to the slope-ratio formulation

$$F_{i+\frac{1}{2}} = f_i + \frac{1}{2}K(r_{i+\frac{1}{2}})(f_i - f_{i-1}), \quad K(r) = \frac{1-\kappa}{2} + \frac{1+\kappa}{2}r \tag{2.18}$$

which fits in the general form (2.12). The next step is to limit $K(r)$ to some function $\phi(r)$ in such a way that the constraints (2.17) are satisfied for all possible values of the slope ratios, whereas for smooth monotone solutions, where $r \simeq 1$, (2.12) still takes the same values as (2.18). Following Koren [4], we define

$$\phi(r) = \max(0, \min(2r, \min(\delta, K(r)))), \quad \delta = 2 \tag{2.19}$$

This definition implies that the slope-ratio interval in which the limiter is switched off, is maximized. The motivation behind (2.19) is to use, as much as possible, the original higher order schemes and to limit only when really needed. However, as far as we know, a unique best choice for all sorts of solution profiles does not exist (see, e.g., LeVeque [8], Sect. 16.2, and Hirsch [3], Ch. 21, for other limiter definitions). Note that for (2.19) no limiting is needed in the interval $\frac{1}{4} \leq r \leq 2\frac{1}{2}$ for $\kappa = \frac{1}{3}$, in the interval $0 \leq r \leq 2$ for κ=1, and in the interval $\frac{1}{2} \leq r \leq \infty$ for $\kappa = -1$. For all other values of r limiting is necessary to satisfy the constraints (2.17). Note that for these values the limiter value $\phi(r)$ coincides with the upper boundary of the positivity region defined by (2.17).

In a similar vein the numerical flux (2.4) can be treated. We find the slope ratio formulation

$$F_{i+\frac{1}{2}} = f_i + \frac{1}{2} K(r_{i+\frac{1}{2}}, r_{i+\frac{3}{2}})(f_i - f_{i-1}), \quad K(r,s) = \frac{1}{6} + r - \frac{1}{6} rs \tag{2.20}$$

Note that now also the forward slope ratio $r_{i+\frac{3}{2}}$ is present. Initially we selected the limiter (2.19) without any modification ($K(r)$ replaced by $K(r,s)$). The corresponding region in (r,s)-space surrounding (r,s)=(1,1), where the original 4-th order scheme is applied, then turns out to be quite large. However, we found that for (2.20) a modification of (2.19) towards a smaller region leads to better results. This modified limiter is given by

$$\phi(r,s) = \max(0, \min(2r, \min(\delta, \min(\frac{1}{6} + r, \max(K(r,s), K(r,s_0)))))), \ \delta = s_0 = 2 \tag{2.21}$$

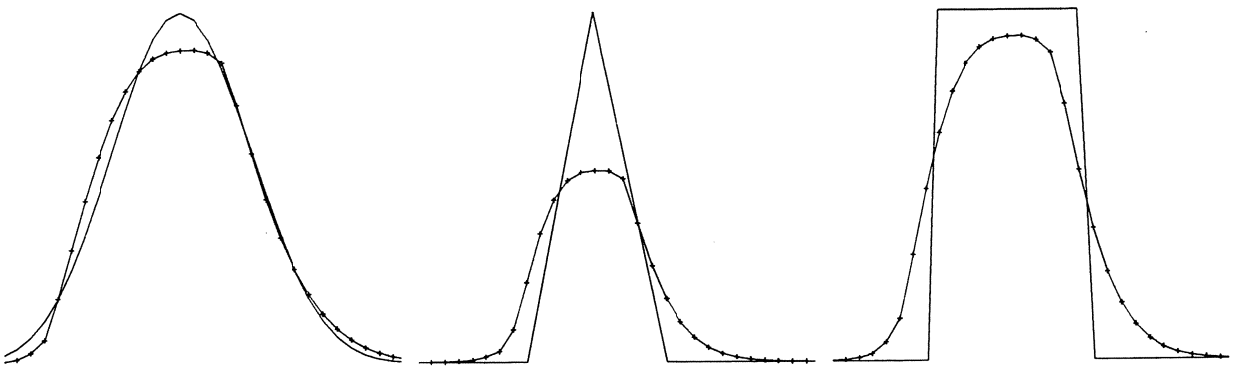


FIGURE 1. 1D-comparison of the four limited 5-point schemes, limited κ=1

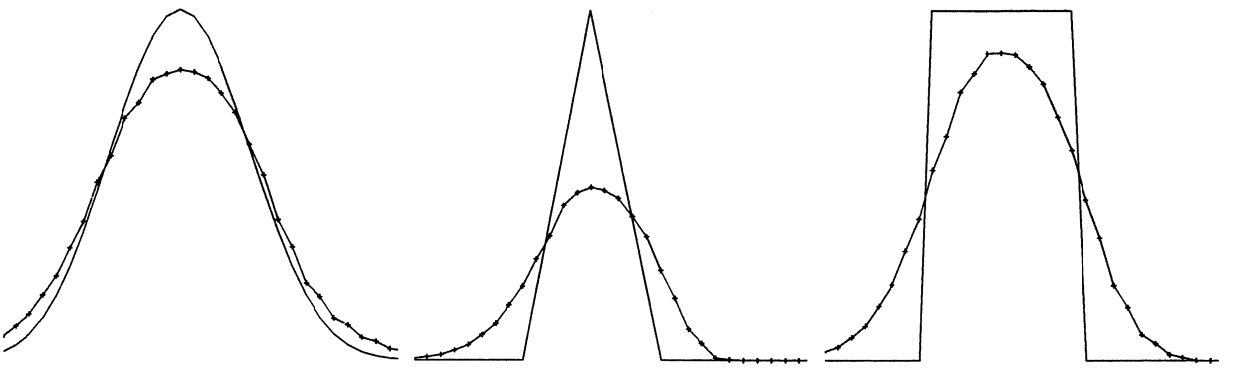


FIGURE 2. 1D-comparison of the four limited 5-point schemes, limited κ=-1

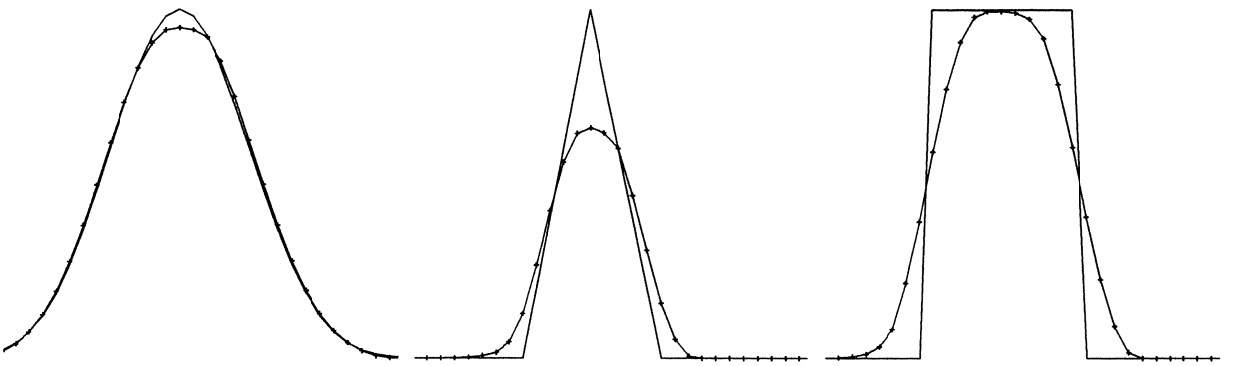


FIGURE 3. 1D-comparison of the four limited 5-point schemes, limited $\kappa = \frac{1}{3}$

To illustrate the effect of the limiting and the difference between the four limited schemes, Figures 3-4 show plots for three computed solution profiles. The true profiles are, in decreasing order of smoothness, the cosine hill $\rho(x,t) = \cos^{10}(x-t-0.5)$, the cone $\rho(x,t) = 1-10|x-t-0.5|$ for $|x-t-0.5| \leq 0.1$ and zero otherwise, and the square $\rho(x,t) = 1$ for $|x-t-0.5| \leq 0.1$ and zero otherwise. The computed profiles have been transported over 50 time steps, using u=1, h=0.02 and Courant number 1.0. The length of the domain is equal to 1 and periodic boundary conditions were used. The explicit 4-stage, 4-th order RK method is used for the time integration. We emphasize that in all cases the numerical errors are dominated by the spatial discretization (smaller step sizes for RK4 give the same plots).

