ACCURACY AND CONSERVATION IN SEMI-LAGRANGIAN TIME INTEGRATION

R J Purser
General Sciences Corporation/National Centers for Environmental Prediction
Washington D.C., U.S.A.

L M Leslie
University of New South Wales
Sydney, Australia

Summary: In its original goal of overcoming the unreasonably restrictive Courant-Friedrichs-Lewy timestep criterion suffered by a traditional Eulerian treatment of advection, the alternative semi-Lagrangian approach to advection has clearly succeeded, leading to more efficient code for integrating the primitive meteorological equations. However, with this freedom comes the new responsibility of ensuring a higher standard of formal accuracy in those methods of time integration that adopt the longer timesteps appropriate to the Lagrangian framework. This paper will primarily focus on ways in which a carefully formulated strategy of grid-to-grid interpolations enables new methods of numerical time integration to exploit the trajectory-based framework provided by the Lagrangian grid, so as to achieve a significant enhancement of the formal accuracy in time. We also describe how the same interpolation strategy allows one, in a very natural and efficient way, to enforce exact conservation laws of mass, and of mass-weighted linear quantities, throughout the entire integration period.

1. INTRODUCTION

Early investigations of the semi-Lagrangian method (for example, by Wiin-Nielsen, 1959; Krishnamurti, 1962; Sawyer, 1963) provided indications that this style of modeling the atmosphere was a feasible way to integrate either filtered (balanced) or primitive equation models, but it was not at that time obvious that the more compliciated semi-Lagrangian approach offered significant computational advantages over the conventional Eulerian approach. Interest in the semi-Lagrangian method for numerical weather prediction was rekindled by Robert and collaborators (Robert 1981, 1982; Robert et al., 1985) who realized that, in combination with the leapfrog semi-implicit technique (Robert, 1969), it provided a way to avoid the stringent restriction on time step implied by the advective Courant-Friedrichs-Lewy (CFL) criterion. The

method could therefore improve the computational efficiency of a forecast model by allowing longer time steps to be taken. Bates and McDonald (1982) showed how the method could be effectively used also with a split-explicit scheme for the adjustment of the fast gravity waves, again, with longer time steps than the Eulerian method would allow, giving rise to a very efficient computational method. The intervening years have seen the development of a number of notable refinements of the semi-Lagrangian method together with its adoption as the method of choice in several operational models around the world. The review article of Staniforth and Côté (1991) remains a good survey of all but the most recent of these developments.

A significant proportion of the dynamics computations in a semi-Lagrangian model is taken up by the grid-to-grid interpolations required at each time step as the data are transferred between the Lagrangian and standard Eulerian grid. Clearly, it is therefore desirable in the case of a semi-Lagrangian model to keep the number of independent trajectories to a minimum, which suggests that every attempt should be made to keep the dynamical variables on a single nonstaggered grid. It is well known that, using conventional numerical differencing procedures, a nonstaggered spatial grid has a tendency to engender a loss of formal accuracy, especially in the geostrophic adjustment process, and is prone also to grid-splitting during any extended integration (Mesinger, 1973). In Purser and Leslie (1988) we showed how such problems with the nonstaggered grid could be very effectively remedied through the use of high-order spatial differencing in conjunction with the periodic application of a low-pass spatial numerical filter. In this way, we were able to achieve very satisfactory results in a limited area barotropic model. Encouraged by this success, we extended the principle of adopting high-order numerics also to the vertical dimension in the multi-level version of this model (Leslie and Purser, 1991), again with measurable and significant improvements in forecasts compared to conventional finite difference Eulerian methods.

In order to gain the full advantages provided by high-order spatial numerics, it is necessary to apply the same philosophy also to the semi-Lagrangian interpolations. This presents no great problem in one dimension, where the standard recipe for producing a Lagrange interpolating polynomial can be easily applied at any desired degree. However, it is apparent that, when applied in two or three dimensions, the computational work

involved in evaluating the cartesian products of such high-degree polynomials increases with the square or cube of the degree chosen. For high-order numerics, this burden can become a disproportionate amount of the total computational load. We were therefore led to consider alternative strategies for the interpolation between grids that would enable the retention of high-order accuracy at a more reasonable cost.

The solution to this problem was found in a dimensional splitting of the interpolation operator into a sequence (or "cascade") of separate one-dimensional interpolations involving a succession of grids, the intermediate grid, or grids, being composed of both Lagrangian and Eulerian coordinates (Purser and Leslie, 1991). This key innovation, in turn, has provided new opportunities for enhancing the semi-Lagrangian technique, as this report will indicate. On the one hand, the cascade interpolation procedure works with equal facility in either direction between the Lagrangian (distorted) and Eulerian (regular) grids, so that it now becomes feasible, and indeed desirable, to adopt a semi-Lagrangian scheme based on *forward* trajectories. This in turn enables the principle of high-order accuracy to be further extended into the time-domain, as we shall discuss in some detail below. On the other hand, the cascade method, by separating out the individual dimensions of the problem, allows the incorporation of exact conservation properties (Rančić, 1995; Leslie and Purser, 1995), at least of mass and other "linear" quantities, directly into the interpolation procedure with little extra cost or inconvenience. Furthermore, this approach enables the surface pressure and the vertical motion in a hydrostatic model to be both handled as diagnostics of the horozontal motion. This feature is more in keeping with the custom adopted in conventional Eulerian models and leads to a significant simplification in the treatment of the vertical aspects of the trajectory definitions. Thus, it is possible now to refute the common claim that practical semi-Lagrangian methods are "unable to conserve anything". We shall provide a description below of a procedure that allows conservation of mass and tracers to be achieved in a natural way in an interpolation procedure of arbitrary order of accuracy.

The introduction of high-order "Adams-Bashforth" (AB) time integration has been shown to produce modest but consistent improvements in forecast model accuracy (Purser and Leslie, 1994), but is achieved at the cost of additional storage requirements. As a first step towards mitigating this storage cost,

the thermal and moisture equations were expressed in terms of quasi-conservative variables whose coupling to the oscillatory gravity modes of the model is, in the Lagrangian frame, indirect. With this partial separation of the moisture and thermal variables from the oscillatory dynamics, one can then use the N-cycle integration method of Lorenz (1971) to retain the high-order temporal accuracy while reducing the storage burden associated with these variables to the theoretical minimum (two fields per variable). In order to effect further storage savings, we observe that, since with the mass-conserving form of continuity, the horizontal velocities explicitly appear only in the horizontal momentum and kinematic equations, these velocities can be eliminated as prognostic variables (with a resulting significant saving in storage) by condensing these four equations into a coupled pair involving a *second-derivative* in time of the trajectory displacements. As shown in Purser and Leslie (1996), such coupled equations can be treated to a high-order of accuracy in time using "Generalized Adams-Bashforth" (GAB) methods possessing exceptionally good characteristics of formal truncation and numerical stability. We end this report with a discussion of a family of "Generalized Runge-Kutta" (GRK) methods, inspired largely by the Lorenz N-cycle methods, by which further economies of storage might be obtained in a semi-Lagrangian framework while retaining in part the property of high-order temporal accuracy.

