# Finite-Volume Methods in Meteorology

Bennert Machenhauer

*Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen, DENMARK*

Eigil Kaas

*University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, DENMARK*

Peter Hjort Lauritzen

*National Center for Atmospheric Research, Boulder, Colorado, P.O. Box 3000, Boulder, CO 80307-3000, USA*

**Abstract**

Recent developments in finite-volume methods provide the basis for new dynamical cores that conserve exactly integral invariants, globally as well as locally, and, especially, for the design of exact mass conserving tracer transport models. The new technologies are reviewed and the perspectives for the future are discussed.

## 1. Introduction

Finite-volume (FV) methods are numerical methods where the fundamental prognostic variable considered is an integrated quantity over a certain finite-control volume. Thus, instead of grid-point values, finite elements or spectral components, *cell-integrated mean values* are considered. In meteorology, *FV methods* are, therefore, frequently referred to as *cell-integrated methods*. Some FV methods include additional prognostic variables to enhance the numerical accuracy. These variables can be higher order moments or point/face values between the control volumes.

In meteorological applications, so far, the control volumes adopted have generally been the conventional grid cells used in most operational prediction models: i.e., quasi-horizontal regular grid cells in cartesian coordinates on map projections of the sphere or regular grid cells in spherical latitude-longitude coordinates. These grid cells are referred

Computational Methods for the Atmosphere and the Oceans
Special Volume (Roger M. Temam and Joseph J. Tribbia, Guest Editors) of
HANDBOOK OF NUMERICAL ANALYSIS, VOL. XIV
P.G. Ciarlet (Editor)

ISSN 1570-8659
DOI 10.1016/S1570-8659(08)00201-9

to as *the Eulerian grid cells*. In the cell-integrated methods, these are complemented by Lagrangian control volumes, which move with the air flow, usually in a quasi-Lagrangian sense, i.e., departing from or arriving at Eulerian grid cells.

Exceptions to the basis of conventional Eulerian grid cells are new operational models based on grids, which are almost uniform on the sphere. Examples are the Massachusetts Institute of Technology general circulation model (ADCROFT, CAMPIN, HILL and MARSHALL [2004]) which is based on the conformal expanded spherical cube, but still has orthogonal coordinates and quadri laterally shaped grid cells, and the German NWP model (MAJEWSKI, LIERMANN, PROHL, RITTER, BUCHHOLD, HANISCH, PAUL, WERGEN and BAUMGARDNER [2002]) that is based on a non-orthogonal icosahedral-hexagonal grid on the sphere. For the sake of simplicity, we shall not go into details with these new grids, which currently is a very active research topic. The same limitation applies to nonuniform grids, such as the one introduced by LI and CHANG (1996). Thus we shall consider only FV methods in conventional grids.

The FV or cell-integrated methods are well suited for the numerical simulation of conservation laws. Before the implementation of FV methods in meteorological modeling, only conservative spatial discretization schemes were developed and used (e.g., ARAKAWA [2000], ARAKAWA and LAMB [1981], BURRIDGE and HASLER [1977], MACHENHAUER [1979], SIMMONS and BURRIDGE [1981]). With these schemes, just the globally integrated discretized time derivative of the invariant quantity in question was zero. Time truncation errors could still cause nonconservation globally. With the introduction of the FV method, the possibility of a conservative full space-time discretization became possible (e.g., MACHENHAUER [1994]). Previously, just global conservation was considered of importance, whereas with the FV methods, local conservation is considered even more important (e.g., MACHENHAUER and OLK [1997]). Conservation laws for mass, total energy, angular momentum, and entropy constitute the fundamental laws for the dynamics and thermodynamics of the atmosphere. Also, potential vorticity is considered a fundamental invariant which should be conserved in an adiabatic friction-free flow. In general, a discretized cell-integrated prognostic equation for a conservative quantity is obtained by integrating the differential flux form of the conservation law in question in space over an Eulerian grid cell and in time over the time-step $\Delta t$. The space integration results in an equation stating that the time rate of change of the total quantity in the grid cell is equal to the sum of fluxes through the cell boundaries. The time integration determines the fluxes through the cell boundaries during the time-step. These fluxes are exact if the integration is performed along exact trajectories ending at the boundaries of the regular Eulerian grid cell (also called *the arrival cell*) at time $t + \Delta t$ and originating from the boundaries of an irregular so-called Lagrangian cell (also called *the departure cell*) at time $t$. With such an exact integration, the integral of the conservative quantity over the arrival cell at time $t + \Delta t$ is equal to the integral over the departure cell at time $t$, plus changes due to sources and sinks, if any. We shall mainly concentrate on conservation of mass, which is the simplest conservation law, as it has no sources or sinks if precipitation and diffusion of mass is neglected. For this conservation law, called the continuity equation, we shall derive the exact prognostic equation (Eq. (1.8) in Section 1.1). Since exact integrations along exact trajectories will be assumed in the derivation, and since no further approximations are being made, this equation is

referred to as *the exact discretized cell-integrated continuity equation*. It implies exact conservation of mass during a time step, both global conservation, i.e., conservation of the total mass in the entire integration area and local conservation, i.e., conservation of the mass in each individual departure cell. During the derivation of the exact discretized cell-integrated continuity equation, it will be demonstrated that there is equivalence between traditional flux-form FV approaches and newer semi-Lagrangian FV methods. In both formulations, one attempts to approximate the same equation.

The general exact discretized cell-integrated continuity equation describes conservation of mass of "moist air," which is the atmospheric air including all its constituents. Corresponding exact continuity equations for the different constituents in the moist air, for example, water vapor or any chemical constituent, are obtained by simply replacing the density of moist air $\rho$ with the density $\rho_q = q\rho$ of the constituent in question, where $q$ is its specific concentration[1]. In meteorological models, the solution to the continuity equation for moist air is of special importance. The solution determines the flow of air mass, which determines the pressure distribution and thus the dynamics of weather systems, especially the development and decay of weather systems. Spurious mass sources due to local nonconservation of mass might thus influence the simulation of weather systems (Machenhauer and Olk [1997]). The solution determines the flow of all constituents in the moist air since they are transported with the air and thus share trajectories with the air. This is important especially in chemical models as spurious changes in the ratios between linearly correlated (in space) concentrations of reacting chemical constituents are avoided (Lin and Rood [1996]). Thus, in meteorological models, a "correct" simulation of the atmospheric dynamics and all kinds of interactions among constituents depends heavily on the accuracy of the numerical solutions to the continuity equations.

In Section 2, the different mass conserving schemes that have been developed for meteorological applications in two dimensions (2Ds) are described in detail. In the different schemes, different approximations are made in the determination of the trajectories and in the integration along the trajectories over the time-step or in the integration over the departure cell. The approximate schemes presented in Section 2 will be compared with the exact solution. It will be shown that all the different schemes conserve mass globally, simply because they are all constructed so that the mass that leaves a certain face of an Eulerian arrival cell during a time-step is exactly gained in the neighboring cell with which the cell face is shared. This, of course, does not guarantee a high level of accuracy as the global conservation may be obtained even with rather inaccurate local fluxes. However, the accuracy with which the local mass conservation is approximated is a real measure of the accuracy of the local transports of the moist air and its constituents. Section 2 will mainly focus on relatively new schemes, most of which are based on (semi-) Lagrangian approaches. For completeness, a short introduction to the more traditional flux-form schemes is presented as well.

Section 3 provides an overview over the general applicability of FV techniques in meteorology. This section is initiated with an example of a complete set of FV

[1] *The specific concentration of a constituent* is the ratio between the mass of the constituent and the mass of the moist air it is mixed into.

prognostic equations that conserve mass, entropy, total energy, and angular momentum in an adiabatic and friction-free atmosphere. Furthermore, Section 3 provides two examples of pioneering mass conserving hydrostatic dynamical cores in spherical geometry, which are based on FV techniques. By a *dynamical core*, we mean a computer code for the numerical integration of the system of meteorological equations governing the dynamics of the atmosphere. Roughly speaking, the dynamical core approximates the solution to the meteorological equations on resolved scales, while parameterizations represent subgrid-scale processes and other processes not included in the dynamical core (THUBURN [2006]). However, in tests of dynamical cores, one includes those dissipation terms, which are needed for smooth and stable integrations. Furthermore, Section 3 includes a discussion of a few remaining issues, such as the so-called mass-wind inconsistency in in-line and off-line FV tracer transport applications, and possibilities of extensions to non-hydrostatic models are briefly discussed. Finally, Section 4 includes a brief summary of the main issues presented in this chapter.

### *1.1. The exact cell-integrated continuity equation*

In this section, an "exact" discretized cell-integrated continuity equation is derived. This is introduced as a pre-requisite and reference for the approximate 2D and three-dimensional (3D) FV schemes to be presented in Sections 2 and 3, respectively. It is exact in the sense explained above. It is derived from assumed exact integrals along assumed exact trajectories, which are determined from given exact 3D fields of density and velocity during a time interval $\Delta t$ from $t$ to $t + \Delta t$. No further assumptions are made, apart from a simplifying one of no vertical shear of the horizontal velocity in each discrete model layer.

Define Eulerian grid cells as the arrival cell indicated to the right in Fig. 1.1 in a cartesian coordinate system $(x, y, h)$ so that the grid length along the $x$-axis is $\Delta x$, the grid spacing along the $y$-axis is $\Delta y$, and $h$ is a terrain following height-based vertical coordinate defined as $h = z - z_s$, where $z$ is the height above mean sea level and $z_s$ is the height of the surface of the Earth. Surfaces with $h$ equal to a constant $h_{k+1/2}$ separate the grid-cell layers in the vertical. The "½" in the index refers to the Lorenz vertical staggering of the variables (LORENZ [1960]). The "half-levels" are located in between "full-levels" $h_k = 1/2\,(h_{k+½} + h_{k-½})$ with integer index $k$, where point values of mass and velocity variables traditionally have been located. Thus, the height difference between the bottom and the top of the Eulerian grid cell centered at level $k$, which is considered in Fig. 1.1, is $\Delta_k h = h_{k+½} - h_{k-½}$.

To derive the FV version of the continuity equation, we need to integrate along exact trajectories ending at the boundaries of the arrival cell at time $t + \Delta t$ and originating from the boundaries of the corresponding departure cell at time $t$. In Fig. 1.1, the departure cell is shown as the irregular cell to the left. Only four of the trajectories are shown in the figure. The exact velocity fields, supposed to be given during the whole time interval $\Delta t$ from $t$ to $t + \Delta t$, determine a trajectory ending at any of the points inside or at the boundaries of the arrival cell. We now define an additional auxiliary *vertical coordinate* $\xi$ for a particle: a Lagrangian vertical coordinate (STARR [1945]), which per definition is constant along its 3D trajectory. We choose the Lagrangian coordinate $\xi$ of a particle,

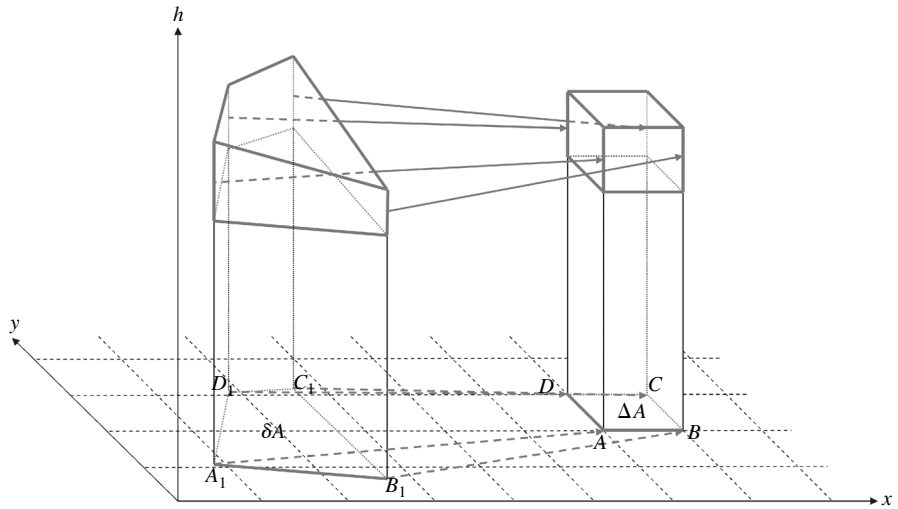


FIG. 1.1 Conceptual sketch showing a cell that is moving with the flow in a Lagrangian model layer during a time-step $\Delta t$. To the left is shown the cell at time $t$ (the so-called departure cell). The horizontal velocity $\vec{V}$ within the model layer is assumed independent of height so that the cell walls, which initially at time $t$ are vertical, remain vertical. The cell ends up at time $t + \Delta t$ as the horizontally regular Eulerian grid cell (the so-called arrival cell) shown in the vertical column to the right. Just four trajectories are shown. The projections on a horizontal plane are shown in more detail in Fig. 1.2. (See also color insert).

that is moving with the 3D flow during the time-step, to be equal to its $h$ value in or at the boundary of the arrival cell. Thus, the trajectories constitute a vertical coordinate system, which is defined only in the time interval from $t$ to $t + \Delta t$. Obviously, in this coordinate system, the vertical velocity of a particle is zero:

$$\dot{\xi} = \frac{\mathrm{d}\xi}{\mathrm{d}t} = 0. \tag{1.1}$$

Here, a simplifying assumption is made, namely that the horizontal wind $\vec{V}$ is independent of height within *the Lagrangian model layer*, i.e., the layer enclosing all the trajectories which are ending inside or at the boundary of the arrival cell. Thus, as indicated in Fig. 1.1, vertical columns that move with the horizontal wind in the layer will remain vertical. Mathematically, it implies a simplifying separation of the vertical and horizontal integrations to be performed in the layer. A column may, of course, still change its thickness $\delta_k h$ due to horizontal convergence or divergence. The trajectories in Fig. 1.2, which are ending at the corners of the arrival cell, originate from the corner of the departure cell. For simplicity of the sketch, it is assumed that the horizontal velocity field is such that the trajectories and lines between neighboring corners in the departure cell are straight, i.e., the vertical faces of the departure cell in Fig. 1.1 are plane. Note that since trajectories ending at the boundaries of the arrival cells are shared by neighboring cells, it follows that the departure cells, as does the arrival cells, fill out the entire integration domain without any cracks in between.

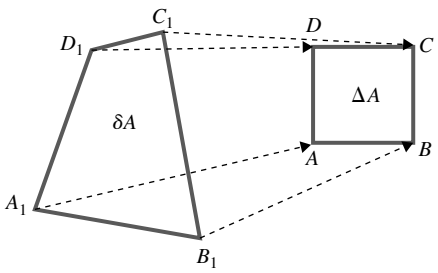


FIG. 1.2 Horizontal projections of the arrival cell ($A$, $B$, $C$, $D$) at time $t + \Delta t$ with area $\Delta A$ and the corresponding upstream departure cell ($A_1$, $B_1$, $C_1$, $D_1$) at time $t$ with area $\delta A$. This figure corresponds to a view from above at the departure and arrival cells in Fig. 1.1.

The differential flux form of the continuity equation in the $\xi$-coordinate system becomes

$$\frac{\partial \rho}{\partial t} = -\nabla_\xi \cdot \rho \vec{V} - \frac{\partial \rho \dot{\xi}}{\partial \xi}, \tag{1.2}$$

where $\rho$ is the density of moist air and $\vec{V}$ is the horizontal velocity. To obtain the continuity equation for a regular vertical column, integrate Eq. (1.2) vertically over the Lagrangian model layer. The result is

$$\frac{\partial \tilde{\rho}_k \, \delta_k h}{\partial t} = -\nabla_\xi \cdot \tilde{\rho}_k \, \delta_k h \vec{V}_k, \tag{1.3}$$

where Eq. (1.1) has been used and $\tilde{\rho}$ is the vertical mean density:

$$\tilde{\rho}_k = \frac{1}{\delta_k h} \int_{\delta_k h} \rho \, \mathrm{d}z.$$

To obtain the cell-integrated continuity equation, integrate Eq. (1.3) horizontally over the area of the arrival grid cell. After application of the Gauss's divergence theorem, we get

$$\Delta A \frac{\partial \overline{\left(\tilde{\rho}_k \, \delta_k h\right)}}{\partial t} = -\sum_{i=1}^{4} \left( < \left(\tilde{\rho}_k \, \delta_k h\right) \vec{V}_k > \cdot \vec{n} \; \Delta l \right)_i, \tag{1.4}$$

where $\Delta A = \Delta x \Delta y$ is the horizontal area of the grid cell and

$$\overline{\left(\tilde{\rho}_k \, \delta_k h\right)} = \frac{1}{\Delta A} \iint_{\Delta x \Delta y} \left(\tilde{\rho}_k \, \delta_k h\right) \, \mathrm{d}x \mathrm{d}y \tag{1.5}$$

is the horizontal mean value of $\tilde{\rho}_k \, \delta_k h$ in the Eulerian grid cell. In Eq. (1.4), $\vec{n}_i$ is a unit vector normal to the $i$th face of the cell pointing outward, and $\Delta l_i$ is the length of the

face equal to either $\Delta x$ or $\Delta y$. $\left(\tilde{\rho}_k\right)_i$, $(\delta_k h)_i$, and $(\vec{V}_k)_i$ are instantaneous values at the cell face $i$, and the angle brackets represent averages in the $x$- or $y$-direction over the cell faces. The next step is to integrate over the time-step $\Delta t$, between $t$ and $t + \Delta t$, which results in

$$\Delta A\left(\overline{\left(\tilde{\rho}_k\ \delta_k h\right)}^{+} - \overline{\left(\tilde{\rho}_k\ \delta_k h\right)}\right) = -\Delta t \sum_{i=1}^{4} \overline{\overline{\left(< \left(\tilde{\rho}_k\ \delta_k h\right)\vec{V}_k >\cdot\ \vec{n}\ \Delta l\right)}}_i \tag{1.6}$$

Here, the plus-sign in superscript indicates the updated value and the double bar refers to the time average over $\Delta t$. Each term on the right-hand side of Eq. (1.6) represents the mass transported through one of the four Eulerian cell faces into the cell during the time-step. Each term involves integrals over the cell face in question and over the time-step. The integral in time over the time-step may be performed in space along the trajectories terminating on the Eulerian cell face in question, cell face AB for instance (see Fig. 1.2). Thus, this term in Eq. (1.6) is computed as a surface integral of $\left(\tilde{\rho}_k\ \delta_k h\right)$ over the area between the Eulerian cell face $AB$, the two backward trajectories, $AA_1$ and $BB_1$ originating from the two end points of the Eulerian cell face and the respective face of the departure cell $A_1B_1$. That is, the mass inflow through the southern (or lower) face in Fig. 1.2 is equal to the integral of $\left(\tilde{\rho}\,\delta_k h\right)$ over the area marked $A_1ABB_1$ in the figure. Writing this integral as $\iint\limits_{A_1B_1BA} \tilde{\rho}_k\,\delta_k h\ \mathrm{d}x\,\mathrm{d}y$, Eq. (1.6) may be rewritten as

$$\begin{aligned}
\overline{\left(\tilde{\rho}_k\ \delta_k h\right)}^{+}\Delta A &= \iint\limits_{ABCD}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y + \iint\limits_{A_1B_1BA}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y \\
&\quad + \iint\limits_{A_1ADD_1}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y + \iint\limits_{D_1DCC_1}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y - \iint\limits_{B_1BCC_1}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y \\
&= \iint\limits_{A_1B_1C_1D_1}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y.
\end{aligned} \tag{1.7}$$

Here, the mass inflows through the remaining three cell faces are included in the second line by similar integrals. The first term on the right-hand side is

$$\Delta A\overline{\left(\tilde{\rho}_k\ \delta_k h\right)} = \iint\limits_{ABCD}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y,$$

i.e., the original mass in the Eulerian grid cell at time $t$. Thus, as illustrated in Fig. 1.2, the sum of the first four terms on the right-hand side of Eq. (1.7), representing the original mass in the Eulerian grid cell, the inflow through the southern, the western, and the northern cell face is compensated partly by the outflow through the eastern cell face, represented by the fifth negative term in Eq. (1.7). The result is the integral on the second right-hand side of Eq. (1.7) that represents the mass in the Lagrangian departure

cell $A_1B_1C_1D_1$. Denoting the departure cell area as $\delta A$ (Fig. 1.2), the result may be written as

$$\iint\limits_{A_1B_1C_1D_1} \left(\tilde{\rho}_k\ \delta_k h\right)\ \mathrm{d}x\,\mathrm{d}y = \overline{\tilde{\rho}_k\ \delta_k h}\,\delta A,$$

and we obtain finally

$$\overline{\left(\tilde{\rho}_k\ \Delta_k h\right)}^{+}\Delta A = \overline{\left(\tilde{\rho}_k\ \delta_k h\right)}\,\delta A. \tag{1.8}$$

This is a prognostic equation predicting the mass in the arrival area at $t + \Delta t$, $\overline{\left(\tilde{\rho}_k\ \delta_k h\right)}^{+}\Delta A$, from the mass in the departure area at time $t$, $\overline{\left(\tilde{\rho}_k\ \delta_k h\right)}\,\delta A$. Note that no information is needed between $t$ and $t + \Delta t$ and recall that in the arrival area (exactly in the center of the area) the Lagrangian model layer coincide with the Eulerian cell so that $\delta_k h = \Delta_k h$. Thus, the right-hand side of Eq. (1.8) can be determined by an integration of $\left(\tilde{\rho}_k\ \delta_k h\right)$ over the departure area and Eq. (1.8) becomes

$$\overline{\left(\tilde{\rho}_k\right)}^{+} = \frac{1}{\Delta_k h\ \Delta A}\iint\limits_{A_1B_1C_1D_1}\left(\tilde{\rho}_k\ \delta_k h\right)\mathrm{d}x\,\mathrm{d}y = \frac{1}{\Delta_k V}\iiint\limits_{\delta_k V}\rho_k\ \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{1.9}$$

or

$$\overline{\left(\tilde{\rho}_k\right)}^{+}\Delta_k V = \iiint\limits_{\delta_k V}\rho_k\ \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{1.10}$$

$\overline{\left(\tilde{\rho}_k\right)}^{+}\Delta_k V$ is the updated mass in the Eulerian arrival grid cell at time $t + \Delta t$. According to Eq. (1.10) it is equal to the mass in the upstream departure cell at time $t$. Thus, the exact discrete cell-integrated continuity equation (Eq. (1.8)) is simply a cell-integrated analog to the well-known grid-point semi-Lagrangian continuity equation (Robert [1969, 1981, 1982]) that presently is used in most operational meteorological models. Contrary to the grid-point version, the cell-integrated equation is inherently mass conservative. It fulfills exactly our definition of a locally mass conserving scheme as the updated mass in an Eulerian arrival grid cell is exactly the mass in the upstream departure cell. It is easily shown by a summation of Eq. (1.8) over the entire integration domain, with assumed periodic lateral boundary conditions, that it also implies global mass conservation. The analogy to the grid-point semi-Lagrangian continuity equation shows that an alternative way to derive Eq. (1.8) would be to set up the mass conservation law directly for FV on a Lagrangian form and then integrate that form over $\Delta t$. The mass in a FV $\delta_k V$ considered at time $t$ is

$$M_{\delta_k V} = \iiint\limits_{\delta_k V}\rho_k\ \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{1.11}$$

The mass conservation law for this FV, which is supposed to move with the flow without any mass flux through its boundaries, is

$$\frac{\mathrm{d}\, M_{\delta_k V}}{\mathrm{d}t} = 0. \tag{1.12}$$

When integrated in time from $t$ to $t + \Delta t$, Eq. (1.12) gives Eq. (1.8), which as shown above leads to Eq. (1.10). The reason for presenting the more complicated derivation starting from the Eulerian flux form of the mass conservation law (1.2) is that some numerical FV schemes, the so-called *flux-form schemes*, are based on the flux form (1.2), whereas others, the so-called *Lagrangian schemes*, are based on the Lagrangian form (1.12). The purpose of the present derivation was to show that in the case of exact trajectories and exact mass integrals over the relevant volumes, the flux-form Eq. (1.2) is equivalent to the Lagrangian form (1.12). When, as it is usually the case, a flux-form scheme becomes different from a Lagrangian scheme, it is due to different approximations to the trajectories defining the departure volume and different approximations to the upstream mass integrals. A measure of accuracy for both types of schemes should therefore be how close they are to the ideal "exact" scheme. That is, how close the approximate departure volume is to the real, exact one and how close the exact mass integral over the exact departure volume is to the approximate mass integral over the approximate departure volume. In other words, how accurate the local mass conservation is.

