JOURNAL OF COMPUTATIONAL PHYSICS 117, 35–46 (1995)

A Positive Finite-Difference Advection Scheme*

W. HUNDSDORFER, B. KOREN, M. VAN LOON, AND J. G. VERWER

CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

Received May 17, 1993; revised July 12, 1994

This paper examines a class of explicit finite-difference advection schemes derived along the method of lines. An important application field is large-scale atmospheric transport. The paper therefore focuses on the demand of positivity. For the spatial discretization, attention is confined to conservative schemes using five points per direction. The fourth-order central scheme and the family of κ-schemes, comprising the second-order central, the second-order upwind, and the third-order upwind biased, are studied. Positivity is enforced through flux limiting. It is concluded that the limited third-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 two up to four. With regard to the demand of positivity, these integration methods turn out to behave almost equally and no best method could be identified.

1. INTRODUCTION

The subject of this paper is the numerical solution of the partial differential equation for linear advection of a scalar quantity w in an arbitrary velocity field $\mathbf{u}$, given by

$$w_t + \nabla \cdot (\mathbf{u}w) = 0. \tag{1}$$

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 $\nabla \cdot (\mathbf{u}w)$, 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 w represents a concentration or density and $\mathbf{u}$ a wind field. 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 to examine a class of positive, finite-difference advection schemes which we consider promising for 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.

For the spatial discretization we confine ourselves to stencils using five points per (spatial) 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 boundary conditions. Second, a 5-point stencil allows orders of consistency up to 4 and comprises a number of potentially interesting spatial discretizations, viz. the second-order central, the second-order upwind, the third-order upwind biased, and the fourth-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 third-order upwind ($\kappa = \frac{1}{3}$) discretization of Van Leer [8] (see also [5]) and the fourth-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 [14].

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 third-order upwind discretization is used for the spatial discretization. Both theoretical and experimental results are presented.

The paper contributes to the state-of-the-art in higher-order

* The research reported belongs to the projects EUSMOG and CIRK which are 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.

0021-9991/95 $6.00

TVD schemes (as reviewed in the introductory section of [10]), by the application of a new limiter function, the presentation of a monotone, fourth-order central scheme for advection, and the investigation of theoretical monotonicity bounds for the time steps to be applied in explicit Runge–Kutta-type calculations.

The paper is organized as follows. Section 2 describes the spatial discretization and the flux limiting in 1D. Positivity of the time integration is discussed in Section 3. The 2D discretization is formulated in Section 4. In Section 5, 2D numerical test examples are presented. In Section 6 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,

$$w_t + f_x = 0, \quad f = uw, \quad u > 0, \tag{2}$$

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

$$\frac{d}{dt} w_i + \frac{F_{i+1/2} - F_{i-1/2}}{h} = 0. \tag{3}$$

Hence, $w_i(t)$ is a continuous time approximation to $w(x_i, t)$ at $x_i = ih$. We interpret $w_i(t)$ as a point value in the finite-difference sense and we suppose a cell–vertex centered grid. $F_{i+1/2}$ is a numerical flux expression that determines the actual semi-discretization. $F_{i+1/2}$ 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, $x_{i+1/2} = (x_{i+1} + x_i)/2$. Throughout Section 2 we suppose u to be constant. 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 (3). In this paper, we confine ourselves to discretizations on 5-point stencils for the reasons outlined in the Introduction. These are the second-order central, the second-order upwind, the third-order upwind biased, and the fourth-order central discretization. Note that to obtain a higher order discretization that fits in (3), 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 [8] and the references therein). The numerical flux for the κ-scheme reads

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

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

$$F_{i+1/2} = f_i + \tfrac{1}{12}(f_i - f_{i-1}) + \tfrac{1}{2}(f_{i+1} - f_i) - \tfrac{1}{12}(f_{i+2} - f_{i+1}). \tag{5}$$

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

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

where $\gamma = 0$ (fourth-order central) or $\gamma = 1$ (third-order upwind biased). The right-hand side is the standard, central approximation to the fourth-order derivative, i.e.,

$$\frac{f_{i-2} - 4f_{i-1} + 6f_i - 4f_{i+1} + f_{i+2}}{h^4} = \frac{\partial^4 f}{\partial x^4}(x_i) + \frac{1}{60} h^2 \frac{\partial^6 f}{\partial x^6}(x_i) + O(h^4). \tag{7}$$

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

$$w_t + f_x = -\frac{1}{12} h^3 \frac{\partial^4 f}{\partial x^4}, \tag{8}$$

and if (7) is used for the fourth-order 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 to the case of the familiar second-order central and first-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. We introduce the trial function $w_i(t) = w_\omega(t) e^{\sigma \omega x_i}$, $\sigma = \sqrt{-1}$. We find

$$w_i(t) = w_\omega(0) e^{-(1/3)\gamma\mu(\cos\xi - 1)^2 t} e^{\sigma\omega(x - u(\xi)t)}, \quad x = x_i, \tag{9}$$

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

$$u(\xi) = \frac{\sin \xi(4 - \cos \xi)}{3\xi} u. \tag{10}$$

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 the short wavelength errors of the central scheme, in a manner prescribed by (9).

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, despite this spurious damping, the third-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 third-order upwind scheme are observed, when compared with the other three. However, with regard to positivity, all four fail.

