JOURNAL OF COMPUTATIONAL PHYSICS 114, 129–145 (1994)

Robust Numerical Methods for 2D Turbulence

FRED H. WALSTEIJN

Institute for Marine and Atmospheric Research, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands

Received August 19, 1992; revised June 4, 1993

A numerical method is called robust if it does not require tuning. Two different approaches to robustness are described. First, existing energy and enstrophy conserving space discretizations are combined with new nonlinear Runge–Kutta schemes that conserve either energy or enstrophy in the inviscid unforced case. This guarantees unconditional stability of long-term integrations, even for the explicit variants. The explicit nonlinear Runge–Kutta schemes are inaccurate (but stable) if the related standard Runge–Kutta formula is unstable. Fortunately, the nonlinear schemes yield a parameter which can be used to signal inaccuracies and to adapt the timestep such that results are accurate. However, as they are based on symmetric space discretizations, these methods still require the *ad hoc* tuning of an artificial viscosity to suppress numerical spatial oscillations. The second approach is to construct robust space discretizations (Jacobians) using the grid-aligned essentially non-oscillatory (ENO) technique. ENO Jacobians based on the conservation form of the vorticity equation generate spikes near stagnation points. Only fifth and lower order ENO Jacobians based on the advection form of the vorticity equation suppress all instabilities and oscillations without tuning if a global Lax–Friedrichs modification is used. A preliminary comparison is made with a tuned Arakawa Jacobian.

CONTENTS

1. *Introduction.*
2. *Conservation of discrete energy and enstrophy.* 2.1. The semi-discrete system. 2.2. Energy or enstrophy conserving time discretizations. 2.3. Concurrent conservation of energy and enstrophy. 2.4. Numerical results.
3. *ENO-Jacobians.* 3.1. Flux form ENO-Jacobians. 3.2. Advection form ENO-Jacobians.
4. *Conclusions.*

1. INTRODUCTION

Robust numerical schemes are proposed for the direct simulation of two-dimensional incompressible turbulence. Following Roe [14], a scheme is called robust if accurate and numerically stable results are obtained without the need to "tune" some of its parameters for each new simulation. Robustness is a desirable property because of the complexity and expense of direct turbulence simulations.

The continuous problem that will be considered is the vorticity equation:

$$\frac{\partial q}{\partial t} + J(\psi, q) = \mathscr{D} + \mathscr{F}, \tag{1}$$

$$q \equiv \nabla^2 \psi + f, \tag{2}$$

on a simply connected spatial domain Ω. Here $\psi(x, y, t)$ denotes the streamfunction, $q(x, y, t)$ is the potential vorticity, while x, y, and t denote the space and time coordinates. The given function $f(x, y)$ can be nonzero in, for example, a quasi-geostrophic (geophysical) context, where it incorporates effects of bottom topography and the spatial variation of the Coriolis parameter, cf. Pedlosky [13]. The "Jacobian" $J(\psi, q)$ represents advection of potential vorticity and is defined as

$$J(\psi, q) \equiv \frac{\partial \psi}{\partial x}\frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial q}{\partial x}.$$

$\mathscr{D}$ denotes a dissipation term (e.g., $\mathscr{D} = \nu \nabla^2 q$), and $\mathscr{F}(x, y, t)$ a given forcing term. For simply connected domains homogeneous Dirichlet boundary conditions can be enforced on the streamfunction:

$$\psi|_{\partial\Omega} \equiv 0. \tag{3}$$

Depending on the order of $\mathscr{D}$ additional "viscous" boundary conditions are required. This paper will mainly focus on the case $\mathscr{F} = 0$ and $\mathscr{D}$ "small", i.e., freely decaying turbulent flows. If $\mathscr{D} = \mathscr{F} = 0$, system (1)–(3) conserves energy E and enstrophy V:

$$E(t) \equiv \int_\Omega \tfrac{1}{2} \|\nabla \psi\|^2 \, d\Omega,$$

$$V(t) \equiv \int_\Omega \tfrac{1}{2} q^2 \, d\Omega.$$

0021-9991/94 $6.00

Due to these conservation properties, energy must cascade to the large scales while enstrophy cascades to the smallest scales (Pedlosky [13]).

In many turbulence simulations the physical viscous length scale cannot be resolved by the numerical model. From a numerical viewpoint this means that effectively an inviscid vorticity equation is simulated, which usually necessitates the introduction of an artificial (numerical) dissipative mechanism. Since energy cascades to the larger scales, statistically accurate results may still be obtained if enough turbulent length scales are reproduced. However, usually the spatial meshwidth cannot be chosen much smaller than the scale of the smallest important eddies; see, for example, Wolff *et al.* [20]. Therefore accurate statistics are only obtained if the numerical scheme reproduces as many scales of motion as possible for a given spatial resolution. Most numerical methods are capable of this only after elaborate tuning, i.e., they are not robust.

Two different approaches are followed in an attempt to obtain robust methods. The first approach, described in Section 2, is to start from symmetric space discretizations that conserve semidiscrete analogues of the energy and enstrophy in the inviscid unforced case. Without at least one conservation property of this type, symmetric schemes are generally unstable and need tuning of an *ad hoc* artificial viscosity to suppress the instability (Sadourny [15]). Important examples of energy and enstrophy conserving space discretizations are the Arakawa [1] Jacobian and its generalizations by Salmon and Talley [16]. The resulting semidiscrete schemes are guaranteed numerically stable and also have correct spectral directions of the semidiscrete enstrophy and inverse energy cascades. If semidiscrete enstrophy is not conserved (but energy is), then the spectral energy distribution will evolve incorrectly, cf. Arakawa [2]. Arakawa also showed that enstrophy conserving schemes violating the energy conservation principle do not have such disadvantages. Therefore discrete enstrophy conservation has a higher priority than energy conservation.

Economical energy and/or enstrophy conserving time discretizations have not been reported in literature yet. The enstrophy conserving implicit midpoint rule mentioned by Arakawa [2] seems too expensive. In geophysical fluid dynamics one commonly uses the nonconservative "leap frog" time-stepping method, stabilized by an *ad hoc* technique such as time averaging or the periodic use of a "one-step" method (Sadourny [15]). No a priori guarantee of stability can be given. By generalizing results of Dekker and Verwer [5], I will construct new explicit energy or enstrophy conserving time discretizations. Combining these methods with any conservative space discretization yields *fully* discrete conservative methods that are guaranteed numerically stable.

Unfortunately, even these methods are not fully robust as they are based on symmetric space discretizations. These run into problems due to the indefinite intensification of vorticity gradients in inviscid turbulent flows (Pedlosky [13]). Oscillations will emerge near large vorticity gradients and ultimately lead to the so-called equipartitioning of enstrophy (Bennet and Haidvogel [3]). To suppress these oscillations an artificial viscosity is needed. For practical applications this is unsatisfactory since the minimally required amount of viscosity must be found by trial and error, while using a safe (large) amount leads to computational inefficiency.

In Section 3 an entirely different, second approach to robustness is described. The essentially non-oscillatory (ENO) technique of Shu and Osher [17] is used to construct "ENO-Jacobians." Only one variant is found that automatically introduces the minimally required damping near large vorticity gradients such that both numerical stability and accuracy are ensured. No "tuning" of an artificial viscosity is needed. A qualitative comparison is made with the Arakawa Jacobian.

2. CONSERVATION OF DISCRETE ENERGY AND ENSTROPHY

2.1. *The Semidiscrete System*

The semidiscrete system is obtained formally by space discretization of (1), (2) and will be denoted as:

$$\mathscr{M}\frac{d\mathbf{q}}{dt} = Q(\mathbf{q}, \boldsymbol{\Psi}, t), \tag{4}$$

$$\mathscr{L}\boldsymbol{\Psi} = \mathscr{M}\mathbf{q} - \mathbf{f}, \tag{5}$$

with $\mathscr{L}$ the discrete Laplacian and $\mathscr{M}$ the "mass matrix." For the semidiscrete variables a vector notation is used. For example, the vector $\boldsymbol{\Psi} \equiv \boldsymbol{\Psi}(t)$ is a continuous function of time and has the gridpoint values or modal amplitudes of the streamfunction as its components. The precise meaning of the components depends, of course, on the type of discretization method actually used.

It is assumed that (4), (5) conserves analogues of both energy and enstrophy in the inviscid unforced case, while an inner product $\langle \cdot, \cdot \rangle$ exists such that these quantities can be written as

$$V = \tfrac{1}{2}\langle \mathbf{q}, \mathscr{M}\mathbf{q} \rangle, \tag{6}$$

$$E = -\tfrac{1}{2}\langle \boldsymbol{\Psi}, \mathscr{M}\mathbf{q} - \mathbf{f} \rangle. \tag{7}$$

Further it is assumed that

$$dV/dt = \langle \mathbf{q}, Q \rangle, \tag{8}$$

$$dE/dt = -\langle \boldsymbol{\Psi}, Q \rangle, \tag{9}$$

and that the inner product satisfies

$$\langle \mathbf{A}, \mathcal{M}\mathbf{B}\rangle = \langle \mathcal{M}\mathbf{A}, \mathbf{B}\rangle \qquad \text{for all} \quad \mathbf{A}, \mathbf{B}.$$

Galerkin methods usually lead to a semidiscrete system with the above properties if the evaluation of the integrals occurring in these methods is exact. The use of quadrature formulas (such as lumping) is allowed if it is done in the way of Salmon and Talley [16].

However, for some types of boundary conditions it is impossible to arrive at a system with all of the above properties. An important example is the viscous system with "no-slip" boundary conditions, i.e., the normal derivative $\partial\psi/\partial n$ prescribed at parts of the boundary (Tezduyar *et al.* [18]). In this case the vorticity equation (4) does not hold for nodes corresponding to the no-slip boundary segments. There the evolution of $\mathbf{q}$ is obtained from (5) instead of (4): the streamfunction equation holds on these segments too. Note that the boundary components of $\mathbf{f}$ then depend on the prescribed $\partial\psi/\partial n$. Because Q is not fully defined at the boundary each nontrivial usage of boundary values of Q in inner products is invalid. So, assumption (8) cannot be expected to hold. Consequently, if no-slip boundary conditions are prescribed, the "enstrophy conserving" Runge–Kutta schemes derived below cannot be applied. The energy conserving schemes are still valid since they do not depend on the boundary values of Q, due to boundary condition (3).

Space discretizations with the abovementioned properties generally yield a $Q(\mathbf{q}, \Psi, t)$ that is differentiable infinitely many times with respect to its arguments. Therefore, the solution vectors $\mathbf{q}(t)$ and $\Psi(t)$ have the same property with respect to time t. This makes high order time discretizations of (4), (5) well founded in a formal sense.

2.2. *Energy or Enstrophy Conserving Time Discretizations*

Conservative nonlinear Runge–Kutta time discretizations for (4), (5) are derived by extending Example 10.3.8 of Dekker and Verwer [5]. They have shown how the coefficients of the classical fourth-order scheme can be modified such that (in the present terminology) enstrophy is conserved. This idea is extended to any standard Runge–Kutta method of arbitrarily high order of accuracy, such that it conserves either energy or enstrophy. Its main application is to explicit schemes but it applies to implicit ones as well.

