EFFICIENT HIGH-ORDER SEMI-LAGRANGIAN METHODS

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

Abstract: The place of semi-Lagrangian methods in numerical weather prediction is now firmly established, but there is a continuing need to improve the numerical efficiency and accuracy of these methods. We describe a number of computational techniques, appropriate to grid point semi-Lagrangian models, by which the effectiveness of this approach may be enhanced through the use of accurate numerical operators applied in computationally efficient ways.

1. INTRODUCTION

In several operational forecasting centres, the semi-Lagrangian treatment of numerical advection has been incorporated within the numerical prediction models in order to achieve computational efficiencies that stem from the longer time steps these methods allow (e.g., Ritchie et al. 1995). In various forms, the method has been in existence since the early days of numerical weather prediction (Wiin-Nielsen 1959, Krishnamurti 1962, Sawyer 1963), but the credit for its recent revival belongs to the late Dr. André Robert, who provided efficient practical algorithms allowing the semi-Lagrangian method to be applied very effectively in the context of the primitive meteorological equations (Robert 1981). He appreciated the fact that, in a semi-implicit model, the time step (and hence, to a certain extent, the numerical efficiency) is restricted by the Courant-Friedrichs-Lewy (CFL) condition for advection, and that this restriction could be lifted by combining the semi-implicit treatment of the fast "gravity waves" with a semi-Lagrangian treatment of advection. Liberation from the conventional CFL restriction and the consequent gain of longer time steps do not come without some costs, since the spatial interpolations required at each time step can be expensive and the temporal truncation errors incurred by using larger time intervals in the integration can significantly exceed the corresponding truncation errors of a comparable Eulerian model. It is also much more difficult to preserve conservation properties in a semi-Lagrangian model. A recent critique of the cost-effectiveness of the semi-Lagrangian method is given in Bartello and Thomas (1996). Here we present some recently developed numerical techniques that may help towards the goal of making a formally accurate semi-Lagrangian model computationally efficient.

Section 2 will discuss the construction of spatially implicit "compact" schemes of spatial differencing which, for a given order of accuracy on a moderately regular grid, achieve remarkably low coefficients of principal truncation error. For evaluations of the derivative staggered with respect to the source data, it emerges that the numerical differentiation is exactly invertible. This property, which remains valid at an arbitrary order of accuracy, can be exploited within a grid-to-grid interpolation scheme of "cascade" form (Purser and Leslie 1991) to ensure, in addition to accuracy, the formal conservation of advected scalar quantities (Purser and Leslie 1995). A description of this method is

given in section 3. The cascade method, unlike other multi-dimensional grid-to-grid interpolators, is not severely handicapped by the source grid being smoothly distorted, as occurs when downstream, or "forward" trajectories are used (Purser and Leslie 1994). In some ways, it is advantageous to adopt this approach instead of the more conventional semi-Lagrangian approach which uses upstream trajectories. With this form it is often easier to incorporate high-order methods of time integration in order to keep the time truncation errors under control, and it is possible to adopt the second-derivative (in time) combination of kinematic and momentum equations for the trajectories themselves. This second derivative form is not only a more economical expression of the trajectory computations; it offers scope for the application of new classes of semi-implicit time integration methods with advantageous characteristics of truncation error and robustness, as we discuss in section 4. Section 5 describes the use of stereographic frames for the representation of trajectory displacements, which eliminates the need to execute explicitly any trigonometric functions in the course of the integration of a semi-Lagrangian model on the surface of the sphere. Conclusions are presented in section 6.

2. COMPACT DIFFERENCING AND QUADRATURE

If we wish to obtain a high order of accuracy for the finite differencing on a regular grid, the most obvious, explicit, centred scheme is not the most accurate one for a given expenditure of work. We shall describe the construction of so-called "Padé" or "compact schemes" for accurate finite-differencing and discuss how these methods may be exploited within a one-dimensional interpolator to obtain an automatic conservation of the quantity being interpolated. The examples will assume that the grids are uniform in their spacing (relative to the appropriate coordinate) and only centred templates are considered.

2.1 Explicit and compact differencing with unstaggered grids.

When we derive the conventional centred unstaggered fourth-order differencing operator for a uniform grid, $x_i = i\delta$, we seek weights, B_j, with $j \in [-2, 2]$ such that, for a smooth function c,

$$d_i = \sum_j B_j c_{i+j} \equiv \left.\frac{dc}{dx}\right|_i + \mathcal{O}(\delta^4). \tag{2.1}$$

One practical way to obtain the weights, B_j, is to require that the scheme be exact for as many as possible of the first few test functions, $c(x)$ that comprise the successive powers of $(x - x_i)/\delta$. In Vandermonde matrix form,

$$\begin{bmatrix} 1, & 1, & 1, & 1, & 1 \\ -2, & -1, & 0, & 1, & 2 \\ 4, & 1, & 0, & 1, & 4 \\ -8, & -1, & 0, & 1, & 8 \\ 16, & 1, & 0, & 1, & 16 \end{bmatrix} \begin{bmatrix} B_{-2}\delta \\ B_{-1}\delta \\ B_0\delta \\ B_1\delta \\ B_2\delta \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{2.2}$$

The solution, written as a vector, $B = (1, -8, 0, 8, -1)^T(1/(12\delta)$, does *not* produce the exact result when applied to the *next* test function, $c(x) = (x - x_i)^5/\delta$, in the series. Instead,

$$(-32, -1, 0, 1, 32) \cdot B = -4/\delta \equiv \left(\frac{d^5c}{dx^5}\right) \varepsilon^{(4)}\delta^4, \tag{2.3}$$

where,

$$\varepsilon^{(4)} = -\frac{1}{30}, \tag{2.4}$$

is the principal coefficient of fourth-order truncation error, characteristic of this scheme.

Suppose that, along an indefinitely long line of a uniform grid, we adopt an alternative scheme,

$$\sum_j A_j d_{i+j} = \sum_{j'} B_{j'} c_{i+j'}, \tag{2.5}$$

with $j \in [-1, 1]$ and $j' \in [-1, 1]$. As before, we should seek coefficients which produce the exact result for as high a degree of generic polynomial as possible. The normalisation of weights A and B is arbitrary, but it is convenient to choose to make the sum of the A_j equal to unity. The requisite constraints, in matrix form, can be written this time in generalised Vandermonde form as:

$$\begin{bmatrix} 1, & 1, & 1; & 0, & 0, & 0 \\ 0, & 0, & 0; & -1, & -1, & -1 \\ 1, & 1, & 1; & 1, & 0, & -1 \\ -2, & 0, & 2; & -1, & 0, & -1 \\ 3, & 0, & 3; & 1, & 0, & -1 \\ -4, & 0, & 4; & -1, & 0, & -1 \end{bmatrix} \begin{bmatrix} A_{-1} \\ A_0 \\ A_1 \\ B_{-1}\delta \\ B_0\delta \\ B_1\delta \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{2.6}$$

and the solution for this, perhaps the best known and most often used compact scheme in meteorological applications (e.g., Navon and Villiers 1987), is the pair,

$$A = (1, 4, 1)^T(1/6), \tag{2.7}$$

$$B = (-1, 0, 1)^T(1/2\delta). \tag{2.8}$$

We apply $-1/\delta$ times the "next" row of the matrix in order to obtain the principal coefficient of truncation error:

$$-(5, 0, 5) \cdot A/\delta - (1, 0, -1) \cdot B = -\frac{2}{3\delta} \equiv \left(\frac{d^5c}{dx^5}\right) \varepsilon^{(4)}\delta^4, \tag{2.9}$$

where,

$$\varepsilon^{(4)} = -\frac{1}{180}, \tag{2.10}$$

a six-fold reduction compared to the explicit formula.

