Method of lines and direct discretization: a comparison for linear advection *

W. Hundsdorfer **, R.A. Trompert

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

Abstract

In this paper two related schemes are compared for two-dimensional linear advection: a scheme based on dimension splitting starting from a fully discrete one-dimensional discretization, and a scheme based on the method of lines. Both schemes use the same four-point, upwind biased stencil in each direction, and both are considered with flux limiting to avoid oscillations and negative values in the solutions. Along with a theoretical discussion, we also give some numerical test results, on uniform grids and with local uniform grid refinement, to compare the accuracy and computational efficiency.

Key words: Hyperbolic PDEs; Linear advection; Method of lines; Flux limiting; Dimensional splitting; Local uniform grid refinement

1. Introduction

There are two standard ways to derive numerical algorithms for multi-dimensional flow problems [9]: either by the method of lines (MOL) approach, where space and time discretization are considered separately, or with dimensional splitting, starting from a fully discrete one-dimensional method. In this paper two such related schemes are compared for the linear advection equation in two space variables,

$$w_t + (aw)_x + (bw)_y = 0, \tag{1.1}$$

with given initial solution and Dirichlet conditions at the inflow boundaries. The subscripts t, x, and y indicate partial derivatives and the velocity field (a, b) may depend on x, y, and t. In many applications this velocity field is supposed to be divergence-free,

$$a_x + b_y = 0. \tag{1.2}$$

We shall consider the linear advection equation with reference to atmospheric air-pollution modeling, where w stands for a vector of concentrations of chemical species advected by the wind field (a, b). In a realistic model a large number of components are involved, and also the z-direction and other processes, such as diffusion, sources, deposition and chemical reactions are included. Yet the efficient solution of the horizontal, two-dimensional, advection part remains a crucial factor, see [3,10]. Along with computational efficiency, other important properties that should be possessed by the advection scheme are: *accuracy*, for good shape preservation and correct phase speeds, and *preservation of positivity*, to avoid instabilities in the chemical reactions. Moreover, for long-term calculations, *mass conservation* is also regarded as desirable.

Solving linear advection equations is a classical problem for which a large number of numerical methods have been developed. An extensive comparison can be found in [19]. Among the finite-difference and finite-volume schemes, the MOL scheme of [5] with flux-limited, third-order upwind-biased space discretization and Runge–Kutta time-stepping performed favourable. Moreover, since the method uses a relatively small stencil it can be combined without any problem with the local uniform grid refinement technique of [16], giving good accuracy, even for sharp fronts, without introducing too many grid points (see [4]). However, the results in [4] also showed that in order to guarantee positivity, the Runge–Kutta method must be carefully selected and the time step should be sufficiently small, without precise theoretical support on the question how small this should be exactly.

We shall compare the MOL scheme of [5] with a closely related direct discretization, where space and time discretization are not separated. Such a scheme can be easily derived for the one-dimensional (1-D) problem $w_t + a w_x = 0$ by imposing the same stencil, four-point upwind-biased, and the same order, namely 3. Flux limiting can be done in the same way as for the MOL scheme.

It will turn out that, for the 1-D model problem, the direct discretization method gives more accurate results at lower computational costs. However, unlike the MOL scheme, it is not easily extended to 2-D. Therefore, the direct scheme will be used in combination with dimensional splitting (see [9]). Although this introduces an extra error term, it will be shown that the direct scheme can be implemented in such a way that the errors remain of the same order as for the MOL scheme while the computational costs are significantly lower.

The contents of this paper is as follows. First, in Section 2 we shall compare the two approaches in one space dimension with flux limiting. The extension to two dimensions is discussed in Section 3. In Section 4 stepsize restrictions for positivity and stability are briefly discussed, and Section 5 contains an error analysis. In Section 6 numerical tests in two dimensions are presented on fixed grids and with local grid refinement. Section 7, finally, contains some concluding remarks.

2. Comparison for 1-D model problem

2.1. Introduction of the methods

Consider the model problem, corresponding to (1.1),(1.2) in one space dimension,

$$w_t + a w_x = 0, \quad a > 0 \text{ constant}, \tag{2.1}$$

for $t \geqslant 0$. Boundary conditions will be ignored for the moment. Third-order upwind-biased spatial discretization, also called the $\kappa = \frac{1}{3}$ scheme [5,18], gives the semi-discrete system

$$\frac{d}{dt} w_i(t) = \frac{a}{h} \left(-\tfrac{1}{6} w_{i-2}(t) + w_{i-1}(t) - \tfrac{1}{2} w_i(t) - \tfrac{1}{3} w_{i+1}(t) \right). \tag{2.2}$$

Here $w_i(t) \approx w(t, x_i)$ with $x_i = ih$. On this system of ordinary differential equations we can now apply a numerical ODE scheme, for instance a Runge–Kutta method, to obtain fully discrete approximations $w_i^n \approx w(t_n, x_i)$ at time levels $t_n = n\tau$, $\tau > 0$ being the stepsize in time. This method of lines procedure has a large flexibility and can be applied in the same way to the 2-D problem (1.1).

We shall compare this MOL scheme with a direct discretization based on the same stencil

$$w_i^{n+1} = \gamma_{-2} w_{i-2}^n + \gamma_{-1} w_{i-1}^n + \gamma_0 w_i^n + \gamma_1 w_{i+1}^n \tag{2.3}$$

with coefficients, depending on the Courant number $\nu = a\tau/h$, given by

$$\gamma_{-2} = -\tfrac{1}{6}\nu(1 - \nu^2), \qquad \gamma_{-1} = \tfrac{1}{2}\nu(2 - \nu)(1 + \nu)$$

$$\gamma_0 = \tfrac{1}{2}(2 - \nu)(1 - \nu^2), \qquad \gamma_1 = -\tfrac{1}{6}\nu(2 - \nu)(1 - \nu).$$

In the following we shall refer to (2.3) as the *direct scheme*. It has also order 3 and it is stable for $\nu \leqslant 1$ (see the results on optimal order schemes of Strang [13]). The method can be derived from the order conditions but also from a semi-characteristic approach, where the characteristics are traced backward in time and Lagrange interpolation is used, either on w or its primitive function. In Section 5 this last approach will be used to derive error bounds for variable velocities. We note already that from the semi-characteristic interpretation it follows that the method becomes more accurate the closer ν is to 1: the global error $w(t_n, x_i) - w_i^n$ can be bounded by $C(1 - \nu)h^3$, with a positive constant C depending on the exact solution.

There is a close connection between (2.2) and (2.3). If we consider the limit $\tau \to 0$ for fixed h and $t_n > 0$, then the approximations w_i^n of (2.3) converge to the solution $w_i(t_n)$ of (2.2). This easily follows by writing (2.3) as

$$\frac{1}{\tau} \left(w_i^{n+1} - w_i^n \right) = \frac{a}{h} \left(-\tfrac{1}{6} w_{i-2}^n + w_{i-1}^n - \tfrac{1}{2} w_i^n - \tfrac{1}{3} w_{i+1}^n \right) + O(\tau), \qquad h > 0, \quad \tau \to 0.$$

So, for h fixed and ν tending to zero the direct scheme becomes the same as a MOL scheme with exact time integration. Note however that there is a clear advantage for the direct scheme since its error will become smaller the larger ν is, up to $\nu = 1$, whereas in the MOL approach the error will grow with increasing ν due to growing temporal inaccuracy of the numerical ODE method.

2.2. Flux limiting

Both methods (2.2) and (2.3) can produce some over- and undershoot, which may lead to negative values. If w stands for a concentration of a chemical species, this should, of course, remain nonnegative. Although the negative values created by the upwind-biased schemes are rather small compared to those created by central difference schemes, they can lead to instabilities when advection is combined with chemical reactions. Schemes with the property that solutions remain nonnegative for arbitrary nonnegative initial values are often called *positive*.

The simplest way to avoid negative values, setting them equal to zero, can lead to a distorted mass balance. Following the approach of Sweby [14], we therefore consider flux limiting. For (2.2) a suitable limiting procedure has been described in detail in [4,5], and the same limiter will be considered here for (2.3).