The standard s-stage Runge–Kutta method is specified by constant coefficients a_{ij}, b_i, c_i, with $1 \leqslant i, j \leqslant s$, and yields for system (4), (5) the approximations $\mathbf{q}^{n+1} \approx \mathbf{q}(t_{n+1})$ and $\Psi^{n+1} \approx \Psi(t_{n+1})$:

$$\mathcal{M}\mathbf{q}^{n+1} = \mathcal{M}\mathbf{q}^n + \sum_{i=1}^{s} b_i \mathbf{F}_i, \tag{10}$$

$$\Psi^{n+1} = \mathcal{L}^{-1}(\mathcal{M}\mathbf{q}^{n+1} - \mathbf{f}), \tag{11}$$

where

$$\begin{aligned}
&\mathbf{F}_i = \Delta t\, Q(\mathbf{q}_i, \Psi_i, t_n + c_i\,\Delta t),\\
&\mathcal{M}\mathbf{q}_i = \mathcal{M}\mathbf{q}^n + \sum_{j=1}^{s} a_{ij}\mathbf{F}_j,\\
&\Psi_i = \mathcal{L}^{-1}(\mathcal{M}\mathbf{q}_i - \mathbf{f}) = \Psi^n + \sum_{j=1}^{s} a_{ij}\mathcal{L}^{-1}\mathbf{F}_j,\\
&t_{n+1} = t_n + \Delta t.
\end{aligned} \tag{12}$$

This scheme is explicit if $a_{ij} = 0$ for $j \geqslant i$. Note that (4), (5) is formally a so-called differential-algebraic system with perturbation index 1 (Hairer *et al.* [8]) for which the Runge–Kutta method (10)–(12) obtains its classical order of accuracy.

All inverses $\mathcal{L}^{-1}\mathbf{F}_j$ satisfy homogeneous Dirichlet boundary conditions, because at each stage the streamfunction Ψ_i must satisfy (3). In the viscous case with no-slip boundary conditions the notation involving $\mathcal{L}^{-1}$ is not very precise because (5) holds for boundary nodes too. So, formally, $\mathcal{L}$ is not invertible in this case. The notation as in (11) should then be interpreted as follows. The interior restriction of (5) determines Ψ^{n+1} completely because the interior restriction of $\mathcal{M}\mathbf{q}^{n+1}$ follows from (10). The unused equations in (5), together with (10), determine $\mathbf{q}^{n+1}$.

The standard scheme (10) is modified to

$$\mathcal{M}\mathbf{q}^{n+1} = \mathcal{M}\mathbf{q}^n + \gamma^n\,\Delta\mathbf{q}^n, \qquad \text{where} \quad \Delta\mathbf{q}^n \equiv \sum_{i=1}^{s} b_i\mathbf{F}_i, \tag{13}$$

which results in

$$\Psi^{n+1} = \Psi^n + \gamma^n\,\Delta\Psi^n,$$

$$\text{where} \quad \Delta\Psi^n \equiv \mathcal{L}^{-1}\,\Delta\mathbf{q}^n = \sum_{i=1}^{s} b_i\mathcal{L}^{-1}\mathbf{F}_i. \tag{14}$$

Obviously for $\gamma^n = 1$ the standard scheme (10) is recovered. The purpose of the form (13) is to make γ^n solution-dependent, such that either enstrophy or energy is conserved. Expressions for γ^n are derived as follows. The discrete energy E^{n+1} and enstrophy V^{n+1} at $t = t_{n+1}$ can be expressed in known quantities and γ^n by using (6), (7), (13), and (14):

$$V^{n+1} = V^n + \gamma^n\{\langle\mathbf{q}^n, \Delta\mathbf{q}^n\rangle + \tfrac{1}{2}\gamma^n\langle\Delta\mathbf{q}^n, \mathcal{M}^{-1}\Delta\mathbf{q}^n\rangle\}, \tag{15}$$

$$\begin{aligned}
E^{n+1} = E^n - \tfrac{1}{2}\gamma^n\{&\langle\Delta\Psi^n, \mathcal{M}\mathbf{q}^n - \mathbf{f}\rangle\\
&+ \langle\Psi^n, \Delta\mathbf{q}^n\rangle + \gamma^n\langle\Delta\Psi^n, \Delta\mathbf{q}^n\rangle\}.
\end{aligned} \tag{16}$$

For *inviscid unforced* flows either $V^{n+1} \equiv V^n$ or $E^{n+1} \equiv E^n$ can be enforced. Each of these requirements yields a quadratic equation for γ^n with nontrivial solution, respectively:

$$\gamma_V^n = -\frac{2\langle \mathbf{q}^n, \Delta \mathbf{q}^n \rangle}{\langle \Delta \mathbf{q}^n, \mathcal{M}^{-1} \Delta \mathbf{q}^n \rangle}, \tag{17}$$

$$\gamma_E^n = -\frac{\langle \Delta \Psi^n, \mathcal{M}\mathbf{q}^n - \mathbf{f} \rangle + \langle \Psi^n, \Delta \mathbf{q}^n \rangle}{\langle \Delta \Psi^n, \Delta \mathbf{q}^n \rangle}. \tag{18}$$

There are also trivial solutions $\gamma^n \equiv 0$, since for this γ^n Eq. (13) reduces to $\mathbf{q}^{n+1} = \mathbf{q}^n$. Note that the applicability of (17) and (18) is not limited to Runge–Kutta type methods.

For *viscous and/or forced* flows the γ^n given in (17) and (18) cannot be used since they would (unjustly) lead to exact enstrophy or energy conservation in these cases also. Alternatives for (17), (18) that are accurate in these cases too are obtained by generalizing the approach of Dekker and Verwer [5]. Combining the identities

$$\langle \mathbf{q}^n, \Delta \mathbf{q}^n \rangle = \sum_{i=1}^{s} b_i \langle \mathbf{q}_i, \mathbf{F}_i \rangle - \sum_{i=1}^{s} b_i \langle \mathbf{q}_i - \mathbf{q}^n, \mathbf{F}_i \rangle,$$

$$\langle \Psi^n, \Delta \mathbf{q}^n \rangle = \sum_{i=1}^{s} b_i \langle \Psi_i, \mathbf{F}_i \rangle - \sum_{i=1}^{s} b_i \langle \Psi_i - \Psi^n, \mathbf{F}_i \rangle$$

with (15) and (16) gives

$$V^{n+1} = V^n + \gamma^n \sum_{i=1}^{s} b_i \langle \mathbf{q}_i, \mathbf{F}_i \rangle - \tfrac{1}{2}\gamma^n \mathcal{Q}_V(\gamma^n), \tag{19}$$

$$E^{n+1} = E^n - \gamma^n \sum_{i=1}^{s} b_i \langle \Psi_i, \mathbf{F}_i \rangle + \tfrac{1}{2}\gamma^n \mathcal{Q}_E(\gamma^n), \tag{20}$$

where

$$\mathcal{Q}_V(\gamma^n) \equiv 2 \sum_{i=1}^{s} b_i \langle \mathbf{q}_i - \mathbf{q}^n, \mathbf{F}_i \rangle - \gamma^n \langle \Delta \mathbf{q}^n, \mathcal{M}^{-1} \Delta \mathbf{q}^n \rangle, \tag{21}$$

$$\mathcal{Q}_E(\gamma^n) \equiv \langle \Psi^n, \Delta \mathbf{q}^n \rangle - \langle \Delta \Psi^n, \mathcal{M}\mathbf{q}^n - \mathbf{f} \rangle + 2 \sum_{i=1}^{s} b_i \langle \Psi_i - \Psi^n, \mathbf{F}_i \rangle - \gamma^n \langle \Delta \Psi^n, \Delta \mathbf{q}^n \rangle. \tag{22}$$

The zeros of the functions $\mathcal{Q}_V$ and $\mathcal{Q}_E$ are, respectively,

$$\gamma_V^n = 1 + \frac{\mathcal{Q}_V(1)}{\langle \Delta \mathbf{q}^n, \mathcal{M}^{-1} \Delta \mathbf{q}^n \rangle}, \tag{23}$$

$$\gamma_E^n = 1 + \frac{\mathcal{Q}_E(1)}{\langle \Delta \Psi^n, \Delta \mathbf{q}^n \rangle}. \tag{24}$$

Therefore Eqs. (19) and (20) become

$$V^{n+1} = V^n + \gamma_V^n \sum_{i=1}^{s} b_i \langle \mathbf{q}_i, \mathbf{F}_i \rangle, \qquad \text{if} \quad \gamma^n = \gamma_V^n, \tag{25}$$

$$E^{n+1} = E^n - \gamma_E^n \sum_{i=1}^{s} b_i \langle \Psi_i, \mathbf{F}_i \rangle, \qquad \text{if} \quad \gamma^n = \gamma_E^n. \tag{26}$$

These still reduce to enstrophy or energy conservation in the inviscid unforced case because then (8), (9) result in

$$\left.\begin{aligned} \langle \mathbf{q}, Q(\mathbf{q}, \Psi, t) \rangle = 0 \\ \langle \Psi, Q(\mathbf{q}, \Psi, t) \rangle = 0 \end{aligned}\right\} \quad \text{for all } \mathbf{q}, \Psi. \tag{27}$$

Also, in this case expressions (17) and (23) are equivalent, as are (18) and (24). However, (23) and (24) may be preferable since they seem less sensitive to rounding errors: their numerator and denominator both contain only terms of $\mathcal{O}(\Delta t^2)$, whereas the numerators in (17), (18) contain $\mathcal{O}(\Delta t)$ terms also (which cancel in exact arithmetic). For viscous and/or forced flows Eqs. (25) and (26) are accurate Runge–Kutta type approximations of the continuous enstrophy and energy behavior, respectively, if $\gamma^n \approx 1$; see also Eq. (29) and Remark 2 below. For viscous unforced flows the semidiscrete system is usually monotonous [5] in the sense that properties (27) change into

$$\left.\begin{aligned} \langle \mathbf{q}, Q(\mathbf{q}, \Psi, t) \rangle \leqslant 0 \\ \langle \Psi, Q(\mathbf{q}, \Psi, t) \rangle \geqslant 0 \end{aligned}\right\} \quad \text{for all } \mathbf{q}, \Psi. \tag{28}$$

Usage of (23) or (24) for γ^n then yields $V^{n+1} \leqslant V^n$ and $E^{n+1} \leqslant E^n$, respectively; i.e., the discrete enstrophy or energy decreases as in the analytic system. However, this is true only if each b_i in (13) is nonnegative and the timestep is limited such that $\gamma^n > 0$, which is a sensible restriction anyway (Dekker and Verwer [5]).

Next, the accuracy of the nonlinear schemes (13), (23) and (13), (24) is considered.

LEMMA. *Each standard Runge–Kutta method* (10)–(12) *that is rth-order accurate for systems of equations satisfies*: $\mathcal{Q}_V(1) = \mathcal{O}(\Delta t^{r+1})$, *as well as* $\mathcal{Q}_E(1) = \mathcal{O}(\Delta t^{r+1})$.