At higher orders of accuracy, the replacement of one wide B-stencil by the two compact A- and B-stencils is even more advantageous in terms of the reduction of truncation error. Let a and b be the (integer) half-widths of the stencils A and B respectively, in grid space units, of a general scheme constructed according to the principles outlined above. Such a scheme has a formal order of accuracy of $2a + 2b$. In this notation, the ordinary explicit fourth-order scheme is denoted $(0, 2)$, our fourth-order

compact scheme is $(1, 1)$, and so on. Figure 1 is a log-log depiction of the relative truncation errors (absolute error divided by magnitude of true value) for a sine wave, plotted against the wavenumber. For a given order of accuracy, the most accurate scheme in the long-wave limit is, in each case, one of those possessing the most compact possible combination of templates.

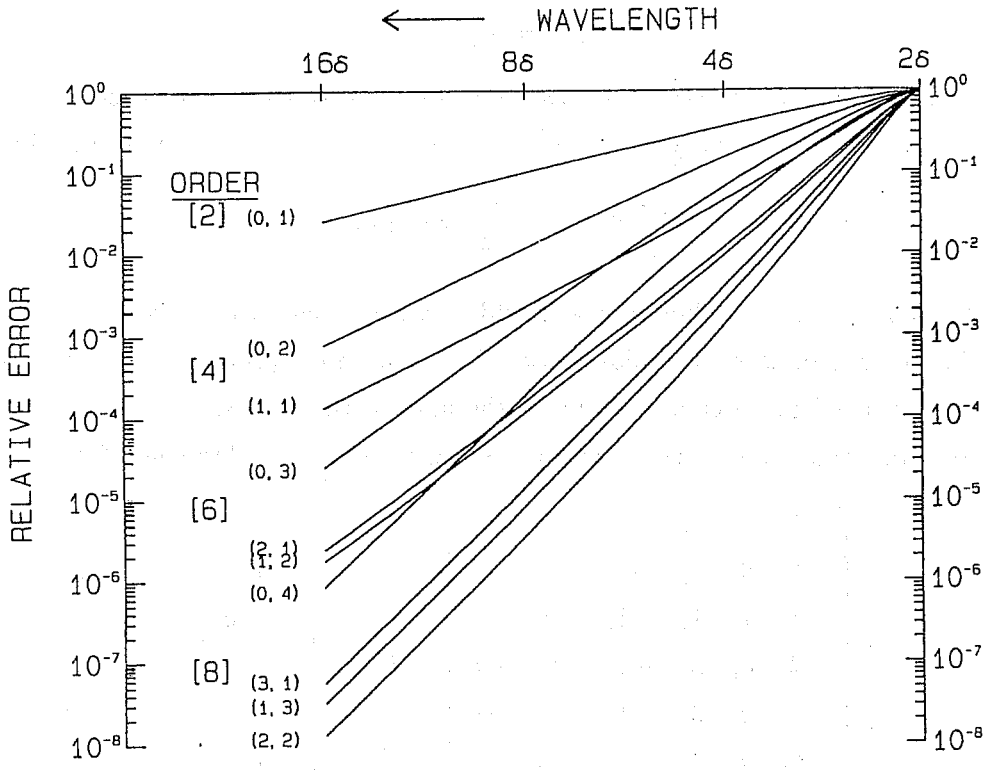

Figure 1. Log-log plot of the relative truncation error of unstaggered centred stencil differencing schemes as a function of the wavenumber of the sine-wave to which they are applied.

In order to solve for the derivative, d, using a compact scheme it is necessary to invert a linear system with a banded matrix $\mathbf{A}$ whose rows are each a copy of the vector A pertaining (in the case where the grid is no longer uniform) to the location that corresponds to that row's index. In the fourth-order example, the matrix $\mathbf{A}$ is tridiagonal. Inverting the linear system involving $\mathbf{A}$ each time "from scratch" is inefficient; instead, we "preprocess" by factorising $\mathbf{A}$:,

$$\mathbf{LDU} = \mathbf{A} \tag{2.11}$$

where $\mathbf{L}$ and $\mathbf{U}$ are lower and upper triangular band matrices with unit main diagonals, and $\mathbf{D}$ is purely diagonal. From $\mathbf{B}$, the band matrix whose rows are the vectors B of this compact scheme, we create the similarly shaped $\hat{\mathbf{B}}$:

$$\hat{\mathbf{B}} = \mathbf{D}^{-1}\mathbf{B} \tag{2.12}$$

and define

$$\hat{\mathbf{L}} = \mathbf{D}^{-1}\mathbf{L}\mathbf{D}, \tag{2.13}$$

which, like $\mathbf{L}$, is a lower band matrix with unit main diagonal elements. Now the computational procedure for finding d from c can be summarised:

$$e = \hat{\mathbf{B}}c, \tag{2.14}$$

$$\hat{\mathbf{L}}f = e, \tag{2.15}$$

$$\mathbf{U}d = f, \tag{2.16}$$

where the two last steps are recursive back-substitutions. The labour involved (not counting the manipulations done beforehand to get $\hat{\mathbf{L}}$, $\mathbf{U}$ and $\hat{\mathbf{B}}$) is equivalent to that required to compute the line of derivatives at the same order of accuracy by the corresponding explicit scheme.

2.2 Invertible differencing and quadrature with staggered grids

The formal numerical relationship in (2.5) between the derivative d and the corresponing "cumulative" c is more symmetrical than in the case of conventional differencing. Can we invert the roles of target and source data in the compact schemes in order to obtain the integrals c from d instead of obtaining derivatives d from c? With the unstaggered (either conventional or compact) schemes described above, the practical answer is no; the matrix $\mathbf{B}$ is never well-conditioned to inversion because, in the ideal generalisation of a grid extending both ways forever, the vector, $c = (\ldots, 1, -1, 1, -1, \ldots)$ is a null-eigenvector of $\mathbf{B}$. For a finite grid, the problem recurs as one of ill-conditioning for structures close to being "two-grid-length". In order to obtain well-conditioned invertible differencing formulae, whether conventional or compact, we must stagger the grid for the derivatives d with respect to the grid for the cumulative values c. Let the bandwidth parameter a be defined as before but, to keep parameter b an integer for the centred stencils B (containing an even number of active elements in each row) of the staggered grid, let b now be half this number of active elements of the generic row. For example, the fourth-order compact scheme, "(1, 1)", for the staggered grids of uniform spacing, δ, has coefficients,

$$A = (1, 22, 1)^T/24, \tag{2.17}$$

$$B = (-1/\delta, 1/\delta)^T. \tag{2.18}$$

For this, and other staggered schemes, a two-grid wave is no longer a null-eigenvector of the matrix $\mathbf{B}$ (the dot-product of a vector of alternating values with the vector of coefficients, B, is now finite). The staggered schemes have one fewer coefficients than their unstaggered counterparts; it would seem, at first, that one would attain one less than the desired order of accuracy using the Vandermonde matrix technique. However, on a uniform grid, the symmetries about the centre of the stencil conspire to make the solution coefficients A and B exact also for what would be the "next" row of the generalised Vandermonde matrix, thereby securing the order of accuracy desired. Figure 2 depicts the log-log plots of relative truncation error for the first few possible staggered schemes on a uniform grid. As before, the most compact combined stencil gives the most accurate scheme at each order. One crucial property of the staggered-grid schemes that we can exploit in a semi-Lagrangian