### *1.2. Longtime step schemes and combinations with semi-implicit time-stepping*

The reason for the recent renewed interest in FV methods in meteorological modeling was the observation of a significant lack of global mass conservation in numerical models using the grid-point version of the semi-Lagrangian scheme unless an unphysical so-called mass-fixer, which restores the total mass globally after each time step, is used. There is an arbitrariness in the way these mass-fixing algorithms repeatedly restore global mass conservation without ensuring any local mass conservation, i.e., without fulfilling a continuity equation for the mass that is transported locally between the Eulerian grid cells of the model each time-step (Machenhauer and Olk [1997]). Without such a mass-fix, a significant drift in the global mass was observed (Bates, Higgins and Moorthi [1995]), and even with a mass-fixer, it seems likely that significant local errors are developed (Machenhauer and Olk [1997]). Nevertheless, the reason for the popularity of the grid-point semi-Lagrangian schemes has been its almost unconditional absolute stability, which in practice eliminates the advective Courant-Fredrichs-Levy (CFL) time-step restriction. This property is utilized in most operational meteorological models in combination with a semi-implicit treatment of the gravity wave terms in the primitive equations, which eliminates the fast wave CFL time-step restriction. Then, in principle, the length of the time-steps in a combined semi-implicit semi-Lagrangian model can be chosen solely based on accuracy considerations. This is extremely important in meteorological models where any gain by an increased time-step can be utilized to increase the realism of parameterized physical processes and/or the spatial resolution of the model

grid. According to general operational experience, such improvements have practically always led to an increase in accuracy. As should be expected from the experience with the grid-point semi-Lagrangian schemes, the recently developed cell-integrated semi-Lagrangian schemes are also (almost) unconditionally stable (Lauritzen [2007]), eliminating in practice the advective CFL time-step restriction. It has furthermore recently been shown that fast waves in cell-integrated semi-Lagrangian models can be stabilized by a combination with a semi-implicit time extrapolation scheme. This has been demonstrated by Machenhauer and Olk [1997] for a simple one-dimensional (1D) mass and momentum or mass and total energy conserving model and Lauritzen, Kaas and Machenhauer [2006] for shallow water models and by Lauritzen, Kaas, Machenhauer and Lindberg [2008] for a complete 3D mass conserving model. An alternative method, which has been used in finite difference grid-point models to stabilize the fast waves, is the so-called split-explicit time-stepping. However, this possibility was abandoned by Machenhauer and Olk [1997] for FV models because when splitting the system of continuous equations into an advective part, which should use large time-steps, and an adjustment gravity wave part, which should use short time-steps, it was found that neither of the sub systems were conserving momentum or total energy. Consequently, these invariants for the full system could not be conserved exactly in any FV version.

As mentioned above, Section 3 describes two mass conserving quasi-hydrostatic dynamical cores, both combined with comprehensive physical parameterization packages. One of these dynamical cores described in Lauritzen, Kaas, Machenhauer and Lindberg [2008] is a semi-implicit version using large time steps for all variables, while the other one described in Lin [2004] and Collins, Rasch, Boville, Hack, Mccaa, Williamson, Kiehl, Briegleb, Bitz, Lin, Zhang and Dai [2004] uses an explicit time-stepping scheme. The latter model uses explicit, relatively small time-steps for the dynamical core but large time-steps for the transport of all tracer species (including water vapor) and for physical parameterizations.

## 2. Transport schemes in one and two dimensions

In meteorological models, a FV method for the continuity is based on the exact cell-integrated continuity equation and obviously it should be approximated as accurately as possible. As discussed in Section 1, the vertical and horizontal problems can be separated in a consistent way considering Lagrangian cells moving with vertical walls along three dimensional trajectories. Consequently, only horizontal integrals of vertically integrated mass distributions are needed in the solution of the continuity equation. So in case of a flux-form Eulerian scheme, the fluxes through the four cell faces can be determined by horizontal integrals (as described in connection with Eq. (1.6)), and for the departure cell-integrated semi-lagrangian (DCISL) scheme, direct integrations over the horizontal departure area approximating the true departure area can be performed (as indicated in Eq. (1.8)). Hence, by using this approach, one can directly apply 2D FV schemes for the 3D problem. Alternatively, flux-form schemes may be extended to 3Ds by including vertical advection through the top and bottom surfaces of Eulerian grid cells. Similarly,

the 3D DCISL scheme would perform a 3D integral over the Lagrangian departure cell. However, following these fully 3D approaches would become very complicated if one aims at a numerically efficient and mass conserving integration.

Because of the general applicability of 2D solutions to the continuity equation, this section, beside basic 1D formulations, is devoted to the fully 2D schemes. Compared with the large number of mass conserving transport schemes published in the general fluid dynamical literature, there are many fewer schemes that have been used or are applicable in real meteorological applications on the sphere. Here, we mostly concentrate on the subset that is potentially applicable in a wide range of atmospheric models. Therefore, descriptions of the vast majority of the hundreds of transport schemes developed in computational fluid dynamics in general are excluded. For a more general review of FV methods, see e.g., LEVEQUE [2002] and EUMARD, GALLOUËT and HERBIN [2000].

Before discussing the different FV schemes used in the atmospheric sciences, it is important to realize which properties a transport scheme ideally should possess. An overview of these properties is provided in Section 2.1. The FV schemes presented in this overview use sub grid representations at time $t$ in order to make the forecast at time $t + \Delta t$. The most frequently used sub grid representations and associated filters ensuring some of the properties listed in Section 2.1 are introduced in Section 2.2. This is followed, in Section 2.3, by an overview of the different types of FV methods applied in 2D problems. Section 2.4 briefly describes some – mostly recent – local mass conservation fixers for semi-Lagrangian models which can be considered closely related to FV semi-Lagrangian schemes. Aiming at enhanced accuracy, Section 2.5 discusses the possibilities to include extra prognostic variables in addition to the cell-mean values. The so-called flux-limiter methods have been popular approaches to maintain attractive shape-preserving properties. A brief introduction to these methodologies, which are complementary to the filtering methods mentioned in Section 2.2, is given in Section 2.6. Finally, Section 2.7 provides some concluding remarks on the basic FV transport schemes in 1 and 2Ds.

### *2.1. Desirable properties*

The equation subject to the toughest requirements is probably the continuity equation for tracers such as moisture, the spatial distribution of which includes sharp gradients. RASCH and WILLIAMSON [1990] have defined seven desirable properties for transport schemes: accuracy, stability, computational efficiency, transportivity, locality, conservation, and shape-preservation. In addition to the seven desirable properties defined by RASCH and WILLIAMSON [1990], even more desirable properties have emerged in the literature, e.g., consistency, compatibility, and preservation of constancy. The perfect scheme would have all the desirable properties listed above under all conditions but, in practice, no single method is advantageous under all conditions.

#### *2.1.1. Accuracy*

The high-accuracy property is, of course, the primary aim for any numerical method, and all the desirable properties listed above, apart from the efficiency requirement, are part of the overall accuracy. Note that for a flow with shocks or sharp gradients, the formal order

of accuracy in terms of Taylor series expansions does not necessarily guarantee a high level of accuracy. Part of the accuracy is also the rate of convergence of the numerical algorithm.

Widely used measures of accuracy in the meteorological community for idealized test cases are the standard error measures $l_1$, $l_2$, and $l_\infty$ (e.g., WILLIAMSON, DRAKE, HACK, JAKOB, SWARZTRAUBER, [1992]):

$$l_1 = \frac{I\left(|\psi - \psi_E|\right)}{I\left(|\psi_E|\right)}, \tag{2.1}$$

$$l_2 = \frac{\left\{I\left[(\psi - \psi_E)^2\right]\right\}^{1/2}}{\left\{I\left[(\psi_E)^2\right]\right\}^{1/2}}, \text{ and} \tag{2.2}$$

$$l_\infty = \frac{\max\left[|\psi - \psi_E|\right]}{\max\left[|\psi_E|\right]}, \tag{2.3}$$

where $I(\cdot)$ denotes the integral over the entire domain, $\psi$ is the numerical solution, and $\psi_E$ is the exact solution if it exists. In case an exact solution does not exist, $\psi_E$ is a high-resolution reference solution. $l_1$ and $l_2$ are the measures for the global "distance" between $\psi$ and $\psi_E$, and $l_\infty$ is the normalized maximum deviation of $\psi$ from $\psi_E$ over the entire domain. In addition to these error measures, the normalized maximum and minimum values of $\psi$ are also used to indicate errors related to overshooting and undershooting.

To evaluate the accuracy of new schemes, several idealized advection test cases have been formulated. The interscheme comparison, however, is often made difficult by the fact that different authors use different test cases and/or different error measures. The test problems can be divided into two categories. Firstly, translational passive advection tests where distributions are transported by prescribed non-divergent winds that, ideally, translate the initial distribution without distorting it; these test cases involve the entire domain. Secondly, deformational test cases which focus on part of the domain such as an initial distribution being deformed by a vortex. Recently, NAIR and JABLONOWSKI [2007] combined these two types of test cases into one.

Probably, the most commonly used idealized test case in the meteorological literature is the solid body rotation of a cosine cone and/or a slotted cylinder. In cartesian geometry, the test case is described in, e.g., ZALESAK [1979] and BERMEJO and STANIFORTH [1992], and the spherical version is test case 1 of the suite of test cases by WILLIAMSON, DRAKE, HACK, JAKOB and SWARZTRAUBER [1992]. The analytic solution to this problem is simply the translation of the initial distribution along a circle in cartesian geometry and a great circle in the spherical case. It is an important part of accuracy that the advection schemes can transport distributions across the singularities of the numerical grids without distortion and imposing severe time-step limitations. DRAKE, HACK, JAKOB, SWARZTRAUBER and WILLIAMSON [1992] suggested that the cosine bell is transported along the equator and across the poles with a slight offset to avoid any symmetry. Note, however, that away from the poles, advection along these great circles is almost along coordinate axis for conventional latitude-longitude grids that, in general, favor the advection scheme.

Passive advection of scalars using the solid body rotation test case only addresses the ability of the scheme to translate a distribution without distorting it. Other commonly used test cases are based on a deformational flow, for example the swirling shear flow test in cartesian geometry considered by DURRAN [1999, Section 5.7.4], which is specified in terms of a periodically reversing time-dependent velocity field. Hence, after one period, the exact solution is the initial distribution. It could, however, be speculated that some errors introduced during the first half period are cancelled when the wind field reverses. Other deformational flow test cases, to which the exact solution is known throughout the time of integration, are defined in SMOLARKIEWICZ [1982] (analytical solution is given in CÔTÉ, STANIFORTH and PUDYKJEWICZ [1987]) and ARMENGAUD and HOURDIN [1999]. The idealized cyclogenesis problem described by DOSWELL [1984], to which the analytic solution is known, has been used for scalar-advection tests by several authors. For example, the non-smooth deformational flow vortex defined on a tangent plane (e.g., RANČIĆ [1992], HÓLM [1995], CÔTÉ, NAIR and STANIFORTH [1999a]). A version was formulated for the sphere by NAIR, SCROGGS and SEMAZZI [2002] and NAIR and MACHENHAUER [2002]. It is a smooth deformational flow test case that consists of two symmetric vortices, one over each pole. This test case has been combined with a translational wind field in NAIR and JABLONOWSKI [2007] to form a test case (where the analytical solution is known) that simultaneously challenges schemes with respect to deformation and translation.

### *2.1.2. Stability*

The stability property ensures that the solution does not "blow up" during the time of integration. Usually, the stability of Eulerian methods is governed by the CFL condition, which in 1D is given by

$$\max \left| \frac{u \Delta t}{\Delta x} \right| \leq 1, \tag{2.4}$$

where $u$ is the velocity and $\Delta x$ the grid interval. Hence, a fluid parcel may not travel more than one grid interval during one time-step. This overly restrictive time-step limitation is usually alleviated in semi-Lagrangian methods and can be replaced by the less severe Lipschitz convergence criterion

$$\left| \frac{\partial u}{\partial x} \right| \Delta t < 1, \tag{2.5}$$

(BENOIT, PUDYKIEWICZ and STANIFORTH [1985]; KUO and WILLIAMS [1990]), which guarantees that parcel trajectories do not cross during one time-step and ensures the convergence of the trajectory algorithm (a multi dimensional extension of Eq. (2.5) is given in BENOIT, PUDYKIEWICZ and STANIFORTH [1985]). Hence, in semi-Lagrangian models, the time-step can be chosen for accuracy and not for stability because of the lenient stability condition.

For global models based on a conventional latitude-longitude grid, the efficiency and stability of the advection schemes are often challenged by the convergence of the meridians near the poles, and special care must be taken in the vicinity of the poles. Alternatively,

the problem can be tackled by using other types of grids that do not have these singularities or at least reduce the effect of them, for example, the icosahedral-hexagonal grid used operationally by the German Weather Service (e.g., Arakawa, Mintz and Sadourny [1968], Williamson [1968], Thuburn [1997], Majewski, Liermann, Prohl, Ritter, Buchhold, Hanisch, Paul, Wergen and Baumgardner [2002]), and the cubed sphere approach originally introduced by Sadourny [1972] which, after having remained dormant for many years, has become a very active research topic (e.g., Iacono, Paolucci and Ronchi [1996], Mesinger and Rančić, Purser [1996], McGregor [1996], Iskandrani, Taylor and Tribbia [1997], Loft, Nair and Thomas [2005]). These grids are more isotropic than conventional latitude-longitude grids, i.e., all cells have nearly the same size, contrarily to latitude-longitude grids, where the areas decrease as aspect ratios increase toward the poles (this effect can, however, be alleviated by using a Gaussian-reduced grid in which the number of longitudes decrease toward the poles).

### *2.1.3. Computational efficiency*

Computing resources are limited and, given the complexity of geophysical fluid dynamics, the algorithms should be computationally efficient in order to allow for high-resolution runs and/or a large number of prognostic variables. Efficiency is, however, hard to measure objectively. One measure for the efficiency of an algorithm is the number of elementary mathematical operations or the total number of floating-point operations per second (FLOPS) used by the algorithm. The advantage of counting FLOPS is that it can be done without using a computer and is, therefore, a machine-independent measure. But the number of FLOPS only captures one of several dimensions of the efficiency issue. The actual program execution involves subscripting, memory traffic and countless other overheads. In addition, different computer architectures favor different kinds of algorithms and compilers optimize code differently. Measuring efficiency in terms of the execution time on a specific platform can be misleading for a user on another computer platform. Weather prediction and climate models are often executed on massively parallel distributed memory computers where the efficiency is partly determined by the amount of communication between the nodes. This becomes increasingly important if the resolution is held fixed while the number of distributed memory processors is increased. Hence, the parallel programmer is concerned about algorithms being local, thus minimizing the need for communication between the nodes. Nevertheless, a very important measure of efficiency is probably the level of simplicity of the algorithm.

Since models include an increasing number of tracers, an important aspect of the efficiency is how much of the transport algorithm can be reused for additional tracers. Obviously, if the entire transport algorithm must be repeated for each additional tracer, such an algorithm would not be attractive in modern transport models that include hundreds of tracers. In semi-Lagrangian models, for example, the computation of trajectories need only be computed once and can be reused for all tracers (e.g., Dukowicz and Baumgardner [2000]).

Thus, the computational cost of a given model depends not only on the number of FLOPS involved in the production of say one model day; it depends also to a high degree on the computer architecture on which the model is run. The optimization of a given model intended for operational application on a given platform is often an extensive

and complicated work for an experienced programmer, and the result will vary with the ingenuity of the programmer. The new algorithms presented here are often developed on an experimental basis by scientists who are not specialized programmers, and therefore, they are usually far from an optimized code suited for operational applications. It is, therefore, not fair to uncritically compare the computational cost of such new FV algorithms with traditional well-optimized algorithms. For this reason, information on the computational costs of the new algorithms are most often not available in the literature. When they are available, they should be considered with the reservations stated here and should only be considered as possible maximum computational costs, which most likely could be reduced for operational applications.

It should be noted that it could be misleading to compare computational efficiency and accuracy of two algorithms at the same spatial and temporal resolution. A scheme might be computationally inefficient at a given resolution compared to other schemes but have an accuracy that other schemes would need a much finer resolution in order to achieve (the opposite situation is, of course, also possible). In other words, ideally one should consider the ratio between computational cost and accuracy when comparing numerical schemes. That would enable one to select the scheme where one pays as little as possible computationally for the highest level of accuracy.

#### *2.1.4. Transportivity and locality*

The transportive and local property guarantee that information is transported with the characteristics and that only adjacent grid values affect the forecast at a given point. For FV schemes, one aspect of the local property is the degree of local mass-conservation that we define as follows. Since the mass enclosed in an area moving with the flow is conserved in the absence of sources and sinks, the degree to which the effective departure area of the numerical scheme coincides with the exact departure area is a measure for the local mass conservation of a given scheme. Another aspect of local mass conservation is the degree to which the reconstruction of the subgrid-scale distribution is local. For example, near sharp gradients, it is important that the gradient is not weakened during the process of reconstructing the subgrid-scale distribution, i.e., the reconstruction should be local.

#### *2.1.5. Shape-preservation*

The shape of a distribution undergoing pure advection should ideally be preserved in the numerical solution. For general velocity fields, the shape of the distribution may be altered in the form of new extrema. In such situations, the numerical scheme should reproduce only the physical extrema without creating spurious numerical extrema. These spurious numerical extrema especially cause problems in situations where the advection scheme produces negative mixing ratios (or concentrations) or when the values are above the maximum possible. Negative mixing ratios or mixing ratios above a physical threshold value are unphysical and would most likely cause a breakdown in physical parameterizations. If a numerical scheme inherently prevents negative undershoots in mixing ratios (or concentrations), it is termed positive-definite (or positivity preserving), if it preserves gradients, then the scheme is monotone, and if artificial oscillations are prevented, it is termed nonoscillatory. All these properties are, of course, interrelated and,

constitutes together the shape-preservation property. The very popular spectral methods are well known for producing "wiggles" (also known as Gibb's phenomena) near sharp gradients and are, therefore, a typical example of a monotonicity-violating and oscillatory numerical method.

*2.1.6. Conservation*

Ideally, all global integral invariants of the corresponding continuous problem should be conserved for any kind of flow. For long simulations, the conservation properties become increasingly important as numerical sources, and sinks can degrade the accuracy and alter global balance budgets significantly over time. Hence, for climate models, the FV methods are very attractive given their inherent conservation properties. However, a numerical model can only maintain a small number of analogous invariant properties constant, and some choice must be made as to which conserved quantities are to be conserved in the numerical model. For a comprehensive discussion of this issue, see THUBURN [2006]. Probably, the most important property to conserve for a continuity equation is the first moment, i.e., mass.

*2.1.7. Consistency*

The consistency property is less frequently discussed in the literature. Notable exceptions are JÖCKEL, VON KUHLMANN, LAWRENCE, STEIL, BRENNINKMEIJER, CRUTZEN, RASCH and EATON [2001] and BYUN [1999]. This property concerns the coupling between the continuity equation for air as a whole and for individual tracer constituents. In the continuous case, the flux-form continuity equation for a constituent with specific concentration $q$,

$$\frac{\partial}{\partial t}(q\rho) + \nabla \cdot \left(\vec{v} q \rho\right) = 0 \tag{2.6}$$

degenerates to

$$\frac{\partial}{\partial t}(\rho) + \nabla \cdot \left(\vec{v}\rho\right) = 0 \tag{2.7}$$

for $q = 1$. This should ideally be the case numerically as well. If the two equations are solved using the same numerical method, on the same grid and using the same time-step, the consistency is guaranteed. However, in reality, in practical applications of FV transport schemes, the settings are often inconsistent in this sense. This is definitely the case in offline tracer transport models. The consistency property, or rather the lack of it (referred to as the mass-wind inconsistency), will be discussed in detail in Section 3.