The following observations can be made. First, the two limited 2-nd order schemes fall behind significantly. This indicates that the higher the order, the better the performance, even for discontinuous profiles like the cone and the square where the limiting is expected to dominate the numerical solution substantially. On the other hand, the results of the limited 3-rd and 4-th order scheme show a surprising resemblance. Here the anticipated advantage of the higher order of the central scheme is not borne out. With regard to the level of accuracy

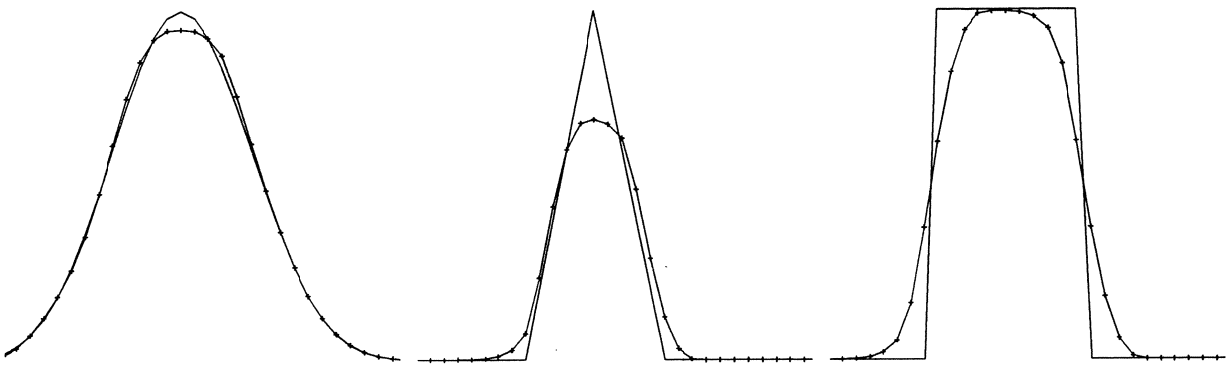


FIGURE 4. 1D-comparison of the four limited 5-point schemes, limited 4-th order central

of these two schemes we notice that they are positive and free of under and overshoot, but at the cost of some smearing for the square wave and quite some clipping for the cone. However, their profiles have not become distorted and the two schemes show very good phase characteristics. If we would halve the mesh width, their accuracy would significantly increase.

Recall that clipping at peaks is a difficulty inherent to any grid based scheme. The limiting, however, makes it worse since at a peak a negative slope ratio is encountered. But even without limiting the upwind scheme suffers from quite some clipping due to its built-in dissipation. With regard to peak preservation the (unlimited) central 4-th order scheme performs significantly better.

The 1D tests indicate that the limited 3-rd order upwind scheme is the most promising one from the four schemes discussed here. In the above experiments its accuracy appeared to be the same as that of the limited 4-th order scheme, but the upwind scheme is slightly cheaper and can be equipped with a simpler inflow/outflow boundary scheme.

3. POSITIVITY OF THE TIME INTEGRATION.

3.1. Preliminaries

In this section we shall discuss the question which explicit Runge-Kutta method can be used efficiently for the semi-discrete system (2.2) with limited upwind fluxes defined by (2.12),(2.19). The main criteria for this are *accuracy* and *positivity* : for reasonable Courant numbers $\nu = |u|\tau/h$ the temporal error should not influence the total error too much and the solutions should remain nonnegative.

If the semi-discrete system is written as

$$\frac{d}{dt}w(t) = g(t, w(t)) \tag{3.1}$$

with vector valued w and g, consecutive approximations $w^n \approx w(t_n)$ at the time levels $t_n = t_0 + n\tau$, $n = 1, 2, \cdots$ are found by computing in each step internal vectors W_i and their

function values $G_i = g(t_{n-1} + \tau\, c_i\,, W_i)$ according to

$$W_i = w^{n-1} + \tau \sum_{j=1}^{i-1} a_{ij}\, G_j\,, \quad i = 1, 2, ..., s, \tag{3.2}$$

followed by

$$w^n = w^{n-1} + \tau \sum_{i=1}^{s} b_i\, G_i\,. \tag{3.3}$$

The method is thus determined by the real coefficients a_{ij}, b_i, c_i and the number of stages s. It can be compactly represented by the array

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array}$$

with lower triangular matrix $A = (a_{ij})$ and with $b = (b_i)$, $c = (c_i)$.

Numerical tests have been carried out on the 1D periodic problem from Section 2 and on a 2D problem, for several methods with $s = 2$, 3 and 4. The methods have order p equal to s, see for instance [2], Sections II.1, II.4, and are given by the following arrays.

$$\begin{array}{c|cc} 0 & & \\ 1/2 & 1/2 & \\ \hline & 0 & 1 \end{array}$$

RK2a

$$\begin{array}{c|cc} 0 & & \\ 1 & 1 & \\ \hline & 1/2 & 1/2 \end{array}$$

RK2b

$$\begin{array}{c|ccc} 0 & & & \\ 1/3 & 1/3 & & \\ 2/3 & 0 & 2/3 & \\ \hline & 1/4 & 0 & 3/4 \end{array}$$

RK3a

$$\begin{array}{c|ccc} 0 & & & \\ 1 & 1 & & \\ 1/2 & 1/4 & 1/4 & \\ \hline & 1/6 & 1/6 & 2/3 \end{array}$$

RK3b

$$\begin{array}{c|cccc} 0 & & & & \\ 1/2 & 1/2 & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \\ \hline & 1/6 & 1/3 & 1/3 & 1/6 \end{array}$$

RK4

The two 2-stage methods are identical for linear problems. The same holds for the two 3-stage methods. Differences in the results are therefore caused by nonlinear phenomena. Note that the semi-discrete system obtained with limiting is highly nonlinear.

We found experimentally that for the *unlimited fluxes*, for which the semi-discrete system is linear, we have stability in 1D for Courant numbers

$$\nu \leq 0.87 \quad \text{for RK2a,b,} \qquad \nu \leq 1.62 \quad \text{for RK3a,b,} \qquad \nu \leq 1.74 \quad \text{for RK4.}$$

For the *limited fluxes* the stability bounds were found to be approximately

$$\nu \leq 1 \quad \text{for RK2a,b,} \qquad \nu \leq 1.25 \quad \text{for RK3a,b,} \qquad \nu \leq 1.4 \quad \text{for RK4.}$$

These values for the limited scheme are only approximately correct since the limited schemes show no very clear-cut transition from small errors to overflow, see the results for minimal values with the 4-stage method given in Table 2.

3.2. Positivity in time

In this subsection we shall briefly discuss some linear and nonlinear theoretical results on positivity. These will be compared with experimental results in the next subsection.

The semi-discrete system (2.15) can be written as

$$\frac{d}{dt} w_i = \gamma_i(w)\,(w_{i-1} - w_i) \tag{3.4}$$

with

$$\gamma_i(w) = \frac{u}{h}(1 + \frac{1}{2}\phi_{i+1/2} - \frac{1}{2r_{i-1/2}}\phi_{i-1/2}). \tag{3.5}$$

It is easily verified that (2.17) implies

$$0 \leq \gamma_i(w) \leq \frac{1}{h}(1 + \delta/2). \tag{3.6}$$

Applying the forward Euler method (RK1) to the system (3.4) gives

$$w_i^{n+1} = w_i^n + \tau\gamma_i(w^n)(w_{i-1}^n - w_i^n) \tag{3.7}$$

and from (3.6) it follows directly that positivity is guaranteed under the condition

$$\nu \leq \nu_0 = \frac{1}{1 + \delta/2}. \tag{3.8}$$

Theoretical bounds which guarantee positivity for higher order methods can be obtained by following the approaches of Shu and Osher [11] on diminution of total variation (TVD) and of Kraaijevanger [6] on contractivity. In the first approach all stages of the Runge-Kutta method are written as convex combinations of forward Euler type steps : introducing

$$\alpha_{ij} \geq 0, \quad \sum_{j=1}^{i-1} \alpha_{ij} = 1, \quad \text{for } i = 2, ..., s+1, \tag{3.9}$$

the method can be written as

$$W_1 = w^{n-1}, \quad W_i = \sum_{j=1}^{i-1}(\alpha_{ij}W_j + \tau\beta_{ij}G_j), \quad i = 2, 3, ..., s+1,$$

$$w^n = W_{s+1}$$

with coefficients

$$\beta_{ij} = a_{ij} - \sum_{k=j+1}^{i-1} \alpha_{ik} a_{kj}, \quad a_{s+1,j} := b_j. \tag{3.10}$$

If all $\beta_{ij} \geq 0$ it can be shown, just as for Euler's method, that we have positivity for Courant numbers

$$\nu \leq \nu_0 \min_{1 \leq j < i \leq s+1} \alpha_{ij}/\beta_{ij}.$$

Here ν_0 is the threshold value for Euler's method, and $\alpha_{ij}/\beta_{ij} = +\infty$ in case $\beta_{ij} = 0$.

Since contractivity results can also be obtained this way, even for all stages, it follows from Theorem 4.2 in [6] that in order to have all $\beta_{ij} \geq 0$ and $\alpha_{ij}/\beta_{ij} > 0$ it is necessary that

$$a_{ij} > 0, \quad b_i > 0 \quad \text{for all } i = 1, 2, ..., s, \quad j = 1, ..., i-1.$$

This condition is not satisfied for the methods RK2a, RK3a and RK4. For the remaining methods RK2b, RK3b it is easy to see that we can achieve the minimum of α_{ij}/β_{ij} to be 1. From the linear results below it follows that this is optimal.