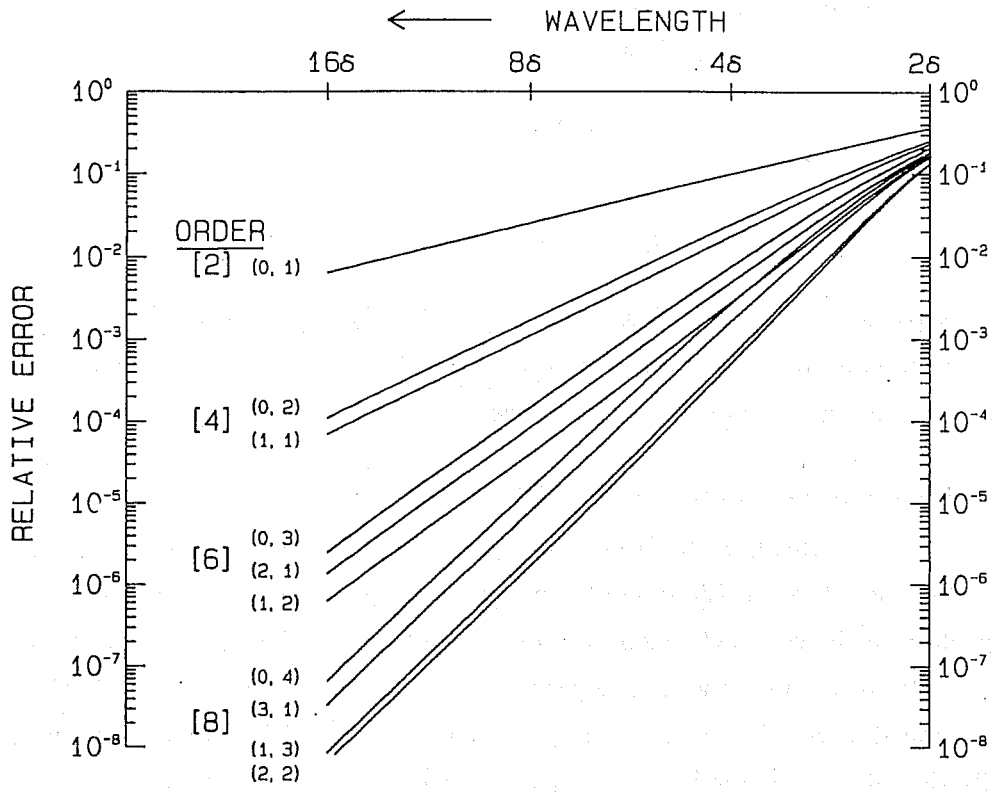

Figure 2. Same as figure 1 but for centred schemes with derivatives staggered with respect to the original data.

model is that, transforming from d to c, we incur no loss of information; the inverse operator is always available to recover the data in its original form. How this is used in practice is the subject of the next section.

3. CONSERVATIVE CASCADE INTERPOLATION USING COMPACT SCHEMES

3.1 Basic cascade

The "cascade" method breaks down the problem of semi-Lagrangian grid-to-grid interpolation into a separate stage for each dimension involved; at a given stage, only along the lines corresponding to the "active" dimension do the data explicitly interact with one-another. Let us denote the grid values of X, Y, and Z, by the sets, $\{\hat{X}\}$, $\{\hat{Y}\}$, $\{\hat{Z}\}$, and assume the full grid to be the Cartesian product of these. In the cascade method's simplest form, proposed by Purser and Leslie (1991), the interpolation of some variable, ψ, from a grid $\{\hat{X}, \hat{Y}, \hat{Z}\}$ to a topologically similar, but relatively distorted, grid, $\{\hat{x}, \hat{y}, \hat{z}\}$, is not delivered piece-meal, one target value at a time (which is the more obvious and conventional procedure), but is delivered whole-sale as the culmination of an orderly sequence (the "cascade") stratified by the space dimensions:

$$\psi(\hat{X}, \hat{Y}, \hat{Z}) \rightarrow \psi(\hat{x}, \hat{Y}, \hat{Z}), \tag{3.1}$$

$$\psi(\hat{x}, \hat{Y}, \hat{Z}) \rightarrow \psi(\hat{x}, \hat{y}, \hat{Z}), \tag{3.2}$$

$$\psi(\hat{x}, \hat{y}, \hat{Z}) \rightarrow \psi(\hat{x}, \hat{y}, \hat{z}). \tag{3.3}$$

The first stage requires us to know (or discover) the coordinate X of each of the intersections that form the grid $(\hat{x}, \hat{Y}, \hat{Z})$. For example, if $(\hat{X}, \hat{Y}, \hat{Z})$ are the Lagrangian grid intersections of a forward-trajectory model, and we know the x, y and z coordinates of these trajectory end points (from the time integration of the dynamics), then a one-dimensional "Newton" iteration at each intersection will suffice to locate all the requisite coodinates, $X(\hat{x}, \hat{Y}, \hat{Z})$. The first stage of the cascade would interpolate, not only ψ, but also y and z. This makes $y(\hat{x}, \hat{Y}, \hat{Z})$ available at the second stage, from which, again by Newton iterations, the values of $Y(\hat{x}, \hat{y}, \hat{Z})$ may then be found. At the second cascade stage, $z(\hat{x}, \hat{Y}, \hat{Z})$ is interpolated to $z(\hat{x}, \hat{y}, \hat{Z})$, and $\psi(\hat{x}, \hat{Y}, \hat{Z})$ is similarly interpolated to $\psi(\hat{x}, \hat{y}, \hat{Z})$ so that, in preparation for the final stage, $Z(\hat{x}, \hat{y}, \hat{z})$ may be found.

It is an important feature of the cascade method that interpolation from a Lagrangian (distorted) grid to the regular Eulerian grid is practically as easy as the converse. Therefore, there remains no impediment to using forward (downstream) trajectories. Janjić (1995) argues that it is physically more justifiable to adopt the downstream choice of semi-Lagrangian construction, but there are practical advantageous also, one of which, the ability to adopt the second-derivative form of the momentum equations, is discussed in section 4. Formally, the solution obtained by treating the x-direction before the y-direction in the interpolation would differ from that obtained by reversing the order. Although experiments to examine this difference have not been carried out, the direct comparison of the cascade interpolation with results from a Cartesian product of conventional Lagrange polynomial interpolators show the differences to be negligible.

3.2 Incorporating conservation.

Although at high-order, N, the cascade method is more efficient than the Cartesian-product interpolation (by which each target is surrounded by a roughly centred box of $N \times N \times N$ source grid points), it does not automatically guarantee that the variable interpolated is conserved in any formal sense. However, there *are* ways to conserve mass and tracers within the cascade by incorporating techniques originally designed for a single dimension (e.g., Emde 1992; Harten et al. 1987). These are probably more satisfactory than *a posteriori* interventions to restore conservation in response to a monitoring of global diagnostics (Navon and de Villiers 1987; Priestley 1993; Gravel and Staniforth 1994) where there is no guarantee that the restoration occurs at the appropriate places. One example of the use of the cascade procedure to achieve conservation was the adaptation by Rančić (1995) of the piece-wise parabolic method (PPM) of Colella and Woodward (1984), Carpenter et al. (1990). Another method, allowing generalisation to any formal order of accuracy, was proposed by Leslie and Purser (1995) and uses the invertible (staggered) compact schemes we described in the previous section. We shall give an outline of this method.

For an extensive quantity, M, (such as mass), we define the corresponding intensive variable, ψ, which is essentially a form of density, to be the "M"-per-coordinate-volume. We express this quantity by generalising the notation for Jacobians:

$$\psi(X, Y, Z) = \frac{\partial M}{\partial(X, Y, Z)}. \tag{3.4}$$

Let us denote the staggered grids which, for bounded coordinate domains, extend half a grid space beyond each end of the original grid, by $\{\hat{X}'\}$, $\{\hat{Y}'\}$, $\{\hat{Z}'\}$. Suppose we now integrate $\psi(\hat{X}, \hat{Y}, \hat{Z})$ with respect to X along all the grid lines, $\{\hat{Y}, \hat{Z}\}$, of constant Y and Z, to the set of staggered locations, $\{\hat{X}'\}$, using the inverse of one of the staggered (and therefore invertible) compact differencing schemes of the previous section. The interpolation that follows is of this field of partial integrals, which it is convenient to denote,

$$\left.\frac{\partial M}{\partial(Y,Z)}\right|_{\{\hat{X}',\hat{Y},\hat{Z}\}} = \int^{\hat{X}'} \frac{\partial M}{\partial(X,Y,Z)} dX \Bigg|_{\{\hat{Y},\hat{Z}\}} . \tag{3.5}$$

At this point, the cascade proceeds by interpolating, not ψ itself, but its partial integral given by (3.5) above, and the target grid is similarly staggered, but in x:

$$\left.\frac{\partial M}{\partial(Y,Z)}\right|_{\{\hat{X}',\hat{Y},\hat{Z}\}} \rightarrow \left.\frac{\partial M}{\partial(Y,Z)}\right|_{\{\hat{x}',\hat{Y},\hat{Z}\}} . \tag{3.6}$$

In the simplest case of a bounded domain with rigid end walls, the interpolated quantity at the two ends of the domain are unchanged; their difference along a given line of $\{\hat{Y}, \hat{Z}\}$ is the conserved total contribution, $\partial M/\partial(Y,Z)$, along this line to the domain integral,

$$M = \int\int \frac{\partial M}{\partial(Y,Z)} dY\, dZ, \tag{3.7}$$

of the substance. We differentiate these partial integrals, along the same grid lines, but now with respect to x instead of X:

$$\frac{\partial M}{\partial(x,Y,Z)} = \frac{\partial}{\partial x}\left(\frac{\partial M}{\partial(Y,Z)}\right)_{(\hat{Y},\hat{Z})} . \tag{3.8}$$

This completes the first stage of the three-stage conservative cascade. The remaining two stages are similar. Figure 3 shows a schematic progression of the conserving cascade in the case of two dimensions.