*2.1.8. Compatibility*

The compatibility property was defined by SCHÄR and SMOLARKIEWICZ [1996] for Eulerian schemes, and the definition is here extended also to include semi-Lagrangian schemes. As the consistency property, it concerns the relationship between continuity equations. Equations (2.6) and (2.7) imply

$$\frac{\mathrm{d}q}{\mathrm{d}t} = 0, \tag{2.8}$$

which states that the constituent mixing ratio is conserved along the characteristics of the flow. Compatible transport is when the discretization of Eq. (2.6) is consistent with the advective form (Eq. (2.8)) so that the predicted mixing ratio $q^{n+1}$, which in a flux-form setting is recovered from $(q\rho)^{n+1}$, is limited by the mixing ratios in the Eulerian cells from which the mass departs. The compatibility property is graphically illustrated in Fig. 2.1.

### *2.1.9. Preservation of constancy*

Another desirable property is the ability of the scheme to preserve a constant tracer field for a non-divergent flow. For traditional semi-Lagrangian methods based on Eq. (2.8), a constant distribution is trivially conserved since the divergence of the velocity field does not appear in the prognostic equation (for a review of traditional semi-Lagrangian methods, see, e.g., STANIFORTH and COTÉ [1991]). In fact, for any velocity field, the traditional semi-Lagrangian method preserves a constant mixing ratio. For FV methods where the divergence appears explicitly since tracer mass, and not mixing ratio, is the prognostic variable, it is not automatic that a constant field is preserved for a nondivergent velocity field. Non-conservation of constant fields may cause error problems and even instability, see Section 3.3.2.

### *2.1.10. Preservation of linear correlations between constituents*

Another desirable property identified by LIN and ROOD [1996] is that a numerical scheme should ideally preserve tracer correlations since correlations carry fundamental information on atmospheric transport. This is particularly important in chemical atmospheric models where the relative concentrations of constituents are crucial for the speed and

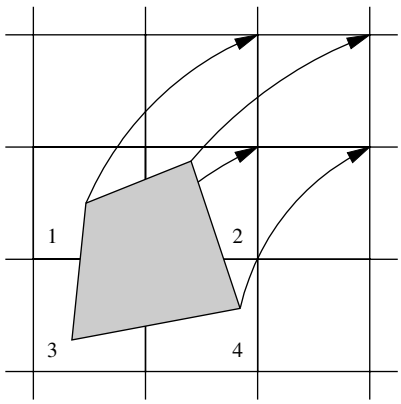


FIG. 2.1 A graphical illustration of the compatibility property. The arrows show the trajectories for the cell vertices. The shaded area is the departure cell that, after one time-step, ends up at the regular grid as depicted by the arrows. A finite-volume scheme predicts the change in total mass in the Eulerian cell $\overline{(q\rho)}^{n+1}$, which is the mass enclosed in the departure cell (shaded area). Since the mixing ratios are preserved along parcel trajectories, the mixing ratio in the arrival cell $\overline{q}^{n+1}$ should be within the range of the mixing ratios at the departure points. For the situation depicted on the figure, the compatibility condition is $\min\left(\overline{q}_1^n, \overline{q}_2^n, \overline{q}_3^n, \overline{q}_4^n\right) \leq \overline{q}^{n+1} \leq \max\left(\overline{q}_1^n, \overline{q}_2^n, \overline{q}_3^n, \overline{q}_4^n\right)$, where $\overline{q}_i$ denotes the average mixing ratio in the cell numbered $i$, $i = 1.4$, on the figure.

balances of chemical reactions. It is possible to construct transport schemes that maintain spatially constant linear correlations between tracers exactly, see, e.g., LIN and ROOD [1996].

### 2.2. *Subgrid-cell distributions of the prognostic variables*

In all FV schemes presented here – flux based as well as Lagrangian types – it is necessary to determine the subgrid-cell distribution from the surrounding cell averages in order to make accurate estimates of the fluxes through the Eulerian cell walls or mass enclosed in the upstream departure cell. Therefore, 1 and 2D reconstructions are discussed before the actual schemes are introduced.

#### 2.2.1. *1D subgrid-cell reconstructions*

Several 1D methods for reconstructing the subgrid distribution have been published in the literature. The simplest subgrid representation is a piecewise constant function followed, in complexity, by a piecewise linear representation (VAN LEER [1974]). Both methods are computationally cheap, optionally monotonic, and positive-definite (the piecewise constant method is shape-preserving by default) but, on the other hand, excessively damping and therefore not suited for long runs at coarser resolutions. To reduce the dissipation to a tolerable level, the subgrid-cell representation must be polynomials of at least second degree. Requirements of computational efficiency put an upper limit to the order of the polynomials used, which explains why the predominant choice is second order.

Let the walls of the $i$th cell be located at $x_i$ and $x_{i+1}$ and denote the cell width $\Delta x_i = x_{i+1} - x_i$. The coefficients of the subgrid-cell reconstruction polynomials are determined by imposing constraints. Apart from the basic requirement of mass conservation within each grid cell, the choice of constraints is not trivial. Probably the simplest parabolic fit is obtained by requiring that the polynomial

$$p_i(x) = (a_0)_i + (a_1)_i\, x + (a_2)_i\, x^2, \quad x \in [x_i, x_{i+1}] \tag{2.9}$$

not only conserves mass in the $i$th grid cell

$$\int_{x_i}^{x_{i+1}} p_i(x)\mathrm{d}x = \Delta x_i \overline{\psi}_i, \quad \psi = \rho, \rho q \tag{2.10}$$

but also in the two adjacent cells:

$$\int_{x_{i+1}}^{x_{i+2}} p_i(x)\mathrm{d}x = \Delta x_{i+1} \overline{\psi}_{i+1}, \tag{2.11}$$

$$\int_{x_{i-1}}^{x_i} p_i(x)\mathrm{d}x = \Delta x_{i-1} \overline{\psi}_{i-1}, \tag{2.12}$$

(Laprise and Plante [1995]). By substituting Eq. (2.9) into Eqs. (2.10), (2.11), and (2.12) and by evaluating the analytic integrals, a linear system results that can easily be solved for the three unknown coefficients $(a_0)_i$, $(a_1)_i$, and $(a_2)_i$ (Laprise and Plante [1995]). When performing this operation for all cells, a global piecewise-parabolic representation is obtained. The method is only locally of second order since it is not necessarily continuous across cell borders. This method is referred to as the piecewise parabolic method 1 (PPM1).

An alternative way of constructing the parabolas, which ensures a globally continuous distribution if no filters are applied, is the piecewise-parabolic method of Woodward and Colella [1984] (hereafter referred to as PPM2). PPM2 has been reviewed in the context of meteorological modeling in Carpenter, Droegemeier, Hane and Woodward [1990]. Instead of requiring that $p_i(x)$ conserves mass in adjacent cells, the constraint is that the polynomial equals prescribed west and east cell-edge values, $p_i^W = p_i(x_i)$ and $p_i^E = p_i(x_{i+1})$, respectively, at the cell edges. $p_i^E$ is computed with a cubic polynomial fit (see Woodward and Colella [1984] for details). For an equidistant grid, the result is

$$p_i^W = \frac{7}{12}\left(\overline{\psi}_{i-1} + \overline{\psi}_i\right) - \frac{1}{12}\left(\overline{\psi}_{i+1} + \overline{\psi}_{i-2}\right), \tag{2.13}$$

(for a nonequidistant grid, see Colella and Woodward [1984]). The east cell-border value, $p_i^E$, is simply an index shift of the formula for $p_i^W$

$$p_i^E = p_{i+1}^W. \tag{2.14}$$

It is convenient to use the cell average, $\overline{\psi}_i$, and $p_i^W$ and $p_i^E$ to define the $i$th parabola, instead of using $(a_0)_i$, $(a_1)_i$, and $(a_2)_i$. The equivalent formula for $p_i(x)$ is given by

$$p_i\left(\xi^x\right) = \overline{\psi}_i + \left(\delta p^x\right)_i \xi^x + \left(p_6^x\right)_i \left[\frac{1}{12} - \left(\xi^x\right)^2\right], \tag{2.15}$$

where $(\delta p^x)_i$ is the mean slope

$$\left(\delta p^x\right)_i = p_i^E - p_i^W, \tag{2.16}$$

$\left(p_6^x\right)_i$ is the "curvature"

$$\left(p_6^x\right)_i = 6\overline{\psi}_i - 3\left(p_i^W + p_i^E\right), \tag{2.17}$$

and $\xi^x$ is the nondimensional position defined by

$$\xi^x = \frac{x - x_i}{\Delta x_i} - \frac{1}{2}. \tag{2.18}$$

PPM2 uniquely defines the parabolas and Eq. (2.15) guarantees that the global subgrid distribution is continuous across cell borders. ZERROUKAT, WOOD and STANIFORTH [2002] found in passive advection tests that using the PPM2 for the subgrid-cell reconstructions (where the parabolas were continuous across cell borders) results in more accurate solutions compared with PPM1 (in which the distribution is not necessarily continuous across cell borders).

Instead of using the PPM1 or PPM2, ZERROUKAT, WOOD and STANIFORTH [2002] derived a piecewise cubic method for the reconstruction of the subgrid-cell distributions. PPM2 is a special case of the piecewise cubic method. Of course, any kind of reconstruction that is mass conserving can be used, e.g., rational functions as used in the transport scheme of XIAO, YABE, PENG and KOBAYASHI [2002] and the parabolic spline method (PSM) recently developed by ZERROUKAT, WOOD and STANIFORTH [2006]. In idealized advection tests, ZERROUKAT, WOOD and STANIFORTH [2007] found that using the PSM for the subgrid-scale reconstructions in their scheme generally leads to more accurate results than when using PPM2. This is despite the fact that, in terms of operation count, PSM is 60% more efficient than PPM2. However, at present, the most widespread subgrid-cell reconstruction method is PPM2.

Without further constraining the coefficients of the parabolas, it is not guaranteed that the subgrid-scale reconstruction preserves monotonicity or positive definiteness (GODUNOV [1959]). A simple monotonic filter was proposed by COLELLA and WOODWARD [1984], and is explained in Fig. 2.2. For local extrema, the filter is similar to the quasi-monotonic filter by BERMEJO and STANIFORTH [1992] for traditional semi-Lagrangian advection schemes, i.e., the subgrid-scale distribution is reduced to a constant when there is a local extrema in the cell averages (Fig. 2.2a). This severe clipping can significantly reduce the accuracy as idealized advection tests have shown (compare CISL-N with CISL-M and CCS-N with CCS-M in Table 2.1). Clearly, one would like to retain the higher order polynomial in the situation depicted in Fig. 2.2a while not altering the treatment of the situation in Fig. 2.2b.

LIN and ROOD [1996] modified the COLELLA and WOODWARD [1984] monotonic filter so that the monotonic filter only applies to undershooting and does not interfere with any of the overshooting (referred to as semi-monotonic filter). The semi-monotonic filter can further be modified so that it only prevents negative undershooting, whereby it becomes a positive-definite filter. Since these filters avoid the severe clipping of overshoots, the application of these filters shows a dramatic increase in accuracy in idealized advection tests compared with the monotonic filter described in the previous paragraph (CISL-P and CCS-P in Table 2.1). Other filters with more relaxed constraints, but which are computationally more efficient, can be found in LIN [2004]. However, all these filters are still not fully satisfactory since they do not interfere with all types of spurious undershooting and overshooting.

As mentioned, the filter should not interfere with local extrema as the one in Fig. 2.2a but still apply the monotonic filter in the situation depicted on Fig. 2.2b (similarly for undershooting). That is what the filter of SUN, YEH, SUN and SUN [1996] for traditional semi-Lagrangian schemes is designed to do. Through a series of logical statements, the filter detects local extrema and does not alter the high-order subgrid-scale reconstruction

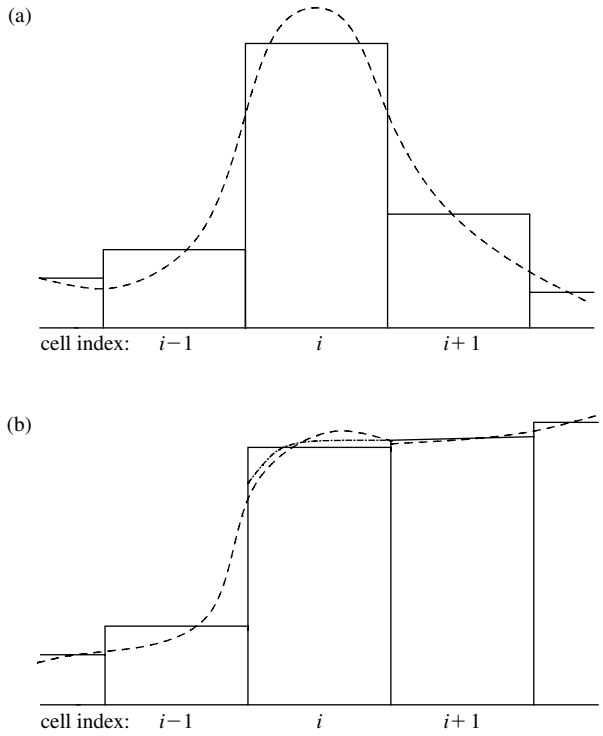


FIG. 2.2 A graphical illustration of the basic monotonic filter of COLELLA and WOODWARD [1984]. Solid lines show the cell averages, and the dashed line is the unmodified piecewise parabolic fit. (a) The situation in which the parabola in cell $i$ is a local extrema. The monotonic filter sets the parabola equal to a constant in cell $i$. (b) The situation when $\overline{\psi}_i$ is in between $(p_E)_i$ and $(p_W)_i$, but is sufficiently close to one of the edge values that the parabola takes values outside the range of the surrounding cell averages, i.e., when $|(p_W - p_E)_i| \geq \left|\left(p_6^x\right)_i\right|$. In this situation, $(p_E)_i$ is reset and the gradient at the east cell wall is set to zero thereby guaranteeing monotonicity of the polynomial in cell $i$ (dash-dotted line) or vice versa.

where these nonspurious extrema are located. As pointed out by NAIR, CÔTÉ and STANIFORTH [1999a], this filter, however, is still unsatisfactory near strong gradients, where the unmodified subgrid-scale distribution exhibits $2\Delta x$ noise. In such a situation, for example, the semi-monotonic filter of LIN and ROOD [1996] or the filter of SUN, YEH and SUN [1996] does not filter the noise satisfactorily (Fig. 2.3). To deal with such situations (and others) while still maintaining non-spurious extrema, ZERROUKAT, WOOD and STANIFORTH [2005] proposed a more advanced filter that, in the situations shown on Figs. 2.2 and 2.3, consecutively reduces the order of the fitting polynomials until none of the spurious overshooting and undershooting depicted on the Figs. 2.2 and 2.3 appear. Then, the severe clipping of physical "peaks" is eliminated and grid-scale noise is removed without introducing excessive numerical damping. Contrarily to the monotonic filter of COLELLA and WOODWARD [1984], this filter can improve the accuracy compared

TABLE 2.1
Error norms for the schemes of ZERROUKAT, WOOD and STANIFORTH [2005] (SLICE), NAIR and MACHENHAUER [2002] (CISL), and NAIR, SCROGGS and SEMAZZI [2002] (CCS) for test case 1 in WILLIAMSON, DRAKE, HACK, JAKOB and SWARZTRAUBER [1992]. $\hat{\alpha}$ is the angle between the axis of solid body rotation and the polar axis of the spherical coordinate system. Hence $\hat{\alpha} = 0$ is solid body rotation along the equator and $\hat{\alpha} = \pi/2$ advection across the poles. The error measures for $\hat{\alpha} = \pi/3$ are from LAURITZEN, KAAS and MACHENHAUER [2006]. "N" denotes no filter, "M" the monotonic filter, and "P" the positive-definite filter used in the respective schemes. Note that the monotonic filter and subgrid-scale reconstructions in SLICE are different from the other schemes (see text for details)

| | $\hat{\alpha} = 0$ | | | $\hat{\alpha} = \pi/2$ | | | $\hat{\alpha} = \pi/3$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Schemes | $l_1$ | $l_2$ | $l_\infty$ | $l_1$ | $l_2$ | $l_\infty$ | $l_1$ | $l_2$ | $l_\infty$ |
| SLICE-N | 0.046 | 0.029 | 0.022 | 0.079 | 0.049 | 0.042 | — | — | — |
| SLICE-M | 0.038 | 0.024 | 0.017 | 0.058 | 0.040 | 0.037 | — | — | — |
| CISL-N | 0.052 | 0.035 | 0.032 | 0.063 | 0.046 | 0.048 | 0.075 | 0.051 | 0.083 |
| CISL-P | 0.025 | 0.025 | 0.031 | 0.059 | 0.045 | 0.048 | 0.043 | 0.082 | 0.076 |
| CISL-M | 0.094 | 0.091 | 0.108 | 0.084 | 0.084 | 0.109 | 0.077 | 0.089 | 0.18 |
| CCS-N | — | — | — | 0.054 | 0.042 | 0.065 | 0.051 | 0.039 | 0.076 |
| CCS-P | 0.036 | 0.034 | 0.042 | 0.051 | 0.041 | 0.065 | 0.033 | 0.034 | 0.077 |
| CCS-M | — | — | — | 0.076 | 0.082 | 0.129 | 0.070 | 0.086 | 0.186 |

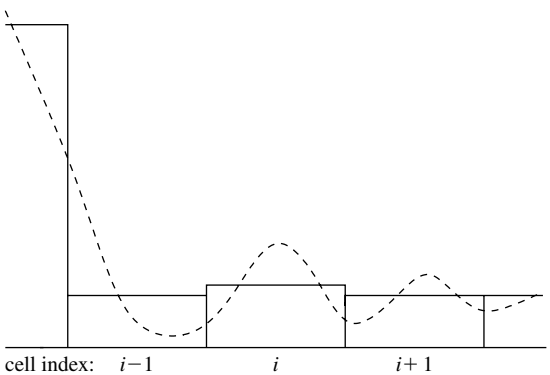


FIG. 2.3 A situation in which the unmodified subgrid-cell reconstruction exhibits strong Gibbs phenomena. The semimonotonic filter of LIN and ROOD [1996] would set the polynomials in cell $i-1$ and $i+1$ equal to the cell average, but would not modify the polynomial in cell $i$ that, in this situation, is a spurious overshoot.

to the unfiltered high-order solution (see Semi-Lagrangian inherently conserving and efficient (SLICE)-N and SLICE-M in Table 2.1). A similar filter has also been developed for the PSM (ZERROUKAT, WOOD and STANIFORTH [2006]).

### *2.2.2. 2D subgrid-cell reconstructions*

As for the 1D case, 2D linear reconstructions exist (e.g., DUKOWICZ and BAUMGARDNER [2000] and SCROGGS and SEMAZZI [1995]), but, in general, they introduce too much numerical damping for meteorological applications. The PPM in 1D can be directly

extended to 2Ds as done by RANČIĆ [1992], i.e., in terms of a fully 2D subgrid-cell reconstruction

$$p_{i,j}(x, y) = (a_1)_{i,j} + (a_2)_{i,j}\, x + (a_3)_{i,j}\, x^2 + (a_4)_{i,j}\, y + (a_5)_{i,j}\, y^2 + (a_6)_{i,j}\, xy + (a_7)_{i,j}\, xy^2 + (a_8)_{i,j}\, x^2 y + (a_9)_{i,j}\, x^2 y^2,\ (x, y) \in [x_i, x_{i+1}] \times \left[y_j, y_{j+1}\right]. \tag{2.19}$$

This fully biparabolic fit involves the computation of nine coefficients, so nine constraints are needed to determine the coefficient values. Apart from the conservation of mass within each cell

$$\iint\limits_{\Delta A_{i,j}} p_{i,j}(x, y)\mathrm{d}x\mathrm{d}y = \overline{\psi}_{i,j}\Delta A_{i,j}, \tag{2.20}$$

the other eight constraints chosen by RANČIĆ were formulated in terms of the four corner values of $p_{i,j}(x, y)$ and the average of $p_{i,j}(x, y)$ along the four cell walls. The corner point scalar values were computed by fitting 2D third-order polynomials using the 16 cell averages surrounding the corner point in question. The average along the cell walls was computed using $\overline{\psi}$ along a line perpendicular to the cell wall in question. For additional details, see RANČIĆ [1992].

The computational cost of the approach taken by Rančić can be reduced significantly by using a quasi-biparabolic subgrid-cell representation. Contrarily to fully biparabolic fits, the quasi-biparabolic representation does not include the "diagonal" terms and simply consists of the sum of two 1D parabolas, one in each coordinate direction. Using the form (Eq. (2.15)) for the parabolas, the quasi-biparabolic subgrid-cell representation is given by

$$p_i\left(\xi^x, \xi^y\right) = \overline{\psi}_{i,j} + \left(\delta p^x\right)_{i,j}\xi^x + \left(p_6^x\right)_{i,j}\left[\frac{1}{12} - \left(\xi^x\right)^2\right] + \left(\delta p^y\right)_{i,j}\xi^y + \left(p_6^y\right)_{i,j}\left[\frac{1}{12} - \left(\xi^y\right)^2\right], \tag{2.21}$$

where $(\delta p^x)_{i,j}$, $\left(p_6^x\right)_{i,j}$, $(\delta p^y)_{i,j}$, and $\left(p_6^y\right)_{i,j}$ are the coefficients of the parabolic functions in each coordinate direction (MACHENHAUER and OLK [1998]). This representation reduces the computational cost of the subgrid-cell reconstruction significantly but, of course, does not include variation along the diagonals of the cells.

By using 1D filters that prevent undershoot and overshoot to the parabolas in each coordinate direction, monotonicity-violating behavior can be reduced but not strictly eliminated in 2Ds. In case of negative values at the cell boundaries of both unfiltered 1D parabolic representations, even larger negative values may be present in one or more of the cell corners when the 1D representations are added. The monotone and positive-definite filters eliminate only the negative values at the boundaries and not the possible negative corner values. As a result, small negative values can appear even after the application of a monotonic filter (e.g., LIN and ROOD [1996] and NAIR and MACHENHAUER [2002]). To eliminate these negative values an additional filter must be applied.