2. CASCADE INTERPOLATION AND CONSERVATION OF MASS

The motivating principle underlying the "cascade" method for grid-to-grid interpolation is most clearly illustrated by comparing the effort needed to perform fourth-order (cubic) interpolation by the conventional construction of the cartesian product of x- and y-interpolators on the one hand with the effort associated with the corresponding cascade method on the other hand. As we see in figure 1a, the conventional approach requires interpolations to each of the four intermediate targets marked "X" before these intermediate values, in the role of source points for the final interpolation, enable the final target value at the location marked "O" to be computed. The computational effort involved is thus roughly proportional to 4^2+4 in this example, but N^2+N for generic Nth-order interpolation in the plane, and a daunting N^3+N^2+N when extending this conventional method to three dimensions. It was this unreasonably explosive growth in the computational cost with formal order

of accuracy of interpolation that provided the original stimulus to the search for a better method - the cascade method (Purser and Leslie, 1991).

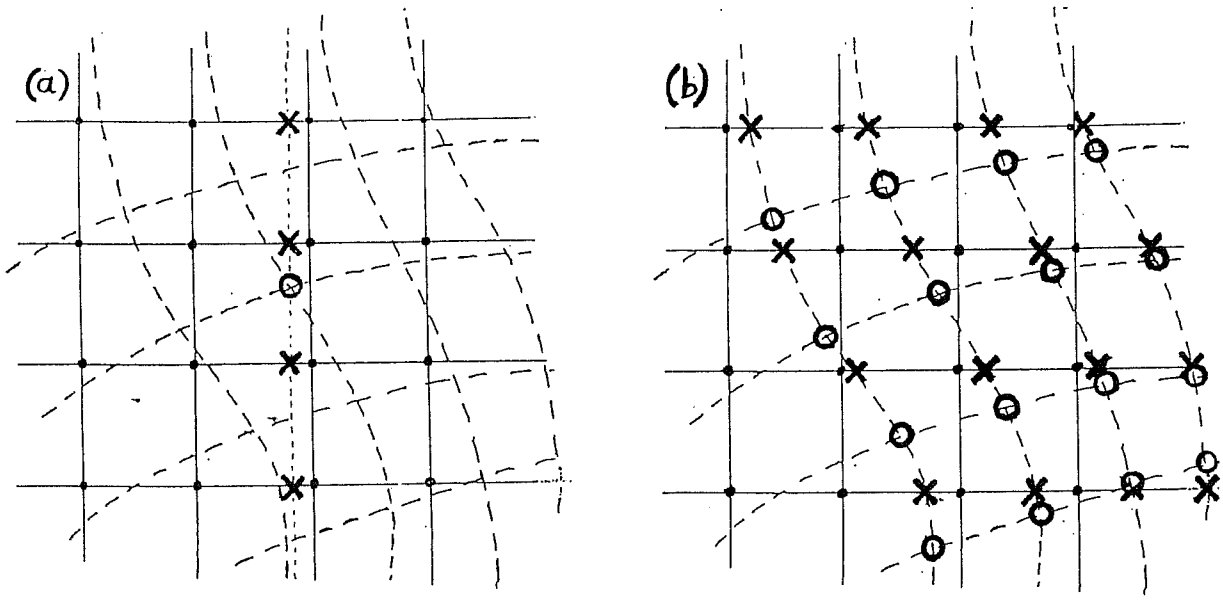

Fig.1 (a) Bi-cubic interpolation to a single target (O) by conventional cartesian-product method.
(b) Cascade interpolation to a Lagrangian grid of points (O) via an intermediate hybrid grid (X).

In Figure 1b are shown the corresponding steps of cubic interpolation but now conducted according to the cascade strategy of splitting the dimensions of the problem into separate steps. In the first step one interpolates, only in one of the dimensions, to target points, again marked "X", that collectively form a hybrid grid - in the case illustrated, a grid of distorted (Lagrangian) X-coordinates and undistorted (Eulerian) y-coordinates. Then the second step, again involving interpolations only in the other dimension, completes the process for a total average cost per target proportional to 4 + 4 computations at fourth-order, or N + N at Nth-order, or 3N at Nth-order in three dimensions.

Symbolically, it is convenient to express the "cascade" of interpolations between successive (possibly hybrid) grids using (lower case) x, y, σ, to denote the Eulerian grid coordinates and (upper case) X, Y, Σ, to denote the corresponding Lagrangian coordinates. While the cascade of figure 1b then symbolically follows the route:

$$(x,y) \longrightarrow (X,y) \longrightarrow (X,Y) \ , \tag{2.1}$$

it is apparent in practice that a reversal of this sequence, that is, starting from the Lagrangian grid and terminating at the Eulerian grid, is equally straightforward and equally inexpensive. For example, in a three-dimensional model employing *forward* trajectories, we might consider the sequence symbolized by:

$$(X,Y,\Sigma) \longrightarrow (x,Y,\Sigma) \longrightarrow (x,y,\Sigma) \longrightarrow (x,y,\sigma) \ . \tag{2.2}$$

We emphasize that the interpolation performed by (2.2) is very hard to execute by the conventional approach without a lot of costly iteration and, mainly for this reason, the latent advantages of forward-trajectory semi-Lagrangian procedures have been largely overlooked.

One important advantage of treating the vertical component of interpolation in the "forward" sense is that it permits surface pressure and vertical motion to be diagnosed directly from horizontal motion once we modify the basic cascade to incorporate the conservation of mass. In a Lagrangian framework, the conservation of mass M can be expressed without the appearance of time derivatives:

$$\frac{\partial M}{\partial(X,Y,\Sigma)} = \text{constant} \ , \tag{2.3}$$

where the constant appropriate to each trajectory is made consistent with the hydrostatic definition for the mass in an infinitesimal volume,

$$dM = (P_s\, dx\ dy\ d\sigma)/g \ . \tag{2.4}$$

The manner in which the cascade may be adapted to incorporate mass continuity is suggested by the "chain-rule" identity:

$$\frac{\partial M}{\partial(x,Y,\Sigma)} = \frac{\partial M}{\partial(X,Y,\Sigma)} \frac{\partial(X,Y,\Sigma)}{\partial(x,Y,\Sigma)} \equiv \frac{\partial M}{\partial(X,Y,\Sigma)} \left.\frac{\partial X}{\partial x}\right|_{Y,\Sigma} \ , \tag{2.5}$$

and others like it. Taking the sequence (2.2) as our cue, we start with mass per unit Lagrangian volume, $\partial M/\partial(X,Y,\Sigma)$.