2.2. *Positive Semi-discretizations*

Scheme (3) is called positive (or non-negative), if for any non-negative initial solution $\{w_i(t_0)\}$ ($w_i(t_0) \geq 0$, $\forall i$) 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 $t \geq t_0$,

$$w_i(t) = 0, \quad w_j(t) \geq 0, \quad \forall j \neq i \Rightarrow \frac{d}{dt} w_i(t) \geq 0. \tag{11}$$

If we check the above four schemes for this condition, their lack of positivity follows immediately. Lack of positivity is of course intimately related to undershoot and 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 $F_{i+1/2}$ satisfies the linear invariance property

$$F_{i+1/2}(\{w_i(t)\}) = \alpha F_{i+1/2}(\{v_i(t)\}) + \beta u. \tag{12}$$

Then $v_i(t)$ is also a solution of scheme (3), so that undershoot is equivalent to overshoot, simply because the graph of the solution of (3) can be folded around and shifted upward and downward in an arbitrary way, according to the transformation $w_i(t) = \alpha v_i(t) + \beta$. For constant velocity u the limiter (introduced below) does not affect this property. Note that with a divergence-free velocity field, the advection problem (1) itself is also linearly invariant. Therefore, a positive scheme that satisfies (12) exhibits no under- and overshoots.

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

$$F_{i+1/2} = f_i + \tfrac{1}{2}\phi_{i+1/2}(f_i - f_{i-1}), \tag{13}$$

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 a nonlinear switch between a high order scheme and a low order, positive one. Note that for $\phi = 0$ the first-order upwind scheme is recovered, which is positive. Following Koren [5], we have adopted the limiting procedure that has been proposed by Sweby [14]. However, other limiting procedures exist that can be followed, too (see, e.g., Hirsch [3], LeVeque [9], and Zalesak [16]).

For the flux-limited form (13) it is straightforward to derive a sufficient condition on the limiter ϕ to guarantee positivity of the semi-discrete scheme [3]. For (13) scheme (3) reads

$$\frac{d}{dt} w_i + \frac{(1 + (1/2)\phi_{i+1/2})(f_i - f_{i-1}) - (1/2)\phi_{i-1/2}(f_{i-1} - f_{i-2})}{h} = 0. \tag{14}$$

Let

$$r_{i-1/2} = \frac{f_i - f_{i-1}}{f_{i-1} - f_{i-2}} \tag{15}$$

and assume that $f_i - f_{i-1} \neq 0$, i.e., $r_{i-1/2} \neq 0$. Then (14) is identical to

$$\frac{d}{dt} w_i + \frac{1}{h}\left[\left(1 + \frac{1}{2}\phi_{i+1/2}\right) - \frac{(1/2)\phi_{i-1/2}}{r_{i-1/2}}\right](f_i - f_{i-1}) = 0. \tag{16}$$

Next, assume that $r_{i-1/2} = 0$. Then (14) is also identical to (16) if we assume, a priori, that $\phi_{i-1/2} = 0$ if $r_{i-1/2} = 0$ (both formulae then yield $(d/dt)\, w_i = 0$, which is sensible in this case). If we now apply the positivity rule (11) to (16), then we immediately conclude that the flux (13) will define a positive scheme if the bracketed term is non-negative. This is true if the limiter values $\phi_{i-1/2}$ and $\phi_{i+1/2}$ satisfy the inequality

$$\frac{\phi_{i-1/2}}{r_{i-1/2}} - \phi_{i+1/2} \leq 2. \tag{17}$$

If we next replace the above a priori assumption by the stronger assumption $\phi_{i-1/2} = 0$ if $r_{i-1/2} \leq 0$, and further suppose that always $\phi_{i-1/2}$, $\phi_{i+1/2} \geq 0$, then (17) is true if $\phi_{i-1/2} \leq 2r_{i-1/2}$.

To sum up, the general numerical flux (13) guarantees a positive semi-discrete solution, if the limiter ϕ satisfies the constraints

$$\phi_{i-1/2} = 0 \quad \text{if } r_{i-1/2} \leq 0,$$
$$0 \leq \phi_{i-1/2}, \phi_{i+1/2} \leq \delta, \quad \phi_{i-1/2} \leq 2r_{i-1/2}, \tag{18}$$

for any constant $\delta > 0$. This constant may serve as a parameter. If we take $\delta = 2$ and, in addition, if we suppose that $\phi_{i-1/2}$ and $\phi_{i+1/2}$ can be uniquely expressed as a function value of the respective slope ratios $r_{i-1/2}$ and $r_{i+1/2}$, then (18) defines the TVD region given in Fig. 1a of Sweby [14] (for the Lax–Wendroff and Beam–Warming methods). 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 Fig. 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 (13) with the original higher order flux forms (4) and (5). First we rewrite (4) to the slope-ratio formulation

$$F_{i+1/2} = f_i + \frac{1}{2} K(r_{i+1/2})(f_i - f_{i-1}),$$
$$K(r) = \frac{1-\kappa}{2} + \frac{1+\kappa}{2} r, \tag{19}$$

which fits in the general form (13). The next step is to limit $K(r)$ to some function $\phi(r)$ in such a way that the constraints (18) are satisfied for all possible values of the slope ratios, whereas for smooth monotone solutions, where $r \simeq 1$, (13) still takes the same values as (19). Following Koren [5], we define

$$\phi(r) = \max(0, \min(2r, \min(\delta, K(r)))), \quad \delta = 2. \tag{20}$$

This definition implies that the slope-ratio interval in which the limiter is switched off, is maximized. The motivation behind (20) is to use, as much as possible, the original higher order schemes and to limit them 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 [9, Section 16.2] and Hirsch [3, Chap. 21] for other limiter definitions). Note that for (20) 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 $\kappa = 1$, 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 (18). Note that for these values the limiter value $\phi(r)$ coincides with the upper boundary of the positivity region defined by (18).