### 2.3. Different schemes in 2Ds

As mentioned above, different approaches can be used to estimate the integral over the departure cell. These can be divided into two main categories:

- Semi-Lagrangian schemes in which the integral over the departure cell is approximated explicitly. These schemes to be described in Section 2.3.1 are referred to as DCISL schemes. DCISL schemes come in two types: fully 2D schemes and cascade schemes in which the approximation of the upstream integral is divided into two steps where each substep applies 1D methods.
- Flux-based schemes in which the fluxes through the Eulerian arrival cell walls are approximated. These schemes are described in Section 2.3.2. As for the DCISL schemes, there are two types of conceptually different schemes of this category: schemes based on a sequential operator splitting (often referred to as time-splitting) and schemes based on direct estimation of the 2D fluxes.

It is important to note, as was also pointed out in the introduction, that DCISL and flux-based FV schemes are conceptually equivalent since they both estimate the mass in the departure cell. However, as will be illustrated, this is, in practice, done in quite different ways.

The following overview of these two categories will mainly focus on recent developments in DCISL schemes since these have not yet been introduced in textbooks or general overview articles. For these schemes, a stability analysis is performed. Furthermore, the level of local mass conservation, i.e., the accuracy of the approximation to the exact departure area in different DCISL schemes and one flux-based method, is investigated.

#### 2.3.1. DCISL schemes

The semi-Lagrangian scheme can either be based on backward or forward trajectories (or equivalently downstream and upstream trajectories), i.e., by considering parcels arriving or departing from a regular grid, respectively. The majority of semi-Lagrangian schemes are based on backward trajectories because it is usually simpler to interpolate/remap from a regular to a distorted mesh. However, forward trajectory cascade schemes and the downstream version of the schemes in LAPRISE and PLANTE [1995] are exceptions to this. The deformed grid resulting from tracking the parcels moving with the flow is referred to as the *Lagrangian grid,* while the stationary and regular grid is referred to as the *Eulerian grid*. The curve resulting from tracking a latitude moving with the flow is referred to as a *Lagrangian latitude*. Similarly for a *Lagrangian longitude*.

The choice of trajectory algorithm is crucial for the accuracy of DCISL schemes. Traditional semi-Lagrangian schemes employ backward trajectories that are computed with an implicit iterative algorithm also known as the second-order implicit midpoint method (see, e.g., CÔTÉ and STANIFORTH [1991]). This trajectory algorithm does not include the acceleration. Several schemes that include estimates of the acceleration in the trajectory computations have been proposed (e.g., HORTAL [2002], MCGREGOR [1993], LAURITZEN, KAAS and MACHENHAUER [2006] – see Section 3).

Using backward trajectories, the 2D discretization of Eq. (1.8) leads to the CISL scheme

$$\overline{\psi}^{+}\Delta A = \overline{\psi}\delta A, \quad \psi = \rho, \rho q, \tag{2.22}$$

where

$$\overline{\psi} = \frac{1}{\delta A}\iint_{\delta A} \psi\,(x, y)\,\mathrm{d}A \tag{2.23}$$

is the integral mean value of $\psi\,(x, y)$ over the irregular departure cell area $\delta A$, and $\overline{\psi}^{+}$ is the mean value of $\psi^{+}\,(x, y)$ over the regular arrival cell area $\Delta A$ (see Fig. 2.4). The approximation of the integral on the right-hand side of Eq. (2.22) employs two steps: firstly, defining the geometry of the departure cell that involves the computation of parcel trajectories; secondly, performing the remapping, i.e., computing the integral over the departure cell using some reconstruction of the subgrid distribution at the previous time-step. The geometrical definition of the departure cell and the complexity of the subgrid-scale distribution are crucial for the efficiency and accuracy of the scheme.

For realistic flows and for time-steps obeying the Lipschitz criterion (see Section 2.1), the upstream cells deform into simply connected but non-rectangular and possibly locally concave geometric patterns. The question is how to integrate $\psi\,(x, y)$ efficiently over such a complex area.

*2.3.1.1. Fully 2D DCISL schemes* In 1D, there is very little ambiguity on how to approximate the upstream cell, but in 2Ds, it is much more complicated and several approaches have been suggested in the literature. In Fig. 2.5, the arrival and departure cells in cartesian geometry for three different DCISL schemes are shown.

Rančić [1992] defines the departure cell as a quadrilateral by tracking backward the cell vertices A, B, C, and D and connecting them with straight lines (Fig. 2.5(a)). The

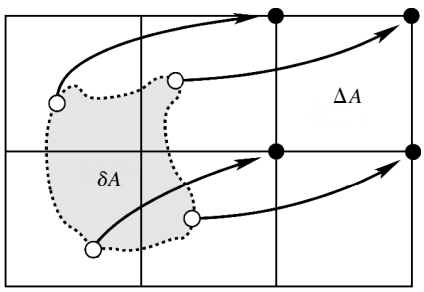


Fig. 2.4 The regular arrival cell with area $\Delta A$ and the irregular departure cell (shaded area) with area $\delta A$ in the continuous case for a generic upstream DCISL scheme. Using the figure of speech in Laprise and Plante [1995], the departure–arrival cell relationship is conceptually equivalent to throwing a fishing net upstream to fetch the mass enclosed into a area that will, after one time-step, end up at the regular mesh. The arrows are the parcel trajectories from the departure points (open circles), which arrive at the regular cell vertices (filled circles).

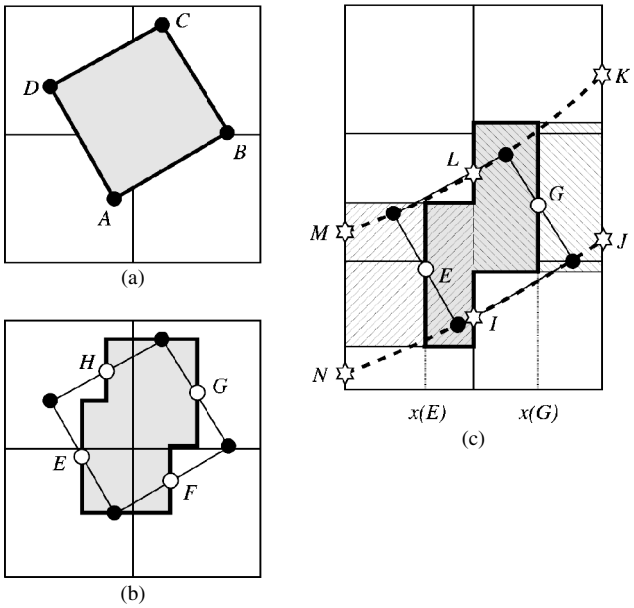


FIG. 2.5 The departure cell (shaded area) when using the scheme of (a) RANČIĆ [1992], (b) MACHENHAUER and OLK [1998] scheme, and (c) the cascade scheme of NAIR, SCROGGS and SEMAZZI [2002]. The filled circles are the departure points, and open circles the midpoints between the departure points, and asterisks are the intermediate grid points which are used to define the intermediate cells in the cascade scheme (crosshatched area).

vertices are not necessarily aligned with the coordinate axis, which leads to some algorithmic complexity for the evaluation of the upstream integral. The integral over the departure area is, in the situation depicted in Fig. 2.5a, decomposed into four subintegrals, i.e., the integral over the areas defined by the overlap between the departure cell and the Eulerian cells. Thus, one has to perform analytic integrals over many possible cases of shapes of subdomains, which makes the computer code rather cumbersome. In addition, the subgrid-scale distribution used by RANČIĆ was a piecewise-biparabolic representation which, being fully 2D, is quite expensive to compute in itself. The combination of the complex geometry of the departure cell and the fully 2D subgrid-cell representation makes the scheme approximately 2.5 times less efficient than the traditional semi-Lagrangian advection scheme using bicubic Lagrange interpolation (RANČIĆ [1992]). This, and the fact that the scheme has not been extended to spherical geometry, has hindered the scheme for widespread use in the meteorological community.

In order to speed up the remapping process, MACHENHAUER and OLK [1998] simplified both the geometry of the departure cell and the subgrid-scale distribution. The departure cell is defined as a polygon with sides parallel to the coordinate axis (Fig. 2.5(b)). The sides parallel to the $x$-axis are at the $y$-values of the departure points, and the sides parallel to the $y$-axis pass through E, F, G, and H, located halfway between the departure points. Hence, the area of the departure cell is identical to the area of the RANČIĆ [1992] scheme. Since the sides of the departure cell are parallel to the coordinate axis, the evaluation of the upstream integral is greatly simplified. By using the

pseudobiparabolic subgrid-scale distribution (see Eq. (2.21)) and accumulated parabola-coefficients along latitudes (see NAIR and MACHENHAUER [2002] for details), the integral over the departure cell can be computed much more efficiently compared to the approach taken by RANČIĆ [1992]. For advection in cartesian geometry, NAIR and MACHENHAUER [2002] reported a 10% overhead with this scheme compared with the traditional semi-Lagrangian scheme.

Note that the departure areas in Fig. 2.5 completely cover the entire integration area without overlaps or cracks, which is crucial to an upstream DCISL scheme, otherwise the total mass is not conserved. For a downwind cell-integrated scheme using forward trajectories, it is, however, not necessary that the arrival cells span the entire domain of integration in order to have global mass conservation. Using the figure of speech of LAPRISE and PLANTE [1995], a downstream cell-integrated scheme is equivalent to throwing dust contained in little buckets (regular departure cells) into the wind and watching it later fall into bins (regular Eulerian cells). Contrary to upstream DCISL schemes, where one integrates over a particular departure cell, a downstream cell-integrated scheme keeps track of the contribution to each regular Eulerian cell from the irregular arrival cells. As long as all the mass in each arrival cell is redistributed with a mass conservative method, mass is conserved even if the neighboring arrival cells overlap. This is taken advantage of in the scheme of LAPRISE and PLANTE [1995], which probably uses the simplest cell geometry of all the schemes presented here. The arrival cell is defined as a rectangle where the edges have the same orientation as the regular cells (see Fig. 2.6). This is achieved by tracing the traverse motion of cell edges and not the cell vertices. Hereby, the arrival cell retains the orthogonality and orientation of the regular departure cell. Note, however, that only two cells share edges, while, if cell vertices are tracked,

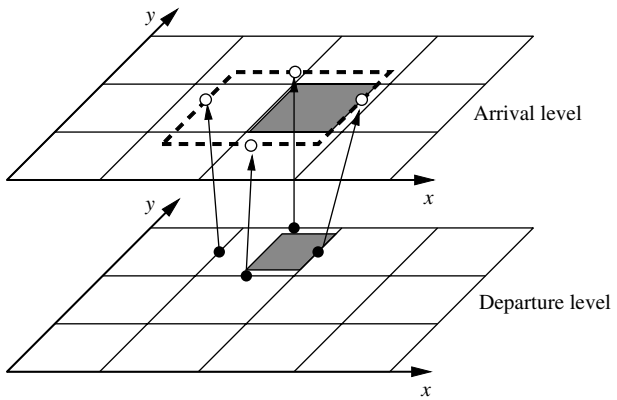


FIG. 2.6 A graphical illustration of the downstream version of the cell-integrated schemes of LAPRISE and PLANTE [1995]. The filled circles are the departure points that are at the edge centers of the regular departure cell. The arrows connect the departure points with the respective arrival points (unfilled circles). The dashed rectangle is the arrival cell which edges have the same orientation as the departure cell. In a downstream cell-integrated scheme, the amount of mass that arrives at a regular Eulerian cell is computed, i.e., the integral over the area (shaded area at the departure level) that arrives at the intersection between the regular Eulerian cell and the arrival cell (shaded area at the arrival level). Similarly for the remaining intersections with Eulerian cells.

cell vertices are shared by four cells. Consequently, one must compute twice as many trajectories compared to a downstream scheme tracking cell vertices.

The actual integral in the downstream scheme of LAPRISE and PLANTE [1995] is not performed at the arrival level since that would require the reconstruction of the subgrid-scale distributions from irregular and overlapping arrival cell averages. The integral is performed at the departure level over the part of the departure cell that, after one time-step, "falls" into a particular regular Eulerian cell (see Fig. 2.6). The intersection between the arrival cell and a particular Eulerian cell is always a rectangular region with sides parallel to the coordinate axis, which simplifies the integration process significantly. Consequently, the downstream scheme of LAPRISE and PLANTE [1995] is approximately twice as fast as the RANČIĆ [1992] scheme, even though both schemes use fully 2D subgrid-scale reconstructions. The schemes of LAPRISE and PLANTE [1995] and RANČIĆ [1992] have not been extended to spherical geometry.

*2.3.1.2. Cascade DCISL schemes* Using the so-called cascade method, originally developed for non-conservative interpolation by PURSER and LESLIE [1991], the 2D upstream integral can also be approximated by splitting it into two 1D steps. The basic idea is to track backward (forward) the Eulerian grid and then apply 1D integrals, firstly along the Eulerian (Lagrangian) longitudes or latitudes and secondly along the Lagrangian (Eulerian) latitudes or longitudes (see Fig. 2.7). To obtain inherent mass conservation, the interpolation in the original cascade interpolation method must be replaced with the PPM (COLELLA and WOODWARD [1984]) or some other mass-conservative remapping method. Contrary to fully 2D DCISL schemes, the cascade approach is equally suited for downstream and upstream trajectories, or equivalently, the 1D remapping methods are equally suited for remapping from a distorted as from a regular 1D grid. However, to facilitate the comparison with fully 2D DCISL schemes, we assume upstream trajectories in the discussion of the cascade schemes, although some of the schemes initially were formulated for downstream trajectories.

The cascade method can be divided into three steps. Firstly, given the departure points, the application of 1D remappings is prepared by computing an intermediate grid. It is crucial to the cascade technique that the intermediate grid is well defined, i.e., that there should not be multiple intersections between Lagrangian latitudes (longitudes) and Eulerian longitudes (latitudes) (see, e.g., Fig. 1 in NAIR, CÔTÉ and STANIFORTH [1999a]). Therefore, in spherical coordinates, it cannot be applied very near the poles. Secondly, a 1D remapping of mass from the regular Eulerian cells to the intermediate grid cells is performed. Thirdly, the mass on the intermediate grid is remapped to the departure cells.

We start out by considering the conservative cascade DCISL scheme of NAIR, SCROGGS and SEMAZZI [2002]. In this scheme, the departure cells are defined as polygons with sides parallel to the coordinate axis as in the MACHENHAUER and OLK [1998] scheme. In each 1D cascade step, the PPM2 is used. Compared to the MACHENHAUER and OLK [1998] scheme, the departure cell geometry is defined somewhat differently (see Fig. 2.5(c)). Two of the sides parallel to the $y$-axis, $x = x(E)$ and $x = x(G)$, are defined as in the MACHENHAUER and OLK [1998] scheme, and the remaining two sides parallel to the $y$-axis are at the Eulerian longitude $x = x_i$. The sides parallel to the $x$-axis are determined from the intermediate Lagrangian grid points $I$, $J$, $K$, $L$, $M$, and

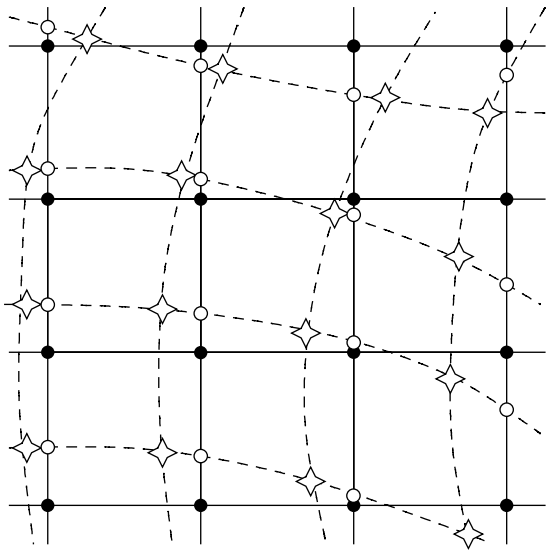

FIG. 2.7 Graphical illustration of the cascade interpolation method introduced by PURSER and LESLIE [1991]. Solid lines are the regular Eulerian grid and the dashed lines are the Lagrangian grid. Here we consider an upstream scheme, hence, the Lagrangian grid end up at the Eulerian grid when moving with the flow over one time-step. The intermediate grid is defined by the crossings between the Eulerian longitudes and the Lagrangian latitudes (unfilled circles). The nonconservative cascade interpolation method proceeds as follows. Perform a 1D interpolation from the Eulerian grid to the intermediate grid, i.e., interpolate along Eulerian longitudes from the filled circles to the unfilled circles. Hereafter interpolate from the intermediate grid to the Lagrangian grid, i.e., from the unfilled circles to the asterisks. Ideally one should interpolate along the curved Lagrangian latitude, but in the original cascade interpolation scheme of PURSER and LESLIE [1991], the $x$-coordinate is used as the position variable for interpolation along the Lagrangian latitude, which corresponds to approximating the Lagrangian latitude with line segments parallel to the $x$-axis.

$N$ defined as $y = \frac{1}{2}[y(I) + y(J)]$, $y = \frac{1}{2}[y(K) + y(L)]$, $y = \frac{1}{2}[y(L) + y(M)]$, and $y = \frac{1}{2}[y(I) + y(N)]$, respectively. The $y$-values of the intermediate points are determined by cubic Lagrange interpolation between the $y$-values of four adjacent departure points along the *Lagrangian latitude* (dashed line in Fig. 2.5(c)). The Lagrange weights for computing the intersections can be efficiently evaluated using the algorithm outlined in the Appendix in PURSER and LESLIE [1991]. The upstream integral is computed by a remapping in the north-south direction from the Eulerian cells to the intermediate cells (crosshatched rectangular regions on Fig. 2.5(c)), followed by a remapping along the *Lagrangian latitudes* from the intermediate cells to the departure cells. Hence, the first remapping is along the Eulerian longitude passing through the Eulerian cell centers. Since the second remapping uses the $x$-coordinate as the independent variable, it is along line segments parallel to the $x$-axis. Without any a priori knowledge of the flow, there is no argument for not reversing the order of the directional sweeps, i.e., first to remap along the Eulerian latitude and then along the Lagrangian longitude. As discussed in some detail in LAURITZEN [2007], the order of the directional sweeps is not symmetric, and hence there is a directional bias built into the cascade approach. However, in neither

of the cascade schemes presented here has this been reported to be a problem. A symmetric version of the cascade scheme can easily be constructed, for example, by alternating between the sweep directions, i.e., by using Lagrangian longitudes and Eulerian latitudes at even time-steps and Lagrangian latitudes and Eulerian longitudes at odd time-steps.

Since the two remappings are 1D and that the intermediate grid can be efficiently computed, the NAIR, SCROGGS and SEMAZZI [2002] scheme is more than twice as efficient as the fully 2D scheme of MACHENHAUER and OLK [1998]. Cascade methods are equally suited for upstream and downstream trajectories. For example, the NAIR, SCROGGS and SEMAZZI [2002] scheme formulated for backward trajectories has also been extended to forward trajectories in NAIR, SCROGGS and SEMAZZI [2003].

On equidistant cartesian grids, the conservative cascade scheme developed by NAIR, SCROGGS and SEMAZZI [2002] is very similar to the one of RANČIĆ [1995], although the way they are presented in the respective articles is very different. RANČIĆ [1995] formulated his scheme without explicit reference to areas by assigning mass to nodes or *mass-points*. The scheme is identical to the PURSER and LESLIE [1991] cascade interpolation but with the two 1D Lagrange interpolation sweeps replaced with PPM2. Although Rančić did not make explicit reference to areas in his formulation, the scheme can, however, be interpreted in terms of areas: in each 1D sweep, the mass nodes represent the mass enclosed in cells with walls located midway between the mass nodes, and the remapping is along line segments which are parallel to the coordinate axis. Hence, by formulating Rančić scheme for upstream trajectories, the only differences between the NAIR, SCROGGS and SEMAZZI [2002] and Rančić scheme are the choice of points for which the trajectories are computed and the order of the 1D sweeps. Where NAIR, SCROGGS and SEMAZZI [2002] track cell vertices as they are transported by the flow, Rančić used cell centers; and where NAIR, SCROGGS and SEMAZZI [2002] remaps first along Eulerian longitudes, the upstream version of Rančić's scheme remaps along the Eulerian latitudes first. Hence, in principle, these schemes are identical and only differing in implementation details when considering a cartesian equidistant grid. However, it is not clear if Rančić's scheme can be extended to non-equidistant grids, and hence it has not been extended to spherical geometry, whereas the NAIR, SCROGGS and SEMAZZI [2002] scheme has been extended to spherical geometry using two different approaches (NAIR, SCROGGS and SEMAZZI [2002] and NAIR [2004]).

In the cascade schemes discussed so far, the second sweep is along Lagrangian latitudes (longitudes) that are defined by line segments parallel to the $x$-axis ($y$-axis). In the continuous case, the Lagrangian latitude (longitude) is a curve, and one should ideally remap mass along such a curve. ZERROUKAT, WOOD and STANIFORTH [2002] refined the approaches described so far by performing the second sweep along a continuous piecewise linear line that more accurately represents the curved Lagrangian latitude (longitude). The scheme is called the SLICE scheme and is described next.