3.3 Remarks.

The conserving cascade is obviously more complicated and typically a little more than two times more expensive to apply than the simplest non-conserving cascade of equivalent order or accuracy. However, when mass is one of the conserved variables treated in this way, an additional bonus is that the vertical motion in a hydrostatic model is obtained automatically. This aspect is discussed in detail in Leslie and Purser (1995).

A limitation of the cascade procedure we have described is that, in its simplest forms, it is not applicable to points close to the polar coordinate singulariries. It is probably this, more than anything else, that has hindered the application of the cascade method to global models. Modifications of the original nonconserving cascade (Nair et al. 1999) seem better able to cope with these problems. An approach being taken at NCEP is to retain the latitude-longitude-based grid away from the poles and to employ a stereographic cartesian "patch" centred over each polar region. Computations carried out within one of these patches are then immune to singularity problems and the computed results at each time step (or more frequently, if required) can be matched and smoothly blended with the standard grid's solution at the outer regions of the patch where the two separate computational domains

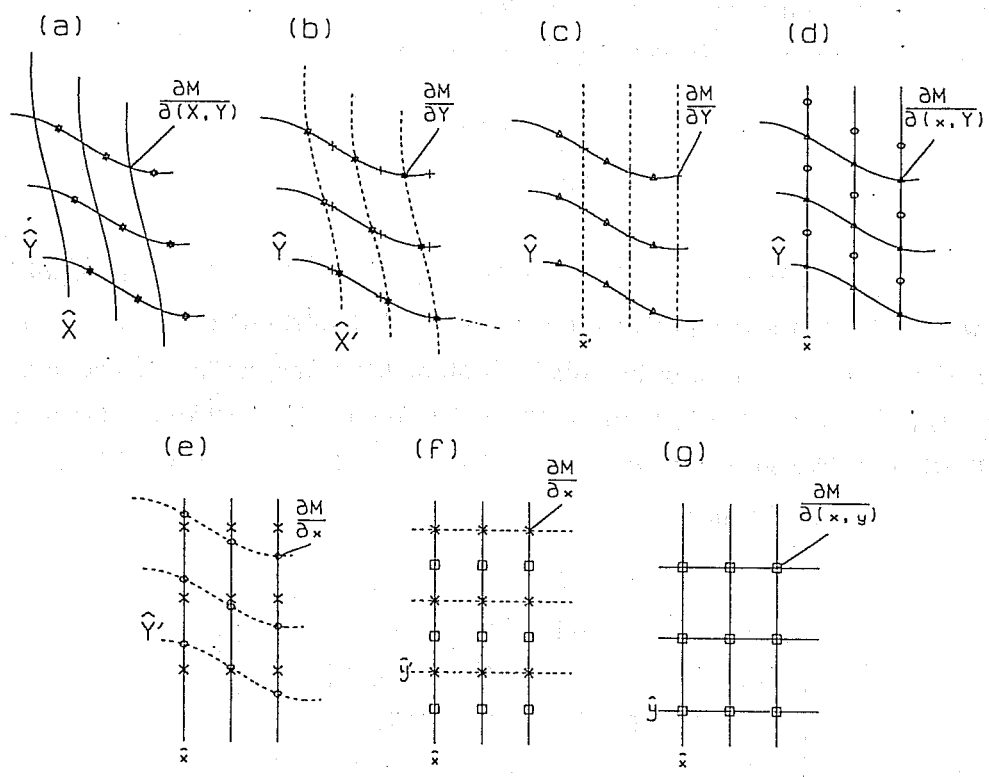

Figure 3. Schematic representation of the successive stages of the conserving casade in two dimensions. Starting at (a) with the coordinate-density of the substance, M, at all the intersections of the solid grid lines,this density is integrated in $\hat{X}$ to the $\hat{X}'$ grid (b). These partial integrals are interpolated to the semi-staggered hybrid $(\hat{x}', \hat{Y})$-grid (c), then differentiate with respect to x at the $(\hat{x}, \hat{Y})$-grid (d) to complete the first stage of the cascade. The panels, (e), (f), (g) illustrate corresponding steps making up the second stage of this cascade, and are the y-direction counterparts to the steps shown in (b), (c), and (d).

overlap. A somewhat similar technique is applicable to the overlaps between quasi-rectangular computational sub-domains when the data of a large horizontal domain are dispersed across processors of a massively parallel computer. Tests at NCEP indicate that no significant spurious numerical artifacts are introduced into the evolving solution using this technique, providing the blending of alternative solutions is carried out in a progressive, smooth, way across a sufficient breadth of overlap. While, formally, the blending of independently derived solutions invalidates the claim of exact conservation, in practice, such errors are hard to detect. The relative extents of the polar and intermediate domains are somewhat arbitrary; some investigation is needed to locate the optimal latitudes for the transition zones, although the present approach places these as close to the poles as possible, in order to reduce the amount of computations required to deal with the blended overlaps.

4. SECOND DERIVATIVE KINEMATICS/MOMENTUM

In a conventional model the horizontal equations on a Cartesian grid might be written,

$$\frac{du}{dt} - fv = F_x, \tag{4.1}$$

$$\frac{dv}{dt} + fu = F_y, \tag{4.2}$$

where F_x and F_y are components of the pressure gradient and frictional force in the horizontal. For a semi-Lagrangian model, we would invoke the kinematic equations,

$$\frac{dx}{dt} = u, \tag{4.3}$$

$$\frac{dy}{dt} = v, \tag{4.4}$$

to integrate the trajectories. In Purser and Leslie (1994) the ability of the cascade interpolation method to handle forward trajectories was put to use in a model that integrated the above momentum and kinematic equations using the third-order Adams-Bashforth scheme, whose attractive properties for numerical models was earlier noted by Durran (1991). Although it is a well known method, we observe that the application of generalised Vandermonde systems, similar to those used in section 2, is one way to derive its coefficients formally:

$$\begin{bmatrix} 1, & 1, & 1; & 0, & 0 \\ 0, & 0, & 0; & -1, & -1 \\ 1, & 1, & 1; & 0, & -1 \\ -4, & -2, & 0; & 0, & -1 \\ 12, & 3, & 0; & 0, & -1 \end{bmatrix} \begin{bmatrix} A_{-2} \\ A_{-1} \\ A_0 \\ B_0\delta t \\ B_1\delta t \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{4.5}$$

for the scheme written in generic notation, using (4.3) as an example with τ as the time step index, δt the time interval:

$$x^{\tau+1} - x^{\tau} = (A_0 u^{\tau} + A_{-1} u^{\tau-1} + A_{-2} u^{\tau-2})\delta t, \tag{4.6}$$

The weights of this scheme, solving (4.5), are $A_0 = 23/12$, $A_{-1} = -16/12$ and $A_{-2} = 5/12$. Table 1 contains the weights of this and other Adams-Bashforth schemes, together with their coefficients of principal truncation error.

TABLE 1. THE CLASSICAL ADAMS-BASHFORTH SCHEMES.