Method (2.3) can be written in conservation form as

$$w_i^{n+1} = w_i^n + \frac{\tau}{h}\left(f_{i-1/2}^n - f_{i+1/2}^n\right). \tag{2.4}$$

Omitting the superscripts n on the right-hand side, i.e. $w_i = w_i^n$, the fluxes $f_{i+1/2}$ are given by

$$f_{i+1/2} = a\left(-\tfrac{1}{6}(1-\nu^2)w_{i-1} + \tfrac{1}{6}(1+\nu)(5-2\nu)w_i + \tfrac{1}{6}(2-\nu)(1-\nu)w_{i+1}\right).$$

Let $d_0 = \tfrac{1}{6}(2-\nu)(1-\nu)$ and $d_1 = \tfrac{1}{6}(1-\nu^2)$. Introducing the ratios

$$\theta_i = \frac{w_i - w_{i-1}}{w_{i+1} - w_i},$$

the flux can be written as

$$f_{i+1/2} = a\left(w_i + (d_0 + d_1\theta_i)(w_{i+1} - w_i)\right). \tag{2.5}$$

For limiting we consider the more general form, with a limiter function ψ,

$$f_{i+1/2} = a\left(w_i + \psi(\theta_i)(w_{i+1} - w_i)\right). \tag{2.6}$$

Since $w_{i+1} - w_i = (w_i - w_{i-1})/\theta_i$, it is easily seen that the total scheme can be written in the form

$$w_i^{n+1} = (1 - \nu\lambda_i)w_i + \nu\lambda_i w_{i-1}$$

with

$$\lambda_i = 1 - \psi(\theta_{i-1}) + \frac{1}{\theta_i}\psi(\theta_i).$$

It is clear that the method is positive iff

$$0 \leqslant \nu\lambda_i \leqslant 1.$$

We want the limiter function ψ to be such that this requirement holds while the original non-limited flux should be reobtained as much as possible, certainly for ratios θ_i close to 1, since this is the generic situation for a smooth profile. As in [4] we consider

$$\psi(\theta) = \max(0, \min(1, d_0 + d_1\theta, \mu\theta)), \tag{2.7}$$

where μ is a positive parameter. With this limiter we have $0 \leqslant \psi(\theta) \leqslant 1$ and $0 \leqslant \psi(\theta)/\theta \leqslant \mu$. Hence

$$0 \leqslant \lambda_i \leqslant 1 + \mu.$$

The condition for positivity, thus, is

$$(1 + \mu)\nu \leqslant 1. \tag{2.8}$$

Note that this condition also implies other nice properties, such as the maximum principle $\min_j w_j^n \leqslant w_i^{n+1} \leqslant \max_j w_j^n$ and nonincreasing total variation (TVD) (see [9]). The semi-discrete system with flux limiting used in [4,5] can be obtained by taking the limit $\tau \to 0$ with $h > 0$ fixed (see Section 3), just as for the schemes without limiting. The resulting semi-discrete system gives nonnegative solutions for any $\mu > 0$ if the time-stepping is done exactly.

Large values of μ give more accurate results, with less clipping of peaks, but at the cost of a smaller allowable Courant number. In the MOL approach of [4] it was found experimentally that $\mu = 1$ performs quite well and is more efficient than $\mu = 3$, for example. The fact that large values of μ are inefficient in the MOL approach is due to the ODE solvers, which require smaller stepsizes, that is, smaller ν, to maintain positivity if μ is increased.

For any fixed $\mu > 0$ condition (2.8) yields a Courant restriction less than 1, which spoils the nicest part $\nu \approx 1$ of the direct scheme, where the errors are smallest. However, not working in the MOL context gives the flexibility to make μ dependent of the Courant number. Taking μ as large as possible within the positivity constraint we get

$$\mu = (1 - \nu)/\nu \quad \text{for } 0 \leqslant \nu \leqslant 1. \tag{2.9}$$

Now the scheme is positive for all $\nu \leqslant 1$ and thus the most accurate region where ν is near 1 can be included. Moreover, for small Courant numbers there is less clipping of peaks compared to fixed $\mu = 1$ (see also van Leer [17, p. 291]).

In Fig. 1 three numerical solutions at $t = 1$ are displayed for the test equation $w_t + w_x = 0$, for $0 \leqslant t \leqslant 1$ and $0 \leqslant x \leqslant 1$, with periodicity conditions at $x = 0, 1$ and exact solution $w(t, x) = \cos(\pi(x - t - \frac{1}{2}))^{100}$. The mesh width was chosen as $h = \frac{1}{50}$. The first two pictures show the

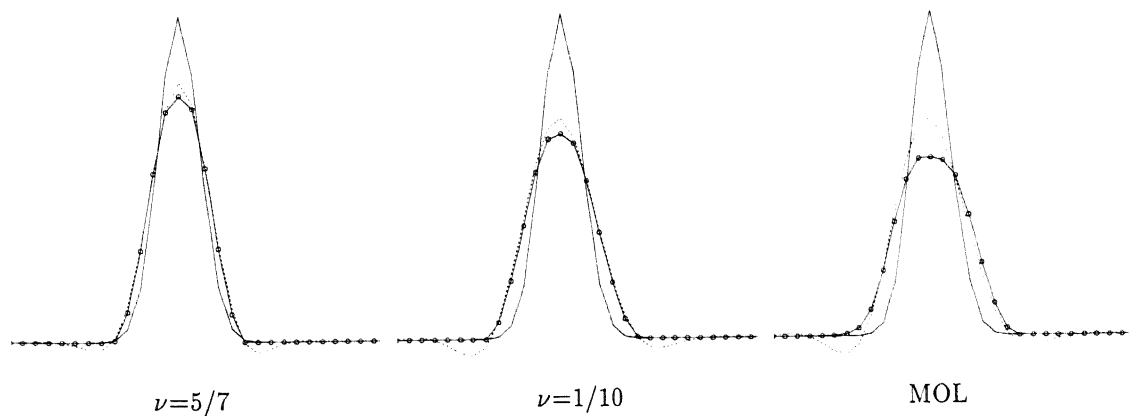$\nu = 5/7$ $\nu = 1/10$ MOL

Fig. 1. Numerical solutions for $(\cos)^{100}$ profile, $h = 1/50$.

results for the direct scheme with limiter value (2.9) at Courant numbers $\nu = \frac{5}{7}, \frac{1}{10}$, respectively. Smaller Courant numbers give almost the same result as $\nu = \frac{1}{10}$. The third picture gives the numerical solution of the MOL approach [4,5] with $\mu = 1$-limiting and the classical fourth-order Runge–Kutta time-stepping with stepsize $\tau = \frac{1}{2}h$; at this stepsize the temporal error is not visible anymore. In all three pictures the dotted lines indicate the corresponding numerical solutions without limiting. We note that the results for the direct scheme with limiter (2.7) and fixed $\mu = 1$ are for small Courant numbers virtually the same as in the third picture for the MOL scheme. So it is clear that making μ dependent on ν gives a considerable improvement of accuracy near sharp peaks.

The limiting procedure does affect the order of accuracy of the method, since even for smooth solutions the limiter will act near extremal values. For the above test equation with smooth periodic solution $w(t, x) = \cos(\pi(x - t - \frac{1}{2}))^2$ we observed convergence with order ≈ 2.5 in the L_1-norm and ≈ 1.8 in the L_∞-norm. This dependence on norm is caused by the fact that the largest errors are introduced at the extremum $x = t + \frac{1}{2}(\mathrm{mod}\ 1)$, and these errors remain confined to a small region near this point, in accordance with Fig. 1.

Numerical comparisons between the MOL scheme and the direct scheme will be presented in Section 6 for 2-D problems. Although it is not our intention to make an extensive comparison with other limited advection schemes (see [19,20] for this), a few remarks are in order on related schemes using the same stencil.

Remark 2.1. Some popular flux-limited schemes can be found in [9,20], expressed in terms of $\phi(\theta)$, related to $\psi(\theta)$ by

$$\psi(\theta) = \tfrac{1}{2}(1 - \nu)\phi(\theta).$$

Typical examples are the MUSCL limiter

$$\phi(\theta) = \max\big(0, \min\big(2, 2\theta, \tfrac{1}{2}(1 + \theta)\big)\big)$$

of van Leer [17], and the "superbee" limiter

$$\phi(\theta) = \max(0, \min(1, 2\theta), \min(\theta, 2))$$

of Roe [11]. In some tests we observed that MUSCL gives significantly more diffusion and clipping of peaks than (2.7),(2.9), whereas the "superbee" limiter gives roughly the same peak values as (2.7),(2.9) and slightly better results for advection of a square wave, but this scheme is inaccurate for smooth solutions with a strong tendency to turn smooth curves into straight lines (see also the figures presented in [20]). It should be noted that these schemes were designed to handle nonlinear hyperbolic systems, whereas here only linear advection is considered.