The remapping procedure used in SLICE is graphically illustrated on Fig. 2.8. As in the scheme of MACHENHAUER and OLK [1998] and the cascade scheme of NAIR, SCROGGS and SEMAZZI [2002], the cell vertices are tracked backward. The corresponding departure points are connected with straight lines to define *Lagrangian longitudes* and *latitudes*. Regular intermediate cells are defined by the intersections between the Lagrangian longitudes and the Eulerian latitudes that pass through the center of the

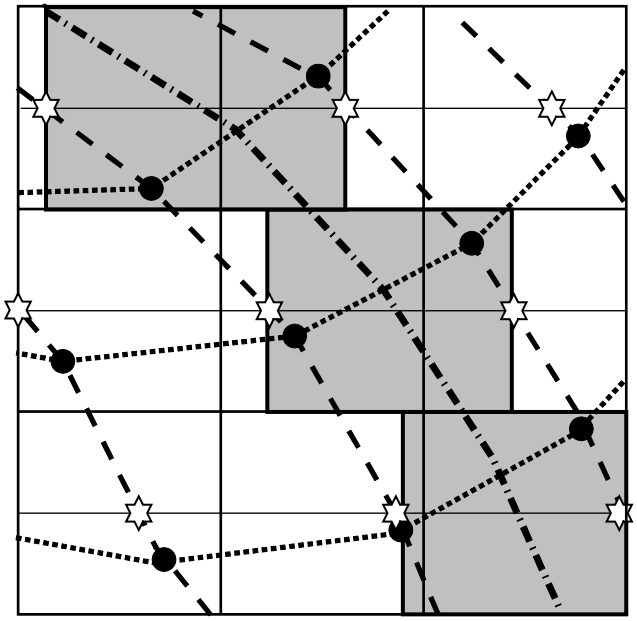

FIG. 2.8 A graphical illustration of the remappings in the Semi-Lagrangian inherently conserving and efficient scheme. The filled circles are the departure points corresponding to the cell vertices. The dotted (dashed) lines are the Lagrangian latitudes defined by connecting the departure points which arrive along the same latitude with straight line segments. The shaded areas are the intermediate areas that are defined by the crossings between the Lagrangian longitudes and the Eulerian latitudes passing through the center of the Eulerian cells (thin lines). The crossings are marked with asterisk. First the mass is remapped from the Eulerian cells to the intermediate cells. The dash-dotted line is the line along which the cumulative distance function is defined, and is used for the second remapping.

cells. Similarly to the NAIR, SCROGGS and SEMAZZI [2002] scheme, the cell averages are mapped from the Eulerian cells to the regular intermediate cells defined by the intersections (see shaded area in Fig. 2.8). As mentioned in the previous paragraph, the remapping from the intermediate cells to the departure cells is quite different from the NAIR, SCROGGS and SEMAZZI [2002] scheme. The second remapping is performed along the Lagrangian longitude that is defined by a continuous piecewise linear line (dash-dot line in Fig. 2.8). A cumulative distance function along the Lagrangian longitude is used to define the Eulerian and Lagrangian north-south cell walls for the second remapping (see ZERROUKAT, WOOD and STANIFORTH [2002] for details). Hereby the independent coordinate for the second sweep is defined along continuous piecewise linear lines that, in principle, are more accurate than line segments parallel to the coordinate axis. However, the intermediate cells have walls parallel to the coordinate axis (east-west walls of shaded on Fig. 2.8). Hence the mass used in the second sweep is only approximately along the piecewise linear Lagrangian longitude (dashed lines on Fig. 2.8). For the 1D remappings, SLICE applies a piecewise cubic method (ZERROUKAT, WOOD and STANIFORTH [2002]) or the PSM (ZERROUKAT, WOOD and STANIFORTH [2007]).

A great potential of cascade schemes is that they may be extended to 3Ds without excessive computational cost and algorithmic complexity. For example, 3D Lagrangian interpolation requires $O(o^3)$ operations, where $o$ is the formal order of accuracy of the interpolator, while cascade schemes require $O(o)$ operations (e.g., PURSER and LESLIE [1991]). Cascade schemes retain their simplicity in higher dimensions, whereas fully higher dimensional DCISL schemes increase rapidly in complexity as the number of dimensions is increased. Cascade interpolation has not only been applied in semi-Lagrangian advection schemes but have also been used for remapping state variables between the regular latitude-longitude grid and the cubed-sphere grid in a conservative and monotone manner (LAURITZEN and NAIR [2007]).

In some situations, the cascade DCISL schemes get a more accurate subgrid-scale representation compared to the fully 2D DCISL scheme of MACHENHAUER and OLK [1998]. The latter scheme uses a 2D reconstruction that does not include variation along the diagonals of the cells. In the cascade schemes, the second sweep is along Lagrangian latitudes (longitudes). Hence, in situations in which the Lagrangian latitudes (longitudes) are sloping toward north-east or north-west and significant variation is along these Lagrangian latitudes (longitudes), the cascade schemes get some of the diagonal variation that is eliminated by the subgrid-scale reconstructions used in MACHENHAUER and OLK [1998]. This is clearly demonstrated when comparing the error measures for the solid body advection for the flow orientation parameter $\hat{\alpha} = \pi/3$ in Table 2.1 (see CCS and CISL). In this situation, the distribution is far away from the poles and the transport is nearly along celldiagonals.

*2.3.1.3. Degree of local mass conservation* To understand to which extent the different DCISL schemes are local, a test case using an analytic flow field involving translation, rotation, and divergence has been constructed. We define the degree of mass "locality" of the schemes as their ability to approximate the domain of the exact upstream departure cell over which the mass is integrated. Note that the domain of dependence is larger than the upstream departure cell area. The domain of dependence is the area from which information is needed to construct the subgrid-cell representations in the Eulerian cells overlapped by the departure cell.

The analytic wind field is given by

$$\begin{aligned} u(x, y) &= u_0 + D_0 x - R_0 y, \\ v(x, y) &= v_0 + D_0 y + R_0 x, \end{aligned} \tag{2.24}$$

where $(u_0, v_0) = 54\frac{\mathrm{m}}{\mathrm{s}} \times (\cos(10^\circ), \sin(10^\circ))$, $D_0 = -0.0023\,/\mathrm{s}$, and $R_0 = 0.0029\,/\mathrm{s}$. The time-step used for the test is $\Delta t = 120\,\mathrm{s}$ and the grid-point spacing is $\Delta x = \Delta y = 5000\,\mathrm{m}$. These values have been estimated from typical forecast values near strong baroclinic developments obtained with the operational $5 \times 5$ km high-resolution limited-area model (HIRLAM) forecasting system run at the Danish Meteorological Institute. However, the HIRLAM $D_0$ and $R_0$ values have been multiplied by a factor 10 in order to visualize the effect of divergence and rotation. The "exact" trajectories and departure cell are shown on Fig. 2.9 ("exact" refers to a 18-digit precise computation of the departure

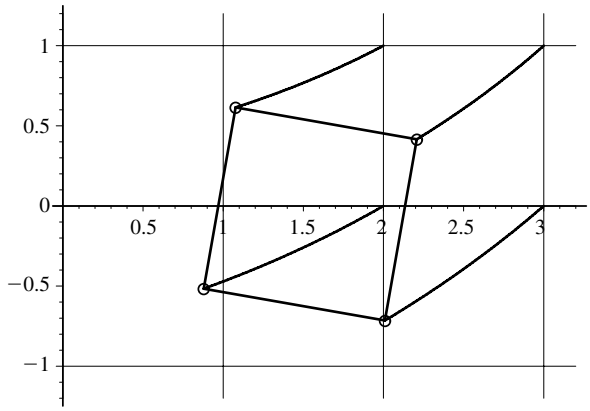


FIG. 2.9 'Exact' departure cell and backward trajectories (curved lines) for the analytic velocity field consisting of a translational, divergent, and rotational part. The unfilled circles are the departure points computed with the trajectory algorithm of LAURITZEN, KAAS and MACHENHAUER [2006]. The values on the $x$- and $y$-axis are in units of 5000 m.

points using a Fehlberg fourth–fifth order Runge-Kutta method). Three error measures are used to measure the degree of local mass conservation:

- The area of the departure cell normalized by the exact departure cell area.
- The area located outside the exact departure cell normalized by the exact departure area.
- The area located inside the exact departure cell normalized by the exact departure area.

Figure 2.10 shows the departure cells of the different DCISL schemes and the exact departure cell for the parameters listed above. For this flow field, the departure cell is a polygon with straight-line walls, and hence the departure area of scheme of RANČIĆ [1992] is exact if exact trajectories are used.

In Table 2.2, the error measures for the degree of local mass conservation are shown for three different DCISL schemes. The trajectory algorithm of LAURITZEN, KAAS and MACHENHAUER [2006] has been used for the computation of the departure points. The deviation from unity of the first error measure (column 1 in Table 2.2) for the MACHENHAUER and OLK [1998] scheme is due to the fact that the departure points are not exact. Had the trajectories been exact, the MACHENHAUER and OLK [1998] scheme would have had the first error measure equal to one. With respect to the first error measure, all DCISL schemes are equally accurate for this particular test case. The cascade schemes are more local than the fully 2D scheme of MACHENHAUER and OLK [1998] in terms of the two remaining error measures. Hence, there is less of the departure cell located outside the exact departure cell and less of the exact departure area that is not included in the schemes departure area. Of the two cascade schemes, the SLICE scheme is most accurate for this particular flow case. Note that the order of the 1D sweeps is reversed in the SLICE scheme compared to the NAIR, SCROGGS and SEMAZZI [2002] scheme.

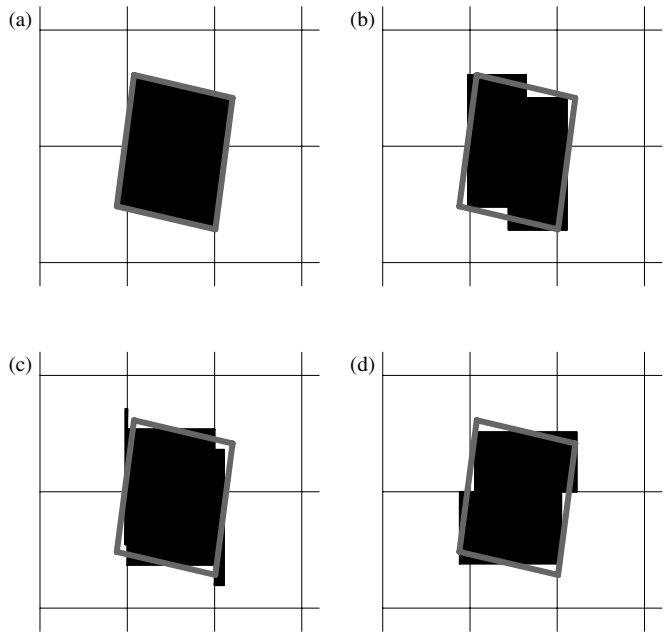


FIG. 2.10 The departure cells (dark area) when using the scheme of (a) RANČIĆ [1992], (b) MACHENHAUER and OLK [1998], (c) the cascade scheme of NAIR, SCROGGS and SEMAZZI [2002] and (d) the cascade scheme of ZERROUKAT, WOOD and STANIFORTH [2002] for the idealized test case for assessing the degree of local mass-conservation. The departure areas are based on the departure points computed with the trajectory scheme of LAURITZEN, KAAS and MACHENHAUER [2006]. The grey lines are the "exact" departure cell walls.

TABLE 2.2
Error measures for the degree of local mass conservation for the schemes of MACHENHAUER and OLK [1998] (CISL), NAIR, SCROGGS and SEMAZZI [2002] (CCS), and ZERROUKAT, WOOD and STANIFORTH [2005] (SLICE) for the analytic flow field described in Section 2

| Scheme | Departure area/(exact departure area) | (Area outside exact departure area)/(exact departure area) | Area mising inside exact departure area)/(exact departure area) |
|---|---|---|---|
| CISL | 1.0009 | 0.1124 | 0.1124 |
| CCS | 1.0009 | 0.0813 | 0.0805 |
| SLICE | 1.0009 | 0.0778 | 0.0769 |

It is important to note that the above example does by no means substitute a general analysis including a statistically large number of departure cells in realistic flows. Therefore, one should not use the analysis to draw general conclusions on the relative accuracy of the three DCISL schemes. For instance, part of the advantage of the NAIR, SCROGGS and SEMAZZI [2002] scheme over the MACHENHAUER and OLK [1998] scheme is, in the case shown, due to the fact that the flow is convergent. Consequently, the departure cell

consists of three rectangles, so there are two "jumps" in the north and south walls, respectively. The MACHENHAUER and OLK [1998] scheme only has one "jump" in the north and south wall. For a divergent flow field, this advantage would no longer appear. Similarly, the direction of the cascade sweeps influences the degree of local mass conservation.

In Section 2.3.2.1, the "locality" of the flux-based scheme of LIN and ROOD [1996] is assessed on the present test case. The results in Fig. 2.11 show the actual areas of information needed to obtain a forecast using that transport scheme and the wind field in Eq. (2.24). It can be seen in this case that the effective departure area is substantially more spread out than those of the DCISL schemes in Fig. 2.10.

*2.3.1.4. Stability analysis* Although the PPM is a widely used numerical method, there has not been performed a Von Neumann stability analysis of that method, as far the

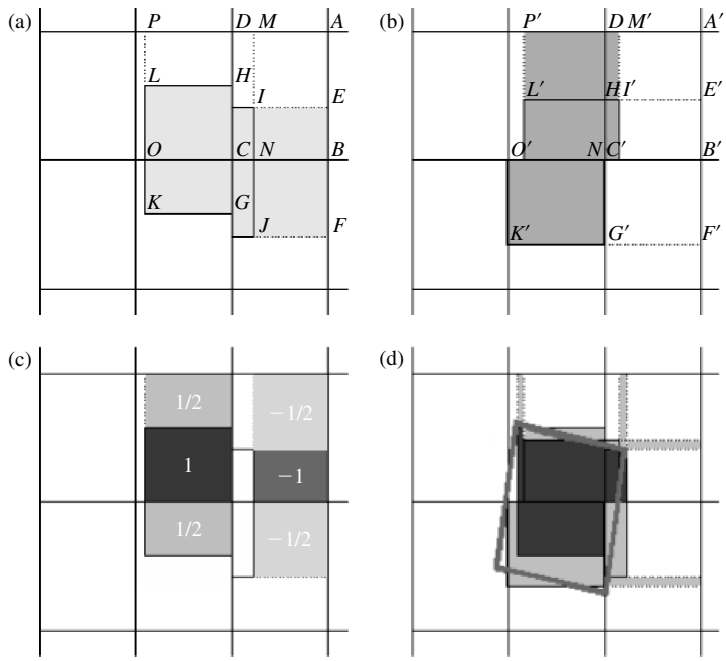


FIG. 2.11 A graphical illustration of the LIN and ROOD [1996] scheme for the idealized test case for assessing the degree of local mass conservation. The arrival cell is the north-eastern most regular grid cell in all plots. The capital letters on (a) and (b) refer to the vertices located south-west of the letter in question except for J' and N' that refer to the vertice to the south-east of the letter in question. The notation $\overline{ABCD}$ will refer to the average value in the cell with vertices at A, B, C, and D. (a) and (c) illustrate $X_C\left[\tfrac{1}{2}\left(\overline{\psi}^n + \overline{\psi}_{AY}\right)\right]$, where $\overline{\psi}_{AY}$ (yellow area) is computed using an advective operator. (a) Following the conceptual illustration of LEONARD, LOCK and MACVEAN [1996], $X_C\left[\tfrac{1}{2}\left(\overline{\psi}^n + \overline{\psi}_{AY}\right)\right]$ is given by $\tfrac{1}{2}\left(\overline{DCOP} + \overline{HGKL}\right) - \tfrac{1}{2}\left(\overline{ABNM} + \overline{EFJI}\right)$. (c) shows the cell averages with weight one (dark blue), half (light blue), minus one (red), and minus half (light red), for the contribution from $X_C$. (b) Similarly for $Y_C$, we get that $Y_C\left[\tfrac{1}{2}\left(\overline{\psi}^n + \overline{\psi}_{AX}\right)\right] = \tfrac{1}{2}\left(\overline{BF'G'C} + \overline{N'J'K'O'}\right) - \tfrac{1}{2}\left(\overline{AE'H'D} + \overline{M'I'L'P'}\right)$ and the green area is $\overline{\psi}_{AX}$. (d) shows the final forecast with the same coloring as in (c). The red rectangle is the exact departure area. (See also color insert).

authors are aware. Rather the stability of the schemes has been demonstrated numerically. Here, a stability analysis of the PPM2 in 1 and 2Ds using DCISL schemes is given and is further detailed in LAURITZEN [2007].

First, consider the 1D situation in which all the DCISL schemes are identical. The stability analysis and the notation used here are similar to that used in the stability analysis of traditional grid-point semi-Lagrangian schemes presented in BATES and MCDONALD [1982]. If we assume a constant flow $u$ (without loss of generality assume $u$ positive), then the west cell wall of cell $i$, located at $x_i = i\Delta x$, departs from

$$(x_i)_* = x_i - u\Delta t. \tag{2.25}$$

Similarly for the right cell wall. Let integer $p$ be such that $(x_i)_*$ is located in between $(i - p - 1)\Delta x$ and $(i - p)\Delta x$, and define

$$\alpha = \frac{u\Delta t}{\Delta x} - p. \tag{2.26}$$

For a constant flow and if a piecewise constant subgrid-cell distribution is used, then the forecast is given by

$$\overline{\psi}_i^+ = (1 - \alpha)\overline{\psi}_{i-p} + \alpha\overline{\psi}_{i-p-1}, \tag{2.27}$$

which is identical to the traditional semi-Lagrangian grid-point scheme using linear Lagrange interpolation under the assumption that the grid-point values represent cell averages. Assume a solution in the form

$$\psi_i = \psi^0\Gamma^n \exp\left(\hat{i}kx\right), \tag{2.28}$$

where $\Gamma$ is the complex amplification factor, $\psi^0$ is the initial amplitude, $k$ is the wave number, and $\hat{i}$ is the imaginary unit. Since a cell-integrated scheme is based on cell averages,

$$\overline{\psi}_i = \int_{i\Delta x}^{(i+1)\Delta x} \psi^0\Gamma^n \exp\left(\hat{i}kx\right)\mathrm{d}x \tag{2.29}$$

must be evaluated and the resulting expression substituted into Eq. (2.28). It may easily be shown that the squared modulus of the amplification factor can be written as

$$|\Gamma|^2 = 1 - 2\alpha(1 - \alpha)(1 - \cos k\Delta x). \tag{2.30}$$

This is the same result as would have been obtained for a traditional semi-Lagrangian grid-point scheme using linear Lagrange interpolation (see, e.g., BATES and MCDONALD [1982]). For all resolvable wavelengths, the scheme is stable $|\Gamma|^2 \leq 1$ as long as $0 \leq \alpha \leq 1$. By definition, $\alpha$ is within that range, and hence the 1D DCISL scheme using a piecewise constant subgrid-scale reconstruction is unconditionally stable.

Using the PPM2 for the subgrid-scale distribution with no filters and assuming a constant wind field, the forecast can be written as

$$\begin{aligned}\overline{\psi}_i^+ = &\frac{1}{12}\alpha^2(1-\alpha)\overline{\psi}_{i-p-3} - \frac{1}{12}\alpha(1+7\alpha)(1-\alpha)\overline{\psi}_{i-p-2}\\ &- \frac{1}{3}\alpha\Big(4\alpha^2-5\alpha-2\Big)\overline{\psi}_{i-p-1} - \frac{1}{3}(1-\alpha)\Big(4\alpha^2-3\alpha-3\Big)\overline{\psi}_{i-p}\\ &- \frac{1}{12}\alpha(1-\alpha)(8-7\alpha)\overline{\psi}_{i-p+1} + \frac{1}{12}\alpha(1-\alpha)^2\,\overline{\psi}_{i-p+2}.\end{aligned} \tag{2.31}$$

By performing a Von Neuman stability analysis, it may be shown, after some algebra, that the squared modulus of the amplification factor can be written as

$$\begin{aligned}|\Gamma|^2 = 1 &+ \frac{2}{9}\alpha^2\Big(4\alpha^2-4\alpha-7\Big)\,(1-\alpha)^2 - \frac{8}{9}\alpha^2\Big(4\alpha^2-4\alpha-5\Big)(1-\alpha)^2\cos k\Delta x\\ &+ \frac{1}{9}\alpha^2\Big(50\alpha^2-50\alpha-39\Big)(1-\alpha)^2\cos^2 k\Delta x\\ &- \frac{2}{9}\alpha^2\Big(19\alpha^2-19\alpha-7\Big)(1-\alpha)^2\cos^3 k\Delta x\\ &+ \frac{1}{9}\alpha^2\Big(14\alpha^2-14\alpha-1\Big)(1-\alpha)^2\cos^4 k\Delta x + \frac{2}{9}\alpha^3(1-\alpha)^3\cos^5 k\Delta x.\end{aligned} \tag{2.32}$$

Since DCISL schemes approximate the integral over the departure area explicitly, the amplification factors in Eq. (2.31) and Eq. (2.32) are not a function of $p$. Figure 2.12 shows $|\Gamma|^2$ for the four shortest wavelengths for the DCISL scheme using piecewise constant subgrid-cell reconstructions and PPM2, and for comparison, the squared modulus of the amplification factor for the traditional semi-Lagrangian scheme using cubic interpolation. Apart from the $2\Delta x$-wave, the higher order subgrid-cell reconstruction leads to a much less damping scheme compared with the lowest order scheme (as expected). The shortest resolvable wave ($2\Delta x$-wave) is, however, severely damped with the DCISL scheme based on PPM2, which might explain why schemes based on PPM2 do not exhibit excessive noise problems near sharp gradients even without applying filters. It can be demonstrated numerically that the scheme is unconditionally stable for all wavelengths when $0 \leq \alpha \leq 1$ (which is satisfied by definition).

The above analysis is directly extended to 2Ds. Assume a constant flow field $(u, v)$ where the velocity components are positive, let $p$ and $q$ be integers such that the southwest vertice of cell $(i, j)$ is located in the Eulerian cell with indices $(i - p - 1, j - r - 1)$ (see Fig. 7 in Bates and Mcdonald [1982]), $\alpha$ is defined in Eq. (2.26) and

$$\beta = \frac{v\Delta t}{\Delta y} - r. \tag{2.33}$$

Here, only the fully 2D schemes of Machenhauer and Olk [1998] and the cascade scheme of Nair, Scroggs and Semazzi [2002] are considered. Note, however, that for a constant flow field the cascade scheme of Nair, Scroggs and Semazzi [2002] and

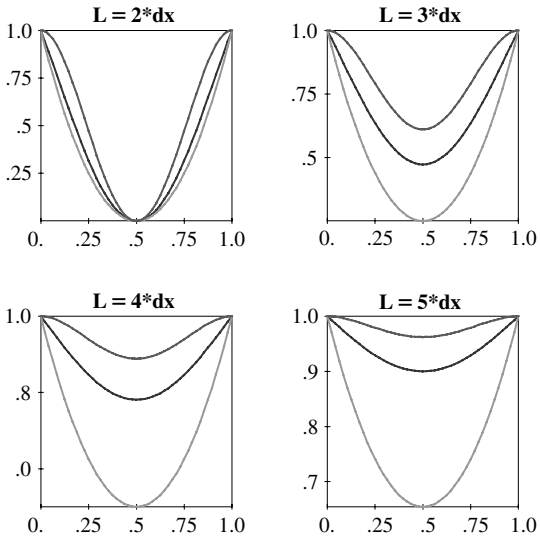


FIG. 2.12 Squared modulus of the amplification factor as a function of $\alpha$ for the (a) $2\Delta x$, (b) $3\Delta x$, (c) $4\Delta x$, and (d) $5\Delta x$ waves. Red and green lines are for the DCISL scheme using PPM2 and piecewise constant subgrid-cell representation, respectively. For comparison, the squared modulus of the amplification factor for the traditional semi-Lagrangian scheme based on cubic Lagrange interpolation (blue line) is shown as well. (See also color insert).