In a similar vein the numerical flux (5) can be treated. We find the slope-ratio formulation

$$F_{i+1/2} = f_i + \tfrac{1}{2} K(r_{i+1/2}, r_{i+3/2})(f_i - f_{i-1}),$$
$$K(r, s) = \tfrac{1}{6} + r - \tfrac{1}{6} rs. \tag{21}$$

Note that now also the forward slope-ratio $r_{i+3/2}$ is present. Initially we selected the limiter (20) 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 fourth-order scheme is applied, then turns out to be quite large. However, we found that for (21), a modification of (20) towards a smaller region leads to better results. This modified limiter is given by

$$\phi(r, s) = \max(0, \min(2r, \min(\delta, \min(\tfrac{1}{6} + r, \max(K(r, s), K(r, s_0)))))),$$
$$\delta = s_0 = 2. \tag{22}$$

To illustrate the effect of the limiting and the difference between the four limited schemes, in [4] we show numerical results for three 1D solution profiles (a cosine hill, a cone, and a square). From these results it appears that the two limited second-order schemes fall behind significantly, indicating 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 third- and fourth-order schemes show a surprising resemblance, but the anticipated advantage of the higher order of the central scheme is not borne out. The 1D tests indicate that the limited third-order upwind scheme is the most promising one from the four schemes discussed here. In the above experiments its accuracy appears to be the same as that of the limited fourth-order scheme, but the upwind scheme is slightly cheaper and can be equipped with a similar inflow/outflow boundary scheme.

3. POSITIVITY OF THE TIME INTEGRATION

3.1. *Preliminaries*

In this section we discuss the question which explicit Runge–Kutta method can be used efficiently for the semi-discrete system (3) with limited upwind fluxes defined by (13), (20). 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

$$\frac{d}{dt} w(t) = g(t, w(t)), \tag{23}$$

with vector-valued w and g, consecutive approximations $w^n \approx w(t_n)$ at the time levels $t_n = t_0 + n\tau$, $n = 1, 2, \ldots$ 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

$$W_i = w^{n-1} + \tau \sum_{j=1}^{i-1} a_{ij} G_j, \quad i = 1, 2, \ldots, s, \tag{24}$$

followed by

$$w^n = w^{n-1} + \tau \sum_{i=1}^{s} b_i G_i. \tag{25}$$

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 [4] 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:

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

RK2a RK2b

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

RK3a 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 non-linear phenomena. Note that the semi-discrete system obtained with limiting is highly non-linear.

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

$$\nu \leq 0.87 \text{ for RK2a,b}, \quad \nu \leq 1.62 \text{ for RK3a,b},$$

$$\nu \leq 1.74 \text{ for RK4}.$$

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

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

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.

3.2. *Positivity in Time*

In this subsection we shall briefly discuss some linear and non-linear theoretical results on positivity. These will be compared with experimental results in the next subsection. The semi-discrete system (16) can be written as

$$\frac{d}{dt} w_i = \gamma_i(w)(w_{i-1} - w_i), \tag{26}$$

with

$$\gamma_i(w) = \frac{u}{h}\left(1 + \frac{1}{2}\phi_{i+1/2} - \frac{1}{2r_{i-1/2}}\phi_{i-1/2}\right). \tag{27}$$

It is easily verified that (18) implies

$$0 \leq \gamma_i(w) \leq \frac{u}{h}(1 + \delta/2). \tag{28}$$

Applying the forward Euler method (RK1) to the system (26) gives

$$w_i^{n+1} = w_i^n + \tau\gamma_i(w^n)(w_{i-1}^n - w_i^n), \tag{29}$$

and from (28) it follows directly that positivity is guaranteed under the condition

$$\nu \leq \nu_0 = \frac{1}{1 + \delta/2}. \tag{30}$$

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 [7] on contractivity. In the first approach all stages of the Runge–Kutta method are written as convex combinations of forward Euler type steps. Introducing

$$\alpha_{ij} \geq 0, \quad \sum_{j=1}^{i-1} \alpha_{ij} = 1, \quad \text{for } i = 2, \ldots, s+1, \tag{31}$$

the method can be written as

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

$$i = 2, 3, ..., s+1, \quad w^n = W_{s+1},$$

with coefficients

$$\beta_{ij} = a_{ij} - \sum_{k=j+1}^{i-1} \alpha_{ik} a_{kj}, \quad \alpha_{s+1,j} := b_j. \tag{32}$$

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 [7] that in order to have all $\beta_{ij} \geq 0$ and $\alpha_{ij}/\beta_{ij} > 0$ it is necessary that

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

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

$$\nu \leq \begin{cases} \dfrac{1}{1+\delta/2} & \text{for RK1, RK2b, RK3b,} \\ 0 & \text{for RK2a, RK3a, RK4.} \end{cases} \tag{33}$$

The nonlinear results are based on worst-case assumptions for all stages. If we assume that $\gamma_i(w)$ in (26) 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"

$$\frac{d}{dt} w_i = c_i (w_{i-1} - w_i), \quad 0 \leq c_i \leq \frac{u}{h}(1 + \delta/2), \tag{34}$$

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 (34) 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 [6, 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

$$\nu \leq \frac{1}{1+\delta/2} \quad \text{for RK1, RK2a,b, RK3a,b, RK4.} \tag{35}$$

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

$$\frac{d}{dt} w_{ij} = \gamma_{ij}(w)(w_{i-1,j} - w_{ij}) + \delta_{ij}(w)(w_{i,j-1} - w_{ij}), \tag{36}$$

see Section 4, and the same conditions (33), (35) as in 1D are obtained if we define

$$\nu = (|u| + |v|)\tau/h. \tag{37}$$

3.3. *Tests on Positivity*

In this subsection some numerical tests on positivity are 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 of δ should be used in the limiter.