Summarizing, we thus have "nonlinear positivity" in 1D if

$$\nu \leq \begin{cases} \dfrac{1}{1+\delta/2} & \text{for RK1, RK2b, RK3b} \\ 0 & \text{for RK2a, RK3a, RK4} \end{cases} \tag{3.11}$$

The nonlinear results are based on worst case assumptions for all stages. If we assume that $\gamma_i(w)$ in (3.4) remains almost the same over the stages, the situation will probably be described more accurately by a linear theory. Therefore, consider the system with "frozen coefficients"

$$\frac{d}{dt} w_i = c_i (w_{i-1} - w_i), \qquad 0 \leq c_i \leq \frac{1}{h}(1 + \delta/2), \tag{3.12}$$

where $c_i = \gamma_i(w(t_n))$ for $t_{n-1} \leq t \leq t_n$. On this system we can apply the linear theory of Bolley and Crouzeix [1]. From their Theorem 2 it can be deduced that we will have positivity for (3.12) under the condition $\nu \leq \nu_0/C$ where ν_0 is the threshold for Euler's method and C is the largest nonnegative number such that the stability function and all its derivatives are nonnegative on the interval $[-C, 0]$. In [5], Theorem 2.2 it was shown that $C = 1$ for any method having order $p = s$. Hence for all methods considered in this section we get the same condition for "linear positivity", namely

$$\nu \leq \frac{1}{1+\delta/2} \quad \text{for RK1, RK2a,b, RK3a,b, RK4.} \tag{3.13}$$

For 2D problems theoretical bounds can be obtained in a similar way. If $u, v > 0$, for example, the semi discrete-system can be written as

$$\frac{d}{dt} w_{ij} = \gamma_{ij}(w)(w_{i-1,j} - w_{ij}) + \delta_{ij}(w)(w_{i,j-1} - w_{ij}), \tag{3.14}$$

see Section 4, and the same conditions (3.11),(3.13) as in 1D are obtained if we define $\nu = (|u| + |v|)\tau/h$.

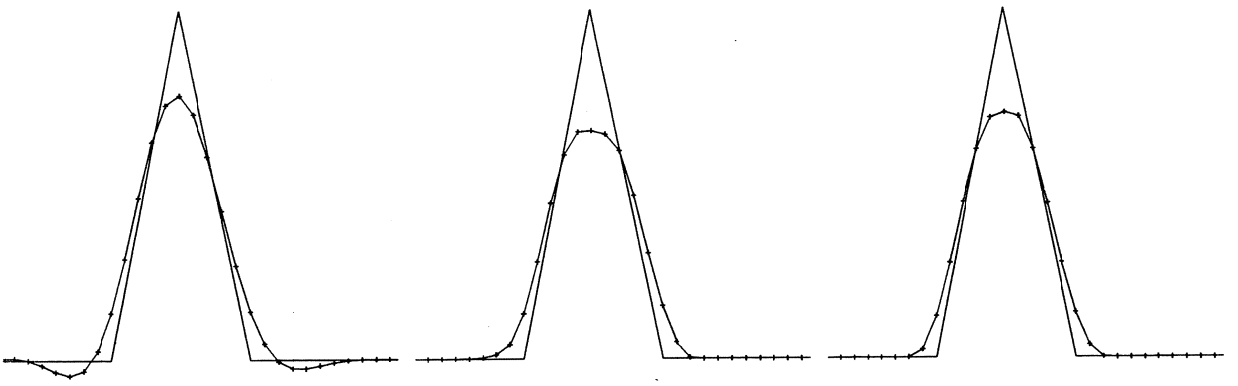


FIGURE 5. Solution for cone profile, $h = 1/50$

3.3. Tests on positivity

In this subsection some numerical tests on positivity will be given in 1D and 2D. The aim is to find out which Runge-Kutta methods are suitable to be combined with flux limiting and which value δ should be used in the limiter.

As said in Section 2, our choice will be $\delta = 2$. We note, however, that for the accuracy of the semi-discrete system it would be preferable to take a larger δ, since this means that the underlying third order scheme is used more often. Especially near peaks this gives a somewhat improved accuracy. In Figure 5 the solutions obtained with RK4 at $\nu = 1/2$ can be found for the cone profile : (a) without limiting, (b) limiting with $\delta = 2$, and (c) limiting with $\delta = 6$. The maximum errors are 0.24 for (a), 0.35 for (b) and 0.30 for (c).

The 1D experiments were carried out on the periodic problem after one "rotation" with the block and cone profile. Due to machine precision, the limiter does not completely avoid negative values (the limiter is not turned on exactly at the moment where it should). On a SUN SPARC workstation with double precision Fortran these negative values are of order of magnitude 10^{-17}. The criterion for positivity is therefore taken as $> -10^{-15}$.

The numerical results in Table 1 for the limiters with $\delta = 2$ and $\delta = 6$ have been obtained for the block profile with $h = 1/100$. The values ν given here are the maximal Courant numbers for which we found nonnegative numerical solutions at time $t = 1$. The cone profile and other h values give positivity for similar Courant numbers (almost the same with $\delta = 6$ and the 4-stage method, exactly the same in all other cases).

	$\delta = 2$	$\delta = 6$
RK2a,b	1	0.5
RK3a,b	0.79	0.39
RK4	1.37	0.78

TABLE 1. ν-values for positivity, 1D, $h = 1/100$

For $\delta = 6$ there is no very clear threshold for the 4-stage method. In the other cases a very distinctive threshold does exist. For example, for the 3- and 4-stage methods with $\delta = 2$ we

found the minima listed in Table 2.

$1/\tau$	RK3a	RK3b	$1/\tau$	RK4
122	$-.15E-02$	$-.32E-03$	69	$-.78E+03$
123	$-.10E-02$	$-.23E-03$	70	$-.98E+01$
124	$-.58E-03$	$-.11E-03$	71	$-.18E-01$
125	$-.15E-03$	$-.25E-04$	72	$-.19E-04$
126	$-.47E-18$	$-.71E-18$	73	$+.11E-06$
127	$-.91E-18$	$-.64E-18$	74	$+.34E-08$
128	$-.51E-18$	$-.80E-18$	75	$+.84E-10$
129	$-.15E-17$	$-.85E-18$	76	$-.20E-18$

TABLE 2. Minimal values for block profile, $h = 1/100$

From Table 1 it can be concluded that the limiter with $\delta = 6$ requires much smaller time steps to maintain positivity than its $\delta = 2$ counterpart. This cancels the better accuracy property of the $\delta = 6$–limiter : if we want to increase accuracy of the $\delta = 2$ results, while maintaining positivity, it turns out to be computationally cheaper to decrease h than to increase δ. For example, the accuracy for $\delta = 6, h = 1/50$ is comparable to the accuracy for $\delta = 2, h = 1/55$, and the latter case produces positive solutions with step size τ almost twice as large as the first case. For this reason, we shall further only consider the flux limiter with $\delta = 2$.

Concerning the choice between the various Runge-Kutta methods, the result of Table 1 is clearly in favor of the RK2 methods, which, after all, are twice as cheap as RK4. On the other hand, RK4 seems better than the RK3 methods. However, the differences vanish if we also take accuracy into account. In order to have a temporal error significantly smaller than the spatial error we should take approximately $\nu \leq 1$ for RK4 and $\nu \leq 1/2$ for the RK2 methods, see Table 3, whereas both RK3 methods give accurate results for Courant numbers equal to their positivity thresholds. Comparing RK4 with $\tau = h$, the two RK3 methods with $\tau = h/1.33$ and the RK2 methods with $\tau = h/2$ (equal work for all five methods), one observes that the methods give comparable errors, with a slight disadvantage for the 2-stage methods.

$1/\tau$	RK2a	RK2b	$1/\tau$	RK3a	RK3b	$1/\tau$	RK4
100	$.38E-1$	$.15E+0$	90	$.34E-1$	$.51E-1$	73	$.35E-1$
150	$.31E-1$	$.33E-1$	100	$.25E-1$	$.24E-1$	80	$.29E-1$
200	$.28E-1$	$.28E-1$	**133**	$.23E-1$	$.24E-1$	**100**	$.24E-1$
300	$.26E-1$	$.26E-1$	200	$.25E-1$	$.25E-1$	200	$.26E-1$

TABLE 3. L_2-errors for cone profile, $h = 1/100$

We did also consider the most simple time integration method RK1, the forward Euler method. The results for this scheme were excellent for the block profile and abominable for all other profiles. There is a very strong tendency to turn all profiles into blocks or staircases. This is caused by the fact that the underlying $\kappa = \frac{1}{3}$–scheme is unstable in combination with forward Euler. Consequently, the limiter is often turned on and all accuracy is lost. We note, however, that the solutions were positive at Courant numbers $\nu \leq 1/2$ for $\delta = 2$, and at $\nu \leq 1/4$ for $\delta = 6$, in agreement with the theoretical prediction (3.8).

It is clear that for the other Runge-Kutta methods neither criterion (3.11) nor (3.13) gives a good agreement with the experimentally found bounds of Table 1.

