Journal of Computational Physics **171,** 373–393 (2001)
doi:10.1006/jcph.2001.6802, available online at http://www.idealibrary.com on IDEAL®

Time Integration of the Shallow Water Equations in Spherical Geometry

D. Lanser, J. G. Blom, and J. G. Verwer

CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
E-mail: debby.lanser@cwi.nl; joke.blom@cwi.nl; jan.verwer@cwi.nl

Received November 6, 2000; revised April 13, 2001

The shallow water equations in spherical geometry provide a prototype for developing and testing numerical algorithms for atmospheric circulation models. In a previous paper we have studied a spatial discretization of these equations based on an Osher-type finite-volume method on stereographic and latitude–longitude grids. The current paper is a companion devoted to time integration. Our main aim is to discuss and demonstrate a third-order, A-stable, Runge–Kutta–Rosenbrock method. To reduce the costs related to the linear algebra operations, this linearly implicit method is combined with approximate matrix factorization. Its efficiency is demonstrated by comparison with a classical, third-order explicit, Runge–Kutta method. For that purpose we use a known test set from literature. The comparison shows that the Rosenbrock method is by far superior.

Key Words: numerical solution of PDEs; atmospheric flow; SWEs in spherical geometry; Osher's scheme; stereographic coordinates, 65M12, 65M20, 65M06, 86-08, 86A10, 86A17.

1. INTRODUCTION

Present day atmospheric circulation models used in weather forecasting and climate research are often discretized by spectral transform methods. These methods are known to provide accurate solutions and to avoid the pole problem, which arises when grid-point methods are used on standard latitude–longitude (lat–lon) grid. However, with the trend toward higher grid resolutions some of the main drawbacks of the spectral transform method become more apparent. These concern the high computational costs of the Legendre transform and the communication overhead for parallel distributed memory computers. Our investigations are directed at grid-point methods, which are expected to provide sufficient spatial accuracy for future fine-grid resolutions.

The current paper is devoted to the spherical shallow water equations (SWEs), which reveal most of the major numerical difficulties associated with the horizontal dynamics found in the full set of primitive equations. The paper is a companion to [13], where we

0021-9991/01 $35.00

examined spatial discretizations based on an Osher-type finite-volume method [15] using the third-order upwind scheme for the constant state interpolation ($(\kappa = \frac{1}{3})$-scheme [20]). This combination provides a solid spatial discretization for the hyperbolic SWEs.

In [13] we proposed a combined lat–lon and stereographic grid to avoid the pole problem that arises when solving the semidiscrete SWEs on a uniform lat–lon grid. In this article a different approach is adopted. Enhancing the grid resolution obviously necessitates an efficient time integration method to keep the solution costs affordable. The aim of the current paper is to demonstrate a third-order, A-stable, Runge–Kutta–Rosenbrock integration method. Rosenbrock methods are linearly implicit and hence require expensive linear system solves. We will show that this disadvantage can be overcome by the technique of approximate matrix factorization, which goes back to the early 1950s with splitting and alternating direction methods; see, e.g., [16]. When combined with this technique, the Rosenbrock method does not only remain third-order consistent and A-stable, but it also becomes cost-effective. We will demonstrate its efficiency by a comparison with a classical third-order explicit Runge–Kutta method using a known SWEs test set from the literature [23]. The comparison shows that the Rosenbrock method is by far superior. In this paper the two integration methods are combined with the upwind spatial discretization from [13]. They can, of course, also be combined with the usual central spatial discretizations.

The paper is organized as follows. In Section 2 we briefly recall the system of SWEs and its linearization. The linearization is used as starting point to analyze stability. In Section 3, the third-order Rosenbrock method and the third-order explicit Runge–Kutta method are discussed. For the explicit method the time step restrictions on the uniform lat–lon and on the combined grid are derived. For the Rosenbrock method with approximate matrix factorization, A-stability is proven. Section 4 describes our numerical experiments, which will demonstrate the qualities of the Rosenbrock method combined with approximate matrix factorization.

2. PRELIMINARIES ON THE SHALLOW WATER EQUATIONS

In this section we briefly recall the system of SWEs in spherical coordinates and its linearization. Assuming Fourier–Von Neumann analysis, the linearized problem is used for the stability analysis. The spherical SWEs describe a pure initial-value problem on the rotating sphere and are defined as follows.

Let $\lambda \in [0, 2\pi)$ denote longitude, $\phi \in [-\frac{\pi}{2}, +\frac{\pi}{2}]$ latitude, and $t \geq 0$ time. Let u be the velocity in the longitudinal direction, v the velocity in the latitudinal direction, and h the height of the free surface above the sphere at sea level, i.e., $h = H + h_s$, where h_s describes the height of underlying mountains. Further, let $\underline{u}$ denote the horizontal velocity field (u, v), f the Coriolis parameter $2\Omega \sin\phi$ with Ω the angular velocity of the Earth, a the radius of the Earth, and g the gravitational constant. Using the flux-form, the two-dimensional SWEs, being composed of a continuity equation and two momentum equations, read [7, 23]

$$\frac{\partial H}{\partial t} + \nabla \cdot (H\underline{u}) = 0, \tag{1}$$

$$\frac{\partial Hu}{\partial t} + \nabla \cdot (Hu\underline{u}) = \left(f + \frac{u}{a}\tan\phi\right)Hv - \frac{gH}{a\cos\phi}\frac{\partial h_s}{\partial \lambda} - \frac{g}{a\cos\phi}\frac{\partial\left(\frac{1}{2}H^2\right)}{\partial\lambda}, \tag{2}$$

$$\frac{\partial Hv}{\partial t} + \nabla \cdot (Hv\underline{u}) = -\left(f + \frac{u}{a}\tan\phi\right)Hu - \frac{gH}{a}\frac{\partial h_s}{\partial \phi} - \frac{g}{a}\frac{\partial\left(\frac{1}{2}H^2\right)}{\partial\phi}, \tag{3}$$

where the divergence operator is defined by

$$\nabla \cdot \underline{u} = \frac{1}{a\cos\phi}\left[\frac{\partial u}{\partial \lambda} + \frac{\partial (v\cos\phi)}{\partial \phi}\right]. \tag{4}$$

The term on the right-hand side in (2) and (3) represent forcing terms. It concerns the Coriolis force, the curvature terms, and the hydrostatic pressure gradient force. Along with the lat–lon coordinate system we apply stereographic coordinates. To save space we here omit the corresponding formulations of the SWEs. In [13] we have studied the spatial discretization of both formulations using the Osher upwind scheme.

2.1. The Linearization

Adopting standard practice, we consider the "frozen" linearized system of the Eqs. (1)–(4) to analyze the stability properties of the integration methods. Let us linearize around a constant state vector $\bar{q} = (\bar{H}, \overline{Hu}, \overline{Hv})^T$, where the upper bar refers to frozen variables. The resulting linearized system then reads

$$q_t + Aq_\lambda + Bq_\phi = Cq, \tag{5}$$

where $q = (H, Hu, Hv)^T$,

$$A = \frac{1}{\tilde{a}}\begin{pmatrix} 0 & 1 & 0 \\ -\bar{u}^2 + g\bar{H} & 2\bar{u} & 0 \\ -\bar{u}\bar{v} & \bar{v} & \bar{u} \end{pmatrix}, \quad B = \frac{1}{a}\begin{pmatrix} 0 & 0 & 1 \\ -\bar{u}\,\bar{v} & \bar{v} & \bar{u} \\ -\bar{v}^2 + g\bar{H} & 0 & 2\bar{v} \end{pmatrix}, \quad \tilde{a} = a\cos\phi, \tag{6}$$

and the force matrix

$$C = \begin{pmatrix} 0 & 0 & \frac{\tan\phi}{a} \\ \frac{-g}{\tilde{a}}\frac{\partial h_s}{\partial \lambda} - \frac{2\tan\phi}{a}\overline{uv} & \frac{2\tan\phi}{a}\bar{v} & \frac{2\tan\phi}{a}\bar{u} + \bar{f} \\ \frac{-g}{a}\frac{\partial h_s}{\partial \phi} + \frac{\tan\phi}{a}(\bar{u}^2 - \bar{v}^2) & -C_{23} & C_{22} \end{pmatrix}.$$