Proof. Apply one step of the standard s-stage Runge–Kutta scheme to the extended system,

$$\frac{d}{dt}\begin{pmatrix} \mathcal{M}\mathbf{q} \\ v \\ e \end{pmatrix} = \begin{pmatrix} Q(\mathbf{q}, \Psi, t) \\ \langle \mathbf{q}, Q(\mathbf{q}, \Psi, t) \rangle \\ -\langle \Psi, Q(\mathbf{q}, \Psi, t) \rangle \end{pmatrix}, \qquad t \geqslant t_n,$$

with initial condition,

$$\begin{pmatrix} \mathbf{q}(t_n) \\ v(t_n) \\ e(t_n) \end{pmatrix} = \begin{pmatrix} \mathbf{q}^n \\ \tfrac{1}{2}\langle \mathbf{q}^n, \mathcal{M}\mathbf{q}^n \rangle \\ -\tfrac{1}{2}\langle \Psi^n, \mathcal{M}\mathbf{q}^n - \mathbf{f} \rangle \end{pmatrix}.$$

This yields

$$\begin{pmatrix} \mathcal{M}\mathbf{q}^{n+1} \\ v^{n+1} \\ e^{n+1} \end{pmatrix} = \begin{pmatrix} \mathcal{M}\mathbf{q}^{n} \\ \frac{1}{2}\langle \mathbf{q}^n, \mathcal{M}\mathbf{q}^n \rangle \\ -\frac{1}{2}\langle \Psi^n, \mathcal{M}\mathbf{q}^n - \mathbf{f} \rangle \end{pmatrix} + \sum_{i=1}^{s} b_i \begin{pmatrix} \mathbf{F}_i \\ \langle \mathbf{q}_i, \mathbf{F}_i \rangle \\ -\langle \Psi_i, \mathbf{F}_i \rangle \end{pmatrix}, \tag{29}$$

with $\mathbf{q}_i$, Ψ_i, and $\mathbf{F}_i$ given by (12). Due to assumptions (8), (9) the exact solutions of the additional equations are

$$v(t) = \tfrac{1}{2}\langle \mathbf{q}(t), \mathcal{M}\mathbf{q}(t) \rangle,$$
$$e(t) = -\tfrac{1}{2}\langle \Psi(t), \mathcal{M}\mathbf{q}(t) - \mathbf{f} \rangle.$$

The Runge–Kutta scheme is assumed to be rth-order accurate, which means that $\mathbf{q}^{n+1}$, Ψ^{n+1}, v^{n+1}, and e^{n+1} can only deviate by $\mathcal{O}(\Delta t^{r+1})$ from the exact solution $\mathbf{q}(t_{n+1})$, $\Psi(t_{n+1})$, $v(t_{n+1})$, $e(t_{n+1})$. This can be exploited as

$$\begin{aligned} v^{n+1} &= v(t_{n+1}) + \mathcal{O}(\Delta t^{r+1}) \\ &\equiv \tfrac{1}{2}\langle \mathbf{q}(t_{n+1}), \mathcal{M}\mathbf{q}(t_{n+1}) \rangle + \mathcal{O}(\Delta t^{r+1}) \\ &= \tfrac{1}{2}\langle \mathbf{q}^{n+1}, \mathcal{M}\mathbf{q}^{n+1} \rangle + \mathcal{O}(\Delta t^{r+1}), \end{aligned}$$

and

$$\begin{aligned} e^{n+1} &= e(t_{n+1}) + \mathcal{O}(\Delta t^{r+1}) \\ &\equiv -\tfrac{1}{2}\langle \Psi(t_{n+1}), \mathcal{M}\mathbf{q}(t_{n+1}) - \mathbf{f} \rangle + \mathcal{O}(\Delta t^{r+1}) \\ &= -\tfrac{1}{2}\langle \Psi^{n+1}, \mathcal{M}\mathbf{q}^{n+1} - \mathbf{f} \rangle + \mathcal{O}(\Delta t^{r+1}). \end{aligned}$$

Together with (29) these relations give

$$\begin{aligned} &\tfrac{1}{2}\langle \mathbf{q}^{n+1}, \mathcal{M}\mathbf{q}^{n+1} \rangle \\ &\quad = \tfrac{1}{2}\langle \mathbf{q}^n, \mathcal{M}\mathbf{q}^n \rangle + \sum_{i=1}^{s} b_i \langle \mathbf{q}_i, \mathbf{F}_i \rangle + \mathcal{O}(\Delta t^{r+1}), \\ &-\tfrac{1}{2}\langle \Psi^{n+1}, \mathcal{M}\mathbf{q}^{n+1} - \mathbf{f} \rangle \\ &\quad = -\tfrac{1}{2}\langle \Psi^n, \mathcal{M}\mathbf{q}^n - \mathbf{f} \rangle - \sum_{i=1}^{s} b_i \langle \Psi_i, \mathbf{F}_i \rangle + \mathcal{O}(\Delta t^{r+1}). \end{aligned}$$

Combining these relations with (19), (20) for $\gamma^n = 1$ proves the lemma. ∎

Since $\langle \Delta\mathbf{q}^n, \mathcal{M}^{-1}\Delta\mathbf{q}^n \rangle = \mathcal{O}(\Delta t^2)$ and $\langle \Delta\Psi^n, \Delta\mathbf{q}^n \rangle = \mathcal{O}(\Delta t^2)$ it follows from (23), (24) and the lemma that

$$\gamma_V^n = 1 + \mathcal{O}(\Delta t^{r-1}), \qquad \gamma_E^n = 1 + \mathcal{O}(\Delta t^{r-1}). \tag{30}$$

These estimates are the sharpest possible because the estimates in the proof of the lemma are sharp. They have been verified by numerical refinement sequences for the semidiscrete system described in Section 2.4. These show that the nonlinear versions of the fourth-order Kutta–Simpson three-eighth rule [5] as well as the classical fourth-order Runge–Kutta scheme lead to (30) with $r = 4$. The fifth-order components of the (4,5)-pairs of England and Fehlberg [11] both yield (30) with $r = 5$.

The order of accuracy of the nonlinear schemes (13), (23) and (13), (24) can be derived from (30) as follows. The Taylor expansion of the local truncation error of a Runge–Kutta scheme with coefficients $\hat{a}_{ij}$, $\hat{b}_i$, $\hat{c}_i$, applied with a stepsize $t_{n+1} - t_n = \widehat{\Delta t}$, is

$$\begin{aligned} &\mathcal{M}\mathbf{q}(t_{n+1}) - \mathcal{M}\mathbf{q}^{n+1} \\ &\quad = \widehat{\Delta t}\left(1 - \sum_{i=1}^{s} \hat{b}_i\right) \mathcal{Q}(\mathbf{q}^n, \Psi^n, t_n) \\ &\qquad + \sum_{i=2}^{\infty} \widehat{\Delta t}^{\,i}\, \mathbf{P}^{(i)}(\hat{a}, \hat{b}, \hat{c}). \end{aligned}$$

Herein $\mathbf{P}^{(i)}$ contains derivatives of the solution at $t = t_n$, and is polynomial in the Runge–Kutta coefficients $\hat{a}_{ij}$, $\hat{b}_i$, $\hat{c}_i$; see, for example, Lambert [11]. The standard Runge–Kutta method (10)–(12) is assumed to be rth-order accurate, which means that the local truncation error is $\mathcal{O}(\Delta t^{r+1})$. The constant coefficients of the standard scheme therefore satisfy

$$\begin{aligned} &\mathbf{P}^{(i)}(a, b, c) = 0, \qquad \text{for} \quad 2 \leqslant i \leqslant r, \\ &1 - \sum_{i=1}^{s} b_i = 0. \end{aligned} \tag{31}$$

The accuracy of the nonlinear modification (13) depends on the definition of t_{n+1}. There are two relevant possibilities (Dekker and Verwer [5]):

1. The standard definition in (12) is retained: $t_{n+1} = t_n + \Delta t$. So, compared to the Taylor expansion of the standard scheme, only the change $\hat{b}_i = \gamma^n b_i$ is made. With (30), (31), and the elementary properties of polynomials we find

$$\begin{aligned} &\mathbf{P}^{(i)}(a, \hat{b}, c) = \mathcal{O}(\Delta t^{r-1}), \qquad \text{for} \quad 2 \leqslant i \leqslant r, \\ &1 - \sum_{i=1}^{s} \hat{b}_i = \mathcal{O}(\Delta t^{r-1}). \end{aligned}$$

The local truncation error is, therefore, $\mathcal{O}(\Delta t^r)$, leading to a global order of accuracy of only $r - 1$.

2. The stepsize is modified by γ^n:

$$t_{n+1} = t_n + \gamma^n \Delta t. \tag{32}$$

Note that the Taylor expansion is done with respect to $\widehat{\Delta t} = \gamma^n \Delta t$. The nonlinear scheme must therefore be rewritten in terms of $\widehat{\Delta t}$ in order to obtain the coefficients as they occur in the expansion. An inspection of (11)–(13) yields simply

$$\hat{a}_{ij} = a_{ij}/\gamma^n = a_{ij} + \mathcal{O}(\widehat{\Delta t}^{r-1}),$$
$$\hat{b}_i = b_i,$$
$$\hat{c}_i = c_i/\gamma^n = c_i + \mathcal{O}(\widehat{\Delta t}^{r-1}).$$

This gives with (31)

$$\mathbf{P}^{(i)}(\hat{a}, \hat{b}, \hat{c}) = \mathcal{O}(\widehat{\Delta t}^{r-1}), \qquad \text{for} \quad 2 \leqslant i \leqslant r,$$
$$1 - \sum_{i=1}^{s} \hat{b}_i = 0,$$

which leads to a local truncation error of $\mathcal{O}(\Delta t^{r+1})$ and a global error of $\mathcal{O}(\Delta t^r)$.

As already concluded earlier by Dekker and Verwer [5] for their enstrophy conserving version of the classical Runge–Kutta scheme, the order of the nonlinearly modified schemes remains high only if the new interpretation (32) is given to the discrete variables $\mathbf{q}^{n+1} \approx \mathbf{q}(t_{n+1})$ and $\Psi^{n+1} \approx \Psi(t_{n+1})$.

The fully discrete nonlinear schemes (13), (23) and (13), (24) exactly conserve enstrophy and energy, respectively, in the inviscid unforced case and are, therefore, unconditionally stable with respect to "enstrophy/energy-norms." This is true even for the explicit variants. However, in practice the timestep should not be taken too large because the explicit nonlinear schemes yield inaccurate (but stable) results as soon as the standard explicit scheme (10) on which they are based would become unstable, see Section 2.4.

Remark 1. In some situations the expressions for γ_E^n can be simplified. The first two terms in (22) cancel if $\mathscr{L}$ is symmetric and the inner product is Euclidian. If these conditions hold, also (18) may be simplified to

$$\gamma_E^n = -\frac{2\langle \Psi^n, \Delta \mathbf{q}^n \rangle}{\langle \Delta \Psi^n, \Delta \mathbf{q}^n \rangle}.$$

Remark 2. The additional local truncation error introduced by the nonlinear modification (13) in the $\mathbf{q}$-evolution is proportional to $|1-\gamma^n|\,\Delta t^2 \equiv \mathcal{O}(\Delta t^{r+1})$ if (32) is used, otherwise it is proportional to $|1-\gamma^n|\,\Delta t \equiv \mathcal{O}(\Delta t^r)$. A disadvantage of the nonlinear schemes seems to be that this error occurs instantly throughout the whole spatial domain even in the case that large temporal discretization errors would be spatially localized in the standard scheme (10). However, any error in (10) has an instantaneous global effect, too, through the streamfunction equations in (11) and (12). Further note that the standard scheme is not accurate either if $|1-\gamma^n|$ is "large"; (23), (24) show that $|1-\gamma^n|$ measures the deviation of the standard scheme's enstrophy/energy evolution (given by (19), (20) with $\gamma^n \equiv 1$) from a direct Runge–Kutta approximation (29) of the semi-discrete enstrophy/energy evolution. If $|1-\gamma^n|$ is large then the standard scheme may even become unstable. By using one of the nonlinear variants (13), (23) or (13), (24) stability can still be guaranteed but a higher accuracy can not (except for accuracy of energy and/or enstrophy evolutions).

Remark 3. Given the asymptotic behavior (30) of the γ^n-parameters it seems straightforward to implement an automatic timestep adaptation that reduces the error that is proportional to $|1-\gamma^n|$ to a specified tolerance at no extra cost. Maximization of the timestep (in order to reach an equilibrium quickly) seems possible by increasing it until, say, $|1-\gamma^n| \approx 0.01$.

Remark 4. For inviscid unforced flows expressions (17), (18) also satisfy (30). This can be proven for any rth order time-discretization using a simplified version of the lemma.