As said in Section 2, our choice is $\delta = 2$. We note, however,

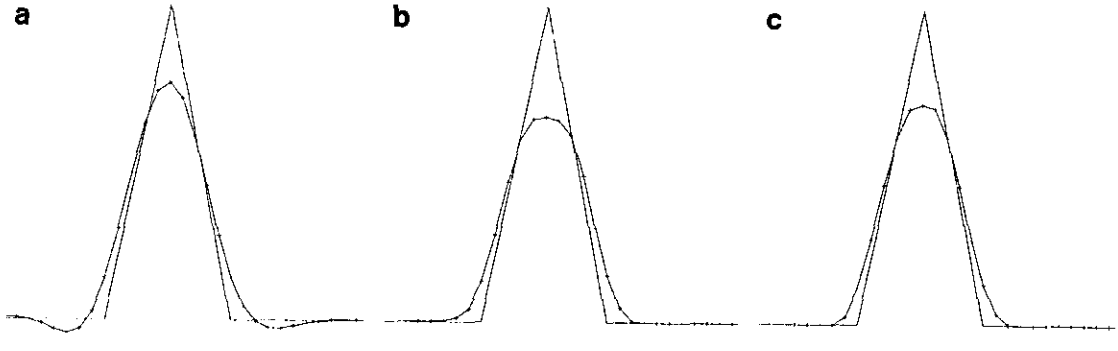

FIG. 1. Solution for cone profile, $h = \frac{1}{50}$.

TABLE I

ν-Values for Positivity, 1D, $h = \frac{1}{100}$

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

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 Fig. 1 the solutions obtained with RK4 at $\nu = \frac{1}{2}$ can be found for the 1D, periodic cone problem from [4]: (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). More 1D experiments are carried out 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 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 I for the limiters with $\delta = 2$ and $\delta = 6$ have been obtained for the block profile with $h = \frac{1}{100}$. The values ν given here are the maximal Courant numbers for which we find 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).

For $\delta = 6$ there is no very clear threshold for the 4-stage method. In the other cases a very distinctive threshold does exist. From Table I 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 $\delta = 6$, $h = \frac{1}{50}$ is comparable to the accuracy for $\delta = 2$, $h = \frac{1}{55}$, and the latter case produces positive solutions with step size τ almost twice as large as the first case. For this reason, in the remainder we only consider the flux limiter with $\delta = 2$.

Concerning the choice between the various Runge–Kutta methods, the result of Table I 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 into account accuracy. In order to have a temporal error significantly smaller than the spatial error we should take approximately $\nu \leq 1$ for RK4 and $\nu \leq \frac{1}{2}$ for the RK2 methods, see Table II, 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 $\tau = h/2$ (the same amount of work for all five methods), one observes that the methods give comparable errors, with a slight disadvantage for the 2-stage methods.

For completeness we still consider the simplest time integration method, RK1, the forward Euler method. The results obtained by this scheme are excellent for the block profile, but 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 $\kappa = \frac{1}{3}$-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 are positive at Courant numbers $\nu \leq \frac{1}{2}$ for $\delta = 2$, and at $\nu \leq \frac{1}{4}$ for $\delta = 6$, in agreement with the theoretical prediction (30).

It is clear that for the other Runge–Kutta methods neither Criterion (33) nor (35) gives a good agreement with the experimentally found bounds of Table I.

Since the 1D experiments are inconclusive for the choice of the Runge–Kutta method, the same positivity test is performed in 2D with $\delta = 2$, a constant velocity field $u = v = -1$ and a uniform grid with mesh width $h = \frac{1}{50}$ in both directions. The initial profile is chosen as the cylinder used in Example 1 in Section 5.1. The output point is $t = 0.25$. The behavior of the methods is different from that in 1D, but not in accordance with the theoretical conditions (33) or (35). Table III gives the Courant numbers $\nu = (|u| + |v|)\tau/h$ needed for positivity.

For the 2- and 3-stage methods again a rapid transition is observed from truly negative values to -10^{-18}. However, for RK4 the minima remain negative, although small in absolute value. 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 $\nu \leq \frac{1}{2}$ can be recommended, due to its simplicity and the fact that it is supported by the non-linear theory. However, in most applications there will be a background concentration, in which case very little undershoot will do no harm. This allows larger Courant numbers, making the higher order methods more attractive, see Table II.