Remark 2.2. The limiter (2.7),(2.9) fits into the general formula (5.13) in [20], attributed to Hain and Lyon. In the numerical tests of [20] this formula was applied with a high-order spatial discretization and Runge–Kutta time-stepping. For the third-order semi-discrete system (2.2) this type of limiting was found to produce better results for small Courant numbers, but for $\nu \geqslant \frac{1}{2}$ the fixed $\mu = 1$ should be maintained. Since we will take relatively large Courant numbers in the numerical experiments, this modification on the MOL scheme is not considered any further here.

3. General linear advection

3.1. Formulation for $w_t + (aw)_x = 0$

As a first step to describe the implementation of the direct discretization for equation (1.1), we consider the one-dimensional counterpart

$$w_t + (aw)_x = 0 \tag{3.1}$$

with variable velocity a. The conservation form of this equation is maintained with

$$w_i^{n+1} = w_i + \frac{\tau}{h}(f_{i-1/2} - f_{i+1/2}), \tag{3.2}$$

where, as before, superscripts n are omitted in the right-hand side. Let $x_{i+1/2} = \frac{1}{2}(x_i + x_{i+1})$. For variable velocities, the ν and a used in (2.6) for the calculation of $f_{i+1/2}$ are replaced by $\nu_{i+1/2} = |\alpha_{i+1/2}|\tau/h$ and $\alpha_{i+1/2} = \alpha(t_n, x_{i+1/2})$, with $\alpha(t, x)$ to be specified later (for the moment we can consider $\alpha(t, x) = a(t, x)$). If $a_{i+1/2} < 0$, the stencil $\{x_{i+1}, x_i, x_{i+1}\}$ used in (2.6) should be reflected around the point $x_{i+1/2}$ to maintain the upwind character. The resulting formula with limiting can be written compactly as

$$f_{i+1/2} = \alpha_{i+1/2}^+ \big(w_i + \psi_{i+1/2}(\theta_i)(w_{i+1} - w_i)\big)$$

$$+ \alpha_{i+1/2}^- \left(w_{i+1} + \psi_{i+1/2}\left(\frac{1}{\theta_{i+1}}\right)(w_i - w_{i+1})\right), \tag{3.3}$$

where $\alpha^+ = \max(0, \alpha)$, $\alpha^- = \min(\alpha, 0)$ and

$$\psi_{i+1/2}(\theta) = \max\big(0, \min\big(1, d_{i+1/2}(\theta), \mu_{i+1/2}\theta\big)\big), \tag{3.4}$$

$$d_{i+1/2}(\theta) = \frac{1}{6}(2 - \nu_{i+1/2})(1 - \nu_{i+1/2}) + \frac{1}{6}(1 - \nu_{i+1/2}^2)\theta, \tag{3.5a}$$

$$\mu_{i+1/2} = (1 - \nu_{i+1/2})/\nu_{i+1/2}. \tag{3.5b}$$

Here (3.4) is the generalization of (2.7), so the method without limiting is reobtained by setting $\psi_{i+1/2}(\theta) = d_{i+1/2}(\theta)$.

By checking all possibilities of the signs of $\alpha_{i+1/2}$ and $\alpha_{i-1/2}$ it can be verified in a straightforward way that positivity of the numerical solutions remains guaranteed with limiting under the condition

$$\max_i \nu_{i+1/2} \leqslant 1. \tag{3.6}$$

In contrast to the constant coefficient case, the maximum principle and the TVD property are now lost, but this is necessary to remain consistent with $w_t + (aw)_x = 0$ for variable velocities a. In view of the form (3.2), mass conservation is still valid.

The corresponding semi-discrete system, which was used in [4,5], is given by

$$\frac{d}{dt}w_i = \frac{1}{h}(f_{i-1/2} - f_{i+1/2}), \tag{3.7}$$

where the fluxes are defined by (3.3),(3.4) with $d_{i+1/2}(\theta)$ and $\mu_{i+1/2}$ replaced by

$$d(\theta) = \tfrac{1}{3} + \tfrac{1}{6}\theta, \qquad \mu = 1. \tag{3.8}$$

Remark 3.1. Formula (2.3) is also the starting point for the flux-limited method $\mathrm{MAPF}_{l=2}$ of Bott [1]. The limiting procedure used in [1], however, is much more complicated than (3.4). It is based on the maximum principle, enforced by modifying the fluxes subsequently in the downstream direction, but the inconsistency caused by this procedure for variable velocities has to be corrected afterwards by adding extra terms. On the other hand, this form of limiting can be applied directly to higher-order methods as well, whereas (3.5) is derived specifically for the four-point stencil used in the present method.

Another popular scheme which is based on (2.3) is the QUICKEST scheme of Leonard presented in [6] without limiting. The limiting procedure derived for the related scheme QUICK in [7] is quite complicated and does not always prevent negative values (see for instance the numerical results in [15]).

Apart from a different limiting procedure, the direct scheme presented here will also have a more accurate treatment for variable velocities, especially in combination with dimensional splitting.

3.2. Formulation for 2-D

The MOL scheme can be applied in a straightforward way in 2-D, simply by considering the right-hand side of (3.7) as a finite-difference approximation to $-(aw)_x$. We note that, also without limiting, the order will be 2 in general. Only for wind fields that are constant in space the order will be 3 (see [4,5] and Section 5).

For the direct scheme a proper 2-D extension is not so easily available. We shall therefore consider this scheme with dimensional splitting. If $w_n(x, y)$ is an approximation to the exact solution at time t_n, the most simple splitting procedure consists of solving on $[t_n, t_{n+1}]$ subsequently

$$w_t^* + (aw^*)_x = 0, \qquad w^*(t_n, x, y) = w_n(x, y) \tag{3.9}$$

$$w_t^{**} + (bw^{**})_y = 0, \qquad w^{**}(t_n, x, y) = w^*(t_{n+1}, x, y), \tag{3.10}$$

giving the next approximation $w_{n+1}(x, y) = w^{**}(t_{n+1}, x, y)$. The fractional steps will be solved with the direct scheme (3.2)–(3.5).

It is well known that this simple splitting procedure is only first-order accurate, unless the wind field is constant, in which case there is no splitting error (see for instance [9]). As we shall see in Section 5, the order can be easily raised to 2 by adapting the wind fields used in the method.

It will be assumed that Dirichlet boundary conditions are given at the inflow boundaries. To avoid inconsistencies in the splitting scheme these will be treated in the following way: in the fractional steps (3.9) and (3.10) the boundaries are integrated along with the interior points, using extrapolation to find missing values of points outside the domain, and after the fractional steps have been performed the given exact boundary values are injected in the new approximation $w_{n+1}(x, y)$.

4. Restrictions on the Courant number

The dimensional splitting complicates the local error analysis, but makes it is easy to determine stepsize restrictions for stability and positivity of the direct scheme. From the 1-D considerations giving (3.6), we know that positivity is guaranteed if

$$\max(|a|, |b|)\tau/h \leqslant 1, \tag{4.1}$$

which should be satisfied over the whole domain. This is also the condition for stability, in the sense of von Neumann, with the non-limited direct scheme for constant velocities. It has been verified experimentally that the same stability condition is valid for the limited direct scheme.

The semi-discrete system used in the MOL approach takes in 2-D the form

$$\frac{d}{dt}w = F(w) + G(w) =: H(w), \tag{4.2}$$

where $F(w)$ and $G(w)$ contain the discretized spatial derivatives in the x- and y-direction, respectively, computed as in (3.7). We will apply Runge–Kutta time-stepping on this system of ODEs. Since both spatial directions are taken together now, we get a condition of the form

$$(|a| + |b|)\tau/h \leqslant C \tag{4.3}$$

with $C \geqslant 0$ depending on the specific Runge–Kutta method and the requirement at hand, either stability or positivity. Several explicit Runge–Kutta methods were applied in [4], without giving a clear preference. We shall briefly discuss two examples.

With the second-order explicit trapezoidal rule

$$w^* = w^n + \tau H(w^n), \qquad w^{n+1} = w^n + \tfrac{1}{2}\tau(H(w^n) + H(w^*)), \tag{4.4}$$

it can be shown, as in [12], that the resulting limited MOL scheme will be positive under the condition (4.3) with $C = \tfrac{1}{2}$. Experiments in [4] indicated that this might be relaxed to $C \approx \tfrac{2}{3}$, while stability was found to hold for $C \approx 1$ (and $C = 0.87$ for the non-limited scheme), but at these stepsizes the temporal error may still be relatively large.