(i) Now integrate with respect to X along Lagrangian grid lines of constant Y and Σ:

$$\frac{\partial M}{\partial(Y,\Sigma)} = \int^{X'} \frac{\partial M}{\partial(X,Y,\Sigma)} dX \ , \qquad (2.6)$$

thus evaluating at each point X′ along such a line the "cumulative" distribution consistent with the "density" distribution with which we began.

(ii) From the Lagrangian grid coordinates X′ at which it is convenient to hold these values, interpolate to the corresponding Eulerian grid coordinates, x′. The cumulative mass recorded at the end of each of the grid lines active at this stage of the cascade is unchanged if the domain is either bounded or periodic. Therefore, conservation of the total mass (obtained by further integrating with respect to Y and Σ) remains unaffected.

(iii) Use the differencing method exactly inverse to the quadrature of step (i), except now with respect to the Eulerian coordinate x, to recover the mass per unit hybrid (x,Y,Σ) grid:

$$\frac{\partial M}{\partial(x,Y,\Sigma)} = \frac{\partial}{\partial x}\left(\frac{\partial M}{\partial(Y,\Sigma)}\right)_{Y,\Sigma} \ . \qquad (2.7)$$

For well-conditioned mutually-inverse operations of quadrature and differencing one is obliged to stagger the grid X′ of cumulative values relative to the standard grid X of density values (and of course, grid x′ is likewise staggered with respect to grid x). Also, at orders of accuracy greater than second, the condition of mutually reciprocal quadrature and differencing operators implies one or both operators become implicit over the entire line. For a given number of computations, greatest accuracy occurs in the case of "compact" operators, as discussed in detail in Purser and Leslie (1994).

The second major stage of the cascade proceeds in a similar manner to yield the mass density of the (x,y,Σ)-grid. To see how we invoke the hydrostatic assumption in the third and final stage to diagnose the new surface pressure, simply integrate this density with respect to Σ along the now vertical coordinate columns:

$$\int_0^1 \frac{\partial M}{\partial(x,y,\Sigma)} \, d\Sigma = \frac{\partial M}{\partial(x,y)} \equiv \frac{P_s(x,y)}{g} \ . \qquad (2.8)$$

Having ascertained the new surface pressure distribution, partial integration of the coordinate-density through each one of these columns to staggered

Lagrangian grid surfaces Σ' reveals the Eulerian vertical coordinate, $\sigma(x,y,\Sigma')$, of the intersection of the column with each such Σ':

$$\int_0^{\Sigma'} \frac{\partial M}{\partial(x,y,\Sigma)}\, d\Sigma = \frac{P_s(x,y)\ \sigma(x,y,\Sigma')}{g} \quad . \tag{2.9}$$

In this way, the vertical elevation σ of the quasi-horizontal Lagrangian surfaces becomes a diagnostic of the procedure, just as the vertical velocity $\dot{\sigma}$ becomes a diagnostic of a conventional hydrostatic Eulerian model. These diagnosed values σ are finally used to complete the cascade interpolation of other dynamic variables back to the standard Eulerian grid. The conservation of quantities other than mass is obtained by proceeding in the same way, creating one-dimensionally cumulative distributions prior to each interpolation step, then differentiating them in the Eulerian target grid. It is pertinent to the development of efficient and accurate time integration methods, described in detail in the following sections, that this quasi-conservative treatment of the continuity of advected quantities by the cascade algorithm obviates the need to deal directly with either horizontal or vertical velocity components for these terms.

Observe that, at every stage of the cascade method, the operations involved can be made accurate to any order desired. However, special procedures must be adopted near the poles of a latitude-longitude grid where, owing to a violation of the necessary condition of "transversality" of the hybrid grid intersections, the simplest cascade strategies break down. Remedial modifications of the conservative cascade in such a grid geometry will be suggested in a future article.

3. HIGH-ORDER TIME INTEGRATION SCHEMES

The cascade method makes it feasible to use forward trajectories. Therefore, it is natural to reconsider the methods by which the model equations are integrated in time. In particular, the forward-trajectory style of semi-Lagrangian modeling makes it relatively straightforward to adopt some of the standard high-order time-integration methods, suitably modified to incorporate a semi-implicit control of the high-frequency gravity modes. This section will briefly characterize two important general classes of numerical integration techniques and will provide a more detailed discussion of two

particular methods relevant to our semi-Lagrangian models: the Adams-Bashforth and the Lorenz (1971) N-cycle schemes.

3.1 One-step and multi-step schemes

For a system of first-order differential equations, such as those used in atmospheric modeling, the traditional methods for numerically integrating them are conventionally (for example, Gear, 1971) classified (perhaps misleadingly) either as "one-step" methods, exemplified by the "Runge-Kutta" (RK) family, or as "multi-step" methods, whose best-known examples are the leapfrog and classical Adams-Bashforth (AB) methods. The RK methods are characterized by a cycle of N relatively simple steps, individually of low-order accuracy, which build upon each other to culminate in a result at the end of the cycle having a higher-order. In contrast, the steps of a multi-step method are each of identical form and each achieves the intended order of accuracy, but at the price of invoking values at earlier time levels of either the state vector or the vector of "force" terms of the coupled equations. Thus, unlike the RK methods, a multi-step method often requires a special starting procedure and is thereafter encumbered by the continuing existence of latent non-physical "computational modes" for which additional storage must be provided. However, one important practical advantage of the multi-step methods in the context of atmospheric modeling is that they can more easily be adapted to accommodate the semi-implicit treatment of the meteorologically uninteresting fastest modes of the model needed to allow the model to run efficiently. This is an especially important consideration in a semi-Lagrangian models whose hallmark is a long time step.