Note that the constant coefficient matrices A, B, and C do not commute, which implies that their eigensystems differ. Consequently, it is not possible to further simplify Eq. (5) to a scalar equation. For our analysis we therefore need the eigenvalue–eigenvector decompositions of A and B. We have $A = X_A E_A X_A^{-1}$ and $B = X_B E_B X_B^{-1}$ with

$$\begin{aligned} X_A &= \begin{pmatrix} 0 & 1 & -1 \\ 0 & \bar{u} + \sqrt{gH} & -\bar{u} + \sqrt{gH} \\ \sqrt{gH} & \bar{v} & -\bar{v} \end{pmatrix}, \\ X_A^{-1} &= \frac{1}{\sqrt{gH}}\begin{pmatrix} -\bar{v} & 0 & 1 \\ \frac{1}{2}(\sqrt{gH} - \bar{u}) & \frac{1}{2} & 0 \\ -\frac{1}{2}(\sqrt{gH} + \bar{u}) & \frac{1}{2} & 0 \end{pmatrix}, \end{aligned} \tag{7}$$

$$X_B = \begin{pmatrix} 0 & 1 & -1 \\ \sqrt{gH} & \bar{u} & -\bar{u} \\ 0 & \bar{v}+\sqrt{gH} & -\bar{v}+\sqrt{gH} \end{pmatrix},$$
$$X_B^{-1} = \frac{1}{\sqrt{gH}} \begin{pmatrix} -\bar{u} & 1 & 0 \\ \frac{1}{2}(\sqrt{gH}-\bar{v}) & 0 & \frac{1}{2} \\ -\frac{1}{2}(\sqrt{gH}+\bar{v}) & 0 & \frac{1}{2} \end{pmatrix}, \tag{8}$$

and

$$E_A = \operatorname{diag}\left(\frac{\bar{u}}{a\cos\phi}, \frac{\bar{u}+\sqrt{gH}}{a\cos\phi}, \frac{\bar{u}-\sqrt{gH}}{a\cos\phi} \right), \tag{9}$$

$$E_B = \operatorname{diag}\left(\frac{\bar{v}}{a}, \frac{\bar{v}+\sqrt{gH}}{a}, \frac{\bar{v}-\sqrt{gH}}{a} \right). \tag{10}$$

Note that both decompositions exist, since our system is hyperbolic. The eigenvalue expressions for A and B are related to well-known physical features. The values containing the $\sqrt{gH}$ term are connected with the so-called gravity waves, while the remaining values are connected with the so-called advective waves. The corresponding wave speeds differ significantly; i.e., the gravity waves run much faster than the advective ones. In practice, these gravity waves need not be resolved, because most meteorologically important motions are close to geostrophic balance, which implies low amplitude gravity waves. In general, unfortunately, these waves dictate the critical time step at which stability can still be guaranteed when using explicit methods. It is for this reason, that we focus on alternative time integration methods.

Following [13], we spatially discretize our system using Osher's scheme [15] with a higher order state interpolation, which yields a second-order method. Assuming a uniform grid, Osher's scheme applied to the constant linear system (5) simplifies to the third-order, ($\kappa = \frac{1}{3}$)-upwind scheme [20] as given below. Consider the cell-centered grid points

$$\lambda_j = \left(j - \frac{1}{2}\right)\Delta\lambda, \quad \Delta\lambda = \frac{2\pi}{N}, \quad \phi_k = -\frac{\pi}{2} + \left(k - \frac{1}{2}\right)\Delta\phi, \quad \Delta\phi = \frac{\pi}{M}, \tag{11}$$

and let the grid function $w_{jk}(t)$ denote the semidiscrete approximation to the solution $q(\lambda_j, \phi_k, t)$ of (5) on this grid. Denote $A^+ = X_A E_A^+ X_A^{-1}$, where $E_A^+ = (|E_A| + E_A)/2$ is obtained from E_A by replacing its negative entries by zero. Introduce analogously B^+ and A^-, B^-, where the positive entries in the eigenvalue matrix are replaced by zero. The semidiscrete, ($\kappa = \frac{1}{3}$)-upwind approximation to (5) on grid (11) can then be written as

$$\frac{d}{dt} w_{jk} = L w_{jk}, \quad L = L_A + L_B + C, \tag{12}$$

where

$$L_A = -(A^+ D_A^+ + A^- D_A^-), \quad L_B = -(B^+ D_B^+ + B^- D_B^-). \tag{13}$$

The operators D_A^+ and D_A^- are the upwind and downwind operators in the longitude direction;

i.e.,

$$D_A^+ w_{jk} = \frac{w_{j-2k} - 6w_{j-1k} + 3w_{jk} + 2w_{j+1k}}{6\Delta\lambda}, \tag{14}$$

$$D_A^- w_{jk} = \frac{-2w_{j-1k} - 3w_{jk} + 6w_{j+1k} - w_{j+2k}}{6\Delta\lambda}. \tag{15}$$

D_B^+ and D_B^- denote their counterparts along latitude. A^+, B^+, etc., are evaluated in each grid cell. For convenience of notation we omit their spatial dependence.

To analyze the semidiscrete system (12), we introduce the harmonic wave solution $w_{jk}(t) = \hat{w}(t)e^{\sigma(w_1\lambda_j + w_2\phi_k)}$, $\sigma = \sqrt{-1}$. An elementary computation yields the ordinary differential equation for the Fourier transform $\hat{w}$

$$\frac{d}{dt}\hat{w} = \hat{L}\hat{w}, \quad \hat{L} = \hat{L}_A + \hat{L}_B + C, \tag{16}$$

where

$$\hat{L}_A = -X_A \hat{E}_A X_A^{-1}, \quad \hat{L}_B = -X_B \hat{E}_B X_B^{-1}. \tag{17}$$

$\hat{E}_A$ and $\hat{E}_B$ are diagonal matrices with entries

$$\hat{e}_A = \frac{1}{3}\frac{|e_A|}{\Delta\lambda}\left((\cos\xi_1 - 1)^2 + \operatorname{sign}(e_A)\sigma(4 - \cos\xi_1)\sin\xi_1\right), \quad \xi_1 = w_1\Delta\lambda, \tag{18}$$

and

$$\hat{e}_B = \frac{1}{3}\frac{|e_B|}{\Delta\phi}\left((\cos\xi_2 - 1)^2 + \operatorname{sign}(e_B)\sigma(4 - \cos\xi_2)\sin\xi_2\right), \quad \xi_2 = w_2\Delta\phi. \tag{19}$$

e_A denotes an eigenvalue of A. Likewise, e_B denotes an eigenvalue of B. A clarifying discussion on the eigenvalues of the ($\kappa = \frac{1}{3}$)-upwind scheme, (18) and (19), can be found in [12].

The stability behavior of any integration method applied to the linear semidiscrete system (12) is governed by its stability behavior for the three-dimensional ODE system in Fourier space (16). By periodicity and symmetry, it suffices to consider ξ_1, ξ_2 in the interval $[-\pi, 0]$. Note that in our notation the dependence of $\hat{w}$ on ξ_1, ξ_2 is suppressed. For an introduction to the theory of Fourier analysis for difference schemes, we refer to [5, 18].

To analyze stability in case of calculations on a combined grid, we also need the linearization and the Fourier decomposition of the SWEs in stereographic formulation. The derivation is similar to the one above and leads to completely equivalent expressions due to the conformal character of the stereographic and lat–lon mapping. Therefore, we only list the counterparts of the eigenvalues expressions.

$$E_{A_{\mathrm{st}}} = \operatorname{diag}(m\bar{U}, m(\bar{U} + \sqrt{g\bar{H}}),\ m(\bar{U} - \sqrt{g\bar{H}})), \tag{20}$$

and

$$E_{B_{\mathrm{st}}} = \operatorname{diag}(m\bar{V}, m(\bar{V} + \sqrt{g\bar{H}}),\ m(\bar{V} - \sqrt{g\bar{H}})), \tag{21}$$

where

$$m(\phi) = \frac{2}{1 + \alpha\sin\phi},$$

and $\bar{U}$ and $\bar{V}$ denote the frozen stereographic velocity components in x_{st}- and y_{st}-direction, respectively.

3. THE RUNGE–KUTTA INTEGRATION METHODS

In this section we discuss the third-order Rosenbrock method and the third-order, explicit Runge–Kutta method. Both integration methods solve general nonlinear ODE systems, $\dot{w} = F(w)$. Note that the semidiscrete system of SWEs fits into this framework. We expect the Rosenbrock method to be an efficient candidate to solve this semidiscrete system, since it permits large time steps. The costs per time step are relatively high though. Therefore, the third-order explicit method is included for comparison.

3.1. The Third-Order Rosenbrock Method

The method is derived from the general two-stage Rosenbrock formula from the stiff ODE field [4, 6],

$$\begin{aligned} w^{n+1} &= w^n + b_1 k_1 + b_2 k_2, \\ Sk_1 &= \tau F(w^n), \\ Sk_2 &= \tau F(w^n + \alpha_{21} k_1) + \gamma_{21} \tau J k_1, \\ S &= I - \gamma \tau J, \end{aligned} \tag{22}$$

where b_1, b_2, α_{12}, γ_{12}, and γ are free parameters which determine the methods specific properties. The numerical solution w^n approximates w at time $t = t_n$, $\tau = t_{n+1} - t_n$ denotes the time step, and $J = F'(w^n)$ is the Jacobian matrix of $F(w)$ at $w = w^n$. When low to moderate accuracy is required, methods of the Rosenbrock type have proven efficient for a variety of stiff ODE applications [6]. For method (22) the order of consistency p is at most 3.