Better temporal accuracy is obtained with the fourth-order classical Runge–Kutta method

$$
\begin{aligned}
w^* &= w^n + \tfrac{1}{2}\tau H(w^n), \\
w^{**} &= w^n + \tfrac{1}{2}\tau H(w^*), \\
w^{***} &= w^n + \tau H(w^{**}), \\
w^{n+1} &= w^n + \tfrac{1}{6}\tau(H(w^n) + 2H(w^*) + 2H(w^{**}) + H(w^{***})).
\end{aligned}
\tag{4.5}
$$

The experimental stability bound for this method given in [4] is condition (4.3) with $C \approx 1.4$ for the limited scheme (and $C = 1.74$ without limiting). However, this fourth-order Runge–Kutta method fails to produce nonnegative solutions in 2-D, although the magnitude of the negative values appears to be small in general, in the order of 10^{-4} or less (see [4] and Section 6). In view of its accuracy, this classical Runge–Kutta method is used in the numerical experiments of Section 6.

5. Local error analysis

5.1. The 1-D error

In this section an error analysis of the direct discretization will be given. Results on the accuracy of the semi-discrete system used in the MOL approach follow by considering $\tau \to 0$. The analysis here will be made for the method without limiting (see also Remark 5.2). Moreover, boundary conditions will be ignored and the wind field will assumed to be constant in time. This last assumption is not essential (see Remark 5.3), but it is convenient for the notation and presentation.

There are two distinct error sources: the splitting of the 2-D problem and the numerical approximation of the 1-D problems in each stage of the splitting procedure. First a bound for the 1-D error will be derived.

So, consider the non-limited scheme (3.3),(3.5), with $\psi_{i+1/2} = d_{i+1/2}$, for the 1-D equation (3.1), with given initial value at $t = 0$ and with velocity a depending on x. Instead of making a Taylor series expansion to find an expression for the local truncation error, the method will be interpreted in such a way that it becomes obvious what the various error contributions are. In this way it will also be clear how the method can be modified in case a higher accuracy is demanded.

Consider a cell $\Omega_i = (x_{i-1/2}, x_{i+1/2})$ and let

$$\bar{w}_i(t) = \frac{1}{h} \int_{\Omega_i} w(t, x) \, dx$$

be the cell average of the exact solution. Further, let $\Omega_i^* = (x_{i-1/2}^*, x_{i+1/2}^*)$ with $x_{i+1/2}^*$ being the departure point at $t = t_n$ of the characteristic $x'(t) = a(x(t))$ that passes through $x_{i+1/2}$ at time t_{n+1}. From a mass balance in moving coordinates it follows that

$$\int_{\Omega_i} w(t_{n+1}, x) \, dx = \int_{\Omega_i^*} w(t_n, x) \, dx = W(t_n, x_{i+1/2}^*) - W(t_n, x_{i-1/2}^*) \tag{5.1}$$

with primitive function

$$W(t, x) = \int_0^x w(t, \xi) \, d\xi.$$

Now, suppose $w_i = w_i^n$ is the numerical solution at $t = t_n$, interpreted as cell average over Ω_i. Define

$$W_{i+1/2} = h \sum_{j \geqslant i} w_j, \tag{5.2}$$

so that

$$w_i = \frac{1}{h}(W_{i+1/2} - W_{i-1/2}).$$

Note that if the cell averages are exact, then also $W_{i+1/2} = W(t_n, x_{i+1/2})$. The numerical formula corresponding to (5.1) is

$$w_i^{n+1} = \frac{1}{h}(W_{i+1/2}^* - W_{i-1/2}^*), \tag{5.3}$$

where $W^*_{j+1/2} \approx W(t_n, x^*_{j+1/2})$. This can be written in flux form as

$$w_i^{n+1} = w_i + \frac{\tau}{h}\left(f_{i-1/2} - f_{i+1/2}\right), \quad f_{i+1/2} = \frac{1}{\tau}\left(W_{i+1/2} - W^*_{i+1/2}\right). \tag{5.4}$$

To determine $W^*_{i+1/2}$ we consider piecewise Lagrange interpolation with polynomials of order 3. Let $x_{i+1/2} - \tau\alpha_{i+1/2}$ be the numerical approximation to $x^*_{i+1/2}$. Further, let $P(x) = P_i(x)$ on Ω_i, with P_i a polynomial of degree 3 satisfying $P_i(x_{j+1/2}) = W_{j+1/2}$ with $j = i - 1, i, i + 1$ and $j = i - 2 \operatorname{sign}(\alpha_{i+1/2})$. The scheme (5.4) is now specified by the definition

$$W^*_{i+1/2} = P(x_{i+1/2} - \tau\alpha_{i+1/2}). \tag{5.5}$$

If $\nu_{i+1/2} \leqslant 1$, we reobtain from (5.4),(5.5), after a straightforward but somewhat tedious manipulation, the same formula for the fluxes $f_{i+1/2}$ without limiting as in Section 3.1.

From this interpretation of the method for variable velocities it is now clear that there are two contributions to the error of the 1-D scheme, namely the error caused by the piecewise polynomial interpolation and the error in the approximation of the departure point $x^*_{i+1/2}$. Starting at time t_n with exact cell averages $w_i = \overline{w}_i(t_n)$, the local truncation error is given by

$$\frac{1}{\tau}\left(\overline{w}_i(t_{n+1}) - w_i^{n+1}\right) = \frac{1}{h}\left(U_{i+1/2} - U_{i-1/2}\right) + \frac{1}{h}\left(V_{i+1/2} - V_{i-1/2}\right), \tag{5.6}$$

with

$$U_{i+1/2} = \frac{1}{\tau}\left(W(t_n, x^*_{i+1/2}) - W(t_n, x_{i+1/2} - \tau\alpha_{i-1/2})\right),$$

$$V_{i+1/2} = \frac{1}{\tau}\left(W(t_n, x_{i+1/2} - \tau\alpha_{i+1/2}) - P(x_{i+1/2} - \tau\alpha_{i+1/2})\right).$$

From the definition of $x^*_{i+1/2}$ it can be seen that

$$x^*_{i+1/2} - \left(x_{i+1/2} - \tau\alpha_{i+1/2}\right)$$
$$= \tau\left(\alpha_{i+1/2} - a(x_{i+1/2})\right) + \tfrac{1}{2}\tau^2 a_x(x_{i+1/2})a(x_{i+1/2}) + \cdots$$
$$= \tau^{q+1}\gamma(x_{i+1/2}) + O(\tau^{q+2}), \tag{5.7}$$

where q and $\gamma(x)$ depend on the choice of $\alpha_{i+1/2}$. If $\alpha_{i+1/2} = a(x_{i+1/2})$, it is obvious that $q = 1$ and $\gamma(x) = \tfrac{1}{2}a_x(x)a(x)$. From (5.7) it follows that

$$U_{i+1/2} = \tau^q \gamma(x_{i+1/2})w(t_n, x_{i+1/2}) + O(\tau^{q+1}),$$

$$\frac{1}{h}\left(U_{i+1/2} - U_{i-1/2}\right) = \tau^q \frac{\partial}{\partial x}\left(\gamma(x_i)w(t_n, x_i)\right) + O(\tau^{q+1}) + O(h^2\tau^q). \tag{5.8}$$

For the error term caused by interpolation it holds that

$$V_{i+1/2} = \frac{1}{4!}h^3(2 - \nu_{i+1/2})(1 - \nu_{i+1/2})\,|\alpha_{i+1/2}|\,(1 + \nu_{i+1/2})\frac{\partial^4}{\partial x^4}W(t_n, x_{i+1/2}) + O(h^4),$$

$$\frac{1}{h}\left(V_{i+1/2} - V_{i-1/2}\right) = O(h^3), \tag{5.9}$$

where the exact form of the $O(h^3)$ term depends on possible sign changes in $\alpha_{i+1/2}$ and $\alpha_{i-1/2}$. Inserting (5.8),(5.9) into (5.6) we thus obtain a bound $O(h^3) + O(\tau^q)$ for the local truncation error.

For constant velocities it is seen from the above that the error can be bounded by $(1 - \nu)O(h^3)$. If we use $\alpha_{i+1/2} = a(x_{i+1/2})$ for variable velocities, the temporal order will be only 1. This can be increased by calculating the departure points $x^*_{i+1/2}$ more precisely. For example, the formula

$$\alpha_{i+1/2} = a(x_{i+1/2}) - \tfrac{1}{2}\tau a_x(x_{i+1/2})a(x_{i+1/2}) \tag{5.10}$$

will give order $q = 2$, and higher orders can be obtained in the same way.