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

The 1D schemes are easily extended to the multi-dimensional case. Here we consider the 2D problem

$$w_t + (uw)_x + (vw)_y = 0. \tag{38}$$

The semi-discretization is the 2D equivalent of (3)

$$\frac{d}{dt} w_i + \frac{F_{i+1/2,j} - F_{i-1/2,j}}{\Delta x} + \frac{F_{i,j+1/2} - F_{i,j-1/2}}{\Delta y} = 0. \tag{39}$$

To save space we present the flux expressions only for the x-

TABLE II

L_2-Errors for Cone Profile, $h = \frac{1}{100}$

$1/\tau$	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

direction. The flux expressions for the y-direction follow in a straightforward way. First suppose $u(x, y, t) \geq 0$. Dropping the subscript j, we then replace (13) by

$$F_{i+1/2} = u_{i+1/2}\left(w_i + \frac{1}{2}\phi_{i+1/2}(w_i - w_{i-1})\right),$$
$$u_{i+1/2} = u\left(\frac{x_{i+1} + x_i}{2}\right), \tag{40}$$

where $\phi_{i+1/2} = \phi(r_{i+1/2})$, with $r_{i+1/2} = (w_{i+1} - w_i)/(w_i - w_{i-1})$, is the limiter value defined by the limiter function (20) with $K(r) = (1 + 2r)/3$. Hence, the only difference with (13) is the variable velocity $u_{i+1/2}$ in front of the bracketed solution-dependent expression. The form (40) is called the state interpolation form.

An alternative is to keep the original form (13), 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 second-order central discretization of u_x:

$$\frac{d}{dt}v_i + \frac{1}{h}\left[\left\{u_{i+1/2}\left(v_i + \frac{1}{2}\phi_{i+1/2}(v_i - v_{i-1})\right)\right\} - \left\{u_{i-1/2}\left(v_{i-1} + \frac{1}{2}\phi_{i-1/2}(v_{i-1} - v_{i-2})\right)\right\}\right] + \beta\alpha^{-1}\frac{u_{i+1/2} - u_{i-1/2}}{h} = 0. \tag{41}$$

In 2D we expect this numerical divergence term to be small for a divergence-free velocity field. Note that the slope ratios $r_{i+1/2}$, and hence the limiter values $\phi_{i+1/2}$, have not changed. For the original flux interpolation formula (13) the counterpart of (41) 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 unphysical. On the other hand, a disadvantage of (40) is that the third-order consistency is lost (in smooth monotone regions, where $\phi_{i+1/2} = K(r_{i+1/2})$). This is illustrated by the modified equation of the state interpolation form which reads

$$w_t + (uw)_x = -\tfrac{1}{24}h^2(wu_{xxx} + 3w_x u_{xx} + 2w_{xx}u_x) + O(h^3). \tag{42}$$

For $u(x, y, t) < 0$ the counterpart of (40) is given by

$$F_{i+1/2} = u_{i+1/2}\left(w_{i+1} + \frac{1}{2}\phi_{i+1/2}(w_{i+1} - w_{i+2})\right),$$
$$u_{i+1/2} = u\left(\frac{x_{i+1} + x_i}{2}\right), \tag{43}$$

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

$$F_{i+1/2} = \max(u_{i+1/2}, 0)F^+_{i+1/2} + \min(u_{i+1/2}, 0)F^-_{i+1/2}, \tag{44}$$

with $F^+_{i+1/2}$ given by (40) and $F^-_{i+1/2}$ by (43). Recall that (44) comprises four different sign cases for the associated semi-discrete scheme (3). For all four cases positivity can be proved by a straightforward application of (11).

We conclude this section with a description of our implementation of inflow/outflow boundary conditions. We hereby suppose 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 (3) is applied for $i \geq 1$, so that only for $i = 1$ an auxiliary variable w_{-1} needs to be introduced for the

TABLE III

ν-Values for Positivity, 2D, $h = \frac{1}{50}$

RK2a	RK2b	RK3a	RK3b	RK4
0.66	0.67	0.86	0.78	<0.1

flux computation $F_{1/2}$ defined by (40). We use the second-order extrapolation $w_{-1} = \max(3w_0 - 3w_1 + w_2, 0)$. Note that this would result in the second-order central discretization at $x = x_1$ if the limiting is switched off and if $u_{1/2} = u_{3/2}$. In the exceptional case of $u_0 \geq 0$ and $u_{1/2} < 0$, $F_{1/2}$ is computed by (43), 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 (3) is applied for $i \geq 0$ and an auxiliary flux computation $F_{-1/2}$ defined by (43) is introduced. $F_{-1/2}$ then uses the auxiliary state variable w_{-1} introduced above. The auxiliary velocity $u_{-1/2}$ is defined by the second-order extrapolation $u_{-1/2} = \min((15u_0 - 10u_1 + 3u_2)/8, 0)$. Assuming a constant velocity and no limiting, the outflow scheme defined this way is just second-order upwind. In the exceptional event of $u_{1/2} > 0$ and $u_0 < 0$, $F_{1/2}$ is computed by (40) which then also uses the auxiliary variable w_{-1}.

5. NUMERICAL 2D EXAMPLES

In this section we apply the positive upwind-RK advection scheme to three 2-space-dimensional example problems. In [15] other tests are performed and a comparison with various other numerical advection schemes is given [15, Chap. 15].