2.3. *Concurrent Conservation of Energy and Enstrophy*

An attempt has been made to construct explicit nonlinear Runge–Kutta methods that conserve both energy and enstrophy in the inviscid unforced case. Conservation of two quantities seems to require the introduction of two solution-dependent coefficients γ_1, γ_2. The investigated schemes all had the form

$$\mathscr{M}\mathbf{q}^{n+1} = \mathscr{M}\mathbf{q}^n + \gamma_1 \Delta \mathbf{q}_1^n + \gamma_2 \Delta \mathbf{q}_2^n, \tag{33}$$

where $\Delta \mathbf{q}_1^n$ and $\Delta \mathbf{q}_2^n$ are different summations of the $\mathbf{F}_i$, while their sum equals the $\Delta \mathbf{q}^n$ given in (13). The standard Runge–Kutta formula (10) is recovered for $\gamma_1 = \gamma_2 = 1$. It can be shown that the discrete enstrophy V^{n+1} and energy E^{n+1} at $t = t_{n+1}$ concurrently satisfy relations similar to (25), (26) if γ_1 and γ_2 satisfy a coupled system of two quadratic scalar equations of the form:

$$\begin{aligned} A_1(\gamma_1)^2 + A_2\gamma_1\gamma_2 + A_3(\gamma_2)^2 + A_4\gamma_1 + A_5\gamma_2 = 0, \\ B_1(\gamma_1)^2 + B_2\gamma_1\gamma_2 + B_3(\gamma_2)^2 + B_4\gamma_1 + B_5\gamma_2 = 0. \end{aligned} \tag{34}$$

As a result, fully discrete energy as well as enstrophy are conserved in the inviscid unforced case. Both diminish if the semidiscrete system is monotonous in the sense of (28).

The derivation and detailed form of system (34) have been omitted since it turns out to become ill-conditioned in actual simulations of the semidiscrete system described in Section 2.4. This was observed for various types of splittings (33) of the fourth- and fifth-order Runge–Kutta

methods mentioned earlier. Therefore, nonlinear schemes of the type (33), (34) do not seem of practical use.

Remark 1. The trivial solution of system (34) is, as anticipated, $\gamma_1 = \gamma_2 = 0$. The more relevant solution $\gamma_1 \approx \gamma_2 \approx 1$ among the three remaining root-pairs has been determined in two ways. By eliminating (for instance) γ_1, a cubic equation for γ_2 is obtained which can be solved analytically. This approach is extremely sensitive to rounding errors. The second method, Newton iteration applied directly on system (34), is less prone to rounding errors but still tends to diverge. All calculations were done with 16 significant decimal digits.

Remark 2. Refinement sequences suggest that if (33), (34) is based on fourth-order Runge–Kutta methods, $\gamma_1, \gamma_2 = 1 + \mathcal{O}(\Delta t^2)$, while fifth-order methods lead to $\gamma_1, \gamma_2 = 1 + \mathcal{O}(\Delta t^3)$. This shows another disadvantage of scheme (33), (34). It will only be second- or third-order accurate, respectively, because a redefinition of the fully discrete variables such as (32) does not seem possible.

2.4. *Numerical Results*

The performance of the nonlinear energy or enstrophy conserving variants of the classical fourth-order Runge–Kutta (RK) scheme has been investigated with a test problem similar to one used earlier by Arakawa [2]. The inviscid unforced system (1)–(3) with $f \equiv 0$ is solved on the unit square $\Omega = (0, 1) \times (0, 1)$. The initial streamfunction is

$$\psi(x, y, t = 0) = -C_0 \sin 8\pi x \{\sin 8\pi y + \tfrac{1}{10} \sin 16\pi y\}, \tag{35}$$

where the positive constant $C_0 \approx 1/1390.8$ is chosen such that initially $|q|_\infty = 1$.

The spatial domain Ω is covered by an equidistant set of gridpoints $\{(x_i, y_j)\}$, with $x_i \equiv i\,\Delta x$ and $y_j \equiv j\,\Delta y$, where Δx and Δy denote the given gridpoint distances. The employed semidiscrete form (4) of the vorticity equation is that derived in Example A of Salmon and Talley [16]. It consists of the second-order Arakawa Jacobian in interior points and a compatible approximation at all boundary points, while the mass matrix $\mathcal{M}$ is the identity operator. Further, for $\mathcal{L}$ the standard five-point (second-order) discrete Laplacian operator was chosen. With these approximations all assumptions of Section 2.1 are valid.

The initial value of the discrete vorticity is obtained by applying the discrete Laplacian to (35). At the boundary $\partial\Omega$ its *initial value* is set to the exact initial value (i.e., zero). The exact vorticity remains zero at the boundary, the discrete approximation however does not. During most of the simulations described in this section the CFL number, defined as

$$\mathrm{CFL} = \max \left\{ \frac{|\psi_y|}{\Delta x} + \frac{|\psi_x|}{\Delta y} \right\} \Delta t,$$

is fixed at the constant value 2. This definition gives a safe overestimate for Arakawa's Jacobian. All experiments are performed on a grid of 257×257 points; i.e., $\Delta x = \Delta y = 1/256$. Similar behavior of the RK methods has been observed on grids of 129×129 and 65×65 points. At each stage of the RK methods the Poisson problem involving $\mathcal{L}$ is solved with a multigrid technique.

The result obtained with the energy conserving RK discretization (13), (24) is given in Fig. 1. It shows the familiar filamentation and corresponding gradient intensification of vorticity (McWilliams [12]). As soon as the filaments cannot be resolved any more, large oscillations occur, ultimately leading to enstrophy equipartitioning (Bennett and Haidvogel [3]). This is reflected by Fig. 2 which shows that the maximum norm $|q|_\infty$ of discrete vorticity increases significantly. The exact inviscid solution keeps this norm constant, of course. Improvement of the numerical behavior will be discussed at the end of this section. Note that in contrast with Fig. 1, the streamfunction field becomes organized in some relatively smooth and large circulation cells (not shown here), i.e., energy indeed cascades mainly to the large scales of motion.

Results obtained with the classical RK scheme (10) itself and the enstrophy conserving variant (13), (23) are similar to the ones shown here. Differences are observed in the evolution of discrete energy and enstrophy (Fig. 3). At $\mathrm{CFL} = 2$ the intrinsic dissipation of the classical RK scheme is significant at short lengthscales. This affects enstrophy much more than energy since the latter is concentrated at large lengthscales. The vertical axis in the energy plot has been magnified to make the slight energy decrease visible. The intrinsic dissipation becomes noticeable as soon as vorticity filaments (and oscillations) are formed (at $t \approx 130$). The energy conserving RK method, of course, fully compensates the slight energy dissipation of the classical scheme, thereby also reducing dissipation of enstrophy. The enstrophy conserving RK scheme compensates the large enstrophy dissipation of the classical scheme with a larger value of its γ'' parameter than needed to conserve energy: γ''_V is usually larger than γ''_E; see Fig. 4. This may explain the slight increase in discrete energy observed in the result of this scheme. It is reminiscent of behavior of the solely enstrophy conserving *space* discretization reported by Arakawa [2].

For all RK schemes shown in Fig. 3, the deviations from conservation diminish with fourth-order convergence for decreasing CFL number. For example, at $\mathrm{CFL} = 0.5$ the energy conserving RK scheme shows a reduction in

enstrophy of less than 0.02% at $t = 460$. Furthermore, the γ^n parameters behaved as $1 + \mathcal{O}(\mathrm{CFL}^3)$: the evolution for $\mathrm{CFL} = 2$ is given in Fig. 4, while at $\mathrm{CFL} = 0.5$ the maximum deviation from unity is less than 4×10^{-5}. This illustrates that the semidiscrete system (4), (5) is C^∞ in t, and that high order time discretizations obtain their theoretical order of accuracy even though the vorticity is *spatially* not smooth at all (Fig. 1).

Both implicit and explicit nonlinear RK methods are stable independent of the choice of Δt (actually only the CFL number is relevant). The question therefore arises if the explicit nonlinear variants still yield accurate results

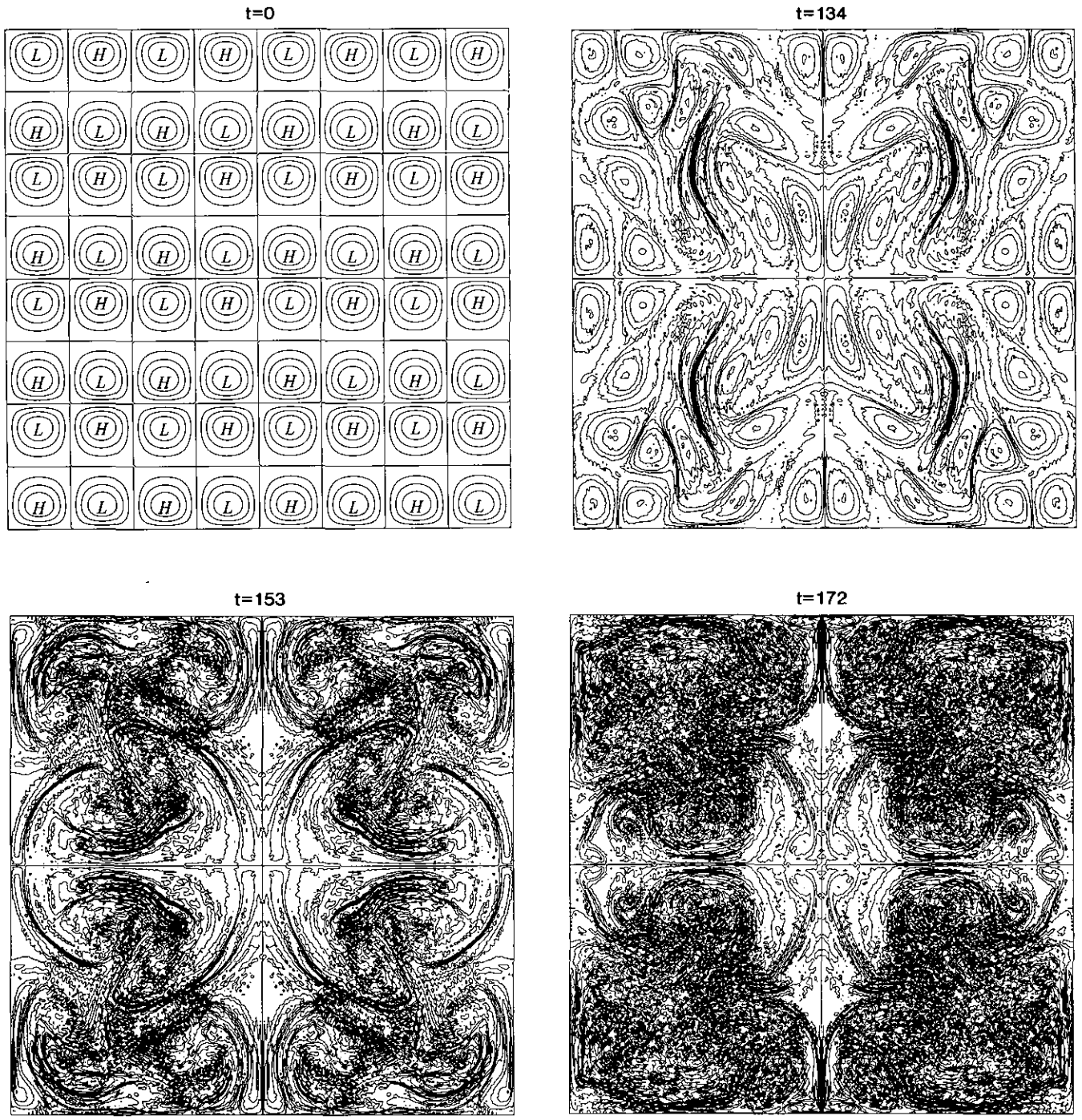


FIG. 1. The inviscid vorticity evolution obtained with Arakawa's Jacobian and the energy conserving nonlinear variant of the classical fourth-order Runge–Kutta scheme. Resolution is 257×257, $\mathrm{CFL} = 2$. The contour interval is $\frac{1}{4}$; the zero contour is included.