For this derivation of the truncation error, w_i^n was interpreted as an approximation to the cell average $\bar{w}_i(t_n)$. If it is considered as an approximation to the point value $w(t_n, x_i)$ the error expression should be slightly modified for variable velocities. Defining $\bar{w}(t, x)$ as the average of $w(t, \cdot)$ over the interval $(x - \tfrac{1}{2}h, x + \tfrac{1}{2}h)$, we have

$$\bar{w} = w + \tfrac{1}{24}h^2 w_{xx} + O(h^4).$$

From $w_t + (aw)_x = 0$, it follows that

$$\bar{w}_t + (a\bar{w})_x = -\tfrac{1}{24}h^2\left((a\bar{w})_{xxx} - (a\bar{w}_{xx})_x\right) + O(h^4),$$

and so this is the differential equation that is approximated with $O(h^3) + O(\tau^q)$ accuracy. Consequently, if $w_i^n = w(t_n, x_i)$, the local truncation error of the scheme with respect to point values is given by

$$\frac{1}{\tau}\left(w(t_{n+1}, x_i) - w_i^{n+1}\right) = \tfrac{1}{24}h^2\left(a_{xxx}w + 3a_{xx}w_x + 2a_x w_{xx}\right)(t_n, x_i) + O(h^3) + O(\tau^q).$$

$$\tag{5.11}$$

Since this leading $O(h^2)$ term is missing in the cell average interpretation it seems this interpretation should be preferred. However, as we will see, in 2-D the $O(h^2)$ term will appear anyway.

Remark 5.1. The above derivation via the primitive function has a similarity with the construction of the PPM scheme of Colella and Woodward [2]. However, in the PPM scheme a larger stencil is used, leading to fourth order but also to a more complicated limiting procedure.

5.2. The splitting error

Equation (1.1) can be written in operator form as

$$w_t + Aw + Bw = 0$$

with

$$Aw = (aw)_x, \qquad Bw = (bw)_y.$$

The exact solution satisfies

$$w(t_{n+1}, x, y) = e^{-\tau(A+B)}w(t_n, x, y),$$

where the exponential of an operator is defined by the power series of the exponential function.

Let us assume for the moment that the 1-D problems are solved exactly. Then the simple splitting procedure (3.9),(3.10) gives

$$w_{n+1}(x,\, y) = \mathrm{e}^{-\tau B}\mathrm{e}^{-\tau A}w_n(x,\, y).$$

If we start on the exact solution $w_n(x,\, y) = w(t_n,\, x,\, y)$, the truncation error is given by

$$\frac{1}{\tau}\bigl(w(t_{n+1},\, x,\, y) - w_{n+1}(x,\, y)\bigr) = \tfrac{1}{2}\tau(AB - BA)w(t_n,\, x,\, y) + \mathrm{O}(\tau^2) \tag{5.12}$$

(see for instance [9, Section 18.2]). By some manipulation it follows that

$$(AB - BA)_w = -\bigl(a_y bw\bigr)_x + \bigl(ab_x w\bigr)_y. \tag{5.13}$$

The error introduced by this simple splitting procedure is of order 1. The accuracy can be increased by interchanging after each step the 1-D subprocesses, giving

$$w_{n+2}(x,\, y) = \mathrm{e}^{-\tau A}\mathrm{e}^{-2\tau B}\mathrm{e}^{-\tau A}w_n(x,\, y).$$

This corresponds to the second-order Strang splitting with stepsize 2τ (see [9]), and it requires the same amount of computational work. In case Strang splitting is used with stepsize τ it becomes twice as expensive as the simple splitting procedure, and this may happen if the stepsize is dictated by other processes, for instance reaction, diffusion or exchange between the horizontal layers. In the next subsection it will be shown that the order of the simple splitting procedure can also be easily raised to 2 by adapting the wind fields used in the method.

The local truncation error in (5.12) is valid for point values. For cell averages and $\mathrm{O}(h^2)$ term should be added to the splitting error, unless a and b are constant. This can be seen by deriving the differential equation that is satisfied by cell averages, just as in the previous subsection but now in 2-D, and comparing this with the 1-D differential equations used in the fractional steps. Hence, the advantage in the order of accuracy that was observed in 1-D for the cell average interpretation no longer holds in 2-D.

5.3. Modified equations and modified wind fields

Under the assumption that the fractional steps (3.9) and (3.10) are solved exactly, it follows from the expression (5.12),(5.13) for the local splitting errors that the *modified equation* for the simple splitting procedure is

$$w_t + (\tilde{a}w)_x + (\tilde{b}w)_y = 0 \tag{5.14}$$

with

$$\tilde{a} = a - \tfrac{1}{2}\tau a_y b, \qquad \tilde{b} = b + \tfrac{1}{2}\tau ab_x.$$

This equation is approximated by splitting with second-order temporal accuracy.

Next, suppose that the 1-D problems are solved by (3.3),(3.5) with velocities $\alpha_{i+1/2,j} = a(x_{i+1/2},\, y_j)$ in the x-direction and $\beta_{i,j+1/2} = b(x_i,\, y_{j+1/2})$ in the y-direction. As we already saw, this will introduce additional first-order temporal errors. If we add these errors of the

fractional steps to the splitting error (5.12) we get a modified equation for the total scheme of the same type as (5.14), but now with

$$\bar{a} = a + \tfrac{1}{2}\tau(a_x a - a_y b), \qquad \bar{b} = b + \tfrac{1}{2}\tau(ab_x + b_y b).$$

So, this equation is approximated by the total scheme with $O(h^2) + O(\tau^2)$ accuracy.

Knowing this, we can now adapt the wind fields used in the method such that the resulting modified equation will become again the original equation (1.1). This is achieved by using the modified velocities $\alpha_{i+1/2,j} = \alpha(x_{i+1/2}, y_j)$ and $\beta_{i,j+1/2} = \beta(x_i, y_{j+1/2})$ in the fractional steps with

$$\alpha = a - \tfrac{1}{2}\tau(a_x a - a_y b), \qquad \beta = b - \tfrac{1}{2}\tau(ab_x + b_y b). \tag{5.15}$$

In conclusion, using these modified wind fields the accuracy of the method will be $O(h^2) + O(\tau^2)$ with respect to equation (1.1).

The same accuracy can be obtained by setting $\alpha = a - \tfrac{1}{2}\tau a_x a$ and $\beta = b - \tfrac{1}{2}\tau b_y b$, to counteract the 1-D errors as in (5.10), together with second-order Strang splitting, but the above modification is a bit simpler.

Remark 5.2. For constant velocities most of the error terms will vanish and the accuracy will become $O(h^3)$. If the variation in the wind field is sufficiently small this third-order term may still dominate the error for given τ and h.

Limiting can be viewed as an intervention on the interpolation procedure. Hence, with limiting the $O(h^3)$ term is expected to become $O(h^p)$ with p approximately in the range 1.8–2.5, depending on the norm used, as observed in Section 2.

Remark 5.3. Although it becomes somewhat more tedious, the same analysis can be performed in case the velocities depend on t, using Taylor expansion to find the splitting error (see for instance [8, formula (2.6)]). The result is that the velocities to be used in (3.3) with the simple splitting procedure should be $\alpha_{i+1/2,j} = \alpha(t_n, x_{i+1/2}, y_j)$ and $\beta_{i,j+1/2} = \beta(t_n, x_i, y_{j+1/2})$ with

$$\alpha = a - \tfrac{1}{2}\tau(a_x a - a_y b - a_t), \qquad \beta = b - \tfrac{1}{2}\tau(ab_x + b_y b - b_t). \tag{5.16}$$

6. Numerical comparison

To compare the accuracy and computational efficiency of the MOL and direct scheme, several numerical tests in 2-D were performed. Two of these are reported here (the conclusions for other tests, taken from [4,15], were more or less the same). In the first problem, with relatively smooth solution, only uniform grids are considered. For the second problem, rotation of a cylinder, we also applied local uniform grid refinement, in the same way as in [4]. For a comparison of the MOL scheme with advection schemes of different type (finite element, spectral, etc.) we refer to [19].

For the MOL scheme the classical fourth-order explicit Runge–Kutta method (4.5) is used for time-stepping. The direct scheme is applied with the modified velocity fields given by (5.15).

The derivatives of the velocities a and b in (5.15) are computed with standard second-order finite differences. Both schemes are considered with and without the appropriate flux limiters. The numerical approximations are regarded as point values at the nodes of the grid.