We analyze the stability properties of our method by applying (22) to the Fourier transformed problem (16). The general, two-stage Rosenbrock method with $p \geq 2$ then yields an amplification factor $R(\tau \hat{L})$, i.e., $\hat{w}^{n+1} = R(\tau \hat{L}) \hat{w}^n$, with $R(z)$ defined as the stability function

$$R(z) = 1 + \frac{2z}{1 - \gamma z} + \frac{\frac{1}{2} z^2 - z}{(1 - \gamma z)^2}. \tag{23}$$

The stability function $R(z)$ yields A-stability for all $\gamma \geq \frac{1}{4}$. In case of the special value $\gamma = \frac{1}{2} + \frac{1}{6}\sqrt{3}$ a third-order, A-stable function is obtained. A-stability is attractive as it implies unconditional stability in the sense of Fourier–Von Neumann for stable linear problems. However, for multidimensional PDE applications as ours solving twice per time step a linear system with the matrix $I - \gamma \tau F'(w^n)$ is rather expensive. Therefore, we will apply approximate matrix factorization. By this technique the numerical algebra costs are substantially reduced, while $p = 3$ and A-stability are still possible.*******

3.1.1. *Approximate Matrix Factorization*

We rewrite the semidiscrete system $\dot{w} = F(w)$ as $\dot{w} = F(w) \equiv F_A(w) + F_B(w)$, where F_A denotes the semidiscrete longitudinal operator extended with the force terms present in Eq. (2) and F_B denotes the semidiscrete latitudinal operator extended with the force terms present in Eq. (3). Hence, F_A and F_B are one-dimensional operators defined along

sets of longitudinal and latitudinal grid lines, respectively. The idea of approximate matrix factorization is to redefine S by

$$S = (I - \gamma\tau J_A)(I - \gamma\tau J_B), \quad J_A = F'_A(w^n), \quad J_B = F'_B(w^n), \tag{24}$$

or, equivalently, J by

$$J = F'(w^n) + \gamma\tau\tilde{J}, \quad \tilde{J} = -J_A J_B. \tag{25}$$

Instead of solving a huge two-dimensional linear system, we thus solve two one-dimensional linear systems, each of which is uncoupled per grid line. The costs per step then amount to two function evaluations for F, one Jacobian evaluation, and one band solve per longitudinal and latitudinal grid line. Since we use the Osher scheme on a stencil of five grid points with three solution components, each Jacobian matrix $F'_A(w^n)$ and $F'_B(w^n)$ consists of a blockband matrix with five blocks of (3×3). Note that $F'_A(w^n)$ is slightly more complex as a consequence of the periodicity in longitudinal direction. The costs per time step are still considerably higher as compared to those of a standard explicit method. However, the Rosenbrock method combined with approximate matrix factorization yields a far more efficient method, as our numerical results will show; see Section 4.

Approximate matrix factorization is reminiscent of the splitting technique already used in more conventional alternating direction methods during the 1950s; see, e.g., [16]. The technique has been used in various other applications since then; see, e.g., [1]. The authors have applied it successfully to large-scale atmospheric transport-chemistry problems, using a second-order method from class (22) [3, 21]. As an iterative technique, approximate matrix factorization has been successfully applied to large-scale transport problems in surface water [10]. A recent survey can be found in [9]. In [11] and references therein, interesting theoretical stability results are given revealing some limitations of approximate matrix factorization in three-dimensional applications.

3.1.2. *Consistency and Stability Properties*

With J defined as in (25), method (22) is third-order consistent for arbitrary $\tilde{J}$ whenever

$$\begin{gathered} b_1 + b_2 = 1, \quad b_2(\alpha_{21} + \gamma_{21}) = \frac{1}{2} - \gamma, \quad b_2\alpha_{21}^2 = \frac{1}{3}, \\ \gamma^2 - \gamma + \frac{1}{6} = 0, \quad b_2\gamma_{21} = -\gamma. \end{gathered} \tag{26}$$

The fifth condition $b_2\gamma_{21} = -\gamma$ results from the matrix factorization. These conditions yield a unique solution which defines the Rosenbrock method

$$\begin{aligned} w^{n+1} &= w^n + \frac{1}{4}k_1 + \frac{3}{4}k_2, \\ Sk_1 &= \tau F(w^n), \\ Sk_2 &= \tau F\left(w^n + \frac{2}{3}k_1\right) - \frac{4}{3}\gamma\tau J k_1, \\ S &= (I - \gamma\tau J_A)(I - \gamma\tau J_B), \end{aligned} \tag{27}$$

with $\gamma = \frac{1}{2} + \frac{1}{6}\sqrt{3}$. For efficiency reasons, the matrix-vector multiplication in the second stage formula is removed by redefining k_2 by $k_2 - \frac{4}{3}k_1$. This gives the following third-order

Rosenbrock method.[1]

$$\begin{aligned} w^{n+1} &= w^n + \frac{5}{4}k_1 + \frac{3}{4}k_2, \\ Sk_1 &= \tau F(w^n), \\ Sk_2 &= \tau F\left(w^n + \frac{2}{3}k_1\right) - \frac{4}{3}k_1, \\ S &= (I - \gamma\tau J_A)(I - \gamma\tau J_B). \end{aligned} \tag{28}$$

In the remainder of this section, we will discuss stability properties of (28) by means of Fourier–Von Neumann analysis. To obtain the linear recurrence relation which governs stability, we apply method (28) to the ODE system (16). Using the notation introduced in Section 2, we find the recurrence relation $\hat{w}^{n+1} = R(\hat{Z}_A, \hat{Z}_B)\hat{w}^n$, where $\hat{Z}_A = \tau(\hat{L}_A + C_A)$, $\hat{Z}_B = \tau(\hat{L}_B + C_B)$, and

$$R(\hat{Z}_A, \hat{Z}_B) = I + \hat{S}^{-1}\left(2\hat{S} + \frac{1}{2}\hat{Z} - I\right)\hat{S}^{-1}\hat{Z}, \tag{29}$$

with $\hat{Z} = \hat{Z}_A + \hat{Z}_B$ and $\hat{S} = (I - \gamma\hat{Z}_A)(I - \gamma\hat{Z}_B)$. Suppose that $\hat{Z}_A$ and $\hat{Z}_B$ are diagonalizable and share well-conditioned eigensystems. We can then proceed with the scalar counterpart of (29), which reads

$$R(z_A, z_B) = 1 + \frac{2z}{(1 - \gamma z_A)(1 - \gamma z_B)} + \frac{\frac{1}{2}z^2 - z}{(1 - \gamma z_A)^2(1 - \gamma z_B)^2}, \tag{30}$$

with $z = z_A + z_B$ and z_A and z_B denoting eigenvalues of respectively $\hat{Z}_A$ and $\hat{Z}_B$. A convenient property of the stability function (30) is that it mimics the A-stability property of the original stability function (23). However, in this case the range of acceptable γ-values of method (28) for which the A-stability property holds is smaller, as is shown in the following theorem.

THEOREM 3.1. *The factorized stability function* (30) *satisfies* $|R(z_A, z_B)| \leq 1$ *for all* z_A, z_B *with* $Re(z_A) \leq 0$, $Re(z_B) \leq 0$ *if and only if* $\gamma \geq \frac{1}{2} + \frac{1}{6}\sqrt{3}$.

Proof. By the maximum modulus theorem, it suffices to consider imaginary values $z_A = ib_1$, $z_B = ib_2$ for arbitrary real numbers b_1, b_2. A simple computation gives $|R(ib_1, ib_2)| \leq 1$ if and only if

$$f(b_1, b_2) \equiv \alpha_1 b_1^2 b_2^2 + \alpha_2\left(b_1^2 + b_2^2\right) + \alpha_3 b_1 b_2 \leq 0, \tag{31}$$

where $\alpha_1 = 3\gamma^4 - 4\gamma^5$, $\alpha_2 = \frac{1}{4} - 2\gamma + 5\gamma^2 - 4\gamma^3$, $\alpha_3 = \frac{1}{2} - 4\gamma + 8\gamma^2 - 4\gamma^3$.

An extremum of the function f is either located at a stationary interior point or at a noninterior point, i.e., for $b_1 \to \pm\infty$ or $b_2 \to \pm\infty$. We first investigate its behavior for $b_1 \to \pm\infty$. In that case f yields

$$\lim_{b_1 \to \pm_\infty} \frac{f(b_1, b_2)}{b_1^2} = \left(\alpha_1 b_2^2 + \alpha_2\right), \quad \forall b_2 \in \mathbb{R}.$$

[1] This method is studied independently in [14] for integrating advection–diffusion problems on sparse grids.

This function is nonpositive for all b_2 when $\alpha_1 \leq 0$ and $\alpha_2 \leq 0$, which yields

$$\gamma \geq \frac{3}{4}. \tag{32}$$

The same result can be derived for $b_2 \to \pm\infty$, since $f(b_1, b_2)$ is symmetric in b_1 and b_2.