Order	η	ηA_0	ηA_{-1}	ηA_{-2}	ηA_{-3}	ηA_{-4}	ε
1	1	1					$\frac{1}{2}$
2	2	3	-1				$\frac{5}{12}$
3	12	23	-16	5			$\frac{3}{8}$
4	24	55	-59	37	-9		$\frac{251}{720}$
5	720	1901	-2774	2616	-1274	251	$\frac{95}{288}$

The enhanced accuracy of schemes of this type is bought at the price of rather substantial storage requirements — essentially four fields of storage for each prognostic equation treated this way in the case of the third-order scheme. This cost may be reduced by combining the momentum and kinematic equations, making a pair of what we refer to as the "second-derivative" form of trajectory equations:

$$\frac{d^2x}{dt^2} - f\frac{dy}{dt} = F_x, \tag{4.7}$$

$$\frac{d^2y}{dt^2} + f\frac{dx}{dt} = F_y. \tag{4.8}$$

In the absence of the Coriolis terms, these equations become separate and may be handled by a generalisation of the Adams-Bashforth family of schemes given by Störmer (1907). Again, the generalised Vandermonde matrix technique can be used to find the coefficients. In the Störmer schemes obtained for vanishing Coriolis, the *three* weights, B_1, B_0, B_{-1}, for $x^{\tau+1}$, x^{τ} and $x^{\tau-1}$, respectively, are $1/\delta t^2$, $-2/\delta t^2$, $1/\delta t^2$. The weights for forcing terms, F_x, and the principal error coefficients for these schemes are listed in table 2. The further generalisation of these second-derivative schemes to incorporate the Coriolis force in a fully implicit way was described by Purser and Leslie (1996) and was based on the fact that the pair, (4.7), (4.8), may be combined as a single *complex* equation,

$$\frac{d^2\chi}{dt^2} + if\frac{d\chi}{dt} = F, \tag{4.9}$$

where $\chi = x + iy$ relative to a local Cartesian frame, and $F = F_x + iF_y$. Once again, it is possible to use the generalised Vandermonde matrix technique, but the evaluations must be repeated for each of the standard latitudes where the momentum/kinematic equations are to be applied, since the matrix itself is a function of the Coriolis parameter. For example, in the simplest "generalised Adams-Bashforth" scheme, GAB1, we may write the Vandermonde system,

$$\begin{bmatrix} 1; & 0, & 0, & 0 \\ 0; & -1, & -1, & -1 \\ if\delta t; & 1, & 0, & -1 \\ 2; & -1, & 0, & -1 \end{bmatrix} \begin{bmatrix} A_0 \\ B_{-1}\delta t^2 \\ B_0\delta t^2 \\ B_1\delta t^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{4.10}$$

To obtain the weights, $A_0 = 1$, $B_{-1} = 1/\delta t^2 - if/\delta t$, $B_0 = -2/\delta t^2$, $B_1 = 1/\delta t^2 + if/\delta t$. The higher order schemes of the same family are obtained by adding earlier forcing terms F and their corresponding weights B in a Vandermonde system with enough rows (based on the successive test functions of the form $\chi = \tau^p$) to make the matrix square.

TABLE 2. THE STÖRMER SCHEMES.

Order	η	ηA_0	ηA_{-1}	ηA_{-2}	ηA_{-3}	ηA_{-4}	ϵ
2	1	1	0				$\frac{1}{12}$
3	12	13	-2	1			$\frac{1}{12}$
4	12	14	-5	4	-1		$\frac{19}{240}$
5	240	299	-176	194	-96	19	$\frac{3}{40}$

An essential practical requirement of a time integration scheme used in a semi-Lagrangian model is that it can be made semi-implicit, at least for the fastest gravity waves (including the external modified "Lamb" wave, which propagates at sonic speed). It is *not* necessary, indeed, it may be undesirable, to treat *all* internal gravity waves implicitly because the higher internal modes cannot objectively be disentangled from what we should regard as "meteorological" structures and an attempt to treat them implicitly would usually degrade the accuracy of their representation. Following Burridge (1975), we recommend that only the very deepest vertical gravity modes be projected out for implicit handling. Each vertical mode can then be associated with a (linearised) equivalent shallow-water system and the

adjustment field derived for the solution of a Helmholtz equation in only the horizontal dimensions. This is considerably more efficient, in a grid point model, than attempting to solve for *all* the vertical modes' adjustment terms. In the case of the generalised Adams-Bashforth schemes, it is possible to organise the calculations in a way that makes the application of the semi-implicit machinery a little more streamlined, as discussed in Purser and Leslie (1996), but a straightforward modal projection should also suffice.

Another class of time integration schemes for the second-derivative trajectory equations, are the "generalised Lorenz N-cycle" schemes introduced by Purser and Leslie (1997) (a variant of the Runge-Kutta-style N-cycle schemes of Lorenz 1971). The goal of these schemes was to obtain a higher formal accuracy in time without incurring an additional storage burden. For purely linear equations, it was possible to find fourth-order in time schemes, but to do so required one of the cycle of N time steps to be negative. The fourth-order of accuracy did not carry over to the nonlinear aspects of the dynamics, so it remains questionable whether such schemes would sufficiently justify the significant practical inconvenience in a full-physics model of having to deal with the negative time step. Even *without* going to the trouble of accommodating higher-order time integration, employing the second-derivative form of momentum and kinematics with only the simplest symmetric time discretization in this framework results in inherently smaller time truncation and greater numerical robustness. While in the idealised experiments of Purser and Leslie (1996, 1997) the higher-order time schemes were slightly beneficial (but more costly to apply) the question of whether it is worthwhile to exploit these methods in a given forecasting of simulation application needs to be re-examined, and the potential advantages balanced against the additional cost and complexity, in each particular context.

5. STEREOGRAPHIC REPRESENTATIONS OF TRAJECTORIES

Polar stereographic frames The numerical description of trajectory displacements in a global semi-Lagrangian model requires some care if this description is to be sufficiently general to apply at all points and to be capable of computationally efficient implementation. It is convenient, in using the second-derivative kinematics/momentum equations, if the coordinates used to describe a local trajectory do not require the introduction of additional terms (a "fictitious force") to correct for the intrinsic curvature of these coordinates; this condition is satisfied if locally cartesian horizontal coordinates are employed, centred on the trajectory's associated Eulerian grid point. It is also desirable that the displacements be expressed in a way that facilitates their spatial interpolation, which the cascade procedure requires; this requires that it be relatively straightforward to transform from one local frame to another, usually nearby one. While there are various choices satisfying these requirements, a particularly useful convention is to use local stereographic frames. As we shall find below, once all the trajectory information is expressed in terms of stereographic map displacements, the need for carrying out any explicit trigonometric computations in the course of the semi-Lagrangian computations is completely avoided. This is another way in which the computational cost of implementing a global semi-Lagrangian model may be minimised.

5.1 Polar stereographic frames.

Let cartesian coordinates $\mathbf{E} \equiv (E_1, E_2, E_3)$ be defined, with respect to which the earth is a unit sphere centered at the origin. Orient the system such that the South and North poles are at $\mathbf{S} \equiv (0, 0, -1)$ and $\mathbf{N} \equiv (0, 0, 1)$ and the prime meridian has $E_1 > 0$ and $E_2 = 0$. Identify the standard north-polar stereographic map coordinates (x, y) of the generic point $\mathbf{E}$ as the values such that the point $(x, y, 0)$ is the intersection of the equatorial plane with the segment $\mathbf{SE}$. As we see from fig. 4, if $r^2 \equiv x^2 + y^2$ and $R^2 = E_1^2 + E_2^2$, we find $R = r/m$ for $m = 1/(1 + E_3)$ and $r = \tan(\phi/2)$, where ϕ is the co-latitude of $\mathbf{E}$. Since $R^2 + E_3^2 = 1$,