In the experiments the solution was computed on uniform Cartesian grids with boundary nodes on the boundary of the spatial domain. For the non-limited schemes, fourth-order extrapolation was used to obtain the necessary values across the boundaries of the domain. This rather high-order extrapolation was chosen so as to obtain the same order of consistency over the entire domain. With limiting, the values outside the domain were computed by constant, first-order extrapolation to avoid nonmonotonicities and to obtain a positive scheme.

The solutions were computed by applying the schemes at all nodes, including those on the boundary. For the direct scheme the exact Dirichlet boundary values were assigned to the inflow boundaries after each complete time step, so that no intermediate boundary conditions are used in the fractional steps. With the Runge–Kutta method this updating at the boundaries is done in all the stages.

6.1. Problem I: mixing of fronts

The first test problem, obtained from [15], is a simple model to describe the mixing of hot and cold air. The initial solution is a horizontal front separating cold air at the upper half of the domain from the warm air at the lower half. The front is twisted by a steady rotational velocity field resembling the cyclonic air motion at low pressure systems observed on the daily-weather maps. The solution is given by

$$w(x, y, t) = \tanh\left(\tfrac{1}{2}x \, \sin(\omega(r)t) - \tfrac{1}{2}y \, \cos(\omega(r)t)\right), \tag{6.1}$$

with domain $-4 \leqslant x, y \leqslant 4$ and

$$\omega(r) = \frac{v(r)}{rv_{\max}}, \quad v(r) = \frac{\tanh(r)}{\cosh^2(r)}, \quad r = \sqrt{x^2 + y^2}. \tag{6.2}$$

The tangential velocity around the center of the domain is denoted by v, and $v_{\max}$ is its maximum which is set as 0.385. The velocity field is defined by

$$a(x, y) = -\omega(r)y, \qquad b(x, y) = \omega(r)x, \tag{6.3}$$

and is shown in Fig. 2. Far away from the center the velocity is rather small, but in the vicinity of this center large velocity components arise. At the center itself the velocity is zero.

This problem was used to subject both schemes to a convergence test. The solution was computed on a 20×20, 40×40, 80×80 and 160×160 grid (these are the number of cells in the grid) from $t = 0$ to $t = 4$, using the time stepsizes 0.25, 0.125, 0.0625 and 0.03125, respectively. The mesh spacing and time stepsizes were chosen such that the maximum Courant number is well below the stability limits of both schemes. Fig. 2 displays the numerical solution at $t = 4$ obtained with the limited direct scheme on a 160×160 grid. The global errors in the L_∞- and L_1-norm are shown in Fig. 3 as a function of the time stepsize τ (which is proportional to h). Logarithmic scaling was applied to the errors and stepsize.

The results in both norms reveal that the direct scheme is more accurate than the MOL scheme on this problem. Reduction of the stepsize τ in the MOL scheme did not improve its
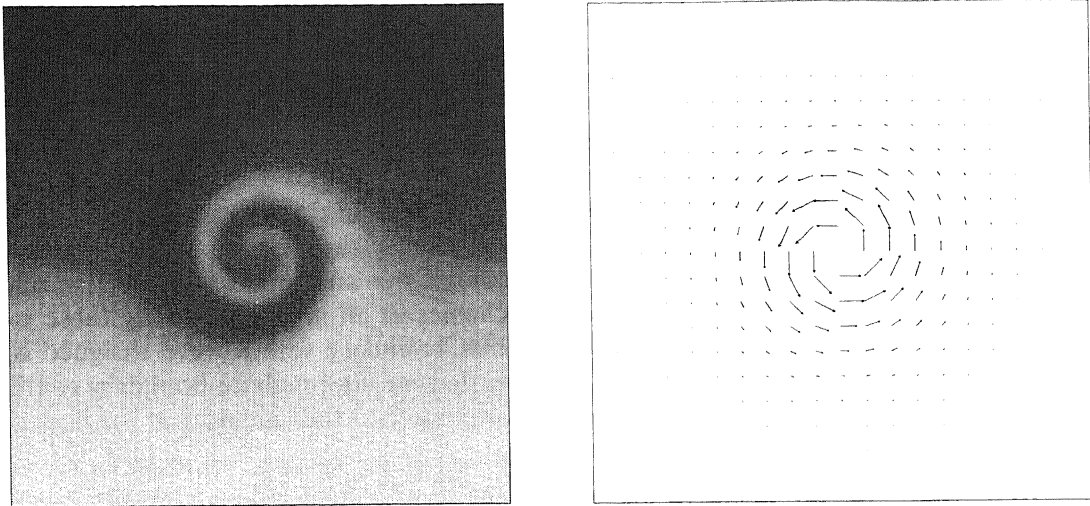
Fig. 2. Solution ($t = 4$) and wind field for Problem I.

accuracy when h was held fixed, so the errors for this scheme are entirely due to the spatial discretization. The fact that the lines in Fig. 3 are curved indicates that the asymptotic region of the error is not reached yet. We estimated the order of convergence using the errors on the 80×80 and 160×160 grids. For both non-limited schemes this estimated order of convergence is approximately 2.4 in both the L_1- and the L_∞-norm. For the limited schemes these numbers are 2.4 in the L_1-norm and 1.8 in the L_∞-norm. There is no significant difference between the MOL and the direct scheme with respect to these orders.
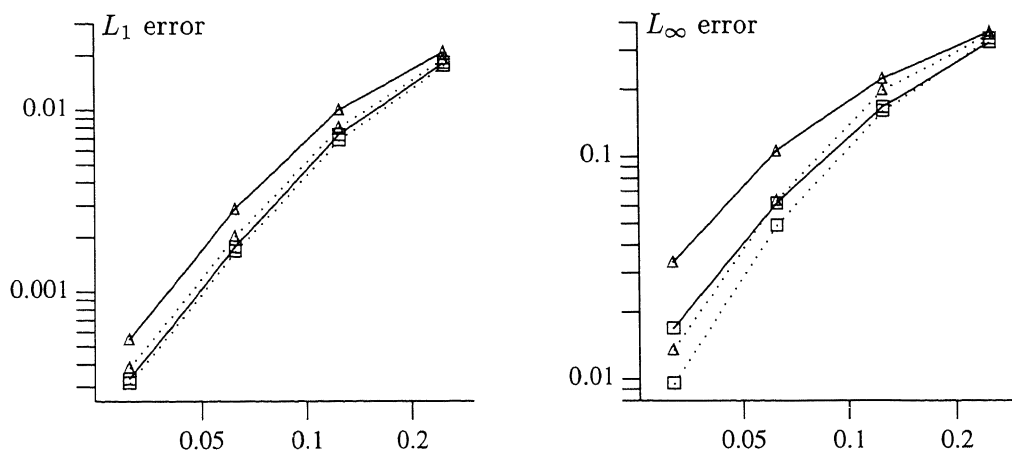

Fig. 3. L_1 and L_∞ errors for Problem I as function of τ ($\square$: direct scheme, $\triangle$: MOL scheme, —: limited, $\cdots$ non-limited).
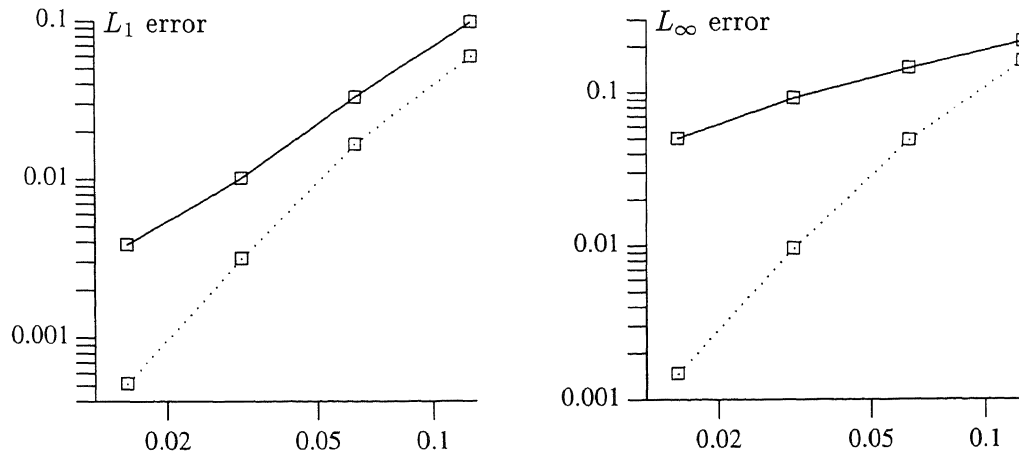
Fig. 4. Errors of direct scheme for Problem I on reduced domain as function of τ (—: limited (first-order extrapolation), $\cdots$: non-limited (fourth-order extrapolation)).