5.1. *Example 1: Solid Body Rotation*

Our first example is concerned with a standard test used by many authors, the so-called Molenkamp–Crowley test or solid body rotation. In Eq. (38) we let $0 \leq x, y \leq 1$ and put $u(x, y, t) = 2\pi(y - \frac{1}{2})$, $v(x, y, t) = -2\pi(x - \frac{1}{2})$. 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 $w(x, y, t) = \Phi(X, Y)$, where

$$\begin{aligned} X &= \cos(2\pi t)(x - \tfrac{1}{2}) - \sin(2\pi t)(y - \tfrac{1}{2}), \\ Y &= \sin(2\pi t)(x - \tfrac{1}{2}) + \cos(2\pi t)(y - \tfrac{1}{2}). \end{aligned} \tag{45}$$

Hence $\Phi(X, Y)$ rotates with period 1 around $(\frac{1}{2}, \frac{1}{2})$ in the clockwise direction. For $\Phi(X, Y)$ we make 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 $t = 0$. For both solutions, one full rotation is carried out on the uniform grid having 80×80 grid cells, using step size $\tau = h/3$ for the RK4 method, which corresponds roughly with a maximal Courant number 2, the Courant number being defined by (37). Note that this maximal value violates the bound given in Section 3.1. However, as the maximum value occurs near the boundary, no instability results since in a sufficiently large neighborhood of the boundary the solution is zero.

The computed solutions at $t = 1$ are shown in Fig. 2. The solutions are positive and accurately centered around the 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 observe the same clipping problem as observed in 1D [4]. Note, however, that in 2D the clipping is 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 has decreased to 0.66. The cone obviously needs a much finer grid. Grids with local refinements suit very well for that purpose; see [4] for solutions obtained on such grids. In [15], the performances of various numerical methods are extensively compared for the solid body rotation with perfectly smooth initial solution.

5.2. *Example 2: Inflow/Outflow Problem*

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

5.3. *Example 3: Spherical Advection over the Poles*

The present example has been borrowed from [13]. It deals with advection on the sphere with (scaled) radius 1. The corresponding advection equation in conservation form is given by

$$\frac{\partial w}{\partial t} + \frac{1}{\cos y}\left[\frac{\partial (uw)}{\partial x} + \frac{\partial (vw \cos y)}{\partial y}\right] = 0, \tag{46}$$

where $x \in [0, 2\pi]$ and $y \in [-\pi/2, \pi/2]$ are the longtitude and latitude coordinates in radians, and u and v are the wind velocities in x- and y-directions. Note that we now consider a pure initial value problem. The κ-scheme applied to (46) on a uniform grid in the (x, y)-plane reads

$$\begin{aligned} \frac{dw_{i,j}}{dt} = \frac{-1}{h \cos y_j}[&F_{i+1/2,j} - F_{i-1/2,j} \\ &+ \cos y_{j+1/2} F_{i,j+1/2} - \cos y_{j-1/2} F_{i,j-1/2}], \end{aligned} \tag{47}$$

with $F_{i+1/2,j}$, $F_{i,j+1/2}$ as before. The relevant 1D Courant numbers are

$$(\nu^x)_{i+1/2,j} = \frac{\tau}{h \cos y_j}|u_{i+1/2,j}|, \quad (\nu^y)_{i,j+1/2} = \frac{\tau}{h}|v_{i,j+1/2}|.$$

A difficulty is that ν^x increases when approaching the two poles, provided the wind field crosses the poles, of course. Such a wind field is given by $u(x, y) = 2\pi \cos x \sin y$, $v(x, y) = -2\pi \sin x$. As in [13] we consider a (cell-centered) 128×64 grid, and the time-integration interval $[0, 1]$ is covered in 5120 steps. This gives a maximal Courant number ≈ 1 near the poles,

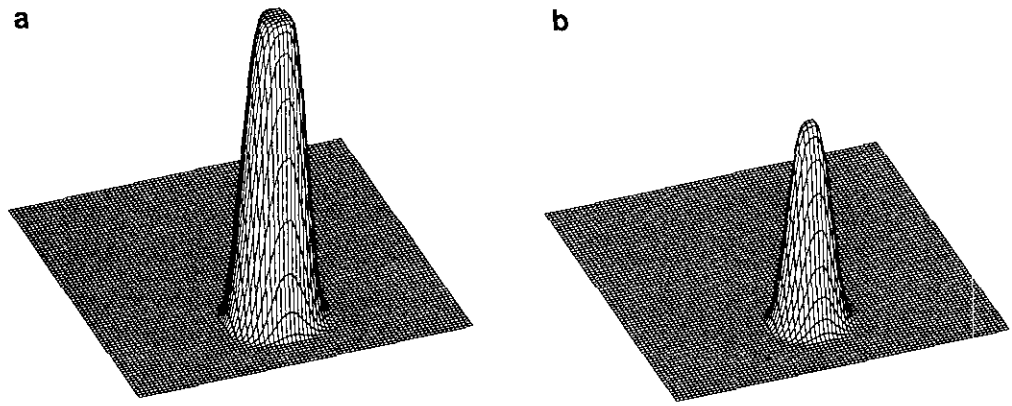

FIG. 2. Computed profiles for the rotational flow field of Example 1: (a) cylinder; (b) cone.

but outside the polar regions we get small Courant numbers and so there limiting with $\delta > 2$ is attractive (see Section 3.3). Along with our limiter with fixed $\delta = 2$ we therefore also consider

$$\delta = 2 \max\left(1, \frac{1-\nu}{\nu+\varepsilon}\right), \tag{48}$$

with a small ε to avoid zero-division. This is applied in a 1D fashion with the above Courant numbers. (Note that inequality (35) is equivalent to $\delta \leq 2((1-\nu)/\nu)$ and we take the maximum of this with 2 to avoid small δ values near $\nu = 1$.)