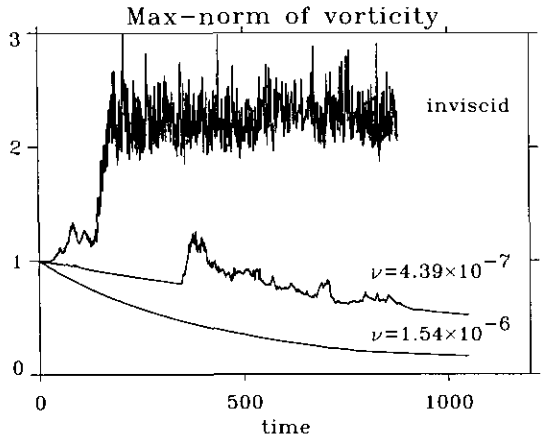


FIG. 2. Evolution of the maximum norm $|q|_\infty$ of vorticity for the inviscid case given in Fig. 1 and for two cases with artificial viscosity added: $\nu = 1.54 \times 10^{-6}$ and $\nu = 4.39 \times 10^{-7}$.

when the CFL number is so large that the standard RK scheme (10) is no longer stable. Repeating the experiment at CFL = 4 and CFL = 8 reveals that this is not the case:

• At CFL = 8 the classical RK scheme increases enstrophy by more than a factor of 10 after the first four timesteps. It remains approximately at this level until a fatal "explosion" occurs at $t \approx 155$. Already after the first four

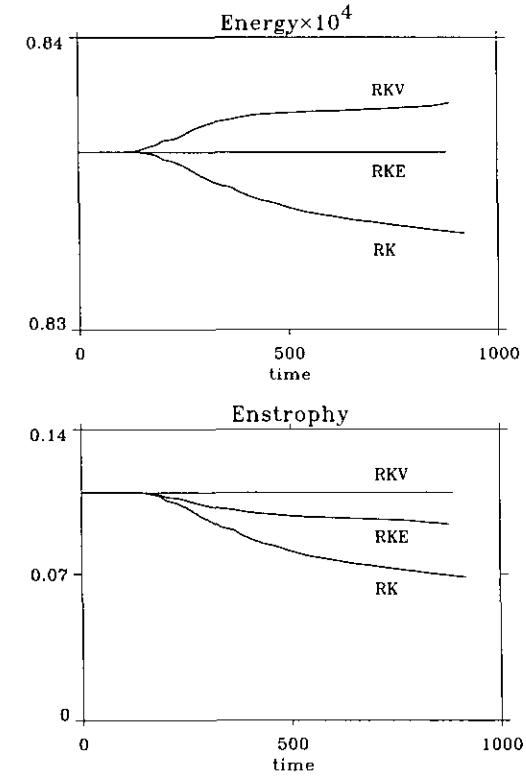


FIG. 3. Energy and enstrophy evolution for the inviscid case given in Fig. 1, obtained with the classical Runge–Kutta scheme (RK), its energy-conserving variant (RKE), and its enstrophy conserving variant (RKV). The vertical scale of the energy plot is magnified.

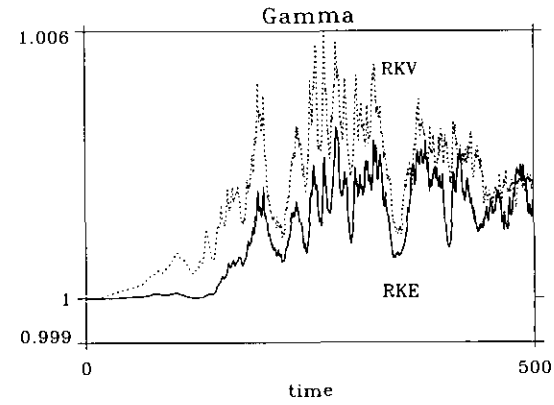


FIG. 4. The γ_V^n (dotted line) and γ_E^n parameters of the enstrophy and energy conserving schemes, respectively, of Fig. 3.

timesteps both nonlinear RK schemes generate a value of their parameter γ^n of approximately -0.03. Then, with alternating signs, each γ^n rapidly approaches zero. This means that the schemes end in artificial steady states. Continuing the integrations at this CFL number is therefore senseless.

• At CFL = 4 the instability of the classical RK scheme is quite mild; it occurs in the form of bounded bursts especially in the vorticity extreme values ($|q|_\infty \leqslant 17$). The energy remains fairly constant (relative deviations are within 1.55%), while the enstrophy decreases globally, showing only some minor bursts; see Fig. 5. The parameters γ^n of both nonlinear RK schemes deviate significantly from unity (Fig. 6). The enstrophy conserving RK scheme displays the largest deviations. This scheme keeps energy nicely constant (relative deviations are within 1.07%) and slightly reduces the magnitude of the bursts in the vorticity extreme values ($|q|_\infty \leqslant 14$). The energy conserving RK scheme initially behaves much like the classical scheme, but at $t \approx 475$ it increases the magnitude of the vorticity bursts by a factor of 5. The enstrophy increases correspondingly in a peculiar

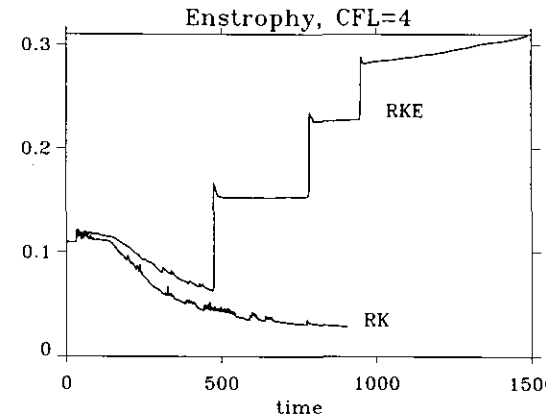


FIG. 5. Enstrophy evolution for the inviscid case of Fig. 1 at CFL = 4, obtained with the classical Runge–Kutta scheme (RK) and its energy conserving variant (RKE).

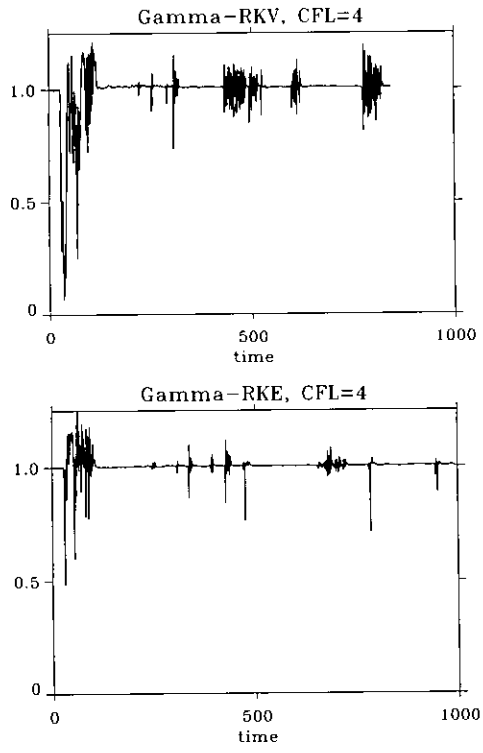


FIG. 6. The γ_V^n and γ_E^n parameters of the enstrophy and energy conserving Runge–Kutta schemes, respectively, (RKV and RKE) at CFL = 4 for the inviscid case of Fig. 1.

step-wise manner (Fig. 5). Each sudden increase of enstrophy coincides with a burst in $|1-\gamma_E^n|$, see Fig. 6. Conservation of energy ensures that enstrophy remains bounded (Sadourny [15]). However, a sufficiently large increase leads to an incorrect spectral distribution of conserved energy (Arakawa [2]).

So, neither the standard nor the nonlinear explicit RK schemes can be applied sensibly at very large CFL numbers. The nonlinear variants suppress the explosive instability of the standard scheme, but are inaccurate at large CFL numbers. Fortunately, this inaccuracy is signalled by large deviations of the γ^n-parameters from unity. In practice, the timestep should be reduced such that $\gamma^n \approx 1$, see also Remarks 2 and 3 in Section 2.2.

It is tempting to use the "CFL = 4"-results to draw an analogy with Arakawa [2] and Sadourny [15], who showed that a *space* discretization that conserves only enstrophy is preferable to a scheme conserving only energy. This conclusion seems to translate to time discretization schemes conserving only a single quantity as well.

Independent of the time discretization large numerical (spatial) oscillations emerge. A common way to suppress these is the introduction of an artificial viscous term in (1):

$$\frac{\partial q}{\partial t} + J(\psi, q) = \nu \nabla^2 q + f, \tag{36}$$

As outlined in the Introduction, the remaining problem is to find the smallest allowable value of viscosity ν for a given spatial resolution Δx. A safe value is of course obtained from a cell-Reynolds condition ($\nu \sim \Delta x$), but this would lead to an extremely large dissipation. Less safe estimates, which yield an order of magnitude smaller viscosity, can be obtained from Henshaw *et al.* [10]. Their estimate is only applicable to Eq. (36) with $f(x, y) \equiv 0$ and in absence of a forcing term. For the initial condition (35) it gives

$$\nu \geqslant \pi^{-2} C_1 \, \Delta x^2 \tag{37}$$

with C_1 a bounded *unknown* constant for which Browning and Kreiss [4] suggest $C_1 \approx 1$.

Simulations of the viscous system with initial condition (35) are described in order to show that bounds such as (37) seem too optimistic and, also, to enable a qualitative comparison with the ENO-Jacobians of the next section.

An additional boundary condition is required by (36). Here the exact vorticity boundary value of the inviscid solution is enforced:

$$q|_{\partial\Omega} \equiv 0. \tag{38}$$

The employed semidiscrete system approximating the viscous problem is similar to the one used for the inviscid problem discussed earlier. The viscous term in (36) has been

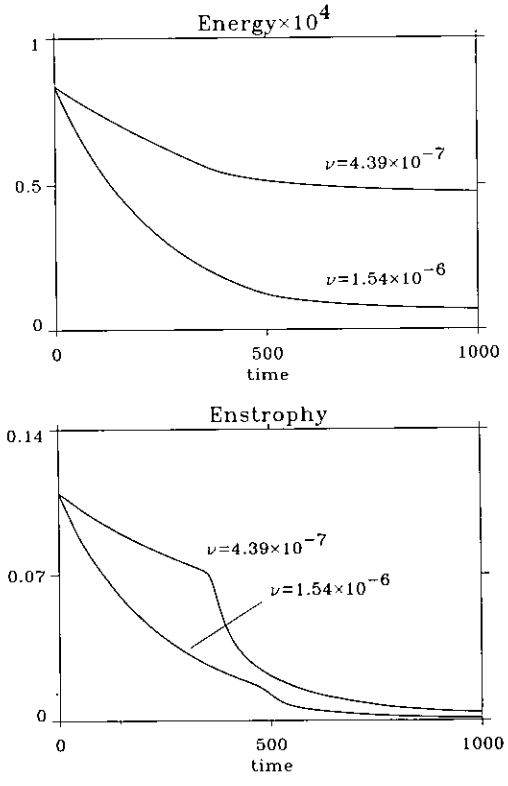


FIG. 7. Evolution of energy and enstrophy at CFL = 2, obtained with the Arakawa Jacobian and the energy conserving Runge–Kutta scheme for two values of the artificial viscosity $\nu = 1.54 \times 10^{-6}$ and $\nu = 4.39 \times 10^{-7}$

discretized by the standard five-point discrete Laplacian $\mathscr{L}$. It can be shown that the resulting semidiscrete system is monotonous in the sense of (28). The time-stepping is performed with the "energy conserving" variant (13), (24) of the classical fourth-order Runge–Kutta scheme. The increased spatial smoothness of the solution results in smaller deviations of γ_E^n from unity at CFL = 2 (less than 3×10^{-4} for the cases described below).