From Fig. 3 it can also be observed that the difference in accuracy between the limited and non-limited versions is larger with the MOL scheme than with the direct scheme. This is probably due to the Courant dependency of the limiter in the direct scheme, see Section 2.

Also CPU times on a Silicon Graphics Indigo workstation were compared. The direct scheme was roughly 2.5 times faster than the MOL scheme. We note that not much effort has been put in optimizing the codes; for instance, the coefficients in (3.5) were recomputed each time step. So, this factor 2.5 gives only a crude indication. The CPU times for the schemes with and without limiting were of similar size.

The correct treatment of boundary conditions is often a matter of concern with splitting methods (see [8]). To test our implementation for the dimensionally splitted direct scheme, the same problem (6.1)–(6.3) was solved on the reduced domain $-1 \leqslant x,y \leqslant 1$, so that the velocities at the boundaries are large and relatively steep parts of the solution enter and leave the region. This in contrast to the previous domain where very little happened at the boundaries. The solution was computed on a 10×10, 20×20, 40×40 and 80×80 grid using the time stepsizes 0.125, 0.0625, 0.03125 and 0.015625, respectively. Both the limited and the non-limited direct scheme were used. As mentioned before, the limited scheme employs first-order and the non-limited scheme fourth-order extrapolation. Fig. 4 shows the global error in the L_∞- and L_1-norm.

Comparing the results obtained with the non-limited scheme on the squares $-4 \leqslant x,y \leqslant 4$ and $-1 \leqslant x,y \leqslant 1$, we see that the errors of these computations on both squares are quite similar for equal stepsize. This shows that our boundary treatment with splitting does not cause a reduction of accuracy. It is also clear that the first-order extrapolation used in the limited scheme will hamper the convergence rate if there is much activity near the boundaries. If more accuracy is required, higher-order extrapolation should be used. Positivity can easily be enforced by setting negative extrapolated values equal to zero. However, small oscillations might be introduced with higher-order extrapolation.
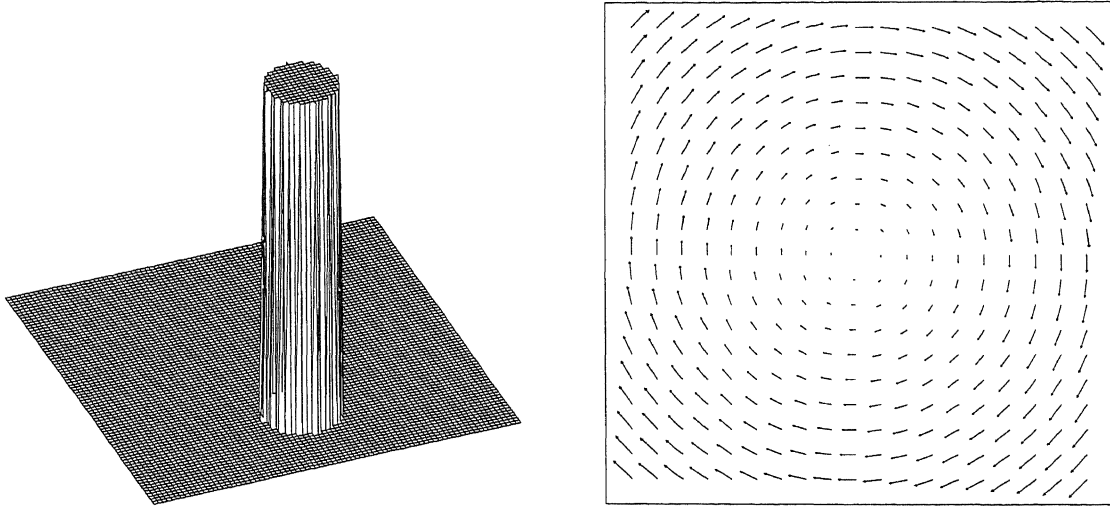
Fig. 5. Exact solution ($t = 0$, 1) and wind field for Problem II.

6.2. Problem II: solid body rotation of cylinder

The so-called Molenkamp–Crowley test or solid body rotation [3,10,19] is the second test problem. A cylinder, given at $t = 0$ by

$$w(x, y, 0) = \begin{cases} 1 & \text{for } \left(x - \tfrac{1}{2}\right)^2 + \left(y - \tfrac{3}{4}\right)^2 \leqslant \tfrac{1}{10}, \\ 0 & \text{otherwise}, \end{cases} \tag{6.4}$$

on the domain $0 \leqslant x, y \leqslant 1$, rotates around the center of the domain in a clockwise manner. After one time unit it will have completed one full rotation. The velocity field is given by

$$a(x, y) = -2\pi\left(y - \tfrac{1}{2}\right), \qquad b(x, y) = 2\pi\left(x - \tfrac{1}{2}\right) \tag{6.5}$$

and is shown in Fig. 5.

Both a uniform grid and an adaptive grid method were applied. The adaptive grid technique is the so-called local uniform grid refinement method. This involves solving PDEs on a series of nested local uniform Cartesian subgrids which become increasingly finer and cover only a part of the domain. On each of these subgrids a new initial boundary value problem is solved separately for one time step in a consecutive order, from coarse to fine. This implies that the subgrids are not patched into the coarser grids but are actually overlaying them. In the version of the method applied here, all grids use the same time stepsize. After a time step is completed on all grids, the solution of the coarser grids is replaced by the solution of the finer grids at corresponding nodes. The fine grid cells are created by subdividing a coarse grid cell in four identical cells. The subgrids created this way are of a staircase-type or, in other words, they have a piecewise polygonal shape. For a more detailed description we refer to [16] and references therein.
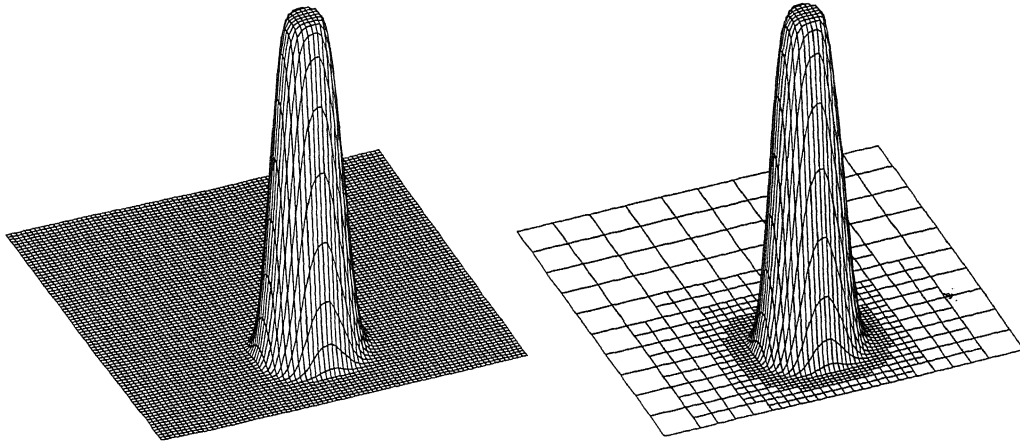
Fig. 6. Solutions ($t = 1$) of direct scheme for Problem II.

The limited MOL and direct scheme were used to compute the solution from $t = 0$ to $t = 1$. The uniform grid was a 80×80 grid and the adaptive grid method involved four grids of which the coarsest was a 10×10 grid. The cells of the finest grid in use in the adaptive grid method are identical to those of the 80×80 grid. The time stepsize was the same for all computations and equal to $\frac{1}{252}$. These time steps were chosen as large as possible with respect to the stability criterion of the direct scheme on the uniform grid. This implies, however, that condition (4.3) is only satisfied with $C = 2$, which is not sufficient in general for stability with the classical Runge–Kutta method. However, the MOL scheme showed no signs of instability during these tests, probably due to the fact that the largest Courant numbers are found at the corners where the boundary conditions play a dominant role and the solution remains virtually zero.

Figs. 6 and 7 show the numerical solution on the uniform and adapted grids. We observe that for both schemes there is no visible difference between the adapted grid and uniform grid solution. Comparing the results of the MOL and the direct scheme, we see that the MOL scheme is somewhat more diffusive. The maximum values of all computed solutions where slightly less than 1, around 0.9998, and the minimum values of the direct scheme solutions were equal to zero while the ones of the MOL scheme solutions were negative and of order 10^{-6}, due to the Runge–Kutta method [4].