The tests are performed with cone- and cylinder-shaped initial profiles (the latter with background concentration). Both profiles have center $(\pi/2, 0)$ and a radius corresponding to seven grid points:

$$r(x,y) = 2\sqrt{(\cos(y)\sin(\tfrac{1}{2}(x-\pi/2)))^2 + (\sin(\tfrac{1}{2}y))^2}, \qquad R = 7\pi/64,$$

$$\text{Cone:} \quad c_0(x,y) = \max(0, 1 - r(x,y)/R),$$

$$\text{Cylinder:} \quad c_0(x,y) = 2 \quad \text{if } r(x,y) \leq R,$$
$$= 1 \quad \text{otherwise.}$$

At time $t = 1$ this profile has completed one full rotation with the trajectory over both poles. The time integration is done through the second-order method RK2b (see Section 3.1). We consider three $\kappa = \frac{1}{3}$ schemes, viz. the scheme without limiting, the scheme with $\delta = 2$-limiting, and the scheme with variable limiter value δ given by (48). In order to compare the results with those in [13] we consider the same error measures and include results obtained by the first-order upwind (donor-cell) scheme. Numerical results are given in Table IV. Along with the error measures of [13] we also give the CPU times on an SGI workstation (single precision Fortran) and the scaled CPU times, denoted by CPU′, with respect to the donor-cell algorithm.

These results are to be compared to the tests in [13] for the Eulerian MPDATA schemes. The MPDATA-1,1,0 scheme corresponds to the donor-cell scheme. The $\kappa = \frac{1}{3}$-schemes appear to be somewhat more accurate and considerably cheaper than the third-order MPDATA schemes.

Comparison with the semi-Lagrangian methods in [13] is favorable for the latter ones (except for mass conservation). There is no CFL restriction with such schemes, so that the small time steps necessary with Eulerian schemes can be avoided. On the other hand, these small time steps are caused by the grid (clustering near the poles), rather than by the problem. Larger

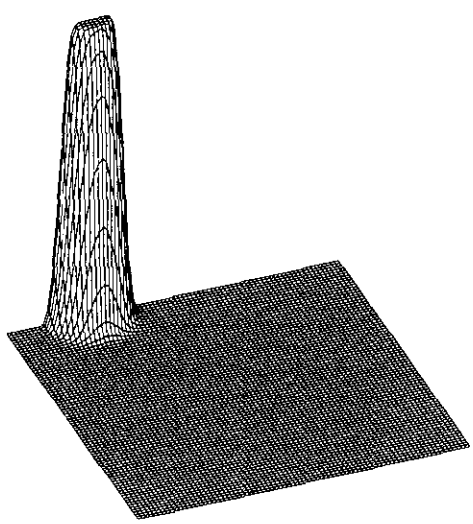

FIG. 3. The computed profile for the semi-rotation of Example 2.

TABLE IV

Results for One Revolution over the Sphere, 5120 Time Steps

	EMIN	EMAX	ERR0	ERR1	ERR2	CPU (min)	CPU'
Cone tests							
Donor-cell	0	−0.83	0.063	0	−0.86	5.8	1
Non-limited $\kappa = \frac{1}{3}$	−0.03	−0.16	0.012	0	−0.11	15.8	2.7
Limited $\kappa = \frac{1}{3}$, $\delta = 2$	0	−0.26	0.015	0	−0.19	29.6	5.0
Limited $\kappa = \frac{1}{3}$, $\delta = 2 \max\left(1, \frac{1-\nu}{\nu+\varepsilon}\right)$	0	−0.20	0.011	0	−0.15	35.3	6.0
Cylinder tests							
Donor-Cell	0	−0.30	0.067	0	−0.023	4.2	1
Non-limited $\kappa = \frac{1}{3}$	−0.03	0.071	0.030	0	0.005	11.4	2.7
Limited $\kappa = \frac{1}{3}$, $\delta = 2$	0	−0.001	0.038	0	0.007	14.1	3.3
Limited $\kappa = \frac{1}{3}$, $\delta = 2 \max\left(1, \frac{1-\nu}{\nu+\varepsilon}\right)$	0	0	0.029	0	0.007	16.1	3.8

Note. The values for the error measures are set to 0 if they approach roundoff (10^{-6}).

time steps can be taken on a reduced grid, where grid cells near the poles are merged. In a forthcoming report these matters will be addressed.

Note that the CPU times for the cone tests and cylinder tests differ significantly. This is due to the fact that with the cone tests no background concentration is present. Due to numerical diffusion very small values w_{ij}^n will arise, which are viewed as underflow values.

For completion we give the formulae for the error measures used here:

$$\text{EMIN} = \frac{\min(w_{i,j}^n) - \min(w_{i,j}^0)}{\max(w_{i,j}^0)},$$

$$\text{EMAX} = \frac{\max(w_{i,j}^n) - \max(w_{i,j}^0)}{\max(w_{i,j}^0)},$$

$$\text{ERR0} = \frac{(\sum_{i,j} \cos(y_j)(w_{i,j}^n - w_{i,j}^0)^2)^{1/2}}{(\sum_{i,j} \cos(y_j))^{1/2} \max(w_{i,j}^0)},$$

$$\text{ERR1} = \frac{\sum_{i,j} \cos(y_j) w_{i,j}^n}{\sum_{i,j} \cos(y_j) w_{i,j}^0} - 1,$$

$$\text{ERR2} = \frac{\sum_{i,j} \cos(y_j)(w_{i,j}^n)^2}{\sum_{i,j} \cos(y_j) w_{i,j}^0)^2} - 1.$$

To conclude, in Figs. 4a,b iso-line plots are given of the cone and cylinder solutions obtained through the limited $\kappa = \frac{1}{3}$-scheme with δ according to (48). For good comparison purposes, we use the same figure layouts as in [13]. The exact cone and cylinder are depicted by dashed isolines. In the graph for the cone, the isoline values considered are 0.1, 0.2, ... up to and including the maximum decimal value found. In the graph for the cylinder with background concentration, this is 1.1, 1.2, ...