Since the 1D experiments were inconclusive for the choice of the Runge-Kutta method, the same positivity test has been performed in 2D with $\delta = 2$ and a constant velocity field $u = v = -1$ and a uniform grid with meshwidth $h = 1/50$ in both directions. The initial profile was chosen as the cylinder used in Example 5.1. The output point was $t = 0.25$. To our surprise the behaviour of the methods was different than in 1D, but again not in accordance with the theoretical conditions (3.11) or (3.13). The following Table 4 gives the Courant numbers $\nu = (|u| + |v|)\tau/h$ needed for positivity.

RK2a	RK2b	RK3a	RK3b	RK4
0.66	0.67	0.86	0.78	< 0.1

TABLE 4. ν-values for positivity, 2D, $h = 1/50$

For the 2- and 3-stage methods again a rapid transition from truly negative values to -10^{-18} was observed. However for RK4 the minima remain negative, though small in absolute value. In Table 5 the minimal values can be found for the 3- and 4-stage methods. The transition for the 2-stage methods is similar as for the 3-stage methods.

$1/\tau$	RK3a	RK3b	$1/\tau$	RK4
104	$-.84E-02$	$-.63E-02$	80	$-.49E-02$
112	$-.18E-02$	$-.46E-02$	120	$-.20E-03$
120	$-.77E-17$	$-.48E-03$	160	$-.18E-08$
128	$-.26E-17$	$-.41E-17$	200	$-.29E-06$
136	$-.20E-17$	$-.20E-17$	240	$-.14E-04$
144	$-.20E-17$	$-.17E-17$	280	$-.20E-05$

TABLE 5. Minimal values for cylinder, $h = 1/50$

The fact that RK4 fails to produce positive solutions for reasonable Courant numbers makes the method less suited than the others in case positivity and mass conservation are crucial. In such a situation the explicit trapezoidal rule RK2b with Courant restriction $\nu \leq 1/2$ can be recommended, due to its simplicity and the fact that it is supported by the nonlinear

theory. However, in most applications there will be a small background concentration, and then very little undershoot will not harm. This allows larger Courant numbers, in which case the higher order methods become more attractive, see Table 3. Since the difference between the various methods turned out to be small anyway, we shall consider further in this paper the well-known classical method RK4.

4. THE 2D FORMULATION OF THE ($\kappa = \frac{1}{3}$)-SCHEME.

We proceed with the 2D problem

$$\rho_t + (u\rho)_x + (v\rho)_y = 0 \tag{4.1}$$

where the divergence free velocity field $\underline{u} = (u, v)$ may now depend on all three coordinates x, y, t. To save space, however, we present the flux expressions only for the x-direction. First suppose $u(x, y, t) \geq 0$. We then replace (2.12) by

$$F_{i+\frac{1}{2}} = u_{i+\frac{1}{2}}(w_i + \frac{1}{2}\phi_{i+\frac{1}{2}}(w_i - w_{i-1})), \quad u_{i+\frac{1}{2}} = \frac{u_{i+1} + u_i}{2} \tag{4.2}$$

where $\phi_{i+\frac{1}{2}} = \phi(r_{i+\frac{1}{2}})$, with $r_{i+\frac{1}{2}} = \dfrac{w_{i+1} - w_i}{w_i - w_{i-1}}$, is the limiter value defined by the limiter function (2.19) with $K(r) = \frac{1+2r}{3}$. Hence, the only difference with (2.12) is the variable velocity $u_{i+\frac{1}{2}}$ in front of the bracketed solution dependent expression. The form (4.2) is sometimes called the state interpolation form.

An alternative is to keep the original form (2.12), by putting $f_i = u_i w_i$ (flux interpolation). It is not clear which form is to be preferred. In both cases the linear invariance property of the advection problem is lost. However, considering the semi-discrete system, when using state interpolation the linear transformation $w_i(t) = \alpha v_i(t) + \beta$ leaves this semi-discrete form unchanged, except for a remainder term which is just the 2-nd order central discretization of u_x:

$$\begin{aligned} \frac{d}{dt}v_i + \frac{1}{h}[\{u_{i+\frac{1}{2}}(v_i + \frac{1}{2}\phi_{i+\frac{1}{2}}(v_i - v_{i-1}))\} - \{u_{i-\frac{1}{2}}(v_{i-1} + \frac{1}{2}\phi_{i-\frac{1}{2}}(v_{i-1} - v_{i-2}))\}] \; + \\ + \quad \beta\alpha^{-1}\frac{u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}}{h} = 0. \end{aligned} \tag{4.3}$$

In 2D we expect this numerical divergence term to be small for a divergence free velocity field. Note that the slope ratios $r_{i+\frac{1}{2}}$, and hence the limiter values $\phi_{i+\frac{1}{2}}$, have not changed. For the original flux interpolation formula (2.12) the counterpart of (4.3) is more complicated, because for flux interpolation the slope ratio expressions do change under the linear transformation. In addition, in this case also the divergence term u_x is discretized by the upwind method and hence concentration dependent limiter values are introduced in the numerical divergence term, which is unnatural. On the other hand, a disadvantage of (4.2) is that the 3-rd order consistency is lost (in smooth monotone regions where $\phi_{i+\frac{1}{2}} = K(r_{i+\frac{1}{2}})$). This is illustrated by the modified equation of the state interpolation form which reads

$$\rho_t + (u\rho)_x = -\frac{1}{12}h^2(\rho_{xx}u_x + 2\rho u_{xxx} + 3\rho_x u_{xx}) + O(h^3) \tag{4.4}$$

For $u(x,y,t) < 0$ the counterpart of (4.2) is given by

$$F_{i+\frac{1}{2}} = u_{i+\frac{1}{2}}(w_{i+1} + \frac{1}{2}\phi_{i+\frac{1}{2}}(w_{i+1} - w_{i+2})), \quad u_{i+\frac{1}{2}} = \frac{u_{i+1} + u_i}{2} \tag{4.5}$$

where $\phi_{i+\frac{1}{2}} = \phi(\frac{1}{r_{i+\frac{3}{2}}})$. For arbitrary velocity $u = u(x,y,t)$ we then get the usual upwind form

$$F_{i+\frac{1}{2}} = \max(u_{i+\frac{1}{2}}, 0)F^+_{i+\frac{1}{2}} + \min(u_{i+\frac{1}{2}}, 0)F^-_{i+\frac{1}{2}} \tag{4.6}$$

with $F^+_{i+\frac{1}{2}}$ given by (4.2) and $F^-_{i+\frac{1}{2}}$ by (4.5). Recall that (4.6) comprises four different sign cases for the associated semi-discrete scheme (2.2). For all four cases positivity can be proved by a straightforward application of (2.10).

We conclude this section with a description of our implementation of inflow/outflow boundary conditions. We work with a vertex centered grid, so the location of a domain boundary always coincides with a grid point. Again it suffices to consider the 1D problem. Suppose that x_0 is the left boundary point. If $u_0 \geq 0$ we then have inflow, with given velocity and state, and otherwise outflow with a given velocity only. In case of inflow, scheme (2.2) is applied for $i \geq 1$, so that only for $i=1$ an auxiliary variable w_{-1} needs to be introduced for the flux computation $F_{\frac{1}{2}}$ defined by (4.2). We use the 2-nd order extrapolation $w_{-1} = \max(3w_0 - 3w_1 + w_2, 0)$ and note that this would result in the 2-nd order central discretization at $x = x_1$ if the limiting would be switched off and $u_{\frac{1}{2}} = u_{\frac{3}{2}}$. In the exceptional event of $u_0 \geq 0$ and $u_{\frac{1}{2}} < 0$, $F_{\frac{1}{2}}$ is computed by (4.5) where w_{-1} does not occur. Hence we then act as if we have an (outflow) Dirichlet condition. Next consider the outflow situation. Then scheme (2.2) is applied for $i \geq 0$ and an auxiliary flux computation $F_{-\frac{1}{2}}$ defined by (4.5) is introduced. $F_{-\frac{1}{2}}$ then uses the auxiliary state variable w_{-1} introduced above. The auxiliary velocity $u_{-\frac{1}{2}}$ is defined by the 2-nd order extrapolation $u_{-\frac{1}{2}} = \min(\frac{15u_0 - 10u_1 + 3u_2}{8}, 0)$. Assuming a constant velocity and no limiting, the outflow scheme defined this way is just 2-nd order upwind. In the exceptional event of $u_{\frac{1}{2}} > 0$ and $u_0 < 0$, $F_{\frac{1}{2}}$ is computed by (4.2) which then also uses the auxiliary variable w_{-1}.

5. Numerical 2D examples.

In this section we apply the positive upwind - RK4 advection scheme to three two-space dimensional example problems.