Figures 2 and 7 show that usage of $C_1 = 1$ ($\nu = 1.54 \times 10^{-6}$) leads to excessive damping of the solution. Using a "tuned" value of approximately 0.3 ($\nu = 4.39 \times 10^{-7}$) gives noticeable oscillations during the filamentation and is probably already too small, see Figs. 2 and 8. Interestingly, the introduction of the artificial viscosity postpones the formation of filaments. In the inviscid case they start to form around $t = 130$. In the

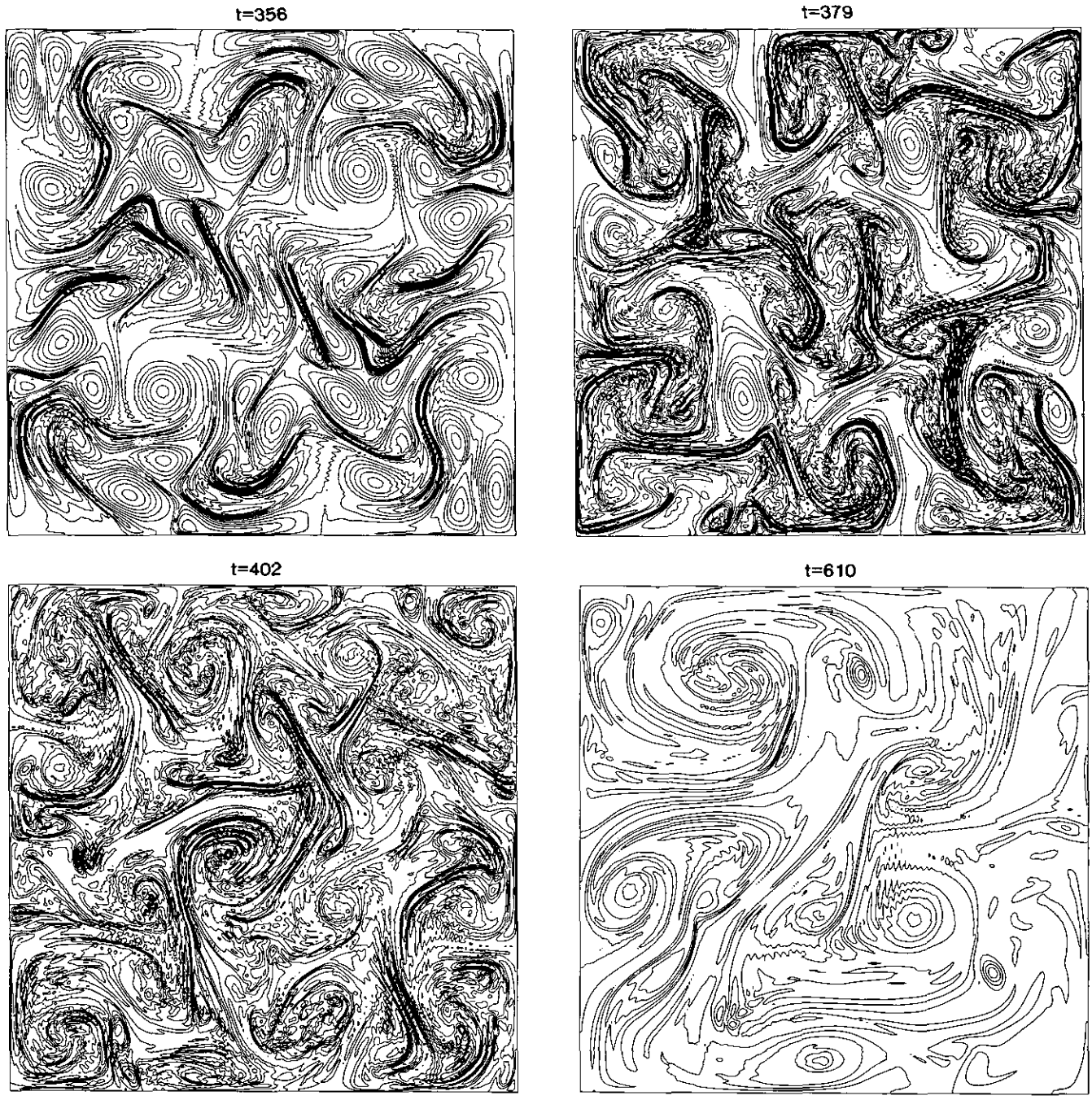


FIG. 8. Vorticity evolution at CFL = 2 obtained with Arakawa's Jacobian, the energy conserving Runge–Kutta scheme, and a "tuned" artificial viscosity $\nu = 4.39 \times 10^{-7}$. Resolution is 257×257. Contour interval is 1/8; zero contour included. The initial condition is that of Fig. 1.

viscous simulation the "cellular" structure of the initial condition (shown in the first panel of Fig. 1) is present roughly up to $t=320$ and is already prone to significant dissipation (Fig. 7). During the filamentation the discrete enstrophy starts to decrease sharply, as expected (Pedlosky [13]). Even for a very small viscosity ν the filamentation is postponed. Therefore this effect is probably caused by the "viscous boundary condition" (38).

In the viscous simulations the antisymmetry of the exact solutions with respect to the lines $x=\frac{1}{2}$ and $y=\frac{1}{2}$ disappears earlier than in the inviscid simulations. This might be caused by the larger amplification of rounding errors in the discrete viscous term. (The discrete systems and the employed multigrid method are symmetric only in exact arithmetic.) The viscous simulation remains point-symmetric with respect to $(x, y)=(\frac{1}{2}, \frac{1}{2})$ for a while, see Fig. 8.

At 65×65 resolution the viscosity cannot be "tuned" such that an acceptable non-oscillatory filamentation is observed. The required viscosity (corresponding to $C_1<0.07$) results in an approximately form-preserving decay of the initial condition's cellular structure. At a resolution of 129×129, a value $C_1\approx 0.15$ is sufficient to obtain results comparable to Fig. 8. This variation of C_1 proportional to $1/\Delta x$ indicates that the Δx^2 in (37) is a bit too optimistic for the second-order Arakawa Jacobian. In summary, even with estimates such as (37) tuning of an artificial viscosity is tedious.

3. ENO-JACOBIANS

The ENO technique introduced by Harten *et al.* [9] is used in an attempt to construct truly robust Jacobians for equidistant rectangular grids. ENO methods are known to yield uniformly high order accurate, oscillation-free approximations of inviscid problems without any tuning, unlike the methods of the previous section and most other numerical techniques. However, this desirable property has been established only for problems where multiple discontinuities are well separated. It is unclear if the method is able to cope with the indefinite steepening and narrowing of vorticity filaments in inviscid two-dimensional turbulence.

Here the grid-aligned ENO variant of Shu and Osher [17] is applied to the vorticity equation. This variant is easy to apply in the multi-dimensional case for which it obtains an arbitrarily high order of accuracy. Two types of high order ENO-Jacobians have been considered, corresponding to the advection and conservation (flux) forms of the vorticity equation. As usual with ENO schemes the accuracy of an "rth order" ENO-Jacobian may degrade to $(r-1)$th order at local extremes of the vorticity or any of its derivatives.

In all results for test problem (35) the employed discretization of the streamfunction equation (2) is identical to that used in the previous section, i.e., the second-order five-point Laplacian is used again. Consequently the differences between the results of this and the previous section are mainly due to the discretization of the Jacobian. The CFL number definition of the previous section has been used as well. The time discretization is the third-order TVD Runge–Kutta scheme as given in Shu and Osher [17], which will not introduce oscillations if the space discretization does not.

3.1. *Flux Form ENO-Jacobians*

The starting point is the inviscid unforced vorticity equation (1) in conservation form,

$$\frac{\partial q}{\partial t}+\frac{\partial uq}{\partial x}+\frac{\partial vq}{\partial y}=0, \tag{39}$$

where

$$u\equiv -\frac{\partial \psi}{\partial y}, \qquad v\equiv \frac{\partial \psi}{\partial x}. \tag{40}$$

The spatial domain Ω is again covered by the equidistant grid $\{(x_i, y_j)\}\equiv\{(i\,\Delta x, j\,\Delta y)\}$. Gridpoint values of the fluxes uq and vq are obtained from the nodal values $q_{i,j}\approx q(x_i, y_j, t)$ and nodal values of the velocity components. The latter are approximated by standard fourth order central differences of (40), for example,

$$u_{i,j}=-\left(\frac{\psi_{i,j-2}-8\psi_{i,j-1}+8\psi_{i,j+1}-\psi_{i,j+2}}{12\Delta y}\right), \tag{41}$$

except in gridpoints adjacent to the boundary where asymmetric fourth-order differences have been used. For example, if the boundary segment $y=0$ corresponds with $j=1$, then at $j=2$:

$$u_{i,2}=-\left(\frac{-3\psi_{i,1}-10\psi_{i,2}+18\psi_{i,3}-6\psi_{i,4}+\psi_{i,5}}{12\Delta y}\right). \tag{42}$$

The normal flux components at the boundary are always zero; therefore the boundary values of q itself are not needed for the determination of the q-evolution in the interior. Note that the second-order discrete streamfunction equation does not depend on these boundary values either. The general grid-aligned conservative approximation of (39) at interior gridpoints is

$$\frac{dq_{i,j}}{dt}+\frac{1}{\Delta x}\left(h^x_{i+1/2,j}-h^x_{i-1/2,j}\right)+\frac{1}{\Delta y}\left(h^y_{i,j+1/2}-h^y_{i,j-1/2}\right)=0. \tag{43}$$

The "numerical fluxes" $h^x_{i+1/2,j}$ and $h^y_{i,j+1/2}$ can be expressed in the nodal values $(uq)_{k,j}$ and $(vq)_{i,k}$, respectively, such that their first divided differences approximate the derivatives in (39) to an arbitrarily high order of accuracy. Here this is accomplished with the ENO technique of Shu and Osher (see [17] for details). It uses solution-dependent h-stencils, selected to contain the smoothest data. To ensure stability, each stencil is forced to include the gridpoint in the local upwind direction. For example, if $h^x_{i+1/2,j}$ is constructed, then the point (x_i, y_j) should be included in its stencil if $u_{i+1/2,j} > 0$; otherwise (x_{i+1}, y_j) should be included. Note that the upwind direction needs to be established at the point $(x_{i+1/2}, y_j)$ which does not coincide with the gridpoints where the velocity components are known. Therefore some form of interpolation is needed to evaluate the sign of the advection velocity at this point. The methods discussed here used

$$\text{sign}(u_{i+1/2,j}) \approx \text{sign}(u_{i,j} + u_{i+1,j}).$$

The basic "ENO-Jacobian" (43) performs poorly in practice. This has been investigated with the fourth- and fifth-order versions applied to the initial condition (35) at low resolution (65×65 gridpoints). Already during the initial smooth stage they produce large spurious vorticity "spikes" near the stagnation points ($u = v = 0$) of the flow. The spikes near the stagnation points at the boundary eventually trigger a fatal instability. Similar instabilities are known to occur in symmetric discretizations of linear variable coefficient problems; see Gottlieb *et al.* [7]. That the fatal instability emerges at the boundary is to be expected because of the reduced number of h-stencils which are available there. This reduces the possibility for introducing intrinsic dissipation.