3.2 Adams-Bashforth schemes and their semi-implicit modification

The Nth-order AB method (abbreviated ABN) applied to an equation or system of equations of the form

$$\frac{d\chi(t)}{dt} = F(\chi,\ t) \quad , \tag{3.1}$$

constructs an increment at the new time step τ according to,

$$\chi^{\tau} = \chi^{\tau-1} + \sum_{k=1}^{N} B_k\, F^{\tau-k}\delta t \quad , \tag{3.2}$$

where

$$\chi^{\tau} \equiv \chi(\tau\delta t) \ , \tag{3.3a}$$

$$F^{\tau} \equiv F(\chi^{\tau}, \ \tau\delta t) \ . \tag{3.3b}$$

The N standard coefficients, B_k, are uniquely defined by the condition that the scheme be accurate to Nth-order in time and are given in Table 1 for N = 2, 3, 4. There are many non-standard, but equivalent, ways of expressing each AB algorithm (Gear, 1971), but a particularly convenient starting point for semi-implicit modification is a two-stage construction,

$$\chi^{\tau} = \chi_a^{\tau-1} + \sum_{k=1} D_k F^{\tau-k}\delta t \ , \tag{3.4a}$$

$$\chi_a^{\tau} = \chi^{\tau} + \sum_{k=0}^{N-1} E_k F^{\tau-k}\delta t \ , \tag{3.4b}$$

in which D and E are related to standard coefficients B by:

$$D_k + E_{k-1} = B_k \ , \ k = 1,..,N \ . \tag{3.5}$$

An implicit modification can be introduced for the fast modes that require it in a very natural way by simply replacing $F^{\tau} \equiv F(\chi^{\tau})$ in (3.4) by $F_a^{\tau} \equiv F(\chi_a^{\tau})$ selectively for just these modes. For example, an implicit treatment of these modes:

$$\chi_a^{\tau} = \chi_a^{\tau-1} + \sum_k C_k F_a^{\tau-k} \ \delta t \ , \tag{3.6}$$

is second-order accurate if,

$$(C_0, C_1, C_2) = \left(\frac{(1+\beta)}{2} \ , \ \frac{(1-2\beta)}{2} \ , \ \frac{\beta}{2}\right) \ ,$$
$$C_k = 0 \ , \ k > 2 \ , \tag{3.7}$$

and is achieved using such a modification of (3.4) and (3.5) when coefficients E_k satisfy the recurrence,

$$E_0 = C_0 \ ,$$
$$E_k = E_{k-1} + C_k - B_k \ , \quad k > 0 \ . \tag{3.8}$$

For schemes, ABN, N = 2, 3, 4, the resulting coefficients D_k and E_k are listed in table 1. The non-negative de-centering parameter β of (3.7) is introduced as a way to regain the "robustness" (explained below) discovered by Durran (1991) to be lacking in the context of a centered semi-implicit AB3 method.

METHOD	η	ηB_1	ηB_2	ηB_3	ηB_4	ηD_1	ηD_2	ηD_3	ηE_0	ηE_1	ηE_2	ηE_3
AB2	2	3	-1			$(2-\beta)$	β		$(1+\beta)$	$(-1-\beta)$		
AB3	12	23	-16	5		$(17-6\beta)$	$(-5+6\beta)$		$(6+6\beta)$	$(-11-6\beta)$	5	
AB4	24	55	-59	37	-9	$(43-12\beta)$	$(-28+12\beta)$	9	$(12+12\beta)$	$(-31-12\beta)$	28	-9

Table 1. Standard, and two-stage, coefficients of Adams-Bashforth schemes.

In practice, the substituted force terms, F_a^τ, cannot be made exactly consistent with χ_a^τ (except by a prohibitive degree of iteration) so one must revert to a linearization of the relevant part of the system about some convenient basic state. Let ν denote the Jacobian operator normalized by the time step δt such that, for any infinitesimal increment of the state-vector, $d\chi$, the corresponding increment of the force-vector, dF, is given by:

$$dF.\delta t = \nu\, d\chi \quad , \tag{3.9}$$

and let the idealized approximation to ν be denoted ν_0. The implied approximation ν_0 is usually based on a linearization, simplified by the neglect of physics, orography and horizontal gradients of the Coriolis parameter, about a stably stratified basic state at rest. Since one need only treat implicitly the very fastest gravity modes, which are found only amongst the modes of deepest vertical profiles, then any components of ν_0 *except* those acting upon these deepest-modes can also be neglected. This is essentially the approach pioneered by Burridge (1975) who treated only the two deepest vertical-modes out of the ten of a ten-level model. The concept of robustness mentioned above refers to the ideal of retaining numerical stability in the semi-implicit algorithm even when the actual Jacobian ν is slightly different than the idealization, ν_0. In practice, it is difficult to comprehensively analyze the robustness of a semi-implicit algorithm for a system of many degrees of freedom; instead we model the problem by considering the simplest case in which there is only one (oscillatory) mode with ν reducing simply to its (imaginary) complex-frequency. We then find that a small positive

coefficient β helps to ensure robust stability. However, the second-order algorithm AB2 is formally unconditionally unstable for pure oscillations even in its explicit form (although it is sometimes used in models possessing other sources of numerical damping).

Apart from the interpolations, the components of one complete time step of the semi-implicit algorithm might be summarized symbolically as follows.

$$\chi^{\tau} = \chi_{*}^{\tau-1} + \sum_{k=1}^{N-1} D_k F_{*}^{\tau-k} \delta t \ , \tag{3.10a}$$

$$R^{\tau} = E_0 F^{\tau} \delta t + \sum_{k=1}^{N-1} E_k F_{*}^{\tau-k} \delta t \ , \tag{3.10b}$$

$$\chi_{a}^{\tau} = \chi^{\tau} + R^{\tau} \ , \tag{3.10c}$$

$$\delta F^{\tau} \delta t = \nu_0 (I - E_0 \nu_0)^{-1} R^{\tau}, \tag{3.10d}$$

$$\delta \chi^{\tau} = E_0 \delta F^{\tau} \delta t \ , \tag{3.10e}$$

$$\chi_{*}^{\tau} = \chi_{a}^{\tau} + \delta \chi^{\tau} \ , \tag{3.10f}$$

$$F_{*}^{\tau} = F^{\tau} + \delta F^{\tau} \ . \tag{3.10g}$$

In (3.10), variables at different time-levels of the same equation are regarded as belonging to the same Lagrangian trajectories. Step (3.10a) denotes the advancement of the array of existing trajectories and the prognostic variables associated with them to the new time level τ. This is immediately followed by an interpolation (via the forward-trajectory cascade method) of the variables at this time level to the standard grid. Simultaneously, using the same interpolation coefficients, those components $F_{*}^{\tau-k}$ of earlier times still needed in subsequent steps, are interpolated also, but now from the existing array of trajectories to the locations implied by a new array of trajectories that terminate at the standard grid at time level τ. Step (3.10b) represents the evaluation of a field of "residuals" (R^{τ}) using the new force terms F^{τ} (which, of course, can only be calculated while the corresponding state variables χ^{τ} are located on the standard grid). Step (3.10c) is the explicit part of the adjustment. The remaining steps are required only for the modes treated implicitly. In practice, (3.10d) translates (in a grid-point model, at least) into the solution of a horizontal elliptic equation for each vertical mode [details are provided in appendix (a)

of Purser and Leslie, 1994], which is why it is so desirable to minimize, subject to stability, the number of these implicitly treated vertical modes in the manner of Burridge (1975).