$$(1 + E_3)^2 r^2 = 1 - E_3^2, \tag{5.1}$$

hence,

$$m = \frac{1 + r^2}{2}. \tag{5.2}$$

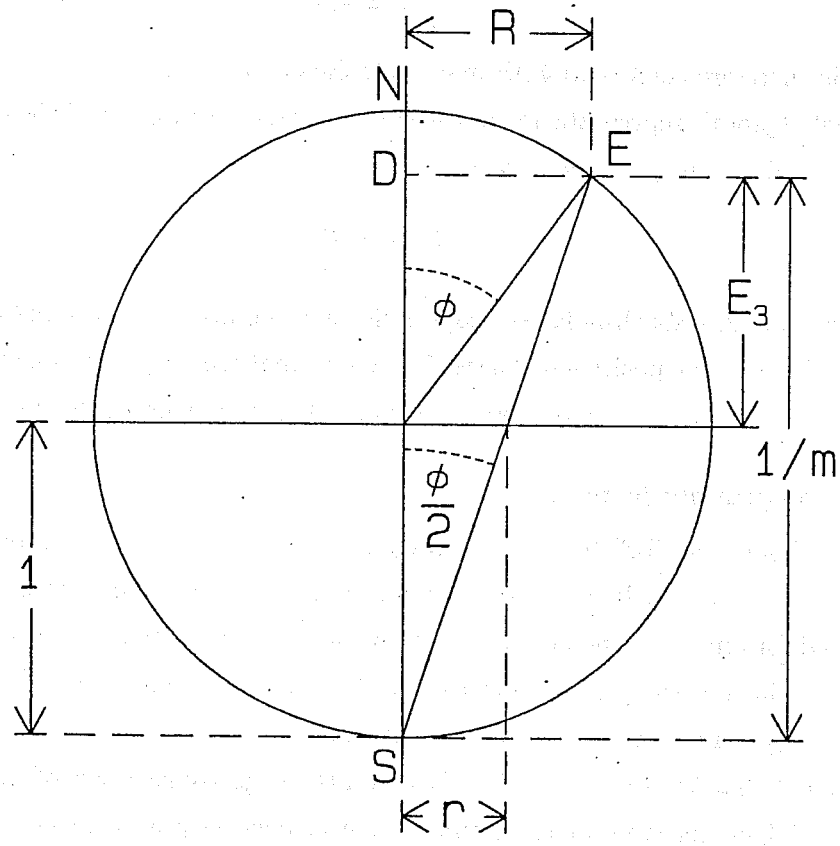

Figure 4. Construction of the polar stereographic mapping.

Therefore, we convert Cartesians, (E_1, E_2, E_3) to polar stereographic coordinates, (x, y) using,

$$\begin{aligned} m &= 1/(1 + E_3), \\ x &= mE_1, \\ y &= mE_2, \end{aligned} \tag{5.3}$$

and convert polar stereographics to Cartesians using (5.2) followed by,

$$\begin{aligned} E_1 &= x/m, \\ E_2 &= y/m, \\ E_3 &= (1-m)/m. \end{aligned} \tag{5.4}$$

Now, taking partials:

$$\begin{bmatrix} \partial E_1/\partial x, & \partial E_1/\partial y \\ \partial E_2/\partial x, & \partial E_2/\partial y \\ \partial E_3/\partial x, & \partial E_3/\partial y \end{bmatrix} = \begin{bmatrix} 1/m - E_1^2, & -E_1E_2 \\ -E_1E_2, & 1/m - E_2^2 \\ -E_1/m, & -E_2/m \end{bmatrix} \tag{5.5}$$

we can verify that,

$$\frac{\partial \mathbf{E}}{\partial x_i} \cdot \frac{\partial \mathbf{E}}{\partial x_j} = \frac{1}{m^2}\delta_{ij},$$

$$\frac{\partial \mathbf{E}}{\partial x_i} \cdot \mathbf{E} = 0.$$

The mapping is therefore conformal with map scale factor m.

A convenient algebraic representation of the stereographic coordinates of the point E is the complex number $E_N \equiv (x, y) \equiv x + iy$. Its reciprocal,

$$E_S = 1/E_N$$

is then the corresponding south-polar stereographic representation of the same point with, again, the prime meridian forming the positive real axis. The conformal property is preserved under compositions by complex analytic functions — a property we exploit below in defining "local" stereographic frames.

5.2 Local stereographic frames.

The local stereographic (LS) frame, with projection "pole" at the location on the trajectory where the acceleration is required, will permit this acceleration to be estimated by temporal finite differencing without the need for any compensating calculations of the coordinate frame's intrinsic curvature. But using local frames in this way presupposes that we be able to transform among them with ease. Here, we review some algebraic machinery to accomplish this.

It is convenient that the North pole be placed on the negative real axis of each local stereographic complex plane. Then the label of the point at the projection pole may be used also to label this local stereographic frame unambiguously. Let A be a non- polar point on the prime meridian with complex representation A_N relative to the frame at the North pole. From the relationship between A's co-latatitude and the radial distance of its image in the polar stereographic map , and using the standard formula for tangents of the sums or the differences of angles, we deduce that another point B on the prime meridian with polar stereographic representation B_N must have a representation relative to frame A of

$$B_A = \frac{B_N - A_N}{1 + B_N A_N}. \tag{5.6}$$

We argue that, by virtue of the transitive property of conformality, under complex analytic functions, the formula above must remain true also for B *not* on the prime meridian, since this formula is manifestly analytic and therefore already respects the conformal mapping property we require. When, in addition, the longitude λ_A of A is also nonzero, it is easily shown (by applying to all points the rotation required to bring the longitude of A back to zero) that the appropriate generalization of the transformation formula is,

$$B_A = \frac{B_N - A_N}{1 + B_N\, A_N^*} e^{-i\lambda_A}. \tag{5.7}$$

Note that the pole ($N_N = 0$) is itself transformed to

$$N_A = -|A_N|. \tag{5.8}$$

Conversely, if the location of the pole, $N_A < 0$ and that of B, B_A, are both given relative to the frame of A, we may convert to the polar stereographic image, B_N, by:

$$B_N = \frac{B_A - N_A}{1 + B_A\, N_A} e^{i\lambda_A}. \tag{5.9}$$

We can combine these formulae to produce the transformation between any pair of LS frames. However, when transforming representations of C between frames A and B, where A, B, C are all closer to each other than any one of them is to the pole, N, a smaller roundoff error is expected by employing a more direct transformation which avoids the errors incurred by subtracting pairs of relatively large, almost equal, magnitudes to extract a much smaller residual. The more direct formula possible when the position of B is known relative to A is:

$$C_B = \frac{C_A - B_A}{1 + C_A\, B_A^*} \varepsilon_{AB}, \tag{5.10}$$

where

$$\varepsilon_{AB} \equiv e^{i\gamma(A,B)} \tag{5.11}$$

for some angle $\gamma(A, B)$, an antisymmetric real-valued function of the two locations A and B. For an interpretation of γ observe that, applying (5.10) to transform A itself, we obtain,

$$A_B = -B_A \varepsilon_{AB} \tag{5.12}$$

and therefore, the rotation factor ε_{AB} simply represents the relative change in orientation of the frames A and B (for example, compared to the geodesic "great circle" that joins them). Observe also that, by differentiating (5.10), if δC_A is an infinitesimal displacement at A, its representation in frame B becomes,

$$\delta C_B = (1 + B_A B_A^*) \delta C_A \varepsilon_{AB}. \tag{5.13}$$

To define algebraically the rotation factor ε_{AB}, observe that the pole must transform by (5.10) to give

$$-|B_N| = -\left(\frac{|A_N| + B_A}{1 - |A_N|\, B_A^*} \right) e^{i\gamma} \tag{5.14}$$

If $|B_N|$ is not already known, it can be determined from

$$|B_N| = \left| \frac{|A_N| + B_A}{1 - |A_N| B_A} \right| \tag{5.15}$$

and hence,

$$e^{i\gamma} = \frac{\left(\frac{|A_N|+B_A}{1-|A_N|B_A^*} \right)}{\left| \frac{|A_N|+B_A}{1-|A_N|B_A^*} \right|}. \tag{5.16}$$

However, in practical calculations, we shall see in the next subsection that there are better computational methods when A, B and C are close together relative to the scale of the earth, as is the case in computations involving short-duration atmospheric trajectories.