ZERROUKAT, WOOD and STANIFORTH [2002] are identical, apart from the order of the polynomial used for the subgrid-scale reconstructions. When using piecewise constant subgrid-scale reconstructions, the explicit forecast formula for all DCISL schemes is given by

$$\begin{aligned}\overline{\psi}^{+}_{i,j} = {} & (1-\alpha)(1-\beta)\overline{\psi}_{i-p,j-r} + \alpha(1-\beta)\overline{\psi}_{i-p-1,j-r} \\ & + \beta(1-\alpha)\overline{\psi}_{i-p,j-r-1} + \alpha\beta\overline{\psi}_{i-p-1,j-r-1}.\end{aligned} \tag{2.34}$$

Again, the formula is equivalent to the forecast for the traditional semi-Lagrangian scheme using bilinear Lagrange interpolation under the assumption that grid-point values represent cell averages (see, e.g., BATES and MCDONALD [1982]). Assuming a solution in the form

$$\psi_{i,j} = \psi^0 \Gamma^n \exp\left[\hat{i}\,(kx + ly)\right], \tag{2.35}$$

where $k$ and $l$ are the components of the wave number vector, then the mean value of the solution over cell $(i, j)$ is given by

$$\overline{\psi}_{i,j} = \int_{j\Delta y}^{(j+1)\Delta y} \left\{ \int_{i\Delta x}^{(i+1)\Delta x} \psi^0 \Gamma^n \exp\left[\hat{i}\,(kx + ly)\right] \mathrm{d}x \right\} \mathrm{d}y. \tag{2.36}$$

Substituting Eq. (2.36) into Eq. (2.34), the complex amplification factor can be written as

$$\Gamma = \left\{1 - \alpha\left[1 - \exp\left(-\hat{i}k\Delta x\right)\right]\right\}\left\{1 - \beta\left[1 - \exp\left(-\hat{i}l\Delta y\right)\right]\right\}\exp\left[-\hat{i}\,(p\Delta x + r\Delta y)\right]. \tag{2.37}$$

It may be easily verified that $|\Gamma|^2 \leq 1$ for $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$, hence the 2D scheme using piecewise constant subgrid-cell reconstructions is unconditionally stable.

For the higher order schemes, Maple software has been used to compute the explicit forecast formulas and for performing the stability analysis. The explicit formula for the forecast when using the scheme of Machenhauer and Olk [1998] and Nair, Scroggs and Semazzi [2002] can be written as a weighted sum

$$\overline{\psi}^{+}_{i,j} = \sum_{h=-2}^{3}\sum_{g=-2}^{3} C_{h,g}(\alpha,\beta)\overline{\psi}^{n}_{i-p-1+h,\,j-r-1+g} \tag{2.38}$$

The coefficients are listed in Tables 2.3 and 2.4 for the Machenhauer and Olk [1998] and Nair, Scroggs and Semazzi [2002] schemes, respectively. The formula for the squared modulus of the amplification factor resulting from the Von Neumann stability analysis is too lengthy to display here. Instead, plots of $|\Gamma|^2$ for selected wave numbers are shown (see Fig. 2.13). See Lauritzen [2007] for further details. The cascade scheme of Nair, Scroggs and Semazzi [2002] is slightly more damping than the fully 2D scheme of Machenhauer and Olk [1998] for the shortest resolvable traverse wave (Fig. 2.12(a)), and the situation is vice versa for the next shortest resolvable traverse wave (Fig. 2.12(b)) as well as longer wavelengths. It has been verified numerically that the squared modulus of the amplification factor is less than or equal to unity for $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$, i.e., the 2D DCISL schemes are unconditionally stable and less diffusive compared to a traditional semi-Lagrangian scheme based on bi cubic interpolation.

*2.3.1.5. Extension of DCISL schemes to spherical geometry* The singularities on the sphere are one of the main challenges for transport schemes formulated on conventional latitude-longitude grids, and most algorithms require a certain amount of "engineering" to tackle the pole problem, which often reduces the efficiency and simplicity of the algorithms. As already mentioned, the number of schemes developed in cartesian geometry is significantly larger than the number of schemes formulated for spherical geometry. For the DCISL schemes discussed here, only the scheme of Machenhauer and Olk [1998], Nair, Scroggs and Semazzi [2002], and Zerroukat, Wood and Staniforth [2002] have been extended to the sphere.

In cartesian geometry, the most accurate approximation to a departure cell, given the departure points, is the polygon resulting from connecting the departure points with straight lines. Similarly, in spherical geometry, the cells defined by connecting the departure points with great circle arcs seem as the optimal choice. But as in cartesian geometry, integrating along the optimal curves leads to complicated and computationally expensive algorithms. Therefore, as in the cartesian case, the area approximation must be simplified.

TABLE 2.3

The coefficients $C_{h,g}$ written in "matrix" format for the explicit forecast formula in case of a constant wind field when using the scheme of MACHENHAUER and OLK [1998] (see Eq. (2.38)). The index $h$ is in the first column and the second index $g$ is in the first row

| $h, g$ | $-2$ | $-1$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| $-2$ | 0 | 0 | $\frac{1}{12}\alpha^2\beta(1-\alpha)$ | $\frac{1}{12}\alpha^2(1-\alpha)(1-\beta)$ | 0 | 0 |
| $-1$ | 0 | 0 | $-\frac{1}{12}\alpha\beta(1-\alpha)(1+7\alpha)$ | $-\frac{1}{12}\alpha(1-\alpha)\times$ $(1-\beta)(1+7\alpha)$ | 0 | 0 |
| 0 | $\frac{1}{12}\alpha\beta^2(1-\beta)$ | $-\frac{1}{12}\alpha\beta(1-\beta)\times$ $(1+7\beta)$ | $-\frac{1}{3}\alpha\beta\times$ $\left(4\alpha^2-5\alpha+4\beta^2-5\beta-1\right)$ | $-\frac{1}{3}\alpha\,(1-\beta)\times$ $\left(4\alpha^2-5\alpha+4\beta^2-3\beta-2\right)$ | $-\frac{1}{12}\alpha\beta\times$ $(1-\beta)\,(8-7\beta)$ | $\frac{1}{12}\alpha\beta\,(1-\beta)^2$ |
| 1 | $\frac{1}{12}\beta^2\,(1-\alpha)\,(1-\beta)$ | $-\frac{1}{12}\beta\,(1-\beta)\times$ $(1-\alpha)\,(1+7\beta)$ | $-\frac{1}{3}\beta\,(1-\alpha)\times$ $\left(4\alpha^2-3\alpha+4\beta^2-5\beta-2\right)$ | $-\frac{1}{3}\,(1-\alpha)\,(1-\beta)\times$ $\left(4\alpha^2-3\alpha+4\beta^2-3\beta-3\right)$ | $-\frac{1}{12}\beta\,(1-\alpha)\times$ $(1-\beta)\,(8-7\beta)$ | $\frac{1}{12}\beta\,(1-\alpha)\,(1-\beta)^2$ |
| 2 | 0 | 0 | $-\frac{1}{12}\alpha\beta\,(1-\alpha)\,(8-7\alpha)$ | $-\frac{1}{12}\alpha\,(1-\alpha)\times$ $(1-\beta)\,(8-7\alpha)$ | 0 | 0 |
| 3 | 0 | 0 | $\frac{1}{12}\alpha\beta\,(1-\alpha)^2$ | $\frac{1}{12}\alpha\,(1-\alpha)^2\,(1-\beta)$ | 0 | 0 |

TABLE 2.4

The coefficients $C_{h,g}$ written in "matrix" format for the explicit forecast formula in case of a constant wind field when using the scheme of NAIR, SCROGGS and SEMAZZI [2002] (see Eq. (2.38)). The index $h$ is in the first column and the second index $g$ is in the first row.

| | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|---|
| $-2$ | $\frac{1}{144}\alpha^2\beta^2(1-\alpha)\times$ $(1-\beta)$ | $-\frac{1}{144}\alpha^2\beta(1-\alpha)\times$ $(1-\beta)\,(1+7\beta)$ | $-\frac{1}{36}\alpha^2\beta(1-\alpha)\times$ $\left(4\beta^2-5\beta-2\right)$ | $-\frac{1}{36}\alpha^2(1-\alpha)(1-\beta)\times$ $\left(4\beta^2-3\beta-3\right)$ | $-\frac{1}{144}\alpha^2\beta(1-\alpha)\times$ $(1-\beta)\,(8-7\beta)$ | $\frac{1}{144}\alpha^2\beta(1-\alpha)\times$ $(1-\beta)^2$ |
| $-1$ | $-\frac{1}{144}\alpha\beta^2(1-\alpha)\times$ $(1-\beta)(1+7\alpha)$ | $\frac{1}{144}\alpha\beta(1-\alpha)\times$ $(1-\beta)(1+7\alpha)\times$ $(1+7\beta)$ | $\frac{1}{36}\alpha\beta(1-\alpha)\times$ $(1+7\alpha)\times$ $\left(4\beta^2-5\beta-2\right)$ | $\frac{1}{36}\alpha(1-\alpha)\times$ $(1-\beta)(1+7\alpha)\times$ $\left(4\beta^2-3\beta-3\right)$ | $\frac{1}{144}\alpha\beta(1-\alpha)\times$ $(1-\beta)\,(8-7\beta)\times$ $(1+7\alpha)$ | $-\frac{1}{144}\alpha\beta(1-\alpha)\times$ $(1-\beta)^2(1+7\alpha)$ |
| $0$ | $-\frac{1}{36}\alpha\beta^2(1-\beta)\times$ $\left(4\alpha^2-5\alpha-2\right)$ | $\frac{1}{36}\alpha\beta(1-\beta)\times$ $\left(4\alpha^2-5\alpha-2\right)\times$ $(1+7\beta)$ | $\frac{1}{9}\alpha\beta\times$ $\left(4\alpha^2-5\alpha-2\right)\times$ $\left(4\beta^2-5\beta-2\right)$ | $\frac{1}{9}\alpha(1-\beta)\times$ $\left(4\alpha^2-5\alpha-2\right)\times$ $\left(4\beta^2-3\beta-3\right)$ | $\frac{1}{36}\alpha\beta(1-\beta)\times$ $\left(4\alpha^2-5\alpha-2\right)\times$ $(8-7\beta)$ | $-\frac{1}{36}\alpha\beta(1-\beta)^2\times$ $\left(4\alpha^2-5\alpha-2\right)$ |
| $1$ | $-\frac{1}{36}\beta^2(1-\alpha)\times$ $(1-\beta)\times$ $\left(4\alpha^2-3\alpha-3\right)$ | $\frac{1}{36}\beta(1-\alpha)(1-\beta)\times$ $(1+7\beta)\times$ $\left(4\alpha^2-3\alpha-3\right)$ | $\frac{1}{9}\beta(1-\alpha)\times$ $\left(4\alpha^2-3\alpha-3\right)\times$ $\left(4\beta^2-5\beta-2\right)$ | $\frac{1}{9}(1-\alpha)(1-\beta)\times$ $\left(4\alpha^2-3\alpha-3\right)\times$ $\left(4\beta^2-3\beta-3\right)$ | $\frac{1}{36}\beta(1-\alpha)(1-\beta)\times$ $\left(4\alpha^2-3\alpha-3\right)\times$ $(8-7\beta)$ | $-\frac{1}{36}\beta(1-\alpha)\times$ $(1-\beta)^2\times$ $\left(4\alpha^2-3\alpha-3\right)$ |
| | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
| $-2$ | $\frac{1}{144}\alpha^2\beta^2(1-\alpha)\times$ $(1-\beta)$ | $-\frac{1}{144}\alpha^2\beta(1-\alpha)\times$ $(1-\beta)\,(1+7\beta)$ | $-\frac{1}{36}\alpha^2\beta(1-\alpha)\times$ $\left(4\beta^2-5\beta-2\right)$ | $-\frac{1}{36}\alpha^2(1-\alpha)(1-\beta)\times$ $\left(4\beta^2-3\beta-3\right)$ | $-\frac{1}{144}\alpha^2\beta(1-\alpha)\times$ $(1-\beta)\,(8-7\beta)$ | $\frac{1}{144}\alpha^2\beta(1-\alpha)\times$ $(1-\beta)^2$ |

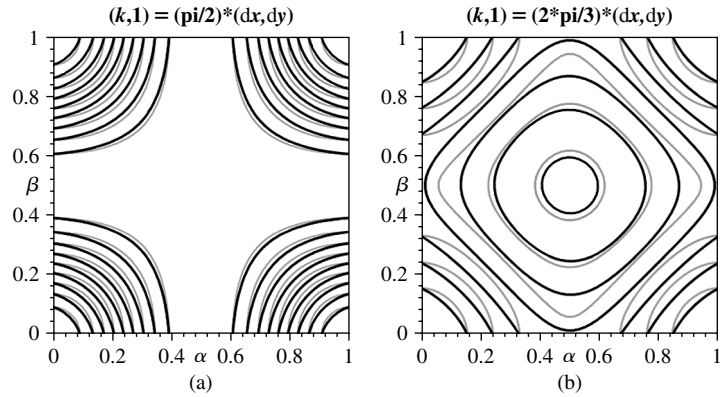


FIG. 2.13 Squared modulus of the amplification factor as a function of $(\alpha, \beta)$ for (a) $(L_x, L_y) = 2\,(\Delta x, \Delta y)$ and (b) $(L_x, L_y) = 3\,(\Delta x, \Delta y)$, where $L_x$ is the wavelength in the $x$-direction and similarly for $L_y$. Black contours are for the scheme of MACHENHAUER and OLK [1998] and grey contours the scheme of NAIR, SCROGGS and SEMAZZI [2002]. The contour-interval is 0.1 and contours start at 0.9 at the corners and decrease toward the center of the plot. The two schemes show similar damping properties.

The MACHENHAUER and OLK [1998] scheme is extended to spherical geometry by using the $\mu$-grid originally introduced by MACHENHAUER and OLK [1996] (NAIR and MACHENHAUER [2002]). The $\mu$-grid is a latitude-longitude grid in which the latitude $\theta$ is replaced by $\mu = \sin(\theta)$ (see Fig. 2.14). This transformation is area preserving, and the $\mu$-grid is essentially a cartesian grid where the latitude grid lines are no longer equidistant. The departure cells are defined as quadrilaterals on the $(\lambda, \mu)$-plane exactly as in cartesian geometry, i.e., the cell walls that in cartesian geometry were parallel to $x$ and $y$ isolines are parallel to the longitudes and latitudes on the $\mu$-grid, respectively. Hence, away from the poles, this transformation is invariant in the sense that the corresponding upstream integrals and departure cells take exactly the same form as in cartesian geometry. Since the algorithm is formally equivalent on the $\mu$-grid, only minor modifications of the algorithm in cartesian geometry are needed away from the poles. In the vicinity of the poles, however, approximating the cells with straight-line walls on the $\mu$-grid is a poor approximation of the cells on the spherical latitude-longitude grid (see, e.g., Fig. 4c in NAIR and MACHENHAUER [2002]). Especially the exact north and south walls are deviating significantly from linearity on the $(\lambda, \mu)$-grid, and hence some "engineering" is needed. Local tangent planes at the poles are introduced for more accurate cell-approximations. The areas in which the tangent planes are used are referred to as the *polar caps*. Ideally, the integration should be performed along straight lines on the tangent planes. Instead, more latitudes are introduced in the polar caps, and the coordinates of the cell vertices on the tangent plane are transformed into $(\lambda, \mu)$-coordinates, and thereafter, the integrals are performed along straight lines in the $(\lambda, \mu)$ plane. In the Lagrangian belt containing the pole point (referred to as the *singular belt*), the algorithm breaks down since the Lagrangian cell containing the Eulerian pole is not well defined (see Fig. 2.14). The total mass inside the singular belt can, however, be computed and is distributed among the cells in a mass conservative way using a regular semi-Lagrangian

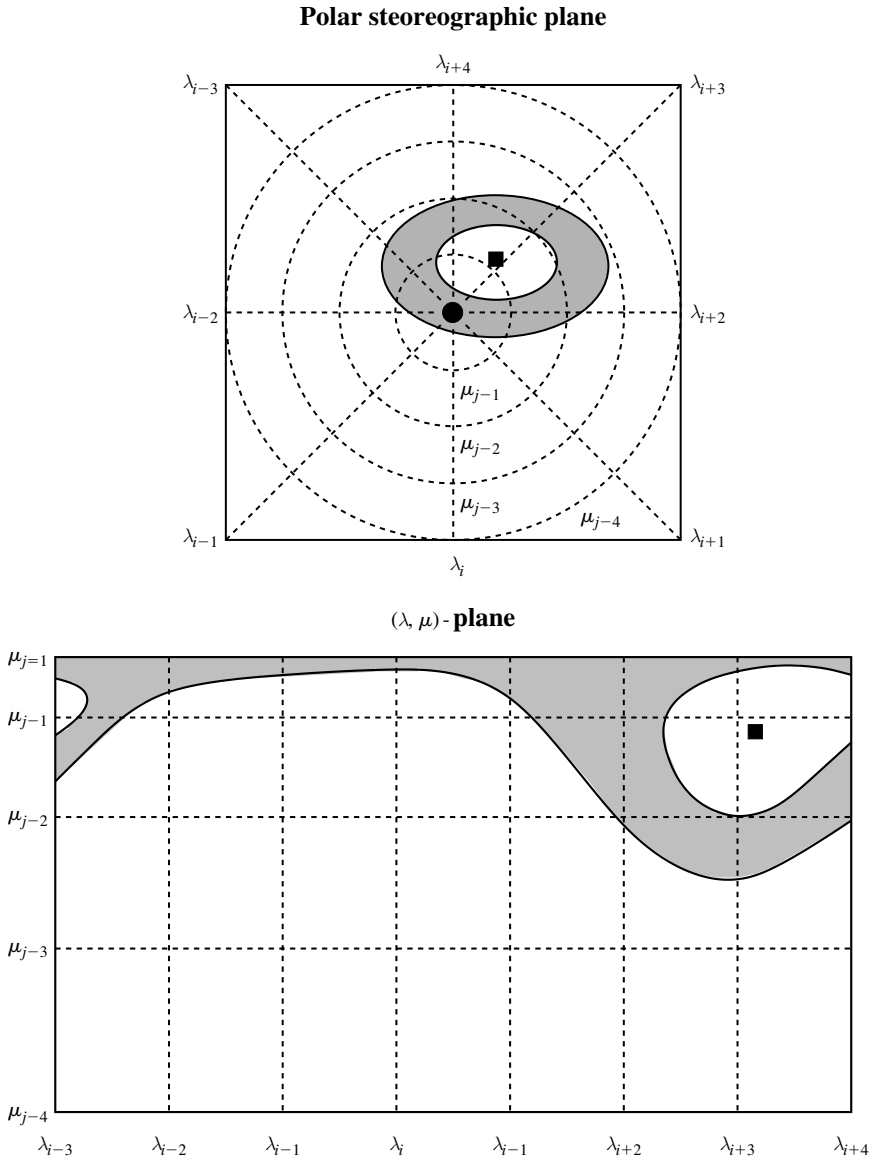


FIG. 2.14 A graphical illustration of the polar cap treatment in the scheme of NAIR and MACHENHAUER [2002]. The upper plot shows the polar stereographic projection of the Eulerian cells (bounded by the dashed lines which are the $\lambda$ and $\mu$ isolines), and the singular belt (shaded region). The singular belt is the set of departure cells bounded by two consecutive Lagrangian latitudes that contain the Eulerian pole point (filled circle). On the lower plot, the Eulerian cells and the singular belt are plotted on the $(\lambda, \mu)$-plane. Note that the pole point (filled circle on upper plot) is the line $\mu_j = 1$ on the lower plot. The filled square is the Lagrangian pole point.

method. The method computes the densities at the approximate departure cell centers using a quasi-bicubic interpolation. These values are used as weights for distributing the total mass in the singular belt among the cells, i.e., the point value of the density at a given departure cell center, normalized by the sum of all the departure cell point values, determines the fraction of the total mass which the cell in question is attributed (see NAIR and MACHENHAUER [2002] for additional details).

The cascade scheme of NAIR, SCROGGS and SEMAZZI [2002] has been extended to the sphere using the $\mu$ grid as well but with two different treatments of the polar cap. The first method used cascade interpolation throughout the spherical domain except for the Lagrangian belts over the Eulerian poles, but the scheme was limited by the meridional Courant number, which must be less than unity in the version presented in NAIR, SCROGGS and SEMAZZI [2002]. Later, the scheme was adapted to large meridional Courant numbers by using the cascade approach away from the polar caps and by using the fully 2D scheme of NAIR and MACHENHAUER [2002] over the polar caps (NAIR [2004]).

Away from the *poles*, the extension of the cascade remapping method from cartesian geometry to the $\mu$-grid is straightforward. The computation of the intermediate grid, or equivalently, the crossings of the Lagrangian latitudes and Eulerian longitudes are computed using cubic Lagrange interpolation on the $\mu$-grid. The mass is transferred to the intermediate grid and from there to the Lagrangian grid, exactly as in the cartesian case, but simply on the $\mu$-grid. Since the meridional Courant number is less than unity, the only problematic zone with ill-defined cells (singular belt) is made up by the cells north of the first Lagrangian latitude that ends up at the first Eulerian latitude circle after one time-step (similarly for the Southern hemisphere). As in the case of the NAIR and MACHENHAUER [2002] scheme, the total mass in the singular belt can be computed and the total mass can be redistributed to the individual cells as explained above. Contrary to the NAIR and MACHENHAUER [2002] scheme, the cascade scheme does not use high-resolution polar belts. Only the singular belts are treated differently from the rest of the domain.

For general applications, this restriction on the meridional Courant number is a severe limitation. NAIR [2004] suggested the use of the efficient cascade method of NAIR, SCROGGS and SEMAZZI [2002] away from the poles and the polar cap treatment of NAIR and MACHENHAUER [2002] in the zones where the cascade method would break down (north of the Lagrangian latitude closest to the Lagrangian pole point and similarly for the Southern hemisphere). Hereby the severe meridional Courant number restriction is alleviated.