3.3 Economizing on storage using the Lorenz N-cycle scheme

A variant of the algorithm (3.10) was tested for the AB3 method by Purser and Leslie (1994) in the semi-Lagrangian version of the Australian Bureau of Meteorology Research Centre (BMRC) regional forecast model (Leslie et al., 1985) and was found to give significantly more accurate forecasts than existing second-order methods. However, as we noted in Leslie and Purser (1995), it is desirable to achieve the high-order temporal accuracy using, if possible, a greater economy of storage. Therefore, we exploited the Lagrangian framework to express the thermal and moisture equations in quasi-conservative form, using specific entropy as the thermal variable. With these two prognostic equations expressed in terms of quasi-conservative variables decoupled (in the Lagrangian frame) from the oscillatory effects of the gravity modes, it becomes possible to employ explicit time integration methods for them. One class of such methods, is the class of high-order, "N-cycle" schemes of Lorenz (1971) which enjoy optimal storage economy. Thus, at least a start can be made in reducing the storage burden by adopting the Lorenz method. Since a generalization of this method is discussed in section 5, it is appropriate to provide a brief summary of the guiding principles here.

Each of the N time steps of the cycle of the Lorenz method superficially resembles the Euler forward method, but istead of the true force terms, a vector of what we shall call "effective force" components is substituted, comprising a weighted average of present and past force vectors. The averaging is done recursively so that the scheme formally can be summarized:

$$G^{\tau} = w^{\tau}F^{\tau}\delta t + (1 - w^{\tau})G^{\tau} , \qquad (3.11a)$$

$$\chi^{\tau+1} = \chi^{\tau} + G^{\tau}. \qquad (3.11b)$$

To start a fresh cycle, say at time level $\tau = 0$, without invoking earlier force contributions, the initial weight of the cycle must be unity, $w^0 = 1$, thus reinstating the true Euler method for the first step. However, nontrivial weights for the subsequent N-1 steps are specially chosen to ensure a result of Nth-order accuracy at the cycle's end, $\tau = N$. The motivation for

this strategy is that it allows the fields G and χ to be incremented in place, thereby keeping the storage requirement minimal. Lorenz shows that Nth-order acuracy for a coupled linear system is achieved when $w^{\tau} = N/(N-\tau)$, $\tau = 1,\ldots,N-1$.

The Lorenz N-cycle method can only be made semi-implicit for fast oscillatory modes with difficulty, which is why we have not attempted to use the Lorenz method directly for trajectory computations. Instead, we have sought further storage economies for the high-order integration of trajectories by first unifying the momentum and kinematic equations, as described in the next section, thereby eliminating the velocity components as prognostic variables.

4. GENERALIZED ADAMS-BASHFORTH METHODS

4.1 **Explicit methods**

Ignoring metric terms, the momentum and kinematic equations of a semi-Lagrangian model can be written,

$$\frac{du}{dt} - fv = -\frac{RT}{P_s}\frac{\partial P}{\partial x} - \frac{\partial \phi}{\partial x} + D_x \equiv F_x \,, \tag{4.1a}$$

$$\frac{dv}{dt} + fu = -\frac{RT}{P_s}\frac{\partial P}{\partial y} - \frac{\partial \phi}{\partial y} + D_y \equiv F_y \,, \tag{4.1b}$$

$$\frac{dx}{dt} = u \,, \tag{4.1c}$$

$$\frac{dy}{dt} = v \,, \tag{4.1d}$$

where D_x, D_y, are parameterized viscous terms, F_x, F_y, are components of the total force densities (but excluding Coriolis forces) and other symbols are conventional. We have already noted that, using Jacobian expressions of continuity, u and v are eliminated from the other governing equations. It is evident that, by combining the momentum and kinematic equations into a pair of coupled equations involving the second derivative in time, the velocity components are eliminated here also. It is convenient to adopt complex number notation for trajectory displacements and for corresponding force terms:

$$\chi \equiv x + iy \,, \tag{4.2a}$$

$$F \equiv F_x + iF_y \,, \tag{4.2b}$$

so that (4.1) are united into a second-order equation governing the complex displacement:

$$\mathcal{D}(\chi) \equiv \frac{d^2\chi}{dt^2} + if\frac{d\chi}{dt} = F \quad . \tag{4.3}$$

It is natural to generalize the form of the Adams-Bashforth schemes of section 3 to accommodate an equation of the form (4.3). For the explicit "Generalized Adams-Bashforth" (GAB) method of Nth-order (GABN), we find that, apart from a choice of normalization, there is only one expression of the construction,

$$A_+\chi^{\tau} + A_0\chi^{\tau-1} + A_-\chi^{\tau-2} = \sum_{k=1}^{N} B_k F^{\tau-k}\delta t \ , \tag{4.4}$$

giving Nth-order accuracy. For N=1, the scheme is quasi-second-order to the extent that exact second-order accuracy is obtained in the case of vanishing Coriolis parameter. The true GAB2 scheme is destabilized by the presence of a Coriolis effect and therefore is disqualified as a practical method, but schemes (4.4) with N = 3 and 4 remain stable and appear to have excellent characteristics. While GAB5 is formally unconditionally unstable when f = 0, the instability is sufficiently weak to allow this scheme be to viable also in practice. In the special limiting case, f = 0, these schemes reduce to the family considered by Størmer (1907). The coefficients for this case are listed for schemes up to fifth-order in table 2. A procedure for the evaluation of the coefficients when $f \neq 0$ is given in Purser and Leslie (1996).

METHOD	η	ηB_1	ηB_2	ηB_3	ηB_4	ηB_5
GAB1	1	1				
GAB3	12	13	-2	1		
GAB4	12	14	-5	4	-1	
GAB5	240	299	-176	194	-96	19

Table 2. Coefficients of GAB schemes when f = 0 (Størmer methods).

4.2 Semi-implicit methods

In order for the GAB schemes to be of practical value in the context of a semi-Lagrangian model we must seek robust semi-implicit modifications. In Purser and Leslie (1996) we show that linear modes about a state of rest of an f-plane model are idealized by the equation,

$$\delta t^2 \mathcal{D}(\chi) = \nu^2 \mathcal{R}e(\chi) \ , \tag{4.5}$$

where, again, ν is the characteristic imaginary "frequency" ($\nu^2 \leq 0$) normalized by a time step δt. A prerequisite for a semi-implicit version of a GAB method is that it be robustly stable when applied to (4.5). In this case we might also demand that an acceptable method be "robustly stable" in the sense that, whenever $|\nu| \leq |\nu_0|$, where ν_0 is the approximation for ν assumed in the basic linearization for the semi-implicit treatment, the stability of the discretization is assured.