5.3 Application of stereographic transformations to trajectories.

We consider the case where the standard computational grid is of latitude-longitude type and one point, A, of this grid is used to define the local stereographic frame. The associated Lagrangian trajectory, also nominally labeled by A and passing through or very close to this Eulerian grid point, is, at some future point B, given by its complex representation relative to A: $B_A = x + iy$. In the forward-trajectory semi-Lagrangian model using cascade interpolation (for example, see Leslie and Purser 1995) we are now faced with several tasks. We shall asume that the order in which the interpolations are carried out is: first latitude (y); then longitude (x).

First, we need to deduce from the fixed polar stereographic position, A_N, of A and from B_A the relative latitude (or some function of it) of B. Then we are in a position to accomplish the first stage of the cascade interpolation by locating the hybrid grid targets, which consist of the intersections of the discrete Lagrangian longitudes with the standard Eulerian latitudes. Second, in order to convert the coordinates of some third point C, or to compare vectors by parallel-transport between A and B we shall need to compute the rotation factor ε_{AB} between the two frames as defined in the previous subsection. Third, we need to compute the relative displacement of (Eulerian) longitude (or some function of it) at the new hydrid grid points. As always, we shall avoid explicitly invoking any trigonometric functions in the actual calculations, for reasons of computational efficiency. This means that, instead of using or computing any angles directly, we shall arrange all calculations in terms of the sines, cosines or tangents of the angles. Thus, for our first goal, we aim to compute the *tangent* of *half* the latitude displacement of B from A.

For convenience, let us measure longitude from the meridian through A and abbreviate A_N by A, B_N by B. Thus, our first objective is to evaluate,

$$\xi^y = -\left(\frac{|B| - |A|}{1 - |A||B|} \right) \equiv -\left(\frac{|B| - A}{1 - A|B|} \right), \tag{5.17}$$

but in a way that does not explicitly require us to subtract A from $|B|$. Writing,

$$B_A \equiv (x, y), \tag{5.18}$$

$$D = 1 - A^* B_A \equiv (1 - Ax, -Ay), \tag{5.19}$$

$$E = 1 + B_A/|A| \equiv (1 + x/A, y/A). \tag{5.20}$$

we note that

$$B = \frac{A + B_A}{1 - A^* B_A} \equiv \frac{AE}{D}. \tag{5.21}$$

We assume we have precomputed,

$$F = \frac{1 - A^2}{2A}. \tag{5.22}$$

From B_A, we get,

$$r^2 = x^2 + y^2 \equiv B_A B_A^*, \tag{5.23}$$

hence the "map factor" at B relative to A (or vice-versa),

$$m = \frac{1 + r^2}{2} \equiv \frac{1 + B_A B_A^*}{2}, \tag{5.24}$$

and the quantity,

$$p = \frac{F - x}{m}. \tag{5.25}$$

Then we find that,

$$|D| = A \left[2m(1 + \frac{p}{A}) \right]^{\frac{1}{2}}, \tag{5.26}$$

$$|E| = \frac{1}{A} \left[2m(1 - Ap) \right]^{\frac{1}{2}}, \tag{5.27}$$

and, since,

$$\xi^y = - \left(\frac{|E| - |D|}{A|E| + \frac{1}{A}|D|} \right), \tag{5.28}$$

multiplying top and bottom by $|E| + |D|$ and simplifying leads to:

$$\xi^y = \frac{-g}{1 + H}, \tag{5.29}$$

where:

$$g = 2x + pr^2, \tag{5.30}$$

and

$$H = \sqrt{(1 + pg)}. \tag{5.31}$$

(Algebraically, $g = 2F - p$, but the definition (5.30) is better for minimizing round-off). Note that, upon interchanging A and $-1/A$, we interchange D and E while leaving F and ξ^y unchanged.

In the previous subsection we showed how the application of the transformation equation (5.10) to the pole led to a calculation of the rotation factor, ϵ_{AB}. Applying (5.16) where B is expressed as in (5.21) we find that

$$\begin{aligned} \epsilon_{AB} &= \frac{\left(\frac{A + B_A}{1 - AB_A^*} \right)}{|B|} \equiv \frac{|ED|}{ED}, \\ &= \frac{2mH}{ED}. \end{aligned} \tag{5.32}$$

But,

$$\frac{ED}{2m} = 1 + pB_A, \tag{5.33}$$

hence,

$$\varepsilon_{AB} = \frac{H}{1 + pB_A}. \tag{5.34}$$

There remains the task of computing at B a function of the longitudinal displacement relative to A. Let $\Re$ and $\Im$ denote the real and imaginary components of the complex argument. As before, we choose for the angle's function the tangent of the half-angle:

$$\xi^x = \frac{\Im(B)}{|B| + \Re(B)}, \tag{5.35}$$

with

$$|B| = \frac{A + \xi^y}{1 - A\xi^y}. \tag{5.36}$$

But an alternative approach is to use,

$$\xi^x = -\frac{t_\varepsilon + t_D}{1 - t_\varepsilon t_D}, \tag{5.37}$$

where

$$t_\varepsilon = \frac{\Im(\varepsilon_{AB})}{1 + \Re(\varepsilon_{AB})}, \tag{5.38}$$

$$t_D = \frac{\Im(D)}{\Re(D)} \equiv \frac{-Ay}{1 - Ax}, \tag{5.39}$$

which avoids ever evaluating $|B|$ itself.

It might be helpful here to summarize these calculations. Assume we begin with the local stereographic components of the displacement B_A. We compute r^2 and the map-factor m. Using the precomputed table of F to look up its value at the standard grid latitude of A we form p, g and H using (5.25), (5.30) and (5.31). We obtain ξ^y from (5.29), ε_{AB} from (5.34). Then, using (5.38) and (5.39) to get t_ε and t_D, we complete these manipulations by applying (5.37) to get ξ^x.

Analogous methods can be used for the two regions of the globe covered by the cartesian discretizations of the two polar stereographic projections, the so-called "polar patches" discussed at the end of section 3. However, in this case the orientations of the LS frames, to line up with the local grid, are chosen to be parallel (along the meridian) to the standard polar stereographic frame. The transformation (5.7) simplifies to,

$$B_A = \frac{B - A}{1 + BA^*}. \tag{5.40}$$

The grid lines in the polar patch delineate the real and imaginary components of the complex polar stereographic representation of the geographical points; these lines therefore form a pair of families of circles on the sphere, each family intersecting tangentially at the opposite pole.

Appropriate increments, ξ^x and ξ^y, in this case are

$$\xi^x = \Re(B - A) \equiv \Re(B_A/D)(1 + AA^*), \tag{5.41}$$

$$\xi^y = \Im(B - A) \equiv \Im(B_A/D)(1 + AA^*), \tag{5.42}$$

where, as in (5.19),

$$D = 1 - B_A A^*. \tag{5.43}$$

The transformation of a third point, C, from frame A to frame B now conforms to (5.10), but with the rotation factor,

$$\varepsilon_{AB} = \frac{D^*}{D}. \tag{5.44}$$

Thus, the case of the polar stereographic grid is actually easier to deal with.