An extremum can also be found in a stationary point of f. Solving for $(\frac{\partial f}{\partial b_1}, \frac{\partial f}{\partial b_2}) = (0, 0)$ yields

$$\begin{aligned} &b_1 = b_2 = 0, && (a)\\ &b_1 = b_2 = b \neq 0 \quad \text{with } b^2 = -\frac{2\alpha_2 + \alpha_3}{2\alpha_1}, && (b)\\ &b_1 = c \neq 0 \quad \text{and} \quad b_2 = -c \neq 0 \quad \text{with } c^2 = -\frac{2\alpha_2 - \alpha_3}{2\alpha_1}. && (c) \end{aligned} \tag{33}$$

We first consider the stationary point $(b_1, b_2) = (0, 0)$, where $f(b_1, b_2) = 0$. Let H_f denote the Hessian determinant in a stationary point $\underline{a}$,

$$H_f(\underline{a}) = \frac{\partial^2 f}{\partial b_1^2}(\underline{a})\frac{\partial^2 f}{\partial b_2^2}(\underline{a}) - \left(\frac{\partial^2 f}{\partial b_1 \partial b_2}(\underline{a})\right)^2.$$

According to, e.g., [19], the function f has a local maximum in $\underline{0}$ if $H_f(\underline{0}) > 0$ and $\frac{\partial^2 f}{\partial b_1^2}(0) < 0$. Taken into account (32), we thus find that f remains nonpositive in a neighborhood of $(b_1, b_2) = (0, 0)$, when γ satisfies

$$\gamma > \frac{1}{2} + \frac{1}{6}\sqrt{3}.$$

This condition is only sufficient. The theorem does not provide a decisive answer when $H_f(\underline{0}) = 0$. In that case a further investigation of the behavior of f in a neighborhood of $\underline{0}$ is necessary. For the γ-values at which $H_f(\underline{0}) = 0$ only $\gamma = \frac{1}{2} + \frac{1}{6}\sqrt{3}$ guarantees nonpositivity of f in a neighborhood of $\underline{0}$. So, for f to be nonpositive, γ should satisfy the following necessary condition

$$\gamma \geq \frac{1}{2} + \frac{1}{6}\sqrt{3}. \tag{34}$$

Finally, we consider the four remaining stationary points of (33). These stationary points only exist when $b^2 > 0$ and $c^2 > 0$. However, these conditions contradict conditions (32) and (34). Therefore, in case that f is nonpositive over $\mathbb{R}^2$, these points do not exist.

Summarizing, f is nonpositive for all $(b_1, b_2) \in \mathbb{R}^2$ iff $\gamma \geq \frac{1}{2} + \frac{1}{6}\sqrt{3}$. ■

This result is of interest in its own, as it shows that for useful values of γ the A-stability property is not lost by the matrix factorization.[2] In general, the matrices $\hat{Z}_A$ and $\hat{Z}_B$ do not commute, so that true unconditional stability for the linearized SWEs cannot be concluded from Theorem 3.1. Note that Theorem 3.1 does provide a necessary condition in this case. The following example will illustrate that for the SWEs and noncommuting matrices $\hat{Z}_A$ and $\hat{Z}_B$, Theorem 3.1 provides a reliable indication for unconditional stability.

[2] In [11] it is pointed out that for a three-term splitting such a result does not exist.

3.1.3. *Example*

We have approximated the maximum value of the amplification operator (29) over the interval $\xi_1, \xi_2 \in [-\pi, 0]$. Calculations are performed at a location near a pole, i.e., at a location where the longitudinal grid size $\Delta\lambda\ a\cos\phi$ on the sphere becomes very small. Locations near the poles are believed to be most critical in relation to stability (the pole problem). The example serves to identify the γ-values at which the Rosenbrock method (28) yields an unconditionally stable method when applied to the linearized SWEs after been spatially discretized with Osher's scheme. For comparison, the same computation will be carried out for the third-order explicit Runge–Kutta method in Section 3.2.2.

Let $\bar{u} = \bar{v} = 30$, $g\bar{H} = 10^5$, $a = 42000000/(2\pi)$ (space and time units are meters and seconds). Choose $\phi = (\pi - \Delta\phi)/2$, i.e., a location close to the north pole. Furthermore, put $\Delta\lambda = \Delta\phi = \pi/128$, which corresponds approximately to a uniform $1.4^\circ \times 1.4^\circ$ grid. Omitting the force matrix C, we have computed accurate estimates of the maximum spectral radius of $R(\hat{Z}_A, \hat{Z}_B)$ for $\tau = 10^i$, $i = 0, 1, 2, 3, 4$ and $\gamma = 0.25, 0.50, 0.75, 0.8, 0.9, 1.0$. The maxima are determined for $-\pi \le \xi_1, \xi_2 \le 0$ using a 100×100 grid. The following table shows these maxima for $\gamma = 0.25, 0.50, 0.75$.

τ	1	10	10^2	10^3	10^4
$\gamma = 0.25$	1.0000	1.0000	1.0008	2.2355	3.2207
$\gamma = 0.50$	1.0000	1.0000	1.0000	1.4014	1.5067
$\gamma = 0.75$	1.0000	1.0000	1.0000	1.0000	1.0000

The table reveals conditional stability for $\gamma = 0.25$ and $\gamma = 0.5$ and indicates unconditional stability for $\gamma = 0.75$. Also for $\gamma = 0.8, 0.9, 1.0$ maxima equal to 1.0 are found. This leads us to conjecture unconditional stability for all $\gamma \ge 0.75$, in line with the result of Theorem 3.1. We believe that the slightly larger value for $\gamma = \frac{1}{2} + \frac{1}{6}\sqrt{3} \approx 0.789$ in this theorem is due to the fact that the requirement for A-stability is more stringent. This property allows eigenvalues to lie in the whole of the left half of the complex plane, which is not the case in practice. Recall that the value $\gamma = 0.75$ also plays a special role for the stability function (30). Inequality (31) implies $\gamma \ge 0.75$ for $|b_1|, |b_2| \to \infty$.

Because the force matrix C can possess eigenvalues with a small positive real part, we have omitted C in the above computation. Note that, since A, B, and C do not share the same eigenvectors, adding the matrix C does not simply mean that the linearized SWEs become unstable. However, maxima slightly larger than 1.0 can occur; see also the example in Section 3.2.2. We assume that the matrix A dictates the stability behavior of system (5), since it grows with the inverse of $\cos\phi$. Note that the entries of C are comparable in size. However, A multiplies the derivative q_λ and C is only a forcing matrix multiplying q.

3.2. Explicit Runge–Kutta Time Stepping

An explicit s-stage Runge–Kutta method applied to system $\dot{w} = F(w)$ has the form

$$w^{n+1} = w^n + \tau \sum_{i=1}^{s} b_i F(W_i), \tag{35}$$

$$W_i = w^n + \tau \sum_{j=1}^{i-1} a_{ij} F(W_j), \quad i = 1, 2, \ldots, s. \tag{36}$$

In combination with central differences for space discretization, the most popular explicit Runge–Kutta method for hyperbolic problems is the classical four-stage method of order four. This higher order method owes its popularity to its imaginary stability boundary of $\sqrt{8}$. In comparison with other explicit methods this boundary is satisfactory and in fact close to the optimal value $s - 1 = 3$ for explict Runge–Kutta methods [8]. However, since we employ upwinding in the space discretization, a different method is chosen.

3.2.1. *Stability Considerations*

Let us consider methods of order $p = s$ for s = 1, 2, 3, 4. When applied to a Fourier transformed problem like (16), such a method yields a polynomial amplification operator $R(\hat{Z})$, $\hat{Z} = \tau\hat{L}$, with $R(z)$ defined by the truncated Taylor series

$$R(z) = \sum_{i=0}^{p} \frac{1}{i!} z^i. \tag{37}$$

Assuming that the most severe time step restriction indeed emerges from the longitudinal operator in the polar region, it makes sense to first examine stability for the longitudinal operator alone. Hence, we take $\hat{L} = \hat{L}_A$. Since our operator is diagonalizable, we are then able to examine stability through the scalar recurrence relation $\hat{w}^{n+1} = R(z)\hat{w}^n$, where

$$z = \frac{\nu_A}{3}((\cos\xi_1 - 1)^2 + \text{sign}(e_A)\sigma(4 - \cos\xi_1)\sin\xi_1), \quad \sigma = \sqrt{-1}, \quad -\pi \le \xi_1 \le 0 \tag{38}$$

with ν_A denoting the one-dimensional CFL number

$$\nu_A = \frac{\tau|e_A|}{\Delta\lambda}, \tag{39}$$

and e_A denoting an eigenvalue of A; see (9). To determine the maximal value of ν_A at which each method is stable, it suffices to draw the z_A-loci which lie inside the stability region of the stability function. Accurate estimates from [12] yield

s	1	2	3	4
ν_A	0	0.87	1.62	1.74
ν_A/s	0	0.43	0.54	0.43

The scaled CFL number ν_A/s, is related to efficiency. Note that explicit Euler ($s = 1$) is not stable. For the other three cases, the scaled CFL numbers ν_A/s are almost equal and close to 0.5. Note that the case $s = 4$ includes the classical four-stage method of order four. At equal costs, third-order methods are slightly more stable.