6. CONCLUSIONS

For the spatial discretization we have considered four (directionally split) 5-point discretizations in conservation form, viz. the second-order central, the second-order upwind, the third-order upwind biased, and the fourth-order central discretization. The first three schemes are well-known members of the family of κ-schemes. Positivity is achieved by flux limiting, using (20) for the three κ-schemes and (22) for the fourth-order scheme. The limited third- and fourth-order discretizations perform equally well and outperform the two limited second-order ones. For general use we recommend the third-order discretization limited by (20). This combination possesses very good shape-preserving properties, in 1D as well as in 2D. No 3D experiments have been carried out, but we expect the behavior in 3D of this combination to be as good as in 2D.

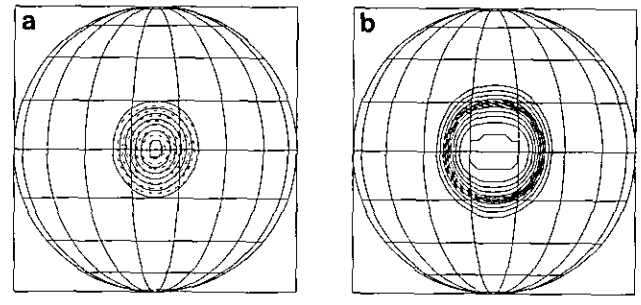

FIG. 4. Isoline distributions, Example 3, limited $\kappa = \frac{1}{3}$-scheme with $\delta = 2 \max(1, (1 - \nu)/(\nu + \varepsilon))$ (solid lines), and exact (dashed lines): (a) cone; (b) cylinder (with background concentration).

For the time integration we have examined a number of explicit RK methods, viz. the second-order method of Runge–Kutta (RK2a), the second-order explicit trapezoidal rule (RK2b), the third-order methods of Heun (RK3a) and Fehlberg (RK3b), and the classical fourth-order method (RK4). 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 behaving about the same and no clearly best method could be identified. For example, we have not found a notable difference in positivity and accuracy/efficiency performance between the second-order explicit trapezoidal rule, which fits in the nonlinear theory, and the classical fourth-order explicit method, which does not fit. For strongly varying solutions the behavior of the combined spatial–temporal scheme is apparently dominated by the spatial discretization where the limiting procedure plays a decisive role.

REFERENCES

1. C. Bolley and M. Crouzeix, *RAIRO Anal. Numér.* **12,** 237 (1978).
2. E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I—Nonstiff Problems,* Comput. Math., Vol. 8 (Springer-Verlag, Berlin, 1987).
3. Ch. Hirsch, *Numerical Computation of Internal and External Flows,* Vol. 2 (Wiley, Chichester, 1990).
4. W. Hundsdorfer, B. Koren, M. van Loon, and J. G. Verwer, Report NM-R9309, CWI, Amsterdam, 1993 (unpublished).
5. B. Koren, in *Numerical Methods for Advection-Diffusion Problems,* edited by C. B. Vreugdenhil and B. Koren, Notes on Numerical Fluid Mechanics, Vol. 45 (Vieweg, Braunschweig, 1993), p. 117.
6. J. F. B. M. Kraaijevanger, *Numer. Math.* **48,** 303 (1986).
7. J. F. B. M. Kraaijevanger, *BIT* **31,** 482 (1991).
8. B. van Leer, in *Large-Scale Computations in Fluid Mechanics,* edited by B. E. Engquist, S. Osher, and R. C. J. Somerville (Am. Math. Soc., Providence, RI, 1985), p. 327.
9. R. J. LeVeque, *Numerical Methods for Conservation Laws,* Lecture Notes in Mathematics, ETH, Zürich (Birkhäuser, Basel, 1992).
10. S. Osher and S. Chakravarthy, in *Oscillation Theory, Computation, and Methods of Compensated Compactness,* edited by C. Dafermos, J. L. Ericksen, D. Kinderlehrer, and M. Slemrod, IMA Vol. Math. Appl., Vol. 2, (Springer-Verlag, New York, 1986), p. 229.
11. C.-W. Shu and S. Osher, *J. Comput. Phys.* **77,** 439 (1988).
12. C.-W. Shu and S. Osher, *J. Comput. Phys.* **83,** 32 (1989).
13. P. K. Smolarkiewicz and P. J. Rasch, *J. Atmos. Sci.* **48,** 793 (1991).
14. P. K. Sweby, *SIAM J. Numer. Anal.* **21,** 995 (1984).
15. C. B. Vreugdenhil and B. Koren (Eds.), *Numerical Methods for Advection-Diffusion Problems,* Notes on Numerical Fluid Mechanics, Vol. 45 (Vieweg, Braunschweig, 1993).
16. S. T. Zalesak, in *Advances in Computer Methods for Partial Differential Equations VI,* edited by R. Vichnevetsky and R. S. Stepleman, IMACS (Baltzer, Basel, 1987), p. 15.