By choosing a normalization of the coefficients in (4.4) that makes $A_+ = 1$, and in analogy to (3.4), we split the GAB scheme first into a two-stage process:

$$\chi^\tau + A_0 \chi_a^{\tau-1} + A_- \chi_a^{\tau-2} = \sum_{k=1} D_k F^{\tau-k} \ \delta t^2 \ , \tag{4.6a}$$

$$\chi_a^\tau = \chi^\tau + \sum_{k=0} E_k F^{\tau-k} \ \delta t^2 \ . \tag{4.6b}$$

The consistency condition,

$$D_k - A_0 E_{k-1} - A_- E_{k-2} = B_k \ , \tag{4.7}$$

must hold for each k. Upon substitution of F_a^τ for F^τ in (4.6), the implicit discretization,

$$\chi_a^\tau + A_0 \chi_a^{\tau-1} + A_- \chi_a^{\tau-2} = \sum_{k=0} C_k F_a^\tau \ \delta t^2 \ , \tag{4.8}$$

is implied, with,

$$C_k = D_k + E_k \ . \tag{4.9}$$

The robustness of various schemes with $C_k = 0$, $k > 3$, were investigated in Purser and Leslie (1996). The best results for $f = 0$ were obtained using,

$$C_1 + C_3 = 0 \ , \tag{4.10a}$$

$$\tfrac{1}{2}(C_0 - C_2) = \delta \ , \tag{4.10b}$$

where δ is a new de-centering parameter analogous to β. For the method, GAB1, robustness is achieved without de-centering, that is, with $C_0 = C_2 = 1/2$ and $C_1 = C_3 = 0$, when $f = 0$. The higher-order GAB schemes are found to require minor de-centering, $\delta \simeq 0.08$, to ensure robustness of their semi-implicit forms.

The substitution of GABN for the ABN schemes in the trajectory calculations essentially halves the storage requirements for these computations. Also, the GAB schemes of a given order are found to be substantially more accurate and more stable than the AB schemes of corresponding order. However, apart from GAB1, which is optimally economical in storage, the higher-order schemes still require extra storage and possess computational modes. The next section describes a recent attempt to remedy this remaining defect by adopting some of the principles embodied in the Lorenz N-cycle schemes.

5. N-CYCLE GENERALIZED RUNGE-KUTTA SCHEMES

5.1 **Explicit methods**

Inspired by the method of Lorenz (1971), we consider the possibility of building an accurate N-cycle scheme for the trajectory equations (4.3) relying on a modification of the scheme GAB1 to provide the basic building block corresponding to Lorenz's use of the Euler method. Since, with $f = 0$, the GAB1 scheme achieves second-order accuracy, we require each step of our N-cycle scheme to be also at least second-order. However, this additional stipulation is too restrictive to allow non-trivial N-cycle methods to be constructed unless we also relax the condition of equality of time steps within the basic cycle. Consider, then, the discretization of the operator $\mathcal{D}$ of (4.3) for non-uniform time steps:

$$(\delta t^2)\mathcal{D}^\tau(\chi) \simeq A_+^\tau \chi^{\tau+1} + A_0^\tau \chi^\tau + A_-^\tau \chi^{\tau-1} \ , \tag{5.1}$$

where the time levels are t^τ and δt is the average time step,

this remaining defect by adopting some of the principles embodied in the Lorenz N-cycle schemes.

5. N-CYCLE GENERALIZED RUNGE-KUTTA SCHEMES

5.1 **Explicit methods**

Inspired by the method of Lorenz (1971), we consider the possibility of building an accurate N-cycle scheme for the trajectory equations (4.3) relying on a modification of the scheme GAB1 to provide the basic building block corresponding to Lorenz's use of the Euler method. Since, with $f = 0$, the GAB1 scheme achieves second-order accuracy, we require each step of our N-cycle scheme to be also at least second-order. However, this additional stipulation is too restrictive to allow non-trivial N-cycle methods to be constructed unless we also relax the condition of equality of time steps within the basic cycle. Consider, then, the discretization of the operator $\mathcal{D}$ of (4.3) for non-uniform time steps:

$$(\delta t^2)\mathcal{D}^\tau(\chi) \simeq A_+^\tau \chi^{\tau+1} + A_0^\tau \chi^\tau + A_-^\tau \chi^{\tau-1} \ , \tag{5.1}$$

where the time levels are t^τ and δt is the average time step, $\delta t = (t^N - t^0)/N$, of the cycle, introduced here to make the coefficients A_α^τ non-dimensional. These coefficients are consistently and symmetrically defined at each τ by,

$$A_\alpha^\tau = \ - \frac{\delta_\alpha}{\delta_- \delta_0 \delta_+} \left(2 \ + \ \text{if } \delta t \ \Delta_\alpha\right) \ , \quad \alpha \in \{-, \ 0, \ +\} \ , \tag{5.2}$$

where,

$$(\delta_-, \ \delta_0, \ \delta_+) \equiv (t^{\tau+1} - t^\tau, \ t^{\tau-1} - t^{\tau+1}, \ t^\tau - t^{\tau-1})/\delta t \ , \tag{5.3a}$$

$$(\Delta_-, \ \Delta_0, \ \Delta_+) \equiv (t^{\tau-1} - \bar{t}^\tau, \ t^\tau - \bar{t}^\tau, \ t^{\tau+1} - \bar{t}^\tau)/\delta t \ , \tag{5.3b}$$

with,

$$\bar{t}^\tau \equiv (t^{\tau-1} + t^\tau + t^{\tau+1})/3 \ . \tag{5.4}$$

Neglecting the Coriolis terms, this discretization becomes second-order accurate only when regarded as applying at time $\bar{t}^\tau$. Since linear interpolation or extrapolation is also a second-order accurate operation, we

can ensure each step of the N-cycle method:

$$G^{\tau} = w^{\tau}(F^{\tau}\delta t^2) + (1 - w^{\tau})G^{\tau-1} \quad , \tag{5.5a}$$

$$A_{+}^{\tau}\chi^{\tau+1} + A_{0}^{\tau}\chi^{\tau} + A_{+}^{\tau}\chi^{\tau-1} = G^{\tau} \quad , \tag{5.5b}$$

is consistently second-order accurate in the absence of a Coriolis effect when w^{τ} is defined:

$$w^{\tau} = \frac{\bar{t}^{\tau} - \bar{t}^{\tau-1}}{t^{\tau} - \bar{t}^{\tau-1}} \quad . \tag{5.6}$$

This is the weight implied by linear interpolation between $F^{\tau}\delta t^2$ (at t^{τ}) and $G^{\tau-1}$ (at $\bar{t}^{\tau-1}$) to an "interpolant", G^{τ} (at $\bar{t}^{\tau}$). As before, in order to break the chain of recursions (5.5a) at the transition between cycles, say at $\tau = 0$, we require $w^{\tau} = 1$, which in turn implies [from (5.6)] that the first and last time step of the basic cycle must be equal. This constraint means that the only non-trivial N-cycle schemes are those with $N \geq 3$. The $N - 2$ "inner" time steps, however, are effectively free parameters completely at our disposal.