Substitution of the maximal wave speed (maximal eigenvalue (18)) into ν_A yields a time step restriction for linear stability. Let $\bar{u} > 0$, then

$$\tau \le \frac{\nu_A\Delta\lambda}{\max|e_A|} = \frac{a\cos(\phi)\nu_A\Delta\lambda}{\bar{u} + \sqrt{gH}}. \tag{40}$$

On a uniform grid ($\Delta\lambda = \Delta\phi$) closest to the poles, $\cos(\phi) \approx \frac{1}{2}\Delta\lambda$, yielding

$$\tau \le \frac{a\,\nu_A}{2(\bar{u} + \sqrt{gH})}\Delta\lambda^2. \tag{41}$$

Consequently, we face a quadratic dependence on the spatial grid size instead of the usual linear one. The quadratic dependence leads to unacceptably small time steps.

3.2.2. *Example*

To illustrate the step size restriction (40), we return to the example of Section 3.1.3. For the data used, (41) yields $\tau \leq 5.8\,\nu_A$. Hence, we find that $\tau \leq 9.4$ for any explicit three-stage, third-order Runge–Kutta method. In our application this step size restriction is very severe.

To check the validity of expression (40) we again compute the maximal spectral radius (see Section 3.1.3) of the amplification operator $R(\hat{Z})$ with $R(z)$ defined by the third-degree polynomial (37). We now distinguish between zero and nonzero force matrix C. The table below yields the maxima for a sequence of time steps τ. The cases Z_{ABC} and Z_{AB} refer to nonzero and zero force matrix C, respectively.

τ	8	9	9.4	10	11
Z_{ABC}	1.015	1.015	1.015	1.201	1.728
Z_{AB}	1.000	1.000	1.000	1.209	1.737

For Z_{AB} the one-dimensional expression appears to be very precise, predicting linear stability for $\tau \leq 9.4$ and error growth for larger time steps. For Z_{ABC} we see nearly equal error growth for the larger time steps. For the smaller ones, we also see a modest growth. This growth is caused by an eigenvalue of $A + B + C$ with a small positive real part.

3.2.3. *Relaxing the Step Size Restriction: A Different Grid Distribution*

As mentioned before, there are several ways to reduce step size limitations. We here recall the grid modifications as used in [13]. We discussed two possible remedies, i.e., longitudinal grid coarsening toward the poles [2, 13, 22] and the use of a different grid structure and coordinate system in the polar regions [13, 17]. The latter approach concerns the construction of a combined grid consisting of two stereocaps on the northern and southern hemisphere, respectively, and a (reduced) lat–lon grid in the intermediate region. Figure 1 visualizes such a grid distribution. In stereographic coordinates the grid distribution on either stereocap is rectangular. The same holds on the intermediate region in lat–lon coordinates.

On both grid types, we can derive a step size restriction for explicit Runge–Kutta methods similar to (40). We first consider a reduced grid. Such a grid is constructed from a uniform lat–lon grid around the equator by halving the amount of grid cells in the longitudinal direction when approaching the poles, whenever the cell width in that direction projected onto the sphere is reduced by a factor of 2. The distance, $a \cos\phi\, \Delta\lambda$, is called the physical cell width. Following (40), the stepsize restriction on a reduced grid yields

$$\tau \leq \frac{a \cos(\phi) \nu_A \Delta\lambda(\phi)}{\bar{u} + \sqrt{gH}}, \tag{42}$$

where $\Delta\lambda(\phi)$ depends on the latitude ϕ, i.e., on the level of reduction. Assuming that the spherical variables, $\bar{H}$, $\bar{u}$, and $\bar{v}$, have the same order of magnitude along the whole domain, the step size restriction is most severe in the area, where the smallest physical cell width is

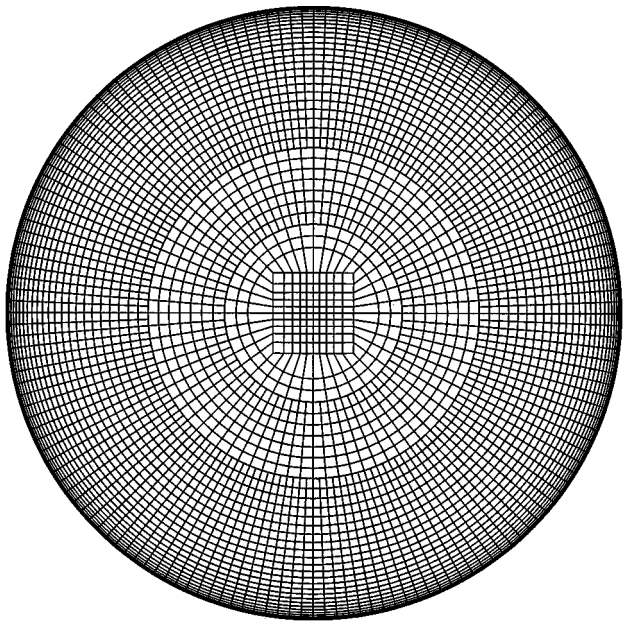

FIG. 1. Projection of a combined grid consisting of a reduced lat–lon grid away from the poles and a stereographic grid at the two polar caps onto the Cartesian (x, y)-plane ($z = 0$). Two reductions were applied.

found. On a global reduced grid this gives

$$\tau \leq \frac{2\pi}{\mathrm{nL}_{\mathrm{nRed}}} \frac{a \cos\left(\frac{\pi-\Delta\phi}{2}\right) \nu_A}{\bar{u} + \sqrt{gH}} = \frac{2\pi\, 2^{\mathrm{nRed}}}{\mathrm{nL}_0} \frac{a \cos\left(\frac{\pi-\Delta\phi}{2}\right) \nu_A}{\bar{u} + \sqrt{gH}}, \tag{43}$$

where nRed denotes the amount of reductions on the northern hemisphere, and nL_0 and $\mathrm{nL}_{\mathrm{nRed}}$ denote the amount of cells in the longitudinal direction after 0 and nRed reductions, respectively.

On a stereographic grid, an analysis similar to Section 3.2.1 can be performed. Again assuming that the step size restriction is most severe in the area with the smallest physical cell width, we find on the combined grid

$$\tau \leq \frac{\sqrt{2}\pi a\, \nu_A \cos\tilde{\phi}}{\mathrm{nL}_{\mathrm{interface}} \max\{|\bar{U} + \sqrt{gH}|, |\bar{V} + \sqrt{gH}|\}}, \tag{44}$$

where $\tilde{\phi}$ is the latitudinal boundary of the (reduced) lat–lon intermediate region of the combined grid and $\mathrm{nL}_{\mathrm{interface}}$ denotes the amount of longitudinal grid points on that boundary. The value $\sqrt{2}\pi a \cos\tilde{\phi}/\mathrm{nL}_{\mathrm{interface}}$ approximates the smallest physical cell width over the sphere after projection of the stereocap onto the globe. $\bar{U}$ and $\bar{V}$ represent the linearized velocity component in x_{st}- and y_{st}-direction, respectively. Note that the stability condition (44) is composed of the two stability conditions found in each dimension, i.e., in the x_{st}- and y_{st}-direction, respectively. Since the matrices $A_{\mathrm{st}} = X_{A_{\mathrm{st}}} E_{A_{\mathrm{st}}} X_{A_{\mathrm{st}}}^{-1}$ and $B_{\mathrm{st}} = X_{B_{\mathrm{st}}} E_{B_{\mathrm{st}}} X_{B_{\mathrm{st}}}^{-1}$ do not share the same eigensystems, each linearized system has to be analyzed separately. In case of atmospheric applications, we expect the gravity waves to dominate the flow; i.e., the quantity $\sqrt{gH}$ is large. Therefore, the step size restriction in stereographic variables is more or less direction independent.

To quantify the relation between the three step size restrictions (41), (43), and (44), we again focus on the example in Section 3.1.3. On the global uniform lat–lon grid, $\Delta\lambda = \Delta\phi = \frac{\pi}{128}$, we have

$$\tau \leq \tau_{\text{uni}} = 5.8\,\nu_A. \tag{45}$$

On the corresponding reduced grid, $\Delta\lambda(0) = \Delta\phi = \frac{\pi}{128}$, when applying three reductions, we have

$$\tau \leq \tau_{\text{red}} = 2^{\text{nRed}}\tau_{\text{uni}} = 8\,\tau_{\text{uni}}. \tag{46}$$

Note that the number of reductions is limited by accuracy; i.e., too much reductions result in a too low grid resolution around the pole to properly represent the fast varying unit direction vectors in this area; see [13]. On the combined grid, we must first position the stereocap; i.e., we have to specify $\tilde{\phi}$. For comparison, $\tilde{\phi}$ is chosen such that the amount of reductions in the intermediate lat–lon region equals the amount of reductions found on the global reduced lat–lon grid; i.e., $\text{nL}_{\text{nRed}} = \text{nL}_{\text{interface}}$. In terms of τ_{uni} we find

$$\tau \leq \tau_{\text{combi}} = \frac{4\sqrt{2}\cos\tilde{\phi}}{\cos\left(\frac{\pi-\Delta\phi}{2}\right)}\tau_{\text{uni}} \approx 34\,\tau_{\text{uni}} \tag{47}$$

with $\tilde{\phi} = \frac{61\pi}{128}$.