The number of nodes and the CPU times were also compared. The adaptive computations with both schemes use approximately only a quarter of the number of nodes of the 80×80 grid. However, the adaptive computation with the direct scheme was only 1.2 times faster than the corresponding 80×80 computation. For the MOL scheme this factor was 1.8. It appears that although fewer nodes are used by the adaptive grid method, no large grains in CPU time are achieved for this test problem. This is caused by the overhead associated with the adaptive grid procedure. Since the direct scheme is considerably faster than the MOL scheme it is obvious that the overhead takes up a relatively larger part of the CPU time than with the MOL scheme. This explains why the gain with the MOL scheme is larger than with the direct scheme.
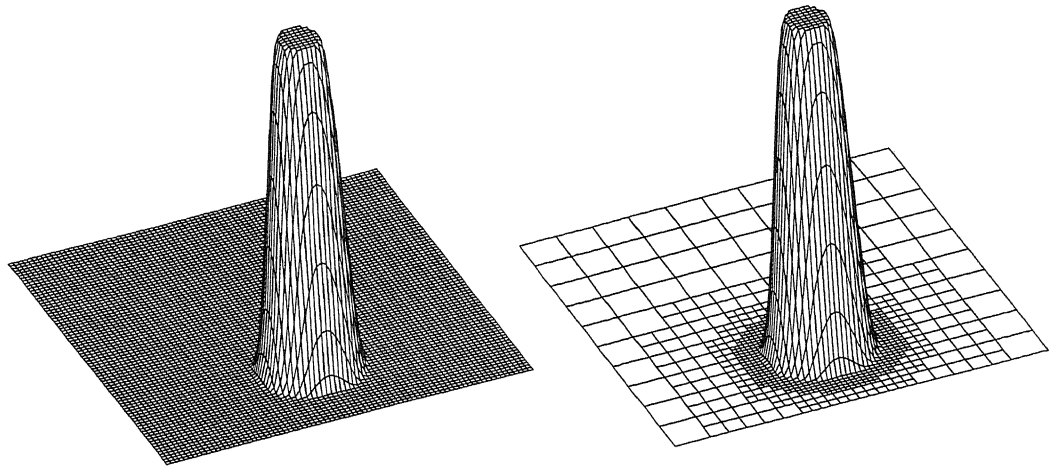
Fig. 7. Solutions ($t = 1$) of MOL scheme for Problem II.

We note that with more levels of refinement the speed-up factor increases. So far, we have not put much effort in minimizing the overhead but there are a number of options open to reduce it, for instance keeping the grids fixed for a number of time steps. We will not elaborate on this any further since it is beyond the scope of this paper. We expect that mesh adaptation will be especially worthwhile to reduce the computing time if chemical reactions have to be solved at each node. In this case the amount of work per grid-point will increase so that the relative overhead becomes smaller. This is in particular so when problems associated with atmospheric air pollution are solved where many chemical components are advected. Adaptive grid methods can in connection to these problems also prove to be useful in reducing the memory requirements.

7. Concluding remarks

From Sections 2 and 5 we know that the direct scheme is more accurate than the MOL scheme for constant velocities, since then splitting errors are absent and the accuracy is determined by the 1-D errors. For slowly varying a and b this advantage is expected to remain valid. The numerical test on Problem I indicates that some gain for the direct scheme is even preserved in case the velocity field has large variations, in particular with limiting, as the Courant-dependent limiter of the direct scheme adds less numerical diffusion to the solution. Only in situations where the solution is much smoother than the velocity field the splitting error is expected to become dominant, resulting in larger errors for the direct scheme.

The main advantage of the direct scheme seems its computational efficiency. Suppose, for example, the MOL scheme would be used with the two-stage explicit trapezoidal rule (4.4) such that condition (4.3) is satisfied with $C = \frac{1}{2}$ to ensure positivity. Then the stepsizes allowed with the direct scheme, according to condition (4.1), are a factor 2 to 4 larger, depending whether the maximal velocities are grid-aligned or diagonal. Moreover, if the dimension of w is large,

say 10 or more, the computational work for one time *step* of the direct scheme will be almost the same as for one Runge–Kutta *stage* in the MOL scheme, since the coefficients to be computed in (3.5) are the same for all components. Hence, for large systems (1.1) a speed-up factor of 4 to 8 is expected with the direct scheme.

The generalization to nonlinear conservation laws seems easier with the MOL scheme [5]. However, for the problems occurring in numerical air pollution modeling, the advection is linear and other physical processes, such as sources and diffusion can be handled by operator splitting. If the diffusion coefficient is small this can be done without introducing significant splitting errors. Diffusion with a large coefficient should be separated anyway, also with the MOL approach, and treated with a different method, for instance implicit, to avoid unreasonable stepsize restrictions.

Finally we note that the extension to advection in 3-D is straightforward, also for the direct scheme with dimensional splitting. Similar as in 2-D, the velocity field can be modified for the direct scheme to obtain second order temporal accuracy.

8. References

[1] A. Bott, Monotone flux limitation in the area-preserving flux-form advection algorithm, *Monthly Weather Rev.* 120 (1992) 2595–2602.

[2] P. Colella and P. Woodward, The piecewise parabolic method (PPM) for gas-dynamical simulations, *J. Comput. Phys.* 54 (1984) 174–201.

[3] Ø. Hov, Z. Zlatev, R. Berkowicz, A. Eliassen and L.P. Prahm, Comparison of numerical techniques for use in air pollution models with non-linear chemical reactions, *Atmos. Envir.* 23 (1989) 967–983.

[4] W. Hundsdorfer, B. Koren, M. van Loon and J.G. Verwer, A positive finite-difference advection scheme applied on locally refined grids, Report NM-R9309, CWI, Amsterdam (1993).

[5] B. Koren, A robust upwind discretization for advection, diffusion and source terms, in: C.B. Vreugdenhil and B. Koren, eds., *Numerical Methods of Advection-Diffusion Problems*, Notes on Numerical Fluid Mechanics 45 (Vieweg, Braunschweig, 1993).

[6] B.P. Leonard, Elliptic systems: finite difference method IV, in: W.J. Minkowycz, E.M. Sparrow, G.E. Schneider and R.H. Pletcher, eds., *Handbook of Numerical Heat Transfer* (Wiley, New York, 1988).

[7] B.P. Leonard, Simple high accuracy resolution program for convective modeling of discontinuities, *Internat. J. Numer. Methods Fluids* 8 (1988) 1291–1318.

[8] R.J. LeVeque, Time-split methods for partial differential equations, Ph.D. Thesis, Department of Computer Science, Stanford University, Stanford, CA (1982).

[9] R.J. LeVeque, *Numerical Methods for Conservation Laws*, Lecture Notes in Mathematics, ETH Zürich (Birkhäuser, Basel, 1992).

[10] G.J. McRea, W.R. Goodin and J.H. Seinfeld, Numerical solution of atmospheric diffusion for chemically reacting flows, *J. Comput. Phys.* 77 (1982) 1–42.

[11] P.L. Roe, Some contributions to the modeling of discontinuous flows, in: Lectures in Applied Mathematics 22 (AMS, Providence, RI, 1985) 163–193.

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

[13] G. Strang, Trigonometric polynomials and difference methods of maximal accuracy, *J. Math. Phys.* 41 (1962) 147–154.

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

[15] P. Tamamidis and D.N. Assanis, Evaluation of various high order accuracy schemes with and without flux limiters, *Internat. J. Numer. Methods Fluids* 16 (1993) 931–948.

[16] R.A. Trompert and J.G. Verwer, A static regridding method for two-dimensional parabolic partial differential equations, *Appl. Numer. Math.* 8 (1991) 65–90.
[17] B. van Leer, Towards the ultimate conservative difference scheme IV: a new approach to numerical convection, *J. Comput. Phys.* 23 (1977) 276–299.
[18] B. van Leer, Upwind-difference methods for aerodynamic problems governed by the Euler equations, in: B.E. Engquist, S. Osher and R.C.J. Sommerville, eds., *Large-scale Computations in Fluid Mechanics* (American Mathematical Society, Providence, RI, 1985) 327–336.
[19] C.B. Vreugdenhil and B. Koren, eds., *Numerical Method for Advection-Diffusion Problems*, Notes on Numerical Fluid Mechanics 45 (Vieweg, Braunschweig, 1993).
[20] S.T. Zalesak, A preliminary comparison of modern shock-capturing schemes: linear advection, in: R. Vichnevetsky and R.S. Stepleman, eds., *Advances in Computer Methods for Partial Differential Equations IV* (IMACS, New Brunswick, NJ, 1987) 15–22.