The SLICE scheme is extended to spherical geometry by using a regular latitude-longitude grid (ZERROUKAT, WOOD and STANIFORTH [2004]). The intermediate grid is computed in spherical coordinates by using the great circle approach of NAIR, CÔTÉ and STANIFORTH [1999b] (for details see Section 2b in this reference), which is more efficient though less accurate than the cubic Lagrange interpolation on the $\mu$-grid used in NAIR, CÔTE and STANIFORTH [2002]. The cascade method breaks down when not all Lagrangian longitudes intersect an Eulerian latitude. Consequently, there are some intermediate cells that are ill defined, i.e., the intermediate cell walls are not both well

defined. Consider the situation in which the western intermediate cell wall exists but not the eastern one, while the east cell wall of intermediate cell $(i + 1, j)$ is well defined. After the first cascade sweep, the total mass between the west wall of intermediate cell $(i, j)$ and the east wall of intermediate cell $(i + 1, j)$ is known. The total mass is split in two and allocated to the nearest Lagrangian mass centers used in the second cascade sweep. This is not accurate but the redistribution is mass conservative. Some of the lost accuracy is recovered with a "post fix" procedure similar to the singular belt treatment in NAIR and MACHENHAUER [2002]. Here, however, the mass in the entire polar cap is redistributed mass-conservatively using Lagrange weights.

For solid body rotation over the poles, the error measures shown in Table 2.1 do not indicate a superior method for treating the pole problem. Even though SLICE uses a subgrid-reconstruction that, under the assumption that no filters are invoked, is one order of magnitude higher than PPM2 used in the scheme of NAIR and MACHENHAUER [2002] and NAIR, SCROGGS and SEMAZZI [2002], SLICE is not superior with respect to error measures $l_1$ and $l_2$. This also suggests that the polar treatment reduces the accuracy of the scheme.

### *2.3.2. Flux-based FV schemes in 2Ds*

Up to the beginning of the 1990s, when RANČIĆ [1992] presented the first semi-Lagrangian DCISL scheme, all FV schemes used in meteorology were flux-based in nature.[2] In flux-based schemes, the prognostic equation for the volume-specific scalar $\psi$ is obtained as a sum of estimates of inward and outward fluxes in the Eulerian grid cell. Generally assuming an Eulerian grid cell to be a polygon with $L$ faces, the differential FV prognostic equation (1.6) for the total "mass" for this particular cell can be written as follows:

$$\overline{\psi}^{n+1} \Delta A = \overline{\psi}^{n} \Delta A + \sum\nolimits_{l=1}^{L} m_l^L, \tag{2.39}$$

where $\overline{\psi}$ is the cell average "density," $n$ is the time step index, $\Delta A$ is the area, and $m_l^L$ is the total inward mass flux integrated over a time step for face $l$. $m_l^L$ is defined negative if the net flow through face $l$ is outward and positive for inward fluxes. The conservation is ensured if $m_l^L$ is unique for face $l$, i.e., the mass that leaves a cell through face $l$ is exactly gained in the neighboring cells sharing face $l$.

As pointed out by HIRSCH [1990], FV schemes of the type in Eq. (2.39) were introduced by GODUNOV [1959], and they were first used in meteorological applications by CROWLEY [1968]. Since then, the schemes have gradually evolved with increasing sophistication, and they have been used extensively in recent decades in both meteorology and oceanography. We will not go into great detail regarding the entire historical development of flux-based FV methods and their application on the sphere. Instead the focus is on some aspects that are important for understanding the methodology and how it has evolved into the most modern schemes.

To introduce the basic ideas behind flux-based FV schemes, consider at first the continuity equation in 1D, $x$, without any source terms. Define the flux convergence

[2]The process of integrating over the departure cell, or equivalently, the remapping or rezoning of mass between two grids, was, however, studied already in the 1970's (e.g., HIRT, AMSDEN and COOK [1974]).

operator

$$X_C\left(u, \Delta t, \Delta x; \overline{\psi}^n\right) = \left[F_w^n + F_e^n\right]/\Delta x \tag{2.40}$$

for a given Eulerian cell with extension $\Delta x$. Indices $w$ and $e$ indicate the left ("western") and right ("eastern") cell boundary, and $F$ is the time-integrated mass flux related to the flow speed $u$, i.e., $F$ corresponds to the flux "$m$" in the multidimensional case (Eq. (2.39)). Since the total fluxes depend on the flow speed, the time-step $\Delta t$, and $\Delta x$, this is also the case for $X_C$. Implicitly $X_C$ also depends on the Eulerian cell averages since they are used for the reconstruction of the subgrid-cell distributions. For schemes where also higher order moments or cell face values are prognostic variables (e.g., PRATHER [1986], XIAO and YABE [2002]), the subgrid-cell representation also depends on these moments or values. In the limit as $\Delta t$ and $\Delta x$ approach zero, the operator $X_C$ divided by $\Delta t$ is the FV approximation to the term $-\partial(u\psi)/\partial x$. Expressed in terms of the flux convergence operator, the 1D version of Eq. (2.39) becomes

$$\overline{\psi}_i^{n+1} = \overline{\psi}_i^n + X_{i_C}(\overline{\psi}^n) \tag{2.41}$$

for the Eulerian grid cell $i$, omitting the obvious dependence on $u$, $\Delta t$, and $\Delta x$ for brevity. The operator $X_C$ redistributes mass between the Eulerian grid cells. By definition, application of the $X_C$ operator does not change the total mass in the integration domain since the left-flux of cell $i$ always cancels the right-flux of cell $i-1$ ($F_{i-1_e} = -F_{i_w}$). In all FV schemes, the fluxes are obtained as integrals – or as approximations to integrals – of the total mass in the length interval being "swept though" the face within the time-step $\Delta t$. This leads to the following "general" equation for $X_C$:

$$X_C\left(\overline{\psi}^n\right) = \frac{1}{\Delta x}\left[\int_{x_w - u_w^* \Delta t}^{x_w} \psi^n \mathrm{d}x - \int_{x_e - u_e^* \Delta t}^{x_e} \psi^n \mathrm{d}x\right]$$
$$\left(= \frac{1}{\Delta x}\left[F_w + F_e\right]\right), \tag{2.42}$$

where $u_w^*$ is the effective advection speed for the area (interval) 'swept though' the left face within the time-step from $n\Delta t$ to $(n+1)\Delta t$. Similarly $u_e^*$ is for the right face. Note, that the integrant $\psi^n$ in (2.42) in each grid cell (i) is an analytic function of $x$ determined as a constant, the cell mean value $\overline{\psi}_i^n$, plus a subgrid-scale deviation from this mean value. The deviation inside grid cell ($i$) is determined by $\overline{\psi}_i^n$ and cell mean values of surrounding cells, in the case of a parabolic representation, of two cells on each side. As mentioned in Section 2.2, the accuracy of FV schemes will depend on how accurate the fluxes are estimated. Assuming exact effective advective speeds, the accuracy of the scheme will therefore depend only on the order of the $\psi$ subcell representation. To maintain mass conservation, obviously $u_{i-1e}^*$ must equal $u_{iw}^*$ for any pair $(i-1, i)$ of grid cells. For flux-based transport schemes, the advective speed used to estimate the mass flux $F_l^n$ is the flow speed at the spatial location of face $l$. Ideally this flow speed should be a time mean value over $\Delta t$. In practice it is evaluated as a local forecast or a simple

extrapolation valid at time $(n + ½)\Delta t$. For a 1D flow with no spatial variations and where Eq. (2.42) is used to determine $X_C$, Eq. (2.41) becomes identical to a 1D DCISL scheme. The reader is referred to Section 1 for a general description of the analogy between DCISL and flux-based schemes. One of the main differences between the flux-based FV schemes and the DCISL schemes is that the departure points is identified from trajectories (i.e., characteristics) estimated iteratively in both time and space to better include the effect of spatial as well as temporal variations in the flow. A similar accurate estimation of the true departure area is generally not part of the flux-based schemes.

The generalization of the 1D FV schemes to 2Ds can be done in two fundamentally different ways: via a more or less direct estimation of the 2D-fluxes – to be described below in Section 2.3.2.3 – or via *operator splitting*. In the operator splitting method Section 2.3.2.1 – the transport problem is split into a combination of operators in each of the two coordinate directions.

*2.3.2.1. Operator splitting* Consider for simplicity only cartesian $x$-$y$ coordinates. In this case, the individual conservative flux convergence operators are in the $x$-direction (Eq. (2.40)) and in the $y$-direction,

$$Y_C\left(v, \Delta t, \Delta y; \overline{\psi}^n\right) = \left[G_s^n + G_n^n\right]/\Delta y, \tag{2.43}$$

where $v$ is the spatially varying speed in the $y$-direction, $\Delta y$ is the grid extension in the $y$-direction while $s$ and $n$ denote the lower ("southern") and upper ("northern") face of the grid cell. $G$ denotes the fluxes related solely to the translations in the $y$-direction.

A simple-minded operator splitting is where the fluxes in each direction are treated independently as simultaneous 1D fluxes:

$$\overline{\psi}_{i,j}^{n+1} = \overline{\psi}_{i,j}^{n} + X_{i,jC}\left(\overline{\psi}^n\right) + Y_{i,jC}\left(\overline{\psi}^n\right), \tag{2.44}$$

where $i$ and $j$ is the spatial index in the $x$- and $y$-direction, respectively. This scheme is inherently mass-conserving, and as noted by Leonard, Lock and Macvean [1996], it is also stable when the flux is calculated using a first-order – or so-called donor cell – method. However, as shown by Leith [1965], a scheme of a type in Eq. (2.44) is unstable when second-order polynomials are used for the subgrid-cell representation in Eq. (2.42) and in the corresponding expression for flux convergence in the $y$-direction,

$$Y_C\left(\overline{\psi}^n\right) = \frac{1}{\Delta y}\left[\int_{y_s - v_s^* \Delta t}^{y_s} \psi^n \mathrm{d}y - \int_{y_n - v_n^* \Delta t}^{y_n} \psi^n \mathrm{d}y\right]$$
$$\left(= \frac{1}{\Delta y}\left[G_s + G_n\right]\right), \tag{2.45}$$

where $v^*$ is the effective advection speed in the $y$-direction.

It is not surprising that the simple-minded update becomes unstable: the effective departure area being split into two separate areas in the upstream $x$ and $y$ directions. What is needed to achieve stability is a sequential flux splitting instead of the simultaneous flux

splitting in (2.44). This means that the transport problem is first solved in one coordinate direction and the resulting field is subsequently transported in the transverse coordinate direction. Thereby the transport through a certain cell wall is determined from both the velocity parallel to the cell wall, by the first transport step, and from the velocity normal to the cell wall, by the second transport step. As a result an unbroken and more realistic departure area is obtained. Inspired by the notation in LEONARD, LOCK and MACVEAN [1996], consider first the intermediate transport problem in the $x$-direction (omitting for simplicity the grid cell indexing):

$$\overline{\psi}_{CX} = \overline{\psi}^{n} + X_C\left(\overline{\psi}^{n}\right). \tag{2.46}$$

By definition, the total mass is conserved after this intermediate forecast. The subsequent second update step in the $y$-direction becomes

$$\overline{\psi}^{n+1}_{CXY} = \overline{\psi}_{CX} + Y_C\left(\overline{\psi}_{CX}\right), \tag{2.47}$$

or equivalently (by inserting Eq. (2.46) into Eq. (2.47)):

$$\overline{\psi}^{n+1}_{CXY} = \overline{\psi}^{n} + X_C\left(\overline{\psi}^{n}\right) + Y_C\left(\overline{\psi}_{CX}\right). \tag{2.48}$$

Here the argument $\overline{\psi}_{CX}$ to the $Y_C$ operator is obtained as integrals of the subgrid-cell representation in the $y$-direction of the $\overline{\psi}_{CX}$ field. Since only conservative operators have been applied, the total mass is unchanged.

The algorithm in Eq. (2.48) introduces a directional bias. Therefore, in practical applications, it has been common procedure to alternate between the directional splitting in (2.48) and the opposite sequential splitting:

$$\overline{\psi}^{n+1}_{CYX} = \overline{\psi}^{n} + Y_C\left(\overline{\psi}^{n}\right) + X_C\left(\overline{\psi}_{CY}\right) \tag{2.49}$$

with

$$\overline{\psi}_{CY} = \overline{\psi}^{n} + Y_C\left(\overline{\psi}^{n}\right). \tag{2.50}$$

The operator split schemes in Eq. (2.48)/Eq. (2.49) are also referred to as *time-split schemes*. Alternatively, instead of alternating Eq. (2.48) and Eq. (2.49), one can, of course, combine the operators and define a spatially symmetric conservative scheme as

$$\begin{aligned}\overline{\psi}^{n+1} &= \tfrac{1}{2}\left(\overline{\psi}^{n+1}_{CXY} + \overline{\psi}^{n+1}_{CYX}\right)\\ &= \overline{\psi}^{n} + X_C\left[\tfrac{1}{2}\left(\overline{\psi}^{n} + \overline{\psi}_{CY}\right)\right] + Y_C\left[\tfrac{1}{2}\left(\overline{\psi}^{n} + \overline{\psi}_{CX}\right)\right].\end{aligned} \tag{2.51}$$

Schemes of the general type Eq. (2.48)/Eq. (2.49)/Eq. (2.51) have been presented in several papers, e.g., TREMBACK, POWELL, COTTON and PIELKE [1987], BOTT [1989, 1992].

There is a fundamental time split error – often referred to as lumpiness – associated with schemes of this type. This splitting error is a consequence of the fact that a constant $\psi$-field is generally not conserved with these schemes as it should in a spatially but divergence free flow (i.e. a flow with $\partial u/\partial x = -\partial v/\partial y \neq 0$). This lake of constancy conservation of the scheme (2.46)–(2.47) arises because $\overline{\psi}_{CX}$ in (2.47) is a result of not only an advection but also a deformation in the $x$-coordinate direction. This is illustrated in Fig. 2.15 for the operator splitting scheme in Eqs. (2.46) and (2.47). The original constant density $\psi_0$ is indicated by the level of shading in the left two columns in the left panel (grid cells with $x$-index, $i-1$ and $i-2$). The effect of the operation (Eq. (2.46)), i.e., $\overline{\psi}_{CX}$, is shown for the cells with $x$-index $i$ (only) as the level of shading. The effective deformation in the $x$-direction for the arrival cell marked $i, j$ is a convergence $-D_0$ leading to increased values with darker shading. Around the arrival cell $i, j-2$, no deformation takes place in the $x$-direction and the shading is therefore unchanged. Finally, in the arrival cell $i, j-3$ deformation (i.e. divergence) is leading to a decreased value and consequently a lighter shading. The horizontal dotted lines in the right panel indicate the limits of the subsequent transport in the $y$-direction through the "southern" and "northern" walls, respectively, of cell $i, j$. Since the $y$-extension of this area is less than one (i.e., divergence $D_0$). However, since the air that is transported into the arrival cell $i, j$ comes from a place with density less than $\psi_0$ the final forecasted value $\overline{\psi}^{n+1}_{CYX}$ (not shown) in cell $i, j$ will end up being less than the original value $\psi_0$ which it should have retained.

The operator splitting error problem leads to serious error growth for transport by deformational flows. Petschek and Libersky [1975] showed that a kind of numerical instability is associated with the time-splitting. In simulations of highly compressible fluids, the splitting error seems to be of less importance (e.g., Woodward and Colella [1984], Carpenter, Droegemeier, Hane and Woodward [1990], Colella [1990]). However, in most geophysical applications, the splitting error must be explicitly dealt with to obtain sufficiently accurate simulations.

Leonard, Lock and Macvean [1996] and Lin and Rood [1996], independently, introduced essentially the same technique to eliminate the splitting error. The following derivation leads to the same expressions as those originally presented by these authors. Here, however, the focus is on the motivation behind the basic idea: ensuring the contribution from flow deformations to the final forecasted value are excluded in the initial transports parallel to the cell faces. The first ingredient is to define those contributions to $X_c \approx -\Delta t\, \partial\big(u\overline{\psi}\big)/\partial x$ and $Y_c \approx -\Delta t\, \partial\big(v\overline{\psi}\big)/\partial y$ that are related to flow deformations in each direction ($TDx \approx -\Delta t\, \overline{\psi}\partial u/\partial x$ and $TDy \approx -\Delta t\, \overline{\psi}\partial v/\partial y$) and those related to advection ($TAx \approx -\Delta t\, u\partial\overline{\psi}/\partial x$ and $TAy \approx -\Delta t\, v\partial\overline{\psi}/\partial y$). For given conservative flux convergences, one can at first define either the deformation or the advection contributions and then determine the other pair using the following relationships:

$$\begin{aligned} X_C\left(\overline{\psi}^n\right) &= TAx + TDx, \\ Y_C\left(\overline{\psi}^n\right) &= TAy + TDy. \end{aligned} \tag{2.52}$$

The advective updates are defined in terms of $TAx$ and $TAy$:

$$\begin{aligned} \overline{\psi}_{AX} &= \overline{\psi}^n + TAx, \\ \overline{\psi}_{AY} &= \overline{\psi}^n + TAy. \end{aligned} \tag{2.53}$$

Lin [2004] defined initially deformations by the following centered approximations (here expressed in cartesian coordinates):

$$\begin{aligned} TDx &= -\Delta t \overline{\psi}^n \frac{u_e^* - u_w^*}{\Delta x}, \\ TDy &= -\Delta t \overline{\psi}^n \frac{u_n^* - u_s^*}{\Delta y}. \end{aligned} \tag{2.54}$$

The associated advective contributions to the forecasts are then defined from the conservative flux-form operators as follows:

$$TAx = X_C - TDx, \tag{2.55}$$

$$TAy = Y_C - TDy, \tag{2.56}$$

respectively. The terms $TAx$ and $TAy$ are referred to as "advective" since they are approximations to $-\Delta t u \partial \overline{\psi}^n / \partial x$ and $-\Delta t v \partial \overline{\psi}^n / \partial y$, respectively.

When initially advective contributions are defined, e.g., as in Lin and Rood [1996], one will typically define them as:

$$\begin{aligned} TAx = \overline{\psi}_{AX} - \overline{\psi}^n &= \frac{1}{\Delta x} \int_{x_w - \overline{u}^* \Delta t}^{x_e - \overline{u}^* \Delta t} \psi^n \mathrm{d}x - \overline{\psi}^n \approx -\Delta t \overline{u}^* \frac{\partial \psi^n}{\partial x}, \\ TAy = \overline{\psi}_{AY} - \overline{\psi}^n &= \frac{1}{\Delta y} \int_{y_s - \overline{v}^* \Delta t}^{y_n - \overline{v}^* \Delta t} \psi^n \mathrm{d}y - \overline{\psi}^n \approx -\Delta t \overline{v}^* \frac{\partial \psi^n}{\partial y}, \end{aligned} \tag{2.57}$$

where $\overline{u}^* = 1/2(u_w^* + u_e^*)$ and $\overline{v}^* = 1/2(v_n^* + v_s^*)$. In this case,

$$\begin{aligned} TDx &= X_C - TAx \\ &= \frac{1}{\Delta x} \left[ \int_{x_w - u_w^* \Delta t}^{x_w} \psi^n \mathrm{d}x - \int_{x_e - u_e^* \Delta t}^{x_e} \psi^n \mathrm{d}x \right] + \overline{\psi}^n - \frac{1}{\Delta x} \int_{x_w - \overline{u}^* \Delta t}^{x_e - \overline{u}^* \Delta t} \psi^n \mathrm{d}x \\ &= \frac{1}{\Delta x} \left[ \int_{x_w - u_w^* \Delta t}^{x_w} \psi^n \mathrm{d}x + \int_{x_w}^{x_e} \psi^n \mathrm{d}x + \int_{x_e}^{x_e - u_e^* \Delta t} \psi^n \mathrm{d}x \right] - \frac{1}{\Delta x} \int_{x_w - \overline{u}^* \Delta t}^{x_e - \overline{u}^* \Delta t} \psi^n \mathrm{d}x \\ &= \frac{1}{\Delta x} \left[ \int_{x_w - u_w^* \Delta t}^{x_e - u_e^* \Delta t} \psi^n \mathrm{d}x - \int_{x_w - \overline{u}^* \Delta t}^{x_e - \overline{u}^* \Delta t} \psi^n \mathrm{d}x \right] \\ &\approx -\Delta t \psi^n \frac{u_e^* - u_w^*}{\Delta x} \end{aligned} \tag{2.58}$$

with a similar expression for $TDy$.

Once the advective update in the $x$-direction $\overline{\psi}_{AX}$ has been calculated from (2.53) and (2.57), one can consider a provisional conservative transport of this update in the $y$-direction:

$$\overline{\psi}_{CYAX} = \overline{\psi}_{AX} + Y_C\left(\overline{\psi}_{AX}\right). \tag{2.59}$$

Note that generally

$$\iint\limits_{totaldomain} \overline{\psi}_{CYAX}\mathrm{d}x\mathrm{d}y = \iint\limits_{totaldomain} \overline{\psi}_{AX}\mathrm{d}x\mathrm{d}y \neq \iint\limits_{totaldomain} \overline{\psi}^{n}\mathrm{d}x\mathrm{d}y, \tag{2.60}$$

where the last integral represents the total mass. To achieve mass conservation, we need to re-add the $TDx$ field:

$$\begin{aligned}\overline{\psi}_{YX}^{n+1} &= \overline{\psi}_{CYAX} + TDx \\ &= \overline{\psi}_{AX} + Y_C\left(\overline{\psi}_{AX}\right) + TDx \\ &= \overline{\psi}_{CX} + Y_C\left(\overline{\psi}_{AX}\right) \\ &= \overline{\psi}^{n} + X_C\left(\overline{\psi}^{n}\right) + Y_C\left(\overline{\psi}_{AX}\right).\end{aligned} \tag{2.61}$$

Like the (2.48) forecast the (2.61) forecast is conservative since it only includes conservative operators. However, the difference is that the field $\overline{\psi}_{AX}$, which is finally transported in the north-south direction into the $i$, $j$-cell, has been advected but not deformed. With $\psi^n \equiv \psi_0$ in (2.61) it is easily verified that

$$\overline{\psi}_{YX}^{n+1} = \psi_0\left(1 - \Delta t\left[\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y}\right]\right) = \psi_0 \quad \text{if } \left[\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y}\right] \equiv 0.$$

Here $\Delta u = u_e - u_n$ and $\Delta v = v_n - v_s$. Thus, a constant is exactly conserved if the discretized divergence $\left[\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y}\right] \equiv 0$.