From (45)–(47), we can conclude that the step size restriction for explicit Runge–Kutta methods is considerably reduced when calculating on a global reduced or combined grid, the latter providing an even better alternative for the uniform lat–lon grid. On grids with a realistic resolution, the alleviation is even more apparent. On a global reduced grid with three reductions and $\Delta\lambda(0) = \Delta\phi = 2\pi/576$, and on a corresponding combined grid, $\tilde{\phi} = \frac{137\pi}{288}$, we find

$$\tau_{\text{red}} = 8\,\tau_{\text{uni}},$$

and

$$\tau_{\text{combi}} = 40\,\tau_{\text{uni}},$$

These are the time step restrictions for the grids on which we will evaluate the time integration methods in the following section.

3.2.4. *The Third-Order Explicit Comparison Method*

In case the step size is limited by stability, a low-order method, e.g., order $p = 2$, will provide sufficient temporal accuracy. However, as seen in Section 3.2.1, order $p = 3$ is slightly more efficient. Therefore, we use the following three-stage, third-order method for the comparison with the Rosenbrock method.

$$w^{n+1} = w^n + \frac{1}{6}\tau F(W_1) + \frac{1}{6}\tau F(W_2) + \frac{2}{3}\tau F(W_3), \tag{48}$$

$$W_1 = w^n, \quad W_2 = w^n + \tau F(W_1), \quad W_3 = w^n + \frac{1}{4}\tau F(W_1) + \frac{1}{4}\tau F(W_2). \tag{49}$$

To avoid an unacceptable workload, these experiments will be done on a combined grid.

4. NUMERICAL EXPERIMENTS: A COMPARISON

In the preceding section we described two Runge–Kutta methods, i.e., the third-order, A-stable, Rosenbrock method combined with approximate matrix, factorization (28), henceforth called Ros3, and the third-order, explicit, Runge–Kutta method (48), henceforth called RK3. For both methods the stability properties for the semi-discrete linearized system of SWEs (12) were investigated.

In this section we intend to show that the Ros3 method with AMF on a uniform grid is far more efficient than RK3 even when this method is applied on a combined grid employing a stereocap to alleviate the step size restriction. We use both methods to integrate the system of ODEs resulting from spatially discretizing the SWEs with Osher's scheme. This finite volume method is discussed in [13]. To judge whether Ros3 with AMF is more efficient than RK3 applied on a combined grid, we also have to consider their relative workload per time step. An estimate of this relative workload is provided, which is confirmed by numerical experiments monitoring execution time.

Both methods are applied to three test cases from the widely acknowledged SWEs test set [23], which was especially developed to validate new numerical methods to be used in circulation models. It concerns Test 2, global, steady-state nonlinear, zonal geostrophic flow, Test 5, zonal flow over an isolated mountain, and Test 6, a Rossby–Haurwitz wave. Test 2 is chosen, because it provides a test with considerable activity in the polar area. Furthermore, it has a known analytic solution without compromising the nonlinearity characteristic to the SWEs. Test 2 is a stationary test case, though. Therefore, to truly test our time integration method, we also consider two nonstationary problems, Test 5 and Test 6. For both cases, no exact solution is known and we have to rely on a high resolution spectral model for reference. These tests describe more realistic atmospheric flow patterns. For example Test 5, resolving a flow around a mountain, is challenging for most numerical solution methods. The other four tests from the SWEs test set, i.e., Tests 1, 3, 4, and 7, will be omitted, since they do not contribute additional information in relation to our efficiency question.

Calculations are performed on two different grids with related resolution. The uniform lat–lon grid has 576 grid points in longitudinal direction and 288 grid points in latitudinal direction, i.e., a $0.625^\circ \times 0.625^\circ$ grid. The combined grid consists of a reduced lat–lon grid for $\phi \in [-\tilde{\phi}, \tilde{\phi}]$ with $\tilde{\phi} = 137\pi/288$ applying three reductions on each hemisphere and two stereocaps. Around the equator the resolution is equal to the resolution found on the uniform grid. By construction, the stereocap contains 18 grid points in x_{st}- and y_{st}-direction. Note that a combined grid has approximately 20% fewer grid points than the corresponding uniform lat–lon grid. The influence on the workload is not significant though, since some additional work is needed for the spatial coupling between the stereocap and the intermediate region. As mentioned before, efficiency mainly depends on the maximal time step allowed by the time integration method and its workload per time step.

In case of the RK3 method the time step is restricted by stability. We determine this time step by trial-and-error and denote it by τ_{RK3}. Note that the discussion on the time step restriction in Section 3 concerned the linearized system of SWEs and thus provides only an estimate for an upperbound for the time step. Analysis of the computational complexity of the Ros3 method with AMF shows that the workload per time step of the Ros3 method is approximately six times as large as the workload per time step of the RK3 method. This value is confirmed by numerical experiments on Tests 2, 5, and 6 monitoring execution time. Therefore, the Ros3 tests are run with time step $\tau_{\text{Ros3}} = 6 \times \tau_{\text{RK3}}$. Next the time step

will be increased to determine the maximal time step at which stability is still obtained and the accuracy is still acceptable.

Besides testing on stability, we measure the accuracy of our solution for each method and time step over a prescribed time period. The accuracy is evaluated by the max-norm of the relative error of the depth of the fluid layer, Rel(H), and the absolute errors of the velocity components in longitudinal and x_{st}-direction, Abs(u, U), and latitudinal and y_{st}-direction, Abs(v, V), i.e.,

$$\begin{aligned}
\mathrm{Rel}(H) &= \max_{i,j} \left| \frac{H_{i,j} - H(\lambda_i, \phi_j)}{H(\lambda_i, \phi_j)} \right|, \\
\mathrm{Abs}(u) &= \max_{i,j} |u_{i,j} - u(\lambda_i, \phi_j)|, \\
\mathrm{Abs}(v) &= \max_{i,j} |v_{i,j} - v(\lambda_i, \phi_j)|,
\end{aligned}$$

and similar expressions for Abs(U) and Abs(V). $H_{i,j}$, $u_{i,j}$ etc. denote the approximate solutions. $H(\lambda_i, \phi_j)$ etc. are the reference solutions, where the solution is exact in the case of Test 2 and given by a high resolution spectral method in the case of Test 5 and Test 6. The high resolution spectral solutions are given on a daily basis.

Besides accuracy and stability, methods can also be tested on their abilities to conserve physical quantities, like energy and enstrophy, which are important for atmospheric flows. We monitored both quantities in the Ros3 runs. The cascade is negligible in all cases, i.e., approximately 0.1% over the prescribed time periods.

4.1. Test 2

Test 2 represents a solid body rotation, where the height field and the velocity components in longitudinal and latitudinal direction read

$$H = h_o - \left(\frac{a\Omega u_0}{g} + \frac{u_0^2}{2g} \right) (-\cos\lambda \cos\phi \sin\alpha + \sin\phi \cos\alpha)^2, \tag{50}$$

$$u = u_0 (\cos\phi \cos\alpha + \sin\phi \cos\lambda \sin\alpha), \tag{51}$$

$$v = -u_0 \sin\lambda \sin\alpha, \tag{52}$$

where h_0 and u_0 are given, $u_0 = 38.6$ m/s and $gh_0 = 2.94 \cdot 10^4$ m^2/s^2. Several orientations are specified; however, we use the one over the poles ($\alpha = \frac{\pi}{2}$). The simulation period is five days. For the RK3 method $\tau_{\mathrm{RK3}} = 111$ s. To reach equal efficiency, we use the Ros3 method with AMF on the uniform grid with time step $\tau = 6 \times \tau_{\mathrm{RK3}} = 666$ s. The computations remain stable. For Ros3 we then increase the time step to $\tau = 1350$ s, which still results in a stable computation. Instability if found for $\tau = 1500$ s. So, the Ros3 method with AMF applied on a uniform grid is more efficient than an explicit method used on a related combined grid. We emphasize, that this grid type already significantly alleviates the time step restriction found on a uniform grid for an explicit method (recall the factor 40 found by linear analysis). We also ran this test with the unfactorized Ros3 method. The computations with this method remained stable independent of the chosen time step.