5.1. Example 1

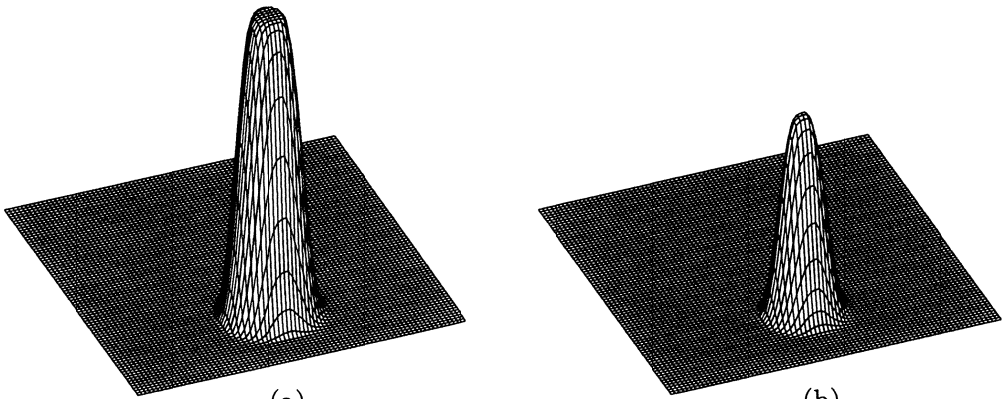
Our first example is concerned with a standard test used by many authors, the so-called Molenkamp-Crowley test or solid body rotation. In equation (4.1) we let $0 \leq x, y \leq 1$ and put $u(x,y,t) = 2\pi(y - 0.5)$, $v(x,y,t) = -2\pi(x - 0.5)$. Note that u is constant in the x-direction and v is constant in the y-direction. For any given function Φ, the solution can be expressed as $\rho(x,y,t) = \Phi(X,Y)$, where

$$X = \cos(2\pi t)(x - 0.5) - \sin(2\pi t)(y - 0.5), \quad Y = \sin(2\pi t)(x - 0.5) + \cos(2\pi t)(y - 0.5) \tag{5.1}$$

Hence $\Phi(X,Y)$ rotates with period 1 around $(\frac{1}{2}, \frac{1}{2})$ in the clockwise direction. For $\Phi(X,Y)$ we have made two choices, viz. a cylinder and a cone with height 1 and radius 0.1, both

centered at $(\frac{1}{2}, \frac{3}{4})$ at $t = 0$. One full rotation has been carried out for both solutions on the uniform grid having 80×80 grid cells, using step size $\tau = \frac{h}{3}$ for the RK4 method (this corresponds roughly with a maximal Courant number 1.0, with the Courant number in 2D defined as on page 15).

The computed solutions at $t = 1$ are shown in Figure 6. Contour lines are shown in Figure 7. We see that the computed solutions are positive and that their profile is accurately centered around the center point $(\frac{1}{2}, \frac{3}{4})$. We consider the accuracy of the cylinder computation very satisfactory (maximum value is 0.999). For the cone we again observe the clipping problem already mentioned in Section 2. Note, however, that in 2D the clipping will be even stronger than in 1D, since the clipping occurs once for every grid line lying under the cone. The position of the top of the computed cone coincides with the center point $(\frac{1}{2}, \frac{3}{4})$, but the maximum value 1.0 has decreased to 0.66. The cone obviously needs a much finer grid.



(a) (b)

FIGURE 6. Computed profiles for the rotational flow field of Example 1
(a) Cylinder and (b) Cone.

5.2. Example 2

In Example 1 there is no real inflow and outflow, since the solution is zero in a neighborhood of the boundary. To test the boundary scheme, we have carried out a semi-rotation with the cylinder on the 80×80 grid around the center point $(\frac{1}{2}, 0)$ for the wind field and starting with the lower boundary point $(\frac{1}{4}, 0)$ as center point for the cylinder. At $t = 0$ we then have a real inflow at $(\frac{1}{4}, 0)$ and at $t = \frac{1}{2}$ a real outflow at $(\frac{3}{4}, 0)$. In this case the step size $\tau = \frac{h}{6}$ corresponds with a maximal Courant number 1.0. Inspection of Figure 8 shows that the boundary scheme works accurately for this example.

5.3. Example 3

In the present example a deformational flow field is used. Like in Example 1 we place the cylinder at $(\frac{1}{2}, \frac{3}{4})$, but now choose the velocities as $u(x, y, t) = x - \frac{1}{2}$ and $v(x, y, t) = -(y - \frac{1}{2})$. For $t > 0$ this results in a stretching of the cylinder along the x-axis into a 'ellipsoidal cylinder', while it moves downward. Then, at $t = 0.5$ the velocities are reversed, which results in a

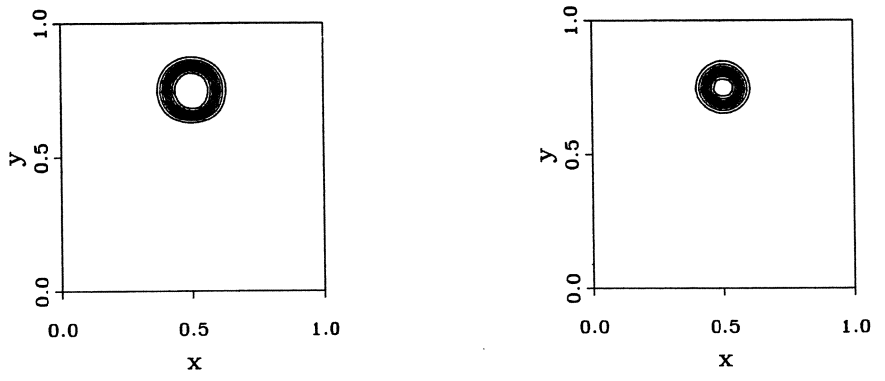


FIGURE 7.
Contour lines for the rotational flow field tests of Example 1 (a) Cylinder and (b) Cone.

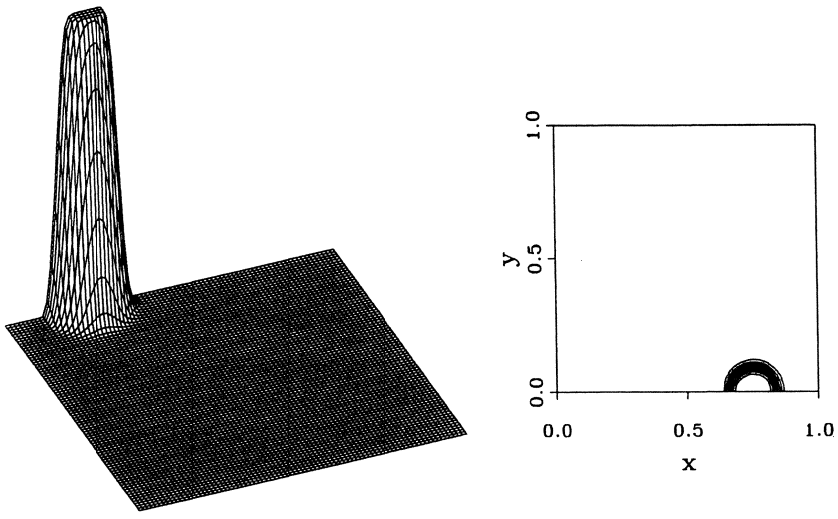


FIGURE 8. The computed profile and contour lines
for the semirotation of Example 2.

restoration of the cylinder at its initial position at $t = 1.0$. Figure 9 shows profiles and contour lines for the computed solution for the 80×80 grid. The step size for RK4 is $\tau = \frac{1}{80}$. The figure reveals that Example 3 is more difficult than the two previous ones. The effect of the deformational flow field is visible. We cannot explain the small cusp. It might have been caused by the sign change in u at $x = \frac{1}{2}$.

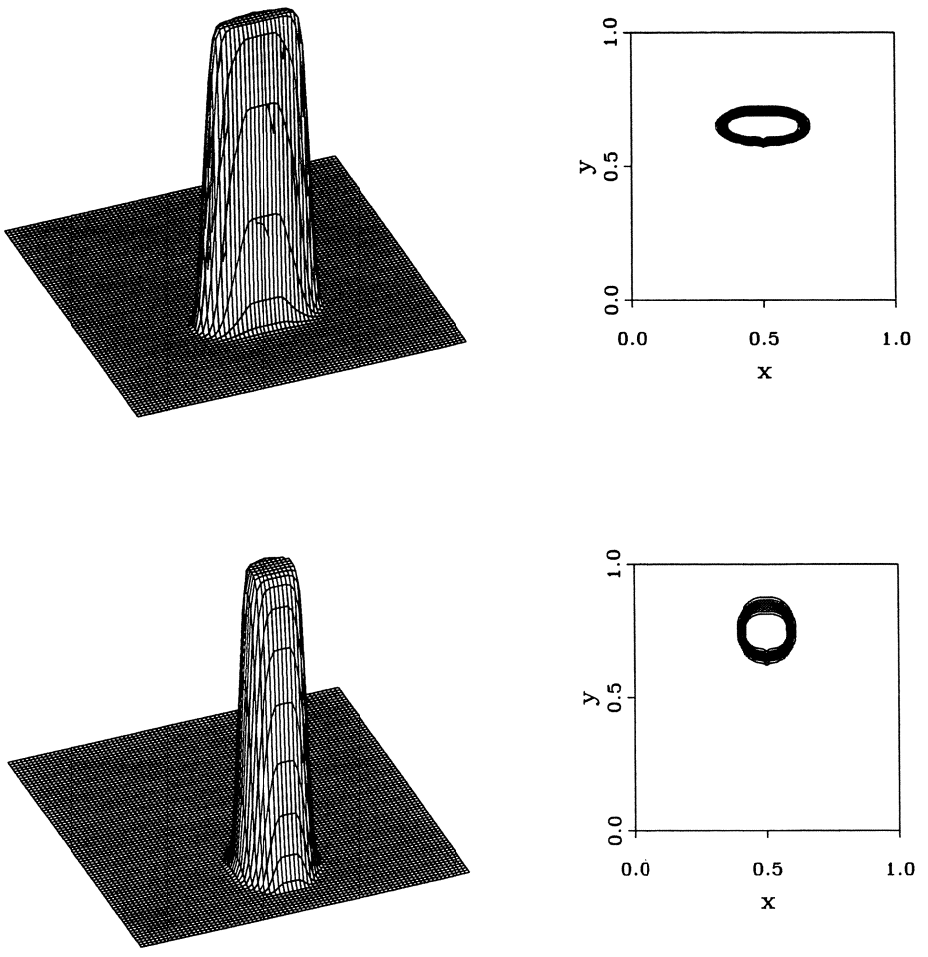


FIGURE 9. Computed profiles and contour lines for the deformational flow field of Example 3 at $t = 0.5$ and $t = 1$.