5.4 Kinematics in a stereographic frame.

Since the map factor m at the origin of a stereographic frame A is $1/2$, the instantaneous velocity u_A and acceration ($\equiv$ specific force) F_A of a particle at the origin are

$$U_A = 2\dot{C_A}|_{C=A}, \tag{5.45}$$

$$F_A = 2\ddot{C_A}|_{C=A}, \tag{5.46}$$

where we extend the complex number representation to these two derivatives. In order to quantify the effect, on the image of the particle trajectory, of the intrinsic map curvature at a generic point, we differentiate (5.10), assuming C_B to be a function of time:

$$\dot{C_B} = \frac{(1+B_AB_A^*)}{(1+C_AB_A^*)^2}\dot{C_A}\varepsilon_{AB}, \tag{5.47}$$

$$\ddot{C_B} = \frac{(1+B_AB_A^*)}{(1+C_AB_A^*)^2}\ddot{C_A}\varepsilon_{AB} - \frac{2B_A^*(1+B_AB_A^*)}{(1+C_AB_A^*)^3}\dot{C_A}^2\varepsilon_{AB}. \tag{5.48}$$

Thus, recalling that the map factor $m = (1+B_AB_A^*)/2$

$$\dot{C_B}|_{C=A} \equiv U_Bm = U_Am\varepsilon_{AB}, \tag{5.49}$$

$$\ddot{C_B}|_{C=A} \equiv (F_B + F_B')m = F_Am\varepsilon_{AB} + 2\dot{C_B}^2A_B^*/2m, \tag{5.50}$$

where the "fictitious' force, F_B', owes entirely to the intrinsic curvature of the coordinates. Combining (5.49) and (5.50), this complex representation of the fictitious force is given by,

$$F_B' = U_B^2A_B^*. \tag{5.51}$$

We can also express this result using vector notation. Use $A_B = x + iy$, $U_B = u + iv$ to define vectors $\mathbf{x} \equiv (x, y)$ and $\mathbf{U} \equiv (u, v)$. Then, expanding the components of (5.51) we find that the vector form $\mathbf{F}'$ of the fictitious force is expressible:

$$\begin{aligned}\mathbf{F}' &= -\mathbf{x}(\mathbf{U}\cdot\mathbf{U}) + 2\mathbf{U}(\mathbf{x}\cdot\mathbf{U}),\\ &= \mathbf{k}\times\mathbf{U}[\mathbf{k}\cdot(\mathbf{x}\times\mathbf{U})] + \mathbf{U}(\mathbf{x}\cdot\mathbf{U}).\end{aligned} \tag{5.52}$$

where

$$\mathbf{k}\times\mathbf{U} \equiv (-v, u), \tag{5.53}$$

$$\mathbf{k}\cdot(\mathbf{x}\times\mathbf{U}) \equiv (xv - yu). \tag{5.54}$$

If, for $\mathbf{x}$, we substitute the quantity, $\nabla(\log m)$, in (5.52), we obtain the corresponding expressions for the fictitious force valid in any conformal mapping framework.

The fictitious force increases linearly with displacement from the stereographic map origin and is quadratic in the true velocity components. When the displacement A_B is just the tiny correction to the location of the trajectory brought about by the numerical time integration scheme's semi-implicit adjustment, the fictitious component to the effective specific force can often be regarded as negligible. However, for other semi-implicit time integration methods it is better to account accurately for the exact force components along any trajectory which does not exactly intersect the LS coordinate origin and the formulae we have derived then come into play.

6. DISCUSSION

We have discussed a variety of numerical approaches, all designed to enable a high degree of formal accuracy in a global semi-Lagrangian model to be achieved reasonably efficiently. For finite differencing in space, we recommend the use of "compact" schemes whenever feasible. For the grid to grid interpolations of a semi-Lagrangian model, we showed in section 3 how the three dimensions of this problem can be separated out; the individual one dimensional steps become relatively straightforward in such a "cascade" interpolation scheme. Also in section 3 we showed how some of the compact schemes discussed in section 2 can play an essential part in achieving formally conserving cascade interpolation schemes without sacrificing the local accuracy which the cascade interpolation method facilitates. The cascade method is practically as easy to use with forward trajectories as it is with backward trajectories, but forward trajectories simplify the application of more accurate time integration schemes. One possibility that looks increasingly attractive for numerical weather prediction is to adopt the "second-derivative" form of the combined momentum and kinematics equations for the trajectories. In section 4 we described some of the options available for treating the trajectories in this way with time integration schemes that allow a semi-implicit treatment of the gravity waves. Trajectory calculations, especially near the poles of a global model, can be complicated by the presence of strong coordinate curvature, and, if done without careful thought, may require the costly evaluations of trigonometrical functions. There are various ways to avoid the trigonometrical penalties; one way we presented in outline is to employ local stereographic frames for the descriptions of all trajectory displacements.

An indirect way of enhancing the efficiency of any global grid point model, given that so much of the computational labour is spent performing "physics" calculations, is to find a grid with a more equitable resolution from region to region. In a framework based on the the latitude–longitude coordinates, the generally recommended procedure is to reduce or skip longitudinal grid points progressively towards the two poles. Like the polar singularities themselves, this introduces some difficulties for a semi-Lagrangian model employing the cascade interpolation method, where it is desirable to have a grid structured as the Cartesian product of simple one-dimensional grids. One solution, which we have alluded to in the context of the polar problem and MPP implementation, is to fragment the global domain into logically rectangular and overlapping grid patches. At the expense of carrying out computations for some overlap regions more than once (on different patches) and blending these solutions smoothly together (with additional interpolations where the respective grids do not exactly

conform), it remains possible to reap the main benefits that the cascade approach offers. It is essentially this approach that is presently being pursued at NCEP in the development of a proposed new global forecasting model. But it may well emerge in the long term, provided that the problems with grid singularies can be satisfactorily dealt with, that one of the alternative global grids, such as those we discuss in another seminar in this volume, are more tidily adapted to high resolution and high-order accurate semi-Lagrangian modeling on the sphere.

Acknowledgements

The author is grateful for the helpful comments of Drs. Zavisa I. Janjić and Henry Juang during the preparation of this manuscript.

References

Bartello, P., and S. J. Thomas, 1996: The cost-effectiveness of semi-Lagrangian advection. *Mon. Wea. Rev.*, **124,** 2883–2897.

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

Carpenter, R. L.,Jr., K. K. Droegemeier, P. R. Woodward, and C. E. Hane, 1990: Applications of the piecewise parabolic method (PPM) to meteorological modeling. *Mon. Wea. Rev.*, **118,** 586–612.

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

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

Emde, K. v. d., 1992: Solving conservation laws with parabolic and cubic splines. *Mon. Wea. Rev.*, **120,** 482–492.

Gravel, S., and A. Staniforth, 1994: A mass-conserving semi-Lagrangian scheme for the shallow-water equations. *Mon. Wea. Rev.*, **122,** 243–248.

Harten, A., B. Engquist, S. Osher, and S. R. Chakravarthy, 1987: Uniformly high order accurate essentially non-oscillatory schemes III. *J. Comput. Phys.*, **71,** 231–303.

Janjić, Z. I., 1995: A note on the performance of the multiply-upstream semi-Lagrnagian advection schemes for one-dimensional nonlinear momentum conservation equation. *Meteor. Atmos. Phys.*, **55,** 1–16.

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

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 step-wise numerical integration. *Mon. Wea. Rev.*, **99**, 644–648.

Nair, R., J. Côté, and A. Staniforth, 1999: Monotonic cascade interpolation for semi-Lagrangian advection. *Quart. J. Roy. Meteor. Soc*, **125**, 197–212.

Navon, I. M., and R. de Villiers, 1987: The application of the Turkel-Zwas explicit large time-step scheme to a hemispheric barotropic model with constraint restoration. **115**, 1036–1051.

Priestley, A., 1993: A quasi-conservative version of the semi-Lagrangian advection scheme. *Mon. Wea. Rev.*, **121**, 621–629.

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 semi-implicit 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.*, **122**, 737–763.

Purser, R. J., and L. M. Leslie, 1997: High-order generalized Lorenz N-cycle schemes for semi-Lagrangian models employing second derivatives in time. *Mon. Wea. Rev.*, **125**, 1261–1276.

Rančić, M., 1995: An efficient conservative, monotonic remapping as a semi-Lagrangian transport algorithm. *Mon. Wea. Rev.*, **123**, 1213–1217.

Ritchie, H., C. Temperton, A. Simmons, M. Hortal, T. Davies, D. Dent, and M. Hamrud, 1995: Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF forecast model. *Mon. Wea. Rev.*, **123**, 1212–1217.

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

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

Störmer, C., 1907: Sur les trajectoires des corpuscules électrisés. *Arch. Sci. Phys. Nat. Genève*, **24**, 5–18, 113–158, 221–247.

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