In addition, the results on the uniform grid are more accurate than their counterparts on a combined one, as can be seen from Fig. 2. The difference in accuracy is not caused by the time integration method, but can be attributed to the higher spatial errors found when calculating on a combined grid; see [13]. Furthermore, increasing the time step for the Ros3

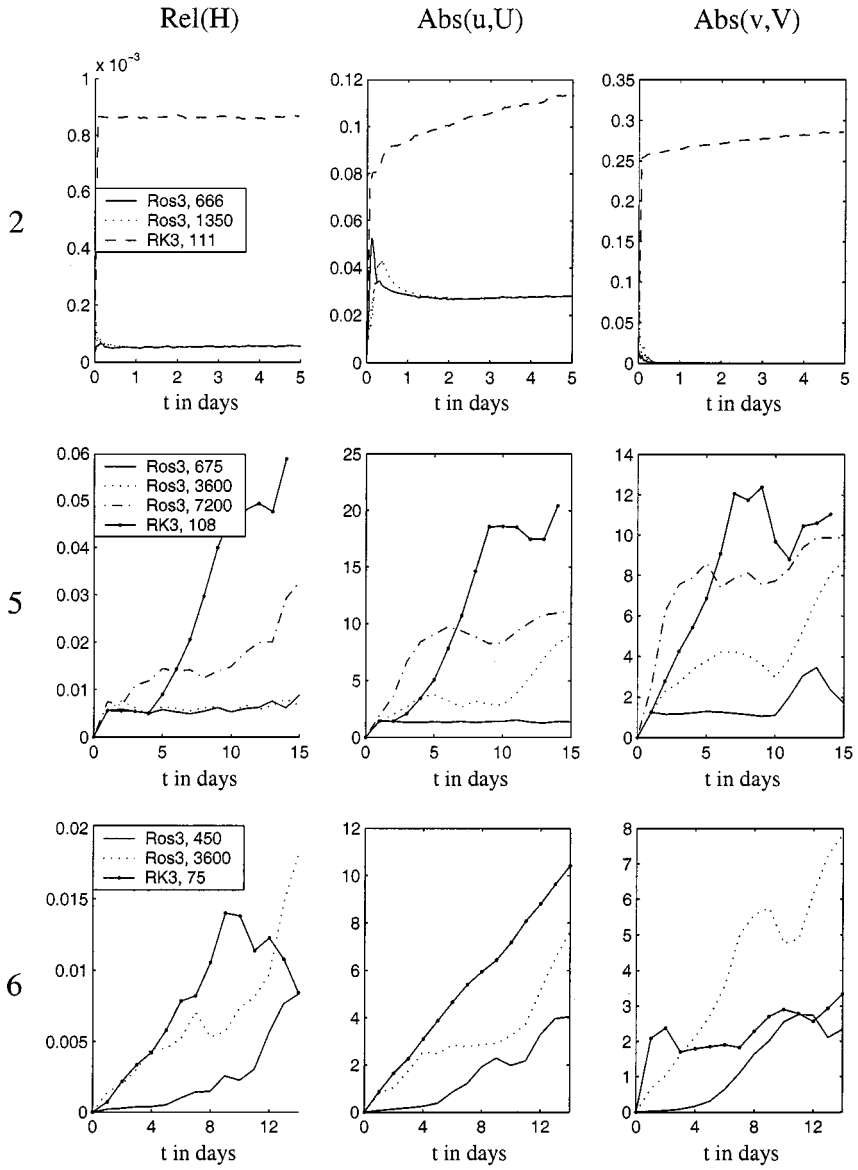

FIG. 2. Max-norm of the relative error in H (first column), absolute error in u, U (second column), and absolute error in v, V (third column) for Test 2 (first row), Test 5 (second row), and Test 6 (third row) found for the two time integration methods (RK3 and Ros3 with AMF) with given time steps. The errors are computed after each time step (Test 2) or on a daily basis (Test 5 and Test 6).

method with AMF does not yield significant accuracy changes. Reducing the resolution on our uniform grid shows that, also in this case, the errors represent spatial ones. Note that for both methods the accuracy is satisfactory.

4.1.1. *A Numerical Order Estimate for the Nonlinear SWE Equations*

Test 2 is also used to illustrate that the Ros3 method with AMF behaves as a third-order method. Calculations are done on a grid with resolution nL = 288 and nP = 144 for varying

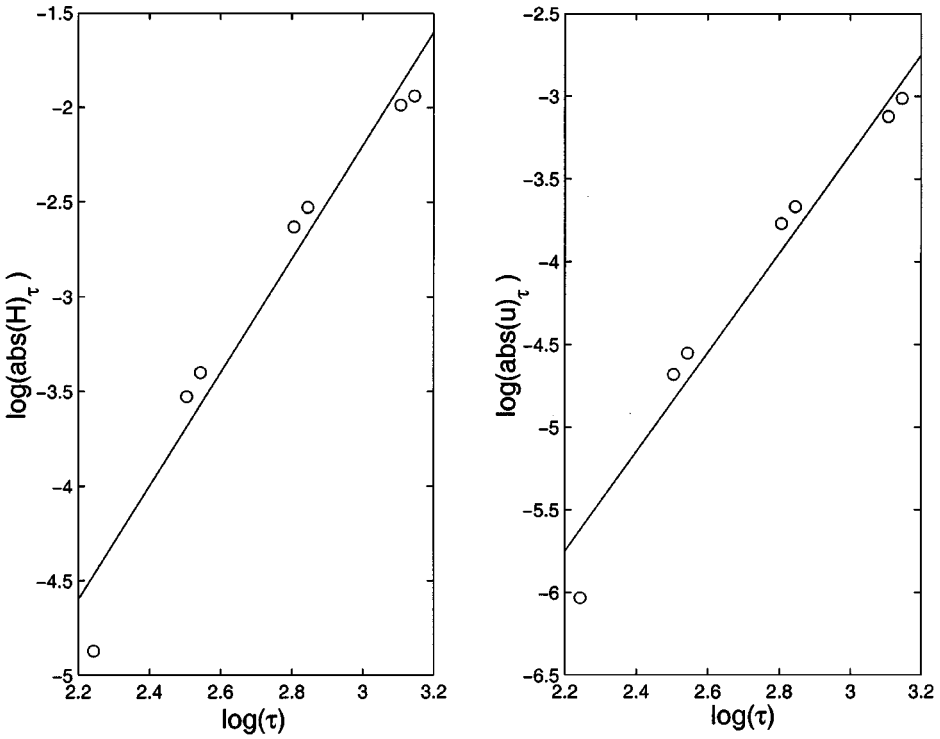

FIG. 3. An order estimate applied to H and u respectively for the Ros3 method with AMF in case of Test 2. The marks "○" denote the $\log(\mathrm{abs}(H)_\tau)$ or $\log(\mathrm{abs}(u)_\tau)$, respectively. The solid lines illustrate the slope for a third-order method.

time steps. As order estimate we use the l_∞-norm of the absolute error

$$\mathrm{abs(var)}_\tau = \max_{i,j} \left| \mathrm{var}^{\tau}_{i,j,t} - \mathrm{var}^{160}_{i,j,t} \right|,$$

where $\mathrm{var}^{\tau}_{i,j,t}$ yields the approximate value of a variable var in gridpoint $x_{i,j}$ at time t calculated with time step τ. We plotted this norm against the time step in a log–log plot for respectively H and u; see Fig. 3. The figure confirms that our method is third-order consistent.

4.2. Test 5

Test 5 consists of a zonal flow parallel to the equator which impinges on a mountain. The initial solution is given by the solid body rotation provided for Test 2 (50)–(52) with $\alpha = 0$, $u_0 = 20$ m/s, and $h_0 = 5960$ m. The surface or mountain height is prescribed by a cone,

$$h_s = h_{s_0}\left(1 - \frac{r}{R}\right), \tag{53}$$

where $h_{s_0} = 2000$ m, $R = \pi/9$, $r^2 = \min[R^2, (\lambda - \lambda_c)^2 + (\phi - \phi_c)^2]$, $\lambda_c = 3\pi/2$, and $\phi_c = \pi/6$. The simulated time period is 15 days.

With regard to efficiency the results lead to conclusions similar to those found for Test 2. The RK3 method is run with a time step $\tau_{\mathrm{RK3}} = 108$ s. The Ros3 method yields

computational stability for $\tau = 675\,\mathrm{s} \approx 6 \times 108\,\mathrm{s}$. Since the reference solution is given on a daily basis, we have to round off the time step to secure that a one-day time period can be taken in an integer number of time steps. The time step for Ros3 can be further increased. Even a time step of 2 h is possible. The results are less accurate though; see Fig. 2. When a time step of 1 h is applied, an error in H of less than 1% is found. For the 2 h time step, we notice an error growth.

Furthermore, we like to comment on the accuracy loss caused by the definition of the mountain height. To prescribe the orography, the test set introduces a cone as given by (53). This choice is a little unfortunate. The surface height is not continuously differentiable over the whole domain. The derivatives $\frac{\partial h_s}{\partial \lambda}$ and $\frac{\partial h_s}{\partial \phi}$ do not exist in the top and on the boundary of the cone. However, to evaluate the force terms of the SWEs (1)–(3) on the right-hand side, these derivatives are needed. To circumvent this problem, we apply second-order central differences to approximate them. Results show an accuracy loss in the cells surrounding the areas, where $\frac{\partial h_s}{\partial \lambda}$ and $\frac{\partial h_s}{\partial \phi}$ are not defined. The test set does not prescribe how the undefined derivatives should be handled. Therefore, we cannot be conclusive about accuracy in these areas. Figure 4 illustrates the relative error of H after 1 day computed with the Ros3 method with AMF on the uniform grid with $\tau = 675$ s. The maximal errors are indeed located close to the circle $(\lambda - \lambda_c)^2 + (\phi - \phi_c)^2 = 0$ and close to the top $(\lambda, \phi) = (\lambda_c, \phi_c)$. Note that the errors remain local over the 1-day time period.