It can be shown that an N-cycle scheme of the kind described above, but constrained to have $N - 1$ steps of "standard" length, $\delta t'$, and only a single time step (the last-but-one of each cycle, say) of non-standard length, $b\delta t'$, has the property of <u>exactly</u> conserving the amplitude of pure oscillatory modes of (4.5) in the absence of the Coriolis term. This implies that, if the parameter b can be "tuned" to eliminate the coefficient of second-order phase-error, the resulting scheme automatically becomes at least fourth-order accurate for the coupled linear equations. By straightforward but laborious algebra, it emerges that this condition is obtained when b is a root of the quartic:

$$(N-1) + (2N-3)b + b^2 + 2b^3 + b^4 = 0 \quad . \tag{5.7}$$

Although the only real roots are negative for each N, this feature does not disqualify the schemes in practice. We choose the least negative root in each case to form the members of our standard family of fourth-order "Generalized Runge-Kutta" (GRK) N-cycle schemes. The parameters, b, of the first few of these schemes, which we denote, GRK3, GRK4, etc., are listed in table 3, together with the coefficients ε of fourth-order truncation error

(defined as in Purser and Leslie, 1996) for these schemes. The corresponding error coefficient for GAB4 is included in the table for comparison.

METHOD	b (IF APPLICABLE)	ε
GRK3	-.682327803828019	-.19781
GRK4	-.611292089333354	-.05925
GRK5	-.579906887754780	-.03465
GRK6	-.562286141036588	-.02542
GRK5*	1.549842058970427 -1.034583655219236 2.448640620176322	+.01180
GAB4		+.07917

Table 3. Time step ratios b and coefficients of fourth-order error

A numerical search of the full parameter spaces of the possible 4-cycle and 5-cycle schemes reveals no new 4-cycle schemes of interest, but several new 5-cycle schemes, among which is one that we denote, GRK5*, that apparently has an even smaller coefficient of fourth-order truncation error, ε, than any of the first few members of the standard family of GRK schemes. The time steps in the basic cycle of GRK5* are in the ratio,

$$1:\ b_1:\ b_2:\ b_3:\ 1 \quad , \tag{5.8}$$

with the parameters b defined also in table 3. While the superiority of GRK5* is suggested by the magnitude of the formal coefficient of principal truncation error, we must remain aware of the limitations of the linearized analysis. Before we can test the new methods in a full atmospheric model, it is necessary to find a way of accommodating the semi-implicit treatment of the fast modes.

5.2 **Semi-implicit methods**

By analogy with the AB and GAB semi-implicit modifications, we can construct a semi-implicit GRK algorithm based on splitting the explicit components into two stages. However, there are two new features: first, the coefficients D and E are now different at each step within the basic cycle; second, the algorithm must include the recursive calculation of effective

force G and allow for the inclusion of the weight w^τ in the formulation of the elliptic equation to be solved at each step for the implicit adjustments. With these features included, the symbolic summary of the semi-implicit GRK algoritm for one entire time step is:

$$A_+^\tau \chi^{\tau+1} + A_0^\tau \chi_*^\tau + A_-^\tau \chi_b^{\tau-1} = D^\tau G_*^\tau \quad , \tag{5.9a}$$

$$\chi_b^\tau = \chi_*^\tau - E^\tau G_*^\tau \ , \tag{5.9b}$$

$$G^{\tau+1} = w^{\tau+1} F^{\tau+1} \delta t^2 + (1-w^{\tau+1}) G_*^\tau \quad , \tag{5.9c}$$

$$R^{\tau+1} = E^{\tau+1}\ G^{\tau+1} \ , \tag{5.9d}$$

$$\chi_a^{\tau+1} = \chi^{\tau+1} + R^{\tau+1} \quad , \tag{5.9e}$$

$$\delta G^{\tau+1} = w^{\tau+1} \nu_0^2 \left[I - E^{\tau+1} w^{\tau+1} \nu_0^2 \right]^{-1} R^{\tau+1} \ , \tag{5.9f}$$

$$\delta \chi^{\tau+1} = E^{\tau+1} \delta G^{\tau+1} \ , \tag{5.9g}$$

$$\chi_*^{\tau+1} = \chi_a^{\tau+1} + \delta \chi^{\tau+1} \quad , \tag{5.9h}$$

$$G_*^{\tau+1} = G^{\tau+1} + \delta G^{\tau+1} \ . \tag{5.9i}$$

As before, δt is the average time step of the cycle. The field, χ_b^τ, is computed in (5.9b) in order to reduce storage requirements. The values of the products $E^\tau w^\tau$ found to give robust stability in each scheme are listed in table 4, in which it assumed that the negative time step of each of the standard schemes, GRKN, is the penultimate step, N-1, of its cycle.

METHOD	E^0	E^1w^1	E^2w^2	E^3w^3	E^4w^4	E^5w^5
GRK3	2.50	0.20	2.50			
GRK4	3.30	1.70	0.20	1.70		
GRK5	4.09	3.00	2.50	0.20	2.00	
GRK6	4.85	4.70	4.70	3.50	0.15	2.50
GRK5*	3.00	0.25	0.25	2.50	2.44	

Table 4. Semi-implicit parameters $E^\tau w^\tau$, for a cycle of each GRK scheme.

5.3 **Practical assessment of schemes**

As with our previous studies, the GRK schemes were tested using a semi-Lagrangian version of the BMRC grid-point regional forecast model using consistently high-order spatial discretization. In this study, the initial fields are the twice-daily analyses from the Australian Bureau of Meteorology's operational archives, for the six month period March 31 to September 30, 1995. The schemes were run in exactly the same manner as described in Purser and Leslie (1994), in which verification of forecasts is carried out only over subdomains in which there are active systems. If the whole domain is used, the differences are usually very small, as they are dominated by the almost identical performance of these highly accurate schemes over areas in which there is little activity. The control model in each example has the same spatial resolution, but a very high temporal resolution (except for the intermittently invoked physics parameterizations).