6. LOCAL REFINEMENT

The purpose of this section is to illustrate once more the positive upwind - RK4 advection scheme, but now applied on locally refined grids. The aim of local refinement is to speed up computations for solutions which require locally a very fine grid, of course under the constraint that the overall accuracy of a corresponding global uniform grid will not be diminished. Our numerical examples serve to show that for the advection problem this goal can be achieved in an effective way. We emphasize that the refinement technique we have used has not been tuned for pure advection problems. The technique is flexible in the sense that it can be used for a wide variety of discretization methods and PDEs. This flexibility is important because our final aim is to solve, in an operator splitting setting, large systems of atmospheric transport problems which also contain diffusion, chemical reaction, and source and sink terms. The refinement technique is borrowed from Trompert & Verwer [14], where it was used for

parabolic equations.

6.1. Outline

Although the local refinement technique is conceptually simple, the implementation involves too many details to discuss here at length. We therefore confine ourselves to an outline and refer to [14] for more details. For clarity we first outline a single level of refinement, i.e., we

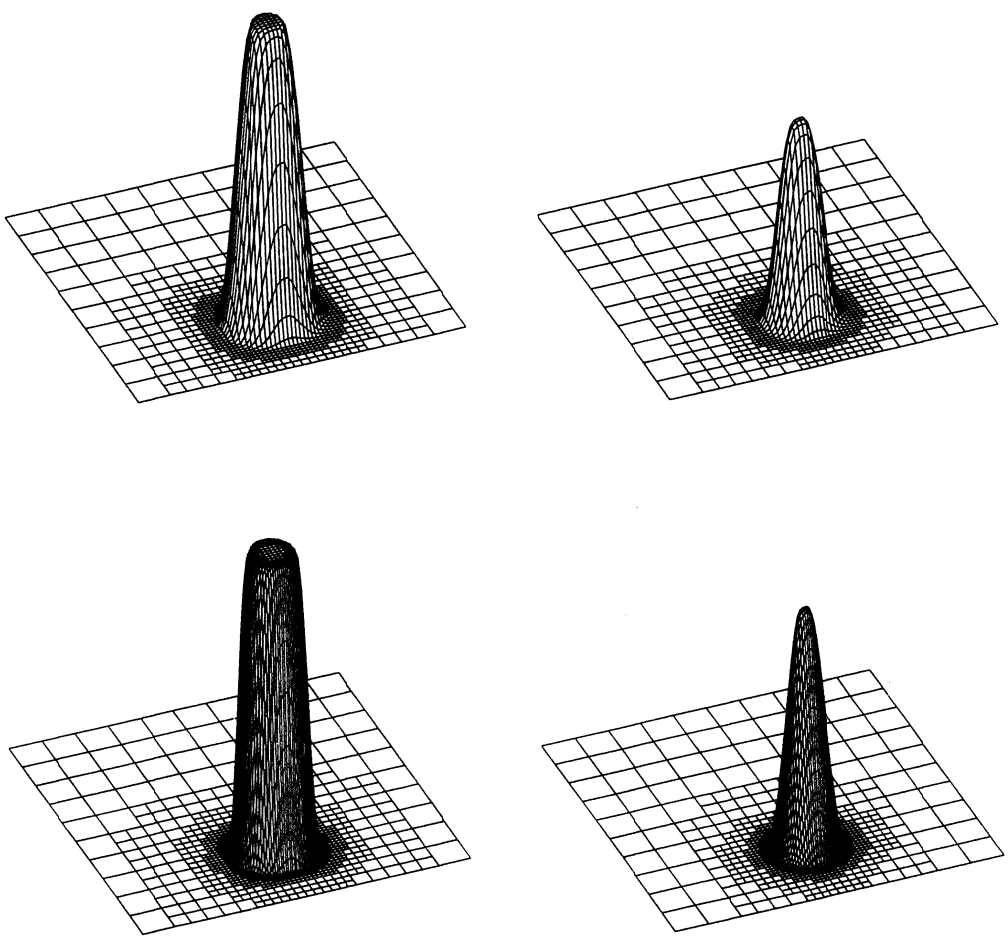


FIGURE 10. Computed profiles for the rotational flow field of Example 1.

consider a coarse base grid which is refined once. The following steps are followed to advance from a time point T to the next time point $T + \Delta t$:

(1) Integration on the coarse grid using a step size Δt. Call the computed values on the coarse grid at time $T + \Delta t$ the new coarse w-values and those at time T the old coarse

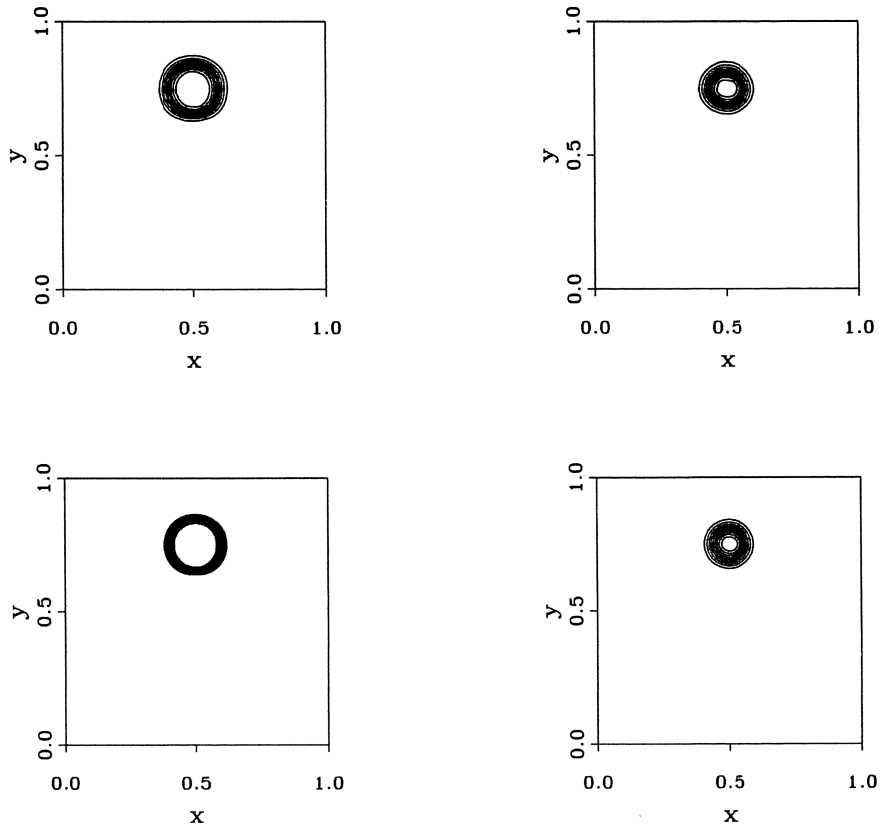


FIGURE 11.
Contour lines for the rotational flow field tests of Example 1. Cylinder at the left and cone at the right. The upper and lower plots belong, respectively, to the 4 and 5-level computation.

w-values.

(2) The coarse grid integration is followed by adaptation of the fine grid. Using the new coarse w-values, it is decided where the new fine grid will be for $T \leq t \leq T + \Delta t$. This is done by invoking a spatial local error indicator and a clustering and buffering algorithm to distribute all untolerable cells over the new fine grid. The fine grid may consist of different disjunct, fine subgrids. The actual refinement is cellular and carried out by bisecting all sides of untolerable coarse grid cells. At this point the non-refined part of the coarse grid is complete and not further processed within the current time interval $T \leq t \leq T + \Delta t$.

(3) The fine grid adaptation is followed by interpolation. We first return to time level T and determine initial values for the new fine grid. If a cell was refined in the previous time interval $T - \Delta t \leq t \leq T$, then we use the available fine grid w-values and interpolation is not needed. If a cell was not refined before, then we interpolate old coarse w-values.