From our results for Test 5 we again conclude that the Ros3 method on a uniform grid is far more efficient than the RK3 method on a corresponding combined grid. We add that

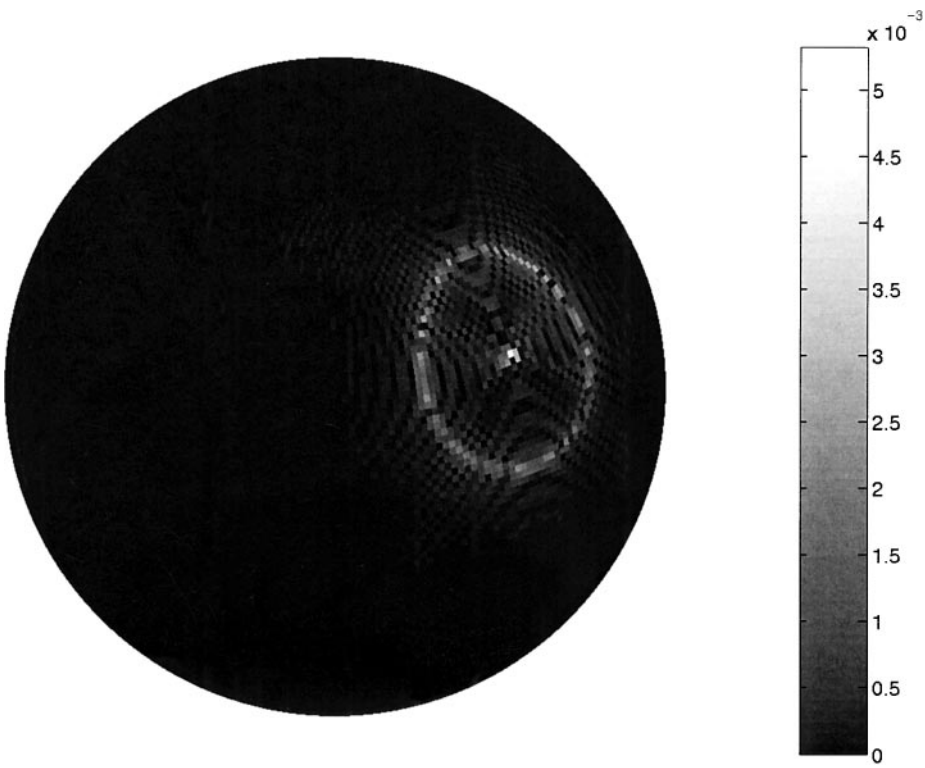

FIG. 4. Relative error of H on a uniform grid in case of Test 5. Calculations are done with the Ros3 method with AMF on a uniform grid with $\tau = 675$ s.

for Test 5 we are not really satisfied with the accuracy found in case of calculations on a combined grid. Numerical experiments show that the accuracy loss on the combined grid is mainly due to the introduction of the stereocaps. When calculating on a global, reduced lat–lon grid the results are much more accurate. We assume that the vorticity waves partly intervene with the interface band and cannot be represented sufficiently accurate. We could avoid this problem by moving the stereocap closer to the poles, however, this would result in a smaller time step.

4.3. Test 6

Test 6 is a Rossby–Haurwitz wave with a simulation period of 14 days. Again, no exact solution is known. Meteorologists consider this test as standard, since similar flow patterns occur in practical applications. A reference solution is provided by a high resolution spectral circulation model.

The time step $\tau_{\mathrm{RK3}} = 75$ s yields computational stability for the explicit RK3 method over the prescribed 14-day period. The Ros3 method with AMF is run for $\tau = 6 \times \tau_{\mathrm{RK3}} = 450$ s. Increasing the time step, computational stability is still found for time step $\tau = 3600$ s. We can conclude, that the Ros3 method is more efficient than the RK3 method on a corresponding combined grid. Again, the results on the uniform grid are more accurate.

5. CONCLUSION

When solving the semidiscrete SWEs on a global uniform lat–lon grid, an explicit time integration method suffers from severe restrictions on the time step (pole problem). This problem can be avoided by applying a suitable spatial grid or by choosing a more stable time integration method, viz. an implicit one. In [13] we proposed the application of a stereographic coordinate system in the polar regions combined with a reduced lat–lon grid in the intermediate region. In this article we considered an alternative time integration method, viz. the third-order Ros3 method with approximate matrix factorization.

We showed that the method is unconditionally stable, when applied to the linearized semidiscrete SWEs system on a uniform grid, provided that the Jacobian matrices of the fluxes in longitudinal and latitudinal direction commute. Furthermore, we showed that, due to the approximate matrix factorization, the method is cost effective. To verify its efficiency, we compared the Ros3 method with AMF to a third-order explict RK3 method applied to the system of ODEs resulting from spatially discretizing our SWEs on a combined grid. Based on Test 2, Test 5, and Test 6 of the SWEs test set, we found that the Ros3 method combined with AMF is far more efficient than the RK3 method even when the latter is applied to the semidiscrete SWEs system on a combined grid, which already significantly alleviates the time step restriction.

ACKNOWLEDGMENTS

This investigations were in part supported by the Research Council for Earth and Lifesciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO), Project Number 750.197.12.

REFERENCES

1. R. M. Beam and R. F. Warming, An implicit finite-difference algorithm for hyperbolic systems in conservation-law form, *J. Comput. Phys.* **22**, 87 (1976).

2. J. G. Blom, W. Hundsdorfer, and J. G. Verwer, *Vectorization Aspects of a Spherical Advection Scheme on a Reduced Grid*, Technical Report NM-R9418 (CWI, Amsterdam, 1997).
3. J. G. Blom and J. G. Verwer, A comparison of integration methods for atmospheric transport-chemistry problems, *J. Comput. Appl. Math.* **126**, 381 (2000).
4. K. Dekker and J. G. Verwer. *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations* (North-Holland, Amsterdam, 1984).
5. B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods* (Wiley, New York, 1995).
6. E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, 2nd ed. (Springer-Verlag, Berlin, 1996).
7. J. R. Holton, *An Introduction to Dynamic Meteorology*, 3rd ed. (Academic Press, San Diego, 1992).
8. P. J. van der Houwen, The development of Runge–Kutta methods for partial differential equations, *Appl. Numer. Math.* **20**, 261 (1996).
9. P. J. van der Houwen and B. P. Sommeijer, Approximate factorization for time-dependent partial differential equations, *J. Comput. Appl. Math.* **128**, 447 (2001).
10. P. J. van der Houwen, B. P. Sommeijer, and J. Kok, The iterative solution of fully implicit discretizations of three-dimensional transport problems, *Appl. Numer. Math.* **25**, 243 (1999).
11. W. Hundsdorfer, *Accuracy and Stability of Splitting with Stabilizing Correction*, Technical Report MAS-R9935 (CWI, Amsterdam, 1999).
12. W. Hundsdorfer, B. Koren, M. van Loon, and J. G. Verwer, A positive finite-difference advection scheme, *J. Comput. Phys.* **117**, 35 (1995).
13. D. Lanser, J. G. Blom, and J. G. Verwer, Spatial discretization of the shallow water equations in spherical geometry, *J. Comput. Phys.* **165**, 542 (2000).
14. B. Lastdrager, B. Koren, and J. G. Verwer, *Time Integration of Advection-Diffusion Problems with a Factorized Rosenbrock Method on Sparse Grids*, Technical Report MAS-R0025 (CWI, Amsterdam, 2000).
15. S. Osher and F. Solomon, Upwind difference schemes for hyperbolic systems of conservation laws, *Math. Comput.* **38**, 339 (1982).
16. D. W. Peaceman and H. H. Rachford, Jr., The numerical solution of parabolic and elliptic differential equations, *J. Soc. Indust. Appl. Math.* **3**, 28 (1955).
17. N. A. Phillips, A map projection system suitable for large-scale numerical weather prediction, *J. Meteor. Soc. Japan* **75**, 262 (1957).
18. R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*, 2nd ed. (Interscience, New York, 1967).
19. R. T. Steeley, *Calculus of Several Variables* (Scott, Foresman, Glenview, IL, 1970).
20. B. van Leer, Upwind-difference methods for aerodynamic problems governed by the Euler equations, in *Large-Scale Computations in Fluid Mechanics*, edited by B. E. Engquist, S. Osher, and R. C. J. Sommerville, AMS Series (Am. Math. Soc., Providence, 1985), pp. 327–336.
21. J. G. Verwer, E. J. Spee, J. G. Blom, and W. Hundsdorfer, A second order Rosenbrock method applied to photochemical dispersion problems, *SIAM J. Sci. Comput.* **20**, 456 (1999).
22. D. L. Williamson, Review of numerical approaches for modelling global transport, in *Air Pollution Modeling and Its Application IX*, edited by H. van Dop and G. Kallos (Plenum, New York, 1992).
23. D. L. Williamson, J. B. Drake, J. J. Hack, R. Jacob, and P. N. Swarztrauber, A standard test set for numerical approximations to the shallow water equations in spherical geometry, *J. Comput. Phys.* **102**, 211 (1992).