Centrally biasing the h-stencils in the way reported by Fatemi *et al.* [6] seems to keep the fourth-order solution bounded, but very large vorticity spikes still occur, with a magnitude of up to $|q|_\infty = 160$.

The local Lax–Friedrichs (LLF) modification of the basic ENO scheme (Shu and Osher [17]) yields nearly enough stabilization. In this variant the presence of a point with where in the present implementation the maximum is taken over the gridpoints in all permitted h^x-stencils around $(x_{i+1/2}, y_j)$, while β was set to 1.1. A similar modification is made in the calculation of the numerical fluxes h^y in y-direction.

With this LLF-fix the fourth-order ENO-Jacobian occasionally generates some small vorticity spikes near stagnation points ($|q|_\infty \leqslant 3$), see Fig. 9. The spikes emerge in a nonregular pattern and therefore induce a nonsymmetric velocity field. This leads to a premature breaking of the antisymmetry of the solution (Fig. 10). At higher resolutions the magnitude of the spikes decreases, see the behavior of the 257×257 result obtained with the fourth-order ENO-LLF Jacobian in Fig. 9. Centrally biasing the h-stencils in the fourth-order method generally leads to twice as large, more persistent vorticity spikes at the boundary.

The fourth-order method is stable for CFL numbers up to 0.8. The fifth-order ENO-LLF Jacobian always generates fatal instabilities at the boundary, even for CFL numbers as low as 0.3.

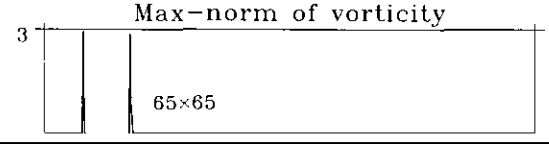
In summary, a local Lax–Friedrichs fix is mandatory for stabilization of flux-form ENO Jacobians, but the resulting schemes are still "fragile" due to problems near stagnation points.

3.2. *Advection Form ENO-Jacobians*

Applying pointwise ENO approximations to the advection form of the vorticity equation,

$$\frac{\partial q}{\partial t} + u\frac{\partial q}{\partial x} + v\frac{\partial q}{\partial y} = 0, \tag{44}$$

leads to fatal instabilities originating at the boundary, similar to those observed in the basic conservation form approximations. Good results are obtained, however, if a



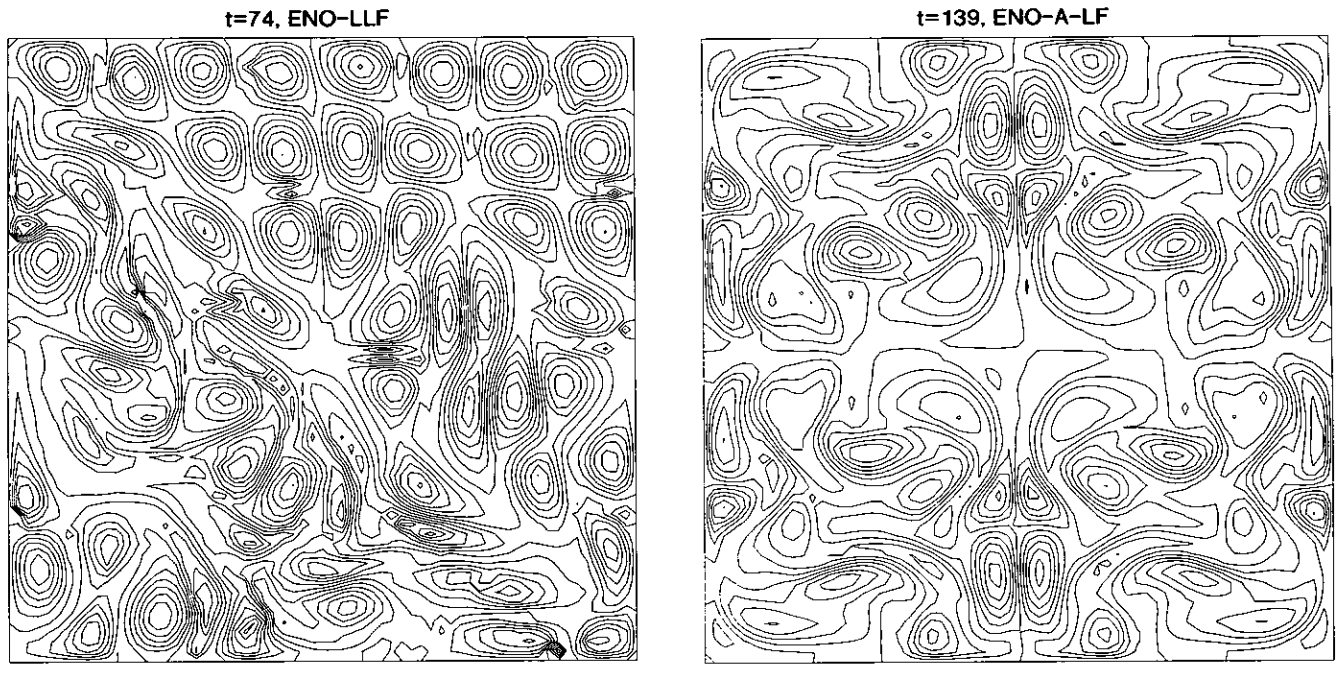


FIG. 10. Vorticity obtained with the fourth-order ENO-LLF and centrally biased ENO-A-LF Jacobians at low resolution (65×65). CFL = 0.3. Contour interval is $\frac{1}{8}$; zero contour included. The initial condition is that of Fig. 1.

global Lax–Friedrichs modification is made. This amounts to (the v-term is treated similarly)

$$u\frac{\partial q}{\partial x} \approx \frac{1}{2}(u_{i,j}+\alpha^x)\frac{\partial q}{\partial x}\bigg|_{i,j} + \frac{1}{2}(u_{i,j}-\alpha^x)\frac{\partial q}{\partial x}\bigg|_{i,j}, \quad (45)$$

where $u_{i,j}$ is again evaluated as in (41), (42), and the q-derivatives in Shu and Osher's [17] ENO way. Each of these derivatives is associated with a sign-definite advection velocity if

$$\alpha^x = \beta\,|u|_\infty \qquad \text{with} \quad \beta \geqslant 1.$$

Again β has been set to 1.1 in the tests described below. To distinguish them from the ENO-Jacobians of the previous section, methods based on (44) will be called ENO-A Jacobians. Note that ENO-A Jacobians need boundary values of q. In general these can be obtained from a pointwise discrete approximation of (44) at the boundary, in which only the tangential derivative is retained. For the present test problem with initial condition (35) this yields the exact value (38).

Both fourth- and fifth-order ENO-A-LF Jacobians show oscillation-free behavior in the low-resolution (65×65) simulation. Unlike the fourth-order ENO-LLF Jacobian, they preserve symmetries in the solution well up to $t \approx 200$, see Fig. 10, and have no problems at the boundary. They are stable and nonoscillatory for CFL numbers up to 0.5. If the "central bias" of Fatemi *et al.* [6] is used in the stencil

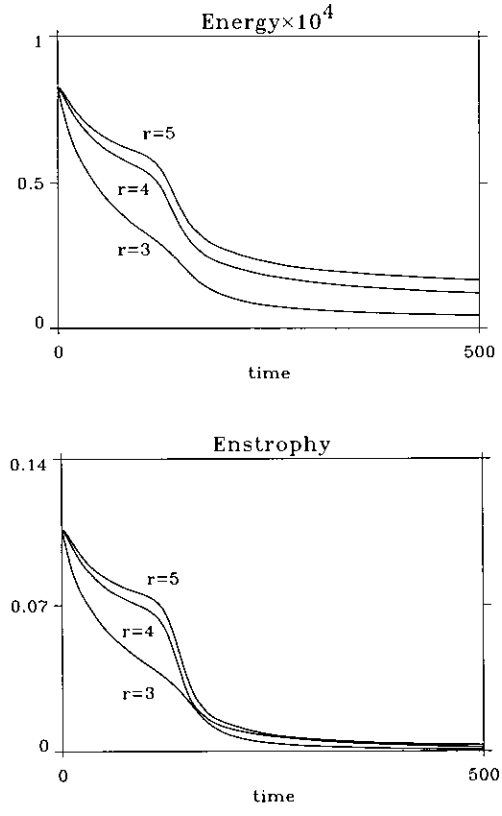


FIG. 11. Discrete energy and enstrophy evolution at low resolution (65×65) for centrally biased ENO-A-LF Jacobians of orders $r = 3, 4, 5$. CFL = 0.5.

construction for the derivatives in (45), then the methods are stable also for slightly larger CFL numbers. However, at higher resolutions this is no longer the case.

Increasing the order of the ENO-A-LF Jacobian to five, significantly reduces its intrinsic dissipation, while stability is retained, cf. Fig. 11. One might be tempted to conclude that a further increase in order would be even more beneficial. Unfortunately, a sixth-order ENO-A-LF Jacobian produces small spikes near the boundary ($|q|_\infty \approx 3$), which destroy the antisymmetry of the solution at an early stage. The failure of uniformly sixth and higher order ENO schemes is investigated further by Walsteijn [19]: at such high orders problems near boundaries are unavoidable.

A high resolution result (257×257) obtained with the fifth-order centrally biased ENO-A-LF Jacobian is given in Fig. 12. The ENO-A-LF Jacobian clearly reproduces all

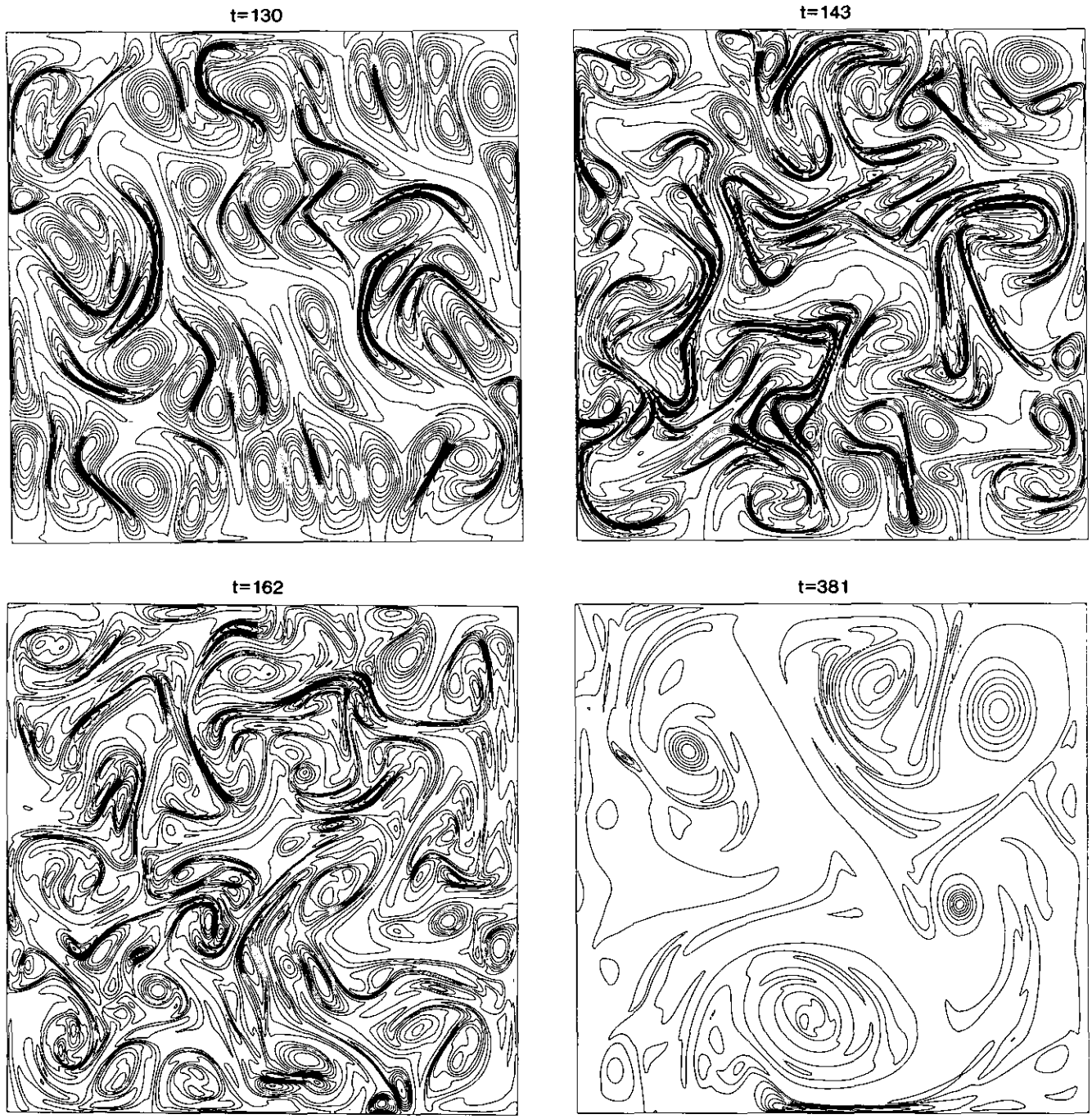


FIG. 12. Vorticity obtained with the centrally biased fifth-order ENO-A-LF Jacobian. Resolution is 257×257, CFL = 0.5. Contour interval is $\frac{1}{8}$; zero contour included. The initial condition is that of Fig. 1.