(4) For $T \leq t \leq T + \Delta t$ we also need to specify boundary values for certain intermediate time points at grid interfaces where fine grid cells abut on coarse cells. These values are prescribed via interpolation and serve as Dirichlet values for the fine grid computation. If an interface coincides with the domain boundary, then boundary conditions are used and interpolation is omitted.

(5) Next we integrate on the fine grid over the interval $T \leq t \leq T + \Delta t$, using either Δt as step size or a smaller value. After this, the new fine grid w-values are injected in the coinciding coarse grid points and all fine grid w-values at $T + \Delta t$ are saved for use in the next time step, as indicated in (3). The solution at time $T + \Delta t$ is now complete.

Multiple levels are handled in a natural, recursive fashion. After each time step of (5) a regridding may take place resulting in a grid of one refinement level higher. In fact, the computational steps follow precisely points (2) - (5) above. Note that all fine grid results at forward time levels are kept in storage and that for step continuation the most accurate results are used that are available. We also note that to keep the solution positive, monotone interpolation is required. In the experiments below linear interpolation is used. The spatial error indicator is based on the curvature and is, together with the actual decision criteria for refinement, similar as in [14].

Different from [14] is that we have used a fixed number of levels, as well as a constant step size. Also different from [14] is that this step size is taken equal for all levels, which implies that the finest level dictates the step size with regard to stability. Using the same step size enables us to interpolate in (4) boundary values at grid interfaces for the intermediate RK stages from the corresponding coarse grid intermediate stages (cf. [15, 16]). The use of larger step sizes at coarser grids, such that at each grid level the same Courant condition is satisfied, will probably be more efficient for pure advection problems. This is not clear when the refinement technique will be used in an operator splitting setting where also other integrations will play a role. Therefore, for the sake of simplicity of presentation, here we keep the step size constant in time and constant over all grid levels.

To measure the gain in efficiency, one may count the total number of nodal points used, accumulated over all levels and all time steps, and compare this total number with its counterpart for the computation at the finest globally uniform grid. We then neglect all overhead costs of the local refinement. For scalar advection problems this is misleading. However, for large systems of atmospheric transport problems this measure is very realistic, since for such systems the nodalwise, nonlinear chemistry computations are already so expensive that this overhead will form a negligible part of the total costs. Note that in this comparison we tacitly assume that the local refinement does not diminish the accuracy of the corresponding global uniform computation. For the experiments below we have observed an average efficiency ratio of approximately 4. For the solution shapes tested, this gain in efficiency is what can be expected in 2D (cf. [17], Sect. 5.2).

Finally we note that a realistic atmospheric transport problem will consist of many components, say about 25 to 50 [19]. For such large systems it may prove useful not to work with a single set of refined grids for all components, but to distinguish between subsets of components for the grid adaptation, so as to achieve a further gain in efficiency. This will lead to extra overhead for additional interpolation tasks, but only for the chemistry computation as the coupling in the system occurs only through the chemistry terms. Hence, the use of more than one set of refined grids will not affect the advection computation.

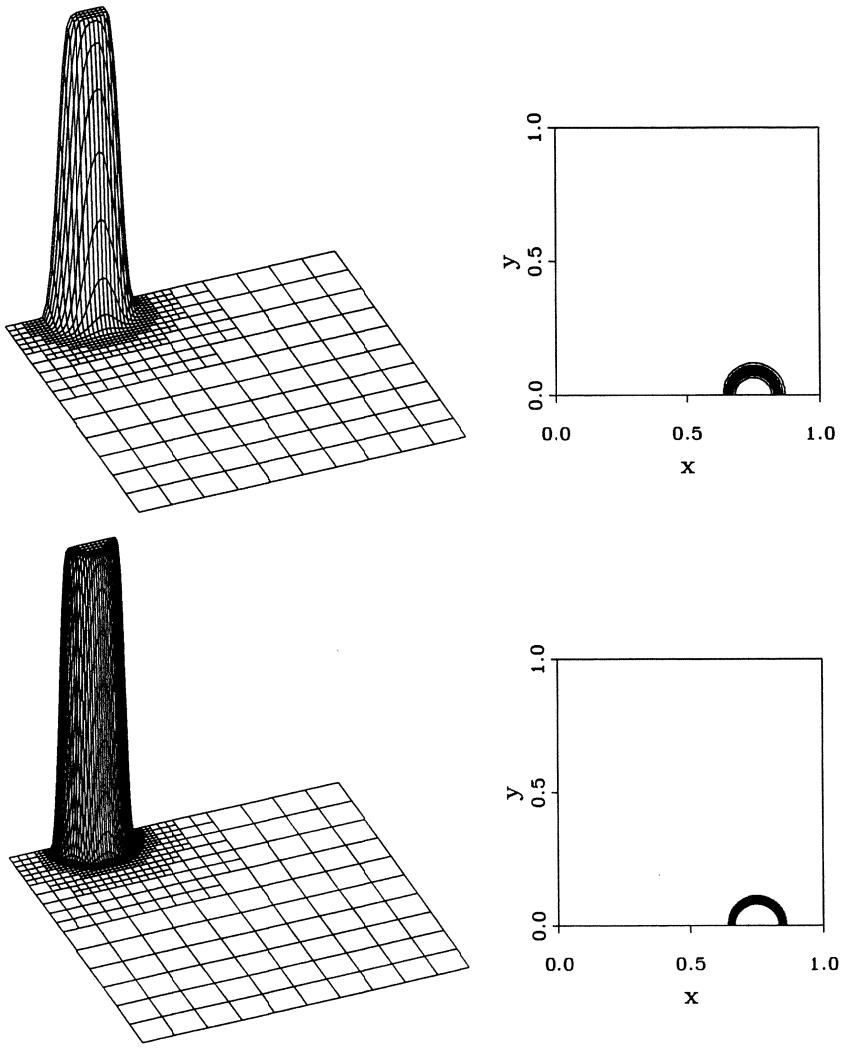


FIGURE 12.
The computed profiles and corresponding contour lines for the semirotation of Example 2.

6.2. Numerical examples

We present results for all three example problems from Section 5 for a 4-level and a 5-level grid of which the coarse base grid is 10×10. Hence the smallest grid size in the 4-level computation is $\frac{1}{80}$ and in this case the step size of the RK4 method is equal to the step sizes taken in Section 5. The accuracies of the 4-level computations should therefore compare with those for the corresponding 80×80 grid computations from Section 5. The Figures 10-14 are the counterparts of Figures 6-9. The figures reveal the virtue of local refinement convincingly. The accuracy of the 4-level results indeed compares very well. Moreover, all 5-level results, where the smallest grid size is $\frac{1}{160}$ and the step size τ is also a factor 2 smaller, show a notable improvement in accuracy.

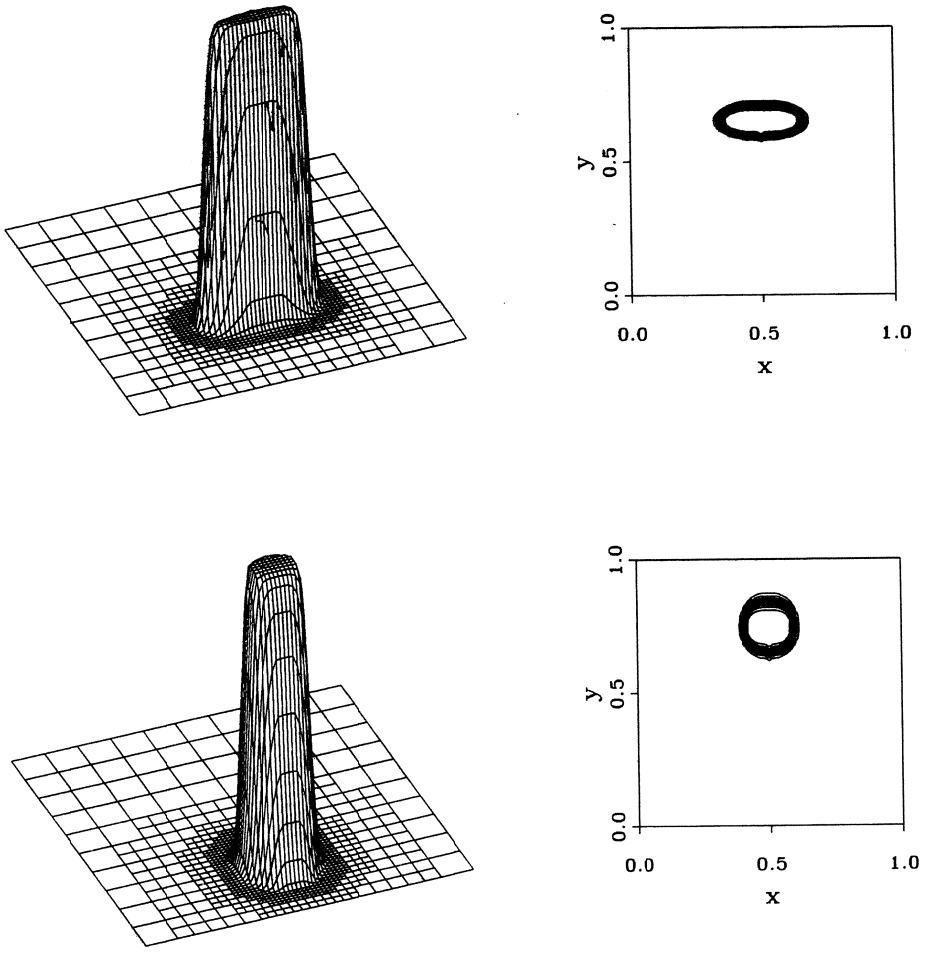


FIGURE 13. The computed profiles and contour lines for the deformational flow field of Example 3 (4 levels).