The RMS differences of the forecast from the control at 48 hours are tabulated for these active sub-domains for the low-order operational model, the GAB1, the GRK models and the fully fourth-order GAB4 scheme. The results at 850 hPa, 500 hPa and 250 hPa are shown for temperature in table 5. The order of increasing accuracy is: operational, GAB1, GRK3, GRK4, GRK5, GRK6, GRK5* and GAB4. These results confirm our hypothesis that GRK schemes provide a superior forecast over quasi-second-order scheme GAB1 (which has the same memory requirements), but despite the small coefficient of principal error of the GAB5* scheme when applied to linear equations, we find in practice that it is not superior to the fully fourth-order GAB4 scheme. However, the latter method places far greater demands on memory.

MODEL	850 hPa	500 hPa	250 hPa
OPERATIONAL	2.307	1.973	2.266
GAB1	1.954	1.665	2.071
GRK3	1.919	1.630	2.048
GRK4	1.896	1.611	2.023
GRK5	1.872	1.584	2.015
GRK6	1.868	1.572	2.002
GRK5*	1.847	1.569	1.974
GAB4	1.820	1.547	1.959

Table 5. RMS errors in 48 hour temperature predictions (K) relative to the control forecasts. The values are averaged over the period March 1 1995 to September 1 1995.

6. CONCLUSIONS

We have described an efficient computational method by which the grid-to-grid interpolations of a semi-Lagrangian model can be performed for either backward (upstream) or forward (downstream) trajectories. An adaptation of the basic method enables mass and other variables to be conserved automatically. By overcoming the obstacle associated with interpolation from forward trajectories, our "cascade" interpolation scheme has cleared the way for the investigation of a new class of semi-Lagrangian algorithms in which both classical and new time integration schemes can be tested in an attempt to reduce the (small) time truncation errors incurred by the relatively long time steps used in a semi-Lagrangian model.

We have discussed the application of the classical AB schemes to increase accuracy and the use of the Lorenz (1971) N-cycle to maintain this accuracy for the thermal and moisture terms at a reduced cost in memory. Improved storage economies are realized by replacing the AB methods for the trajectories by GAB methods in which the velocities become merely diagnostic variables. The high-order GAB methods are the most accurate schemes we have examined, but still require extra memory. In an attempt to reap the benefits of high temporal accuracy at an almost optimal economy of storage, we have developed a new family of GRK schemes inspired by the Lorenz N-cycle schemes. In their semi-implicit forms, these methods are shown to produce improvements

over the second-order schemes of comparable storage requirements, and to approach the accuracy of the truly fourth-order GAB4 scheme in practice.

Further work will be directed to extending these techniques to the global domain and to incorporation of nonhydrostatic effects.

ACKNOWLEDGMENTS

Support for this work was provided by the UCAR Visiting Scientist Program and by ONR Grant N00014-94-1-0556.

REFERENCES

Bates, J R, and A McDonald, 1982: Multiply-upstream, semi-Lagrangian advective schemes: Analysis and application to a multilevel primitive equations model. Mon Wea Rev, 110, 1831-1842.

Burridge, D M, 1975: A split-semi-implicit reformulation of the Bushby-Timpson 10-level model. Quart. J. Roy. Meteor. Soc., 101, 777-792.

Durran, D R, 1991: The third-order Adams-Bashforth method: an attractive alternative to the leapfrog time integration scheme. Mon. Wea. Rev., 119, 702-720.

Gear, C W, 1971: Numerical Initial Value Problems in Ordinary Differential Equations. Prentice Hall, 253pp.

Krishnamurti, T N, 1962: Numerical integration of primitive equations by a quasi-Lagrangian advective scheme. J Appl Meteor, 1, 508-521.

Leslie, L M, G A Mills, L W Logan, D J Gauntlett, G A Kelly, M J Manton, J L McGregor and J M Sardie, 1985: A high resolution primitive equations model for operations and research. Aust. Meteor. Mag., 33, 11-35.

Leslie, L M, and R J Purser, 1991: High-order numerics in an unstaggered three-dimensional time-split semi-Lagrangian forecast model. Mon Wea Rev, 119, 1612-1623.

Leslie, L M, and R J Purser, 1995: Three-dimensional mass-conserving

semi-Lagrangian scheme employing forward trajectories. Mon Wea Rev, 123, 2551-2566.

Lorenz, E N, 1971: An N-cycle time-differencing scheme for stepwise numerical integration. Mon. Wea. Rev., 99, 644-648.

Mesinger, F, 1973: A method for construction of second order accuracy difference schemes permitting no false two-grid interval wave in the height field. Tellus, 25, 444-458.

Purser, R J, and L M Leslie, 1988: A semi-implicit semi-Lagrangian finite-difference scheme using high-order spatial differencing on a nonstaggered grid. Mon Wea Rev, 116, 2069-2080.

Purser, R J, and L M Leslie, 1991: An efficient interpolation procedure for high-order three-dimensional semi-Lagrangian models. Mon Wea Rev, 119, 2492-2498.

Purser, R J, and L M Leslie, 1994: An efficient semi-Lagrangian scheme using third-order time integration and forward trajectories. Mon Wea Rev, 122, 745-756.

Purser, R J, and L M Leslie, 1996: Generalized Adams-Bashforth time integration schemes for a semi-Lagrangian model employing the second-derivative form of the horizontal momentum equations. Quart J Roy Meteor Soc (to appear).

Rančić, M, 1995: An efficient conservative, monotonic remapping as a semi-Lagrangian transport algorithm. Mon Wea Rev, 123, 1212-1217.

Robert, A, 1969: The integration of a spectral model of the atmosphere by the implicit method. Proc of WMO/IUGG Symp on NWP, Tokyo, Japan Meteorological Agency, VII.19-VII.24.

Robert, A, 1981: A stable numerical integration scheme for the primitive meteorological equations. Atmos Ocean, 19, 35-46.

Robert, A, 1982: A semi-Lagrangian and semi-implicit numerical integration scheme for the primitive meteorological equations. J Met Soc Japan, 60, 319-325.

Robert, A, T L Yee and H Ritchie, 1985: A semi-Lagrangian and semi-implicit numerical integration scheme for multi-level atmospheric models. Mon Wea Rev, 113, 388-394.

Sawyer, J S, 1963: A semi-Lagrangian method of solving the vorticity advection equation. Tellus, 15, 336-342.

Staniforth, A, and J Côté, 1991: Semi-Lagrangian integration schemes for atmospheric models - a review. Mon Wea Rev, 119, 2206-2223.

Størmer, C, 1921: Méthodes d'intégration numérique des équations différentielles ordinaires. C R Congr Intern Math Strasbourg, 243-257.

Wiin-Nielsen, A, 1959: On the application of trajectory methods in numerical forecasting. Tellus, 11, 180-196.