features of turbulent flows in a nonoscillatory way: filamentation and intensification of vorticity gradients, vortices merging into larger scale phenomena, and also the survival of some small-scale intense "coherent structures." It automatically introduces enough dissipation near filaments to prohibit the emergence of numerical oscillations and enstrophy equipartitioning. The persistence of coherent vortices results in an almost constant value of $|q|_\infty$ during the entire simulation. Note that at any resolution the filamentation starts at roughly the same moment as in the *inviscid* simulation with the Arakawa Jacobian. During filamentation, symmetry is lost early if the CFL number is close to the stability limit. A point-symmetry with respect to $(x, y) = (\frac{1}{2}, \frac{1}{2})$ can still be recognized in the first panel of Fig. 12. This point-symmetry is reminiscent of the Arakawa Jacobian with an artificial viscosity.

The energy and enstrophy evolutions of the ENO-A-LF Jacobian are given in Fig. 13. Comparing with Fig. 7 shows that the intrinsic (numerical) dissipation of the ENO Jacobian is smaller than the amount of Laplacian viscosity that is needed to "tune" the Arakawa Jacobian. Of course, if the Arakawa Jacobian would be tuned with a hyperviscosity (McWilliams [12]) then the dissipation of this scheme would not be so large. Furthermore, a thorough quantitative comparison is not possible because the time at which filamentation starts is very different.

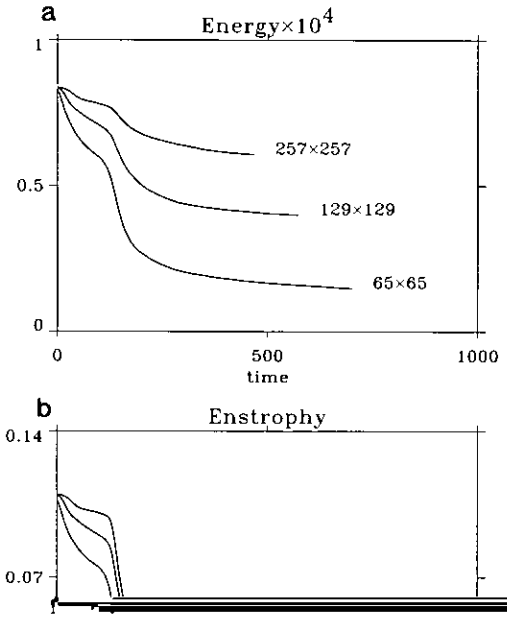


Figure 13a also illustrates that the magnitude of intrinsic viscosity of ENO Jacobians is solution-dependent. The energy dissipation rate increases temporarily when vorticity gradients become large. Comparing the energy evolutions at different resolutions reveals that the energy reduction at a fixed large time t diminishes only slowly. A high order "convergence" to the initial energy level is not possible due to the properties of the continuous system (1)–(3). Although energy cascades mainly to the large scales of motion, this process must be accompanied by a small energy flux to the smallest scales (Pedlosky [13]). For example, usage of the numerical truncation wavenumber of $\mathcal{O}(1/\Delta x)$ in a highly schematic energy redistribution argument of Pedlosky [13, Eq. (3.27.23)] suggests that a fraction of $\mathcal{O}(\Delta x^2)$ of the initial energy ends up at the smallest lengthscale (where it would be dissipated rapidly). Figure 13a shows that the actual fraction is even larger.

A disadvantage of the ENO-A-LF Jacobian is its high computational cost. At identical spatial resolutions a single evaluation of the centrally biased fifth-order ENO-A-LF Jacobian is approximately 7.5 times more expensive than the second-order Arakawa Jacobian without viscous terms. This ratio has been established on a scalar computer and is consistent with the ratio of rough operation count estimates. Note that the Arakawa Jacobian has been implemented precisely as in Arakawa [1, Eq. (45) with $\zeta_{i,j}$ omitted]. The cost of ENO Jacobians varies approximately linearly with their order of accuracy. Without a global Lax–Friedrichs fix the cost is halved. Implementations of ENO Jacobians can be arranged such that they are fully vectorizable. Nevertheless, a relative slowdown by an additional factor of up to 2 has been observed on a vector computer. This is probably caused by the complexity of the (vectorizable) loops.

Of course, the cost of a complete simulation is determined by other aspects too. First, the solution of the second-order discrete streamfunction equation by N_c typical multigrid W-cycles (that use two "red–black" Gauss–Seidel sweeps on each grid) is approximately $2.5N_c$ times more expensive than a single evaluation of Arakawa's Jacobian. The required number N_c of W-cycles depends on the admitted residual magnitude. Second, in order to obtain a given accuracy a high order ENO Jacobian may require fewer gridpoints than the second-order Arakawa technique (cf. Browning and Kreiss [4]). Finally, the ENO variants

Numerical schemes that conserve fully discrete energy and/or enstrophy in the inviscid unforced case are guaranteed unconditionally stable. Explicit schemes with this property can be constructed by combining existing conservative space discretizations with new nonlinear Runge–Kutta time discretizations. Generalizations of the nonlinear Runge–Kutta scheme of Dekker and Verwer [5] conserve enstrophy. A further extension leads to energy conserving nonlinear Runge–Kutta schemes. An extension attempting to conserve both energy and enstrophy in the time discretization failed, because it is ill-conditioned in practice.

In tests, the enstrophy conserving nonlinear variant of the classical fourth-order Runge–Kutta scheme performs slightly better than both the energy conserving variant and the non-conservative classical scheme itself. The main advantage is that the nonlinear modifications yield unconditional stability in an explicit scheme. However, if such explicit schemes are applied at very large CFL numbers (for which the standard scheme "explodes"), then they introduce large errors and may lead to an artificial steady state. Fortunately, the accuracy of these schemes can be monitored by inspecting the deviations of their γ-parameters from unity, and reducing the timestep such that $|1-\gamma|$ is of the order of the required relative accuracy.

Although nonlinear stability is guaranteed, these schemes still require the trial and error tuning of artificial viscosity to suppress spatial oscillations in nearly inviscid simulations (as they are based on symmetric space discretizations).

The ENO technique of Shu and Osher [17] has been applied to construct new grid-aligned ENO Jacobians. All investigated ENO Jacobians need to be stabilized by local or global Lax–Friedrichs (LLF or LF, respectively) modifications. ENO-LLF Jacobians based on the conservation form of the vorticity equation then still suffer from instabilities at the domain boundary and interior stagnation points. The best performing Jacobian is the ENO-A-LF Jacobian, based on the advection form of the vorticity equation. (Ironically, this method does not conserve any quantity.) It is the only ENO Jacobian that produces nonoscillatory, spike-free results. However, this is true only if its order of accuracy is not higher than 5, otherwise spikes will occur near the closed domain boundary. Similar order restrictions have been established (theoretically and from numerical experiments) by Walsteijn [19] for ENO techniques applied to problems with an inflow boundary. With the above order limitation, ENO-A-LF Jacobians enable "inviscid" simulations without the manual introduction or tuning of an artificial viscosity.

ACKNOWLEDGMENTS

This investigation was supported by the Dutch National Research Programme on Global Air Pollution and Climate Change, project number 850025. Computations were partly performed on the CRAY Y-MP at the Academic Computing Centre (SARA), Amsterdam, The Netherlands. Use of these computing facilities was sponsored by the Stichting Nationale Supercomputerfaciliteiten (National Computing Facilities Foundation, NCF) with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research, NWO).

REFERENCES

1. A. Arakawa, *J. Comput. Phys.* **1**, 119 (1966).
2. A. Arakawa, "Finite-Difference Methods in Climate Modeling," in *Physically-Based Modelling and Simulation of Climate and Climatic Change—Part I*, edited by M. E. Schlesinger (Kluwer Academic, Dordrecht, 1988), p. 79.
3. A. F. Bennet and D. B. Haidvogel, *J. Atmos. Sci.* **40**, 738 (1983).
4. G. L. Browning and H.-O. Kreiss, *Math. Comput.* **52**, 369 (1989).
5. K. Dekker and J. G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations* (North-Holland, Amsterdam, 1984), pp. 194, 265.
6. E. Fatemi, J. Jerome, and S. Osher, *IEEE Trans. Computer-Aided Design* **10**, 232 (1991).
7. D. Gottlieb, S. A. Orszag, and E. Turkel, *Math. Comput.* **37**, 293 (1981).
8. E. Hairer, C. Lubich, and M. Roche, *The Numerical Solution of Differential-Algebraic Systems by Runge–Kutta Methods*, Lecture Notes in Mathematics, Vol. 1409 (Springer-Verlag, Berlin, 1989), p. 23.
9. A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy, *J. Comput. Phys.* **71**, 231 (1987).
10. W. D. Henshaw, H. O. Kreiss, and L. G. Reyna, *Theoret. Comput. Fluid Dynamics* **1**, 65 (1989).
11. J. D. Lambert, *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem* (Wiley, Chichester, 1991), pp. 167, 185.
12. J. C. McWilliams, *J. Fluid Mech.* **146**, 21 (1984).
13. J. Pedlosky, *Geophysical Fluid Dynamics*, 2nd ed. (Springer-Verlag, New York, 1987), pp. 169, 246, 262.
14. P. L. Roe, *Ann. Rev. Fluid Mech.* **18**, 337 (1986).
15. R. Sadourny, *J. Atmos. Sci.* **32**, 680 (1975).
16. R. Salmon and L. D. Talley, *J. Comput. Phys.* **83**, 247 (1989).
17. C.-W. Shu and S. Osher, *J. Comput. Phys.* **83**, 32 (1989).
18. T. E. Tezduyar, J. Liou, D. K. Ganjoo, and M. Behr, *Int. J. Numer. Methods Fluids* **11**, 515 (1990).
19. F. H. Walsteijn, "Essentially Non-oscillatory (ENO) Schemes," in *Numerical Methods for Advection Diffusion Problems*, edited by C. B. Vreugdenhil and B. Koren (Vieweg, Braunschweig, 1993).
20. J.-O. Wolff, E. Maier-Reimer, and D. J. Olbers, *J. Phys. Oceanogr.* **21**, 236 (1991).