7. SUMMARY OF MAIN CONCLUSIONS

For the directionally-split spatial discretization we have examined four 5-point discretizations in conservation form, viz. the 2-nd order central, the 2-nd order upwind, the 3-rd order upwind biased, and the 4-th order central discretization. The first three schemes are well-known members of the family of κ-schemes. Positivity is achieved by flux limiting, using (2.19) for the three κ-schemes and (2.21) for the 4-th order scheme. The limited 3-rd and 4-th order discretization perform equally well and clearly outperform the two limited 2-nd order ones. For general use we recommend the 3-rd order discretization limited by (2.19). This combination possesses very good shape preserving properties, also in 2D. The state-interpolation 2D combination, provided with a boundary scheme for inflow/outflow boundaries, has been described in Section 4.

For the time integration we have examined a number of explicit RK methods, viz. the 2-nd

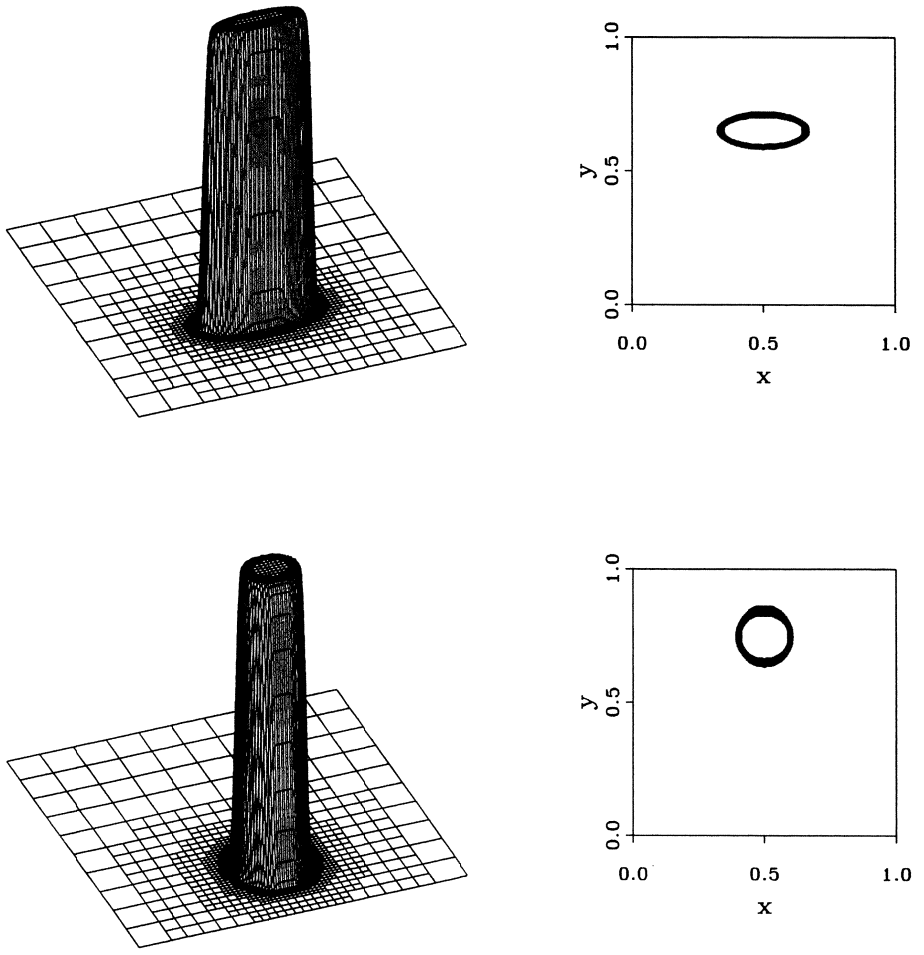


FIGURE 14.
The computed profiles and contour lines for the deformational flow field of Example 3 (5 levels).

order method of Runge, the 2-nd order explicit trapezoidal rule, the 3-rd order methods of Heun and Fehlberg, and the classical 4-th order method. We have tested analytical results on positivity from the linear theory of [1] and the nonlinear theory of [11, 12]. Our tests indicate that for the current application both theories are of limited practical value. With regard to positivity, all methods tested turn out to behave almost equally and no truly best method could be identified. For example, we have not found a notable difference in accuracy/efficiency performance between the 2-nd order explicit trapezoidal rule, which fits in the nonlinear theory, and the classical 4-th order explicit method, which does not fit. The behaviour of the combined spatial-temporal scheme apparently is dominated by the spatial discretization where the limiting procedure plays a decisive role.

The results shown in Section 6 on local refinement support our view that for large-scale

atmospheric flow problems this technique may help significantly to reduce CPU time, of course without diminishing accuracy. For the examples shown an average reduction factor of about 4 was observed.

REFERENCES

[1] C. Bolley and M. Crouzeix. Conservation de la positivité lors de la discrétisation des problèmes d'évolution paraboliques. *R.A.I.R.O Analyse Numérique*, 12: 237 – 245, 1978.

[2] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving ordinary differential equations I – nonstiff problems.* Springer Series in Computational Mathematics 8. Springer–Verlag, Berlin, 1987.

[3] Ch. Hirsch. *Numerical computation of internal and external flows, Volume 2.* Wiley, Chichester, 1990.

[4] B. Koren. A robust upwind discretization method for advection, diffusion and source terms. In C.B. Vreugdenhil and B. Koren, editors, *Notes on Numerical Fluid Mechanics.* Vieweg, Braunschweig, to appear.

[5] J.F.B.M. Kraaijevanger. Absolute monotonicity of polynomials occurring in the numerical solution of initial value probems. *Numer. Math.*, 48: 303 – 322, 1986.

[6] J.F.B.M. Kraaijevanger. Contractivity of Runge–Kutta methods. *BIT*, 31: 482 – 528, 1991.

[7] B. van Leer. Upwind–difference methods for aerodynamic problems governed by the Euler equations. In B.E. Engquist, S. Osher, and R.C.J. Somerville, editors, *Large–scale computations in fluid mechanics*, pages 327 – 336. AMS Series, American Mathematical Society, Providence, RI, 1985.

[8] R.J. LeVeque. *Numerical methods for conservation laws.* Lecture Notes in Mathematics ETH Zűrich. Birkhäuser Verlag, Basel, 1992.

[9] G.J. McRae, W.R. Goodin, and J.H. Seinfeld. Numerical solution of the atmospheric diffusion equation for chemically reacting flows. *J. Comput. Phys.*, 45: 1 – 42, 1982.

[10] R.B. Rood. Numerical advection algorithms and their role in atmospheric transport and chemistry models. *Reviews of Geophysics*, 25: 71 – 100, 1987.

[11] C.-W. Shu and S. Osher. Efficient implementation of essentially non–oscillatory shock–capturing schemes. *J. Comput. Phys.*, 77: 439 – 471, 1988.

[12] C.-W. Shu and S. Osher. Efficient implementation of essentially non–oscillatory shock–capturing schemes II. *J. Comput. Phys.*, 83: 32 – 78, 1989.

[13] P.K. Sweby. High resolution schemes using flux–limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.*, 21: 995 – 1011, 1984.

[14] R.A. Trompert and J.G. Verwer. A static–regridding method for two–dimensional parabolic partial differential equations. *Appl. Numer. Math.*, 8: 65 – 90, 1991.

[15] R.A. Trompert and J.G. Verwer. Analysis of the implicit Euler local uniform grid refinement method. *SIAM J. Sci. Comput.*, 14: 259 – 278, 1993.

[16] R.A. Trompert and J.G. Verwer. Runge–Kutta methods and local uniform grid refinement. *Math. Comp.*, 1993. to appear.

[17] J.G. Verwer and R.A. Trompert. Analysis of local uniform grid refinement. Report NM–R9211, CWI, Amsterdam, 1992. to appear in *Appl. Numer. Math.*

[18] S.T. Zalesak. A preliminary comparison of modern shock–capturing schemes: linear advection. In R. Vichnevetsky and R.S. Stepleman, editors, *Advances in Computer Methods for Partial Differential Equations VI.*, pages 15 – 22. IMACS, 1987.

[19] Z. Zlatev and J. Wasniewski. Large scale computations in air pollution modelling. Report UNIC–92–05, Scientific Computing Group, Danmarks EDB–Center for Forskning og Uddannelse, 1992.