As for Eq. (2.48), a directional bias is introduced by Eq. (2.61). This can be compensated by averaging the update $\overline{\psi}_{YX}$ in Eq. (2.61) with the equivalent update $\overline{\psi}_{XY}$ in the opposite direction leading to the symmetric expression:

$$\begin{aligned}\overline{\psi}^{n+1} &= \tfrac{1}{2}\left(\overline{\psi}_{YX} + \overline{\psi}_{XY}\right) \\ &= \overline{\psi}^{n} + X_C\left[\tfrac{1}{2}\left(\overline{\psi}^{n} + \overline{\psi}_{AY}\right)\right] + Y_C\left[\tfrac{1}{2}\left(\overline{\psi}^{n} + \overline{\psi}_{AX}\right)\right].\end{aligned} \tag{2.62}$$

An alternative and slightly cheaper approach than Eq. (2.62) is to alternate between $\overline{\psi}_{YX}$ and $\overline{\psi}_{XY}$ in each second time-step.

Considering the forecast in Eq. (2.61) as an example, it is noteworthy to observe that the contributions to the divergence term $-\Delta t\overline{\psi}(\Delta u/\Delta x + \Delta v/\Delta y)$ are determined from field values at different locations for each of the two directions. It is obvious that a contribution comes from $\overline{Y}_C(\overline{\psi}_{AX})$ at a location upstream in the $y$-direction, and due

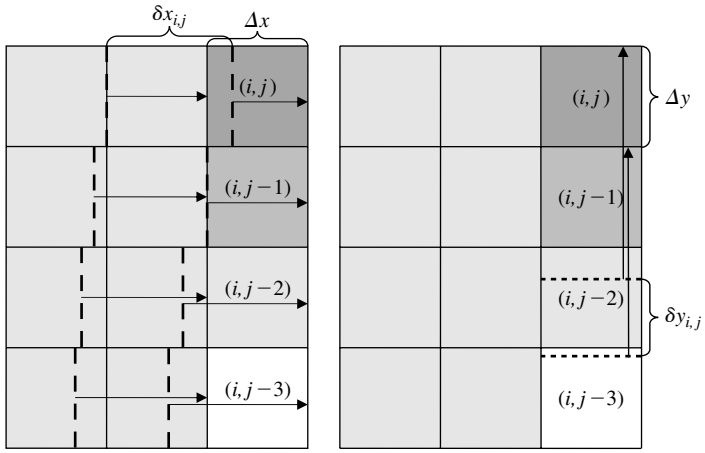


FIG. 2.15 Schematic illustration of the mass conservative but not constancy conserving time splitting in Eqs. (2.46)/(2.47) for a nondivergent flow. The left panel illustrates the intermediate forecast of $\overline{\psi}_{CX}$ for the $i$th column of grid cells. The upstream departure areas $\delta x = \Delta x + \Delta t\left(u_e^* - u_w^*\right)$ arriving in the $i$th column are indicated with dashed lines. The shading in column $i$ indicates the level of flow deformation $(-(\delta x - \Delta x)\,/\,(\Delta t \Delta x))$ related to the flow in the $x$-direction *only*, with dark shading indicating strong "convergence" and light shading "divergence." The right panel illustrates the final forecast (Eq. (2.47)) of $\psi_{CXY}^{n+1}$ for grid cell $(i,\ j)$. Here, the upstream departure area $\delta y_{i,\,j} = \Delta y + \Delta t\left(v_e^{*i,\,j} - v_w^{*i,\,j}\right)$ is illustrated with dotted lines. The shading in the right panel is identical to that in the left, i.e., the final forecast in $(i,\ j)$ is not indicated with shading. It is obvious, however, that it would be lighter than the original shading in the two grid columns $j-1$ and $j-2$. Thus, the scheme is not constancy conserving. Note that since flow is nondivergent, we have

$$D \approx -\left(\delta x_{i,\,j} - \Delta x\right)\,/\,(\Delta t \Delta x) - \left(\delta y_{i,\,j} - \Delta y\right)\,/\,(\Delta t \Delta y) = 0.$$

to the definition of $\overline{Y}_C$ ($\overline{\psi}_{AX}$) the ultimate origin of this contribution is upstream in both the $y$- and $x$-directions as it should be according to the "exact solution". Contrary to this, the term $TDx$ gives a contribution to the divergence term that is based on field values at a different location: if $TDx$ is chosen as the primary definition and the definition in Eq. (2.54) is used, it is based on the value $\overline{\psi}^n$ at the location of the arrival Eulerian cell, and if $TAx$ is chosen as the primary definition and (2.58) apply, it is based on the value $\overline{\psi}^n$ at a location $\Delta t \overline{u}^*$ upstream in the $x$-direction. Similar arguments apply to $\overline{\psi}_{XY}$, and therefore the forecast in Eq. (2.62) includes some (small) contributions from the $\overline{\psi}^n$ field either at the location of the Eulerian arrival cell (when $TDx$ and $TDy$ are chosen as primary definition) or upstream in each of the two directions (when $TAx$ and $TAy$ are primaries). Fig. 2.11 shows the actual departure area for a forecast using the transport schemes by LIN and ROOD [1996] and the wind field in Eq. (2.24). It can be seen that for this particular case, there is a considerable spread out of the departure area relative to the DCISL schemes in Fig. 2.10. A careful inspection of Fig. 2.11 shows that the net departure area is displaced systematically towards "south-east" as compared to the "exact departure area". This is related to curvature of the trajectories in Fig. 2.9. As for the DCISL schemes the shift of mass inside grid cells, due to calculated sub-grid-scale deviations from the cell average value $\overline{\psi}^n$ in cells crossed by the boundary of the departure area, may result in a larger or a smaller effective departure area from which mass is contributed to $\overline{\psi}^{n+1}$,

compared to the departure area shown in Fig. 2.11. However, in any case the global mass is conserved and the sub-grid-scale distributions serve to reduce any artificial damping.

The approaches by LEONARD, LOCK and MACVEAN [1996] and LIN and ROOD [1996] to eliminate the operator – or time – splitting error may have been inspired by the earlier techniques introduced by BOTT [1993] and EASTER [1993]. BOTT [1993] applied flux limiters (see Section 2.6) to obtain monotonicity-preserving transport, but otherwise, the basic idea is the same as explained above. One fundamental difference is, however, that the BOTT [1993] scheme does not permit time-steps exceeding the CFL criterion such as the schemes by LEONARD, LOCK and MACVEAN [1996] and LIN and ROOD [1996].

EASTER [1993] introduced an alternative way of eliminating the splitting error in the original positive-definite scheme by BOTT [1989]. His approach is equivalent to that by LEONARD, LOCK and MACVEAN [1996] and LIN and ROOD [1996]: the 1D conservative transport in the first direction of the operator splitting will generally change the fluid density due to fluid deformations and due to pure advection. For tracer mixing ratio, only pure advection along tracer gradients can change the value. By eliminating the deformational part, one can isolate mixing ratio transport and subsequently during the transport in the second direction include the deformational contribution in a consistent way: the estimated total deformation in each of the two directions for a given cell is based on the local flow around this cell.

Schemes of the type presented above have been quite popular in recent years (e.g., RASCH and LAWRENCE [1998], LIN [2004]) and have and will be been implemented into atmospheric models ranging from meso-scale models (e.g., SKAMAROCK, KLEMP, DUDHIA, GILL, BARKER, WANG and POWERS [2007]) to general circulation models GCMs (e.g., ADCROFT, CAMPIN, HILL and MARSHALL [2004]. Note, however, that by far most applications have been offline, i.e., passive advection of tracers in models using a different scheme for the solution of the continuity equation in the dynamical core. For such models, the need for special attention in relation to the mass-wind inconsistency problem – see Section 3 – is often even more important than in the online models mentioned here.

As shown by LIN and ROOD [1996], the flux-based schemes derived above will normally lead to conservation of linear correlations between the mixing ratios $q_a$ and $q_b$ of two tracers $a$ and $b$, i.e., if $\overline{q}_b^n = \alpha \overline{q}_a^n + \beta$ then $\overline{q}_b^{n+1} = \alpha \overline{q}_a^{n+1} + \beta$ where $\alpha$ and $\beta$ are constants. This is because the flux convergences and transports of the types of Eqs. (2.42), (2.45), and (2.54) satisfy the general linear relationships:

$$\begin{aligned} Z(\psi + \beta) &= Z(\psi) + Z(\beta), \\ Z(\alpha\psi) &= \alpha Z(\psi). \end{aligned} \tag{2.63}$$

The above relationships will in general also apply to DCISL schemes based on, e.g., the PPM method.

The conservation of linear relationships between different tracers is an attractive feature in, e.g., chemical modeling because it prevents artificial chemical reactions in idealized situations where the mixing ratio within a domain of one tracer can be expressed as a linear function of another. Note, however, that for schemes where the upstream subgrid-cell representation is forced positive-definite or monotonic, Eq. (2.63) is generally not fulfilled. It should also be mentioned that no such thing as linear relationships exist

between tracers in nature. Generally, the flux-based schemes (and the DCISL schemes) do, therefore, not conserve the local relative concentrations or mixing ratios between two tracers. In particular, this is the case in regions where one tracer has a reasonably smooth behavior while the other is dominated by a sharp variation of the spatial gradient.

In practice, the mixing ratio for a tracer is used as prognostic variable in many flux-based schemes. HOURDIN and ARMENGAUD [1999] used a scheme quite similar to that by LIN and ROOD [1996], where a few requirements on the spatial behavior of the mixing ratio were sufficient to ensure both monotonicity and positive definiteness. Generally, schemes based on mixing ratio should conserve a constant mixing ratio, although enforcement of positive definiteness and monotonicity in such schemes may deteriorate correlations between mixing ratios. Note that schemes based solely on mixing ratio will generally not conserve the total mass of the tracer unless care is taken to conserve total mass of the air in a consistent way. If flux-limiters (see Section 2.6) or a priori constraints on the subcell representation are applied solely to mixing ratios, one will lose mass conservation unless special additional constraints are imposed. In meteorological models the atmospheric air contain more or less water vapour so *specific concentration* (see definition page 3) is used instead of *mixing ratio*, defined relative to dry air.

*2.3.2.2. Stability of operator-split, flux-based schemes* The operator-split, flux-based schemes have often been subject to a CFL criterion (i.e., $\Delta t < \Delta x/u^*$) in 1D. One can identify two main reasons for this:

- The upstream subgrid-cell representation needed to estimate the face fluxes was defined locally from the grid cells neighboring the target Eulerian cell. This means that longer time steps led to extrapolation (and not aggregation) of information and hence the CFL criterion. It was not considered to apply "semi-Lagrangian" thinking.
- Some schemes are quite heavily hooked up on localized flux limiters. This makes it difficult to generalize into far upstream constraints. Examples are the otherwise popular and accurate schemes by BOTT [1989, 1992, 1993].

As pointed out by LEONARD [1994], there are, however, no immediate scientific reasons to limit the integration domain to the neighboring grid cells for schemes where the face fluxes are based on pure upstream integrals of "mass".[3] If the subgrid-cell information is defined everywhere without extrapolation, the face fluxes consist of interpolation/aggregation of information and they will be stable as shown in the 1D PPM case for the DCISL schemes (Section 2.3.1.4).

The stability of low-order versions of 2D flux-based semi-Lagrangian schemes was cursorily investigated by LIN and ROOD [1996].[4] Later, LAURITZEN [2007] made a detailed stability analysis of both higher and lower order versions of the LIN and ROOD

[3] However, depending on the machine architecture and the actual scheme it may in some case be more efficient to require that characteristics depart from the immediate neighbouring cells because then there is no need for searching.

[4] Note that equations (A.5) and (B.1) in LIN and ROOD [1996] are missing some terms, but that the correct formulas have the same generic form. See the Appendix in LAURITZEN [2007].

[1996] class of schemes as well as a conceptual analysis to explain the results. The LIN and ROOD [1996] class of schemes is given by Eq. (2.62), and the different schemes in the class are formed by varying the order of the 1D operators that are applied. The flux-form operators applied to the terms in the square brackets in Eq. (2.62) are referred to as *outer* operators, whereas the operators used for $\overline{\psi}_{AX}$ and $\overline{\psi}_{AY}$ are referred to as the *inner* operators. When the *inner* and *outer* operators differ, LAURITZEN [2007] showed in a linear Von Neumann stability analysis that increased damping (or weak instability) may result, but this spurious damping disappears when the operators are identical (similarly for phase errors). This is due to the fact that for Courant numbers larger than unity there can be contributions to the forecast not originating from the Lagrangian departure area as they physically should. If the operators are identical and under the assumptions applied in a Von Neumann stability analysis, the LIN and ROOD [1996] scheme becomes formally identical to the NAIR, SCROGGS and SEMAZZI [2002] and ZERROUKAT, WOOD and STANIFORTH [2002] schemes, and therefore only includes information from the departure cell. For more details, see LAURITZEN [2007].

*2.3.2.3. Explicit estimation of the 2D fluxes* Most flux-based schemes have used the technique of time- or operator splitting described in Section 2.3.2.1, and the approach has proven to be very efficient and economic. There will, however, always be a slight inconsistency since the splitting prevents integration of the exact departure area. To reduce these problems, several papers (e.g., DUKOWICZ and RAMSHAW [1979], SMOLARKIEWICZ [1984], BELL, DAWSON and SHUBIN [1988], COLELLA [1990], DUKOWICZ and KODIS [1987], SMOLARKIEWICZ and GRABOWSKI [1990], RASCH [1994], LEONARD [1994], HÓLM [1995], DUKOWICZ and BAUMGARDNER [2000]) have investigated the possibility of constructing fully 2D flux-based schemes. In these schemes, one aims directly at an estimate of the transport in the "cross-directions", which was taken care of by the sequential approach in the operator split methods.

The fully 2D flux-based schemes are similar to fully 2D DCISL schemes since they – more or less directly – are based on estimates of integrals over upstream areas to obtain the mass interactions with all the neighboring grid cells. One example is the scheme by HÓLM [1995] in cartesian geometry, which for a given grid cell is based on four unique fluxes in the $x$-direction, the $y$-direction, and the two cross-directions (see Fig. 2.16).

The scheme proposed by RASCH [1994] appears to be somewhat different. This scheme is based on an upwind biased stencil of points that are used to define an upstream spatial interpolation of the same type as that in semi-Lagrangian models.[5] However, making use of certain symmetry rules in the upstream polynomials, this interpolation can be formulated in the traditional flux form (Eq. (2.39)) for a given Eulerian grid cell. In other words, constraints on the polynomial coefficients ensure that the implied fluxes are unique for each face.

Contrary to the operator-split flux-form FV schemes, it is complicated to circumvent the CFL criterion for fully 2D flux-based schemes. Referring to the discussion in Section 1 on the analogies between flux-form and DCISL schemes, a fully 2D semi-Lagrangian flux-form scheme would in fact be the same as a DCISL scheme.

[5] Note, however, that the scheme by RASCH [1994] is Eulerian and therefore subject to a CFL criterion.

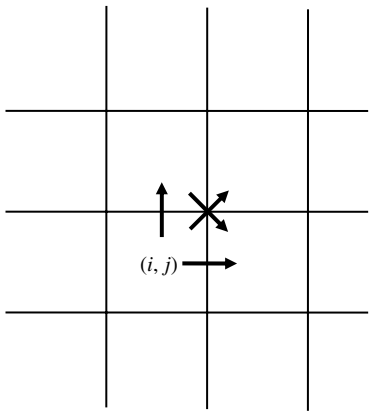


FIG. 2.16 Schematic illustration of the four unique fluxes needed in fully 2D flux-based finite-volume schemes.

### 2.4. *Locally mass conserving semi-Lagrangian grid point methods*

Recently, a few locally mass conserving upstream or downstream interpolating semi-Lagrangian grid point schemes have been proposed in the literature. Mass conservation has been achieved via local modifications of the polynomial interpolations in such a way that the total mass is always conserved. Effectively the prognostic grid point variable in such schemes is the average density in Eulerian grid cells, and therefore these schemes can be considered special types of semi-Lagrangian FV schemes.

We have already – in Section 2.3.1.1 – mentioned the scheme by LAPRISE and PLANTE [1995] where a downstream semi-Lagrangian scheme was modified along these lines. The scheme by RASCH [1994] – although Eulerian – is an example of an upstream mass conserving scheme based on modifications of the polynomial coefficients. More recently, KAAS [2008], COTTER, FRANK, REICH [2007] and REICH [2007] have proposed upstream and downstream grid point semi-Lagrangian schemes which are locally mass conserving. The basic idea is to modify the upstream or downstream polynomial interpolation coefficients.

For an upstream traditional semi-Lagrangian scheme – following KAAS [2008] – these coefficients can be considered area (or volume in the 3D case) weights transferring information from Eulerian grid points to the different irregularly spaced neighboring semi-Lagrangian departure points. The original weights in this remapping are modified by that fraction, which ensures that the sum of the weights given off by a given Eulerian grid point to all the surrounding departure points is equal to the unique area (volume) represented by this grid point. Hereby a local mass conservation is achieved when the prognostic variable is density. The forecasted densities (in the arrival Eulerian grid points), including the effects of divergence, are equal to the modified upstream interpolated values divided by the unique area (volume) represented by the arrival Eulerian grid point.

For downstream schemes as in COTTER, FRANK, REICH [2007] and REICH [2007], the procedure is opposite and somewhat more tricky since it is the distribution weights that are modified. Also REICH [2007] discusses the possibility of introducing modified upstream semi-Lagrangian weights.

The domain of dependence of locally mass conserving semi-Lagrangian schemes is comparable to that in pure DCISL schemes and depends on the actual polynomial accuracy used for the upstream interpolations/subgrid-cell representations. However, the degree of local mass conservation is higher in the DCISL schemes since the departure cell area is close to the true departure area, while in the locally mass conserving semi-Lagrangian schemes, mass is extracted from grid cells in a larger domain.

### *2.5. Additional prognostic variables*

To improve the subgrid-cell representation needed to estimate the mass fluxes, one may introduce additional prognostic variables. VAN LEER [1977] (scheme IV) and PRATHER [1986] used traditional second-order polynomials to represent the spatial distribution and used both gradients and curvatures as additional prognostic variables to define these polynomials. This allowed for the formulation of a formally very accurate scheme that conserved second-order moments. Note, however, the arguments by THUBURN [2006] that it is not desirable to conserve second moments since one can only conserve the resolved part. Furthermore, the PRATHER [1986] scheme is very computationally demanding both in terms of CPU and memory requirements. Therefore, it has not been popular in "real" applications.

More recently, schemes have been introduced (e.g., XIAO and YABE [2002]), where not only the cell mean values but also the values and gradients at the cell interfaces are prognostic variables. These additional prognostic variables have the same role as the moments introduced by PRATHER [1986]: the reduction of the loss of information (damping) associated with the spatial remappings that are fundamental to all FV schemes. The proposed new schemes are highly accurate as the scheme by PRATHER [1986], but at significantly reduced computational cost, particularly in terms of CPU usage. The scheme has been further improved and generalized to 2 and 3Ds using directional splitting (XIAO, YABE, PENG and KOBAYASHI [2002], PENG, XIAO, OHFUCHI and FUCHIGAMI [2005]). To describe the basics behind the new so-called conservative semi-Lagrangian schemes based on rational functions (CSLR) schemes, consider transport in 1D and assume that we know the cell mean value $\overline{\psi}$ and the "west" and "east" interface values, $\psi_w$ and $\psi_e$. From this information, one has three degrees of freedom to construct the subgrid-cell representation at a given time step $n$. One possible choice of functions could of course be the PPM. However, XIAO, YABE, PENG and KOBAYASHI [2002] found that rational functions with second-order polynomials gave better results at less numerical cost. However, the rational functions used have a built-in singularity that causes problems unless special care is taken. In the CSLR scheme, this singularity can appear when a local maximum or minimum is transported. According to XIAO, YABE, PENG and KOBAYASHI [2002], the problem can be dealt with by introducing a small machine-dependent constant that prevents division by zero at the singularity. In the 2D case, the update of the $\overline{\psi}$ values is performed as standard flux-form integrals, Eqs. (2.42) and (2.45), of the rational

functions. The cell interface values are updated using standard semi-Lagrangian upstream interpolation based on the rational functions followed by the relevant change related to the divergence of the flow. The new schemes are stable and efficient, but, of course, they will be more memory demanding since an additional prognostic variable is introduced. It is anticipated that the basic idea behind the CSLR schemes is so powerful that it will be adopted in many future integration schemes used in atmospheric models.

We see no fundamental problems in applying the powerful technique of including cell interface values as additional prognostic values in fully 2D and cascade DCISL schemes although the authors are not aware of any specific attempts along this direction.

### *2.6. Flux limiters*

FV schemes based on polynomial unfiltered subcell representations do not, in general, fulfill requirements such as positive definiteness and monotonicity. In particular, numerical oscillations often develop near discontinuities or large variability in gradients. In Section 2.2.1, it was described how it is possible to introduce different filters or constraints on the subgrid-cell representations to reduce or eliminate these problems. In most cases – with the filter by ZERROUKAT, WOOD and STANIFORTH [2005] as an exception – the applications of such filters tend to reduce the accuracy of the schemes because of the implied clippings and smoothings of the subcell scale polynomials. We can denote these filters *a priori* filters because they are introduced before the estimation of fluxes or the upstream cell integrations. It is, however, also possible to introduce a posteriori corrections – often referred to as flux limiters – of the fluxes to ensure fulfillment of the desired properties. This type of flux corrected transport (FCT) filters was introduced by BORIS and BOOK [1973] and by ZALEZAK [1979]. The basic idea behind the classical FCT is to perform a local mixing of the fluxes obtained from a high-order scheme (which is accurate but violate the desired properties) with fluxes from a low-order highly diffusive scheme (which fulfill the properties), e.g., a simple so-called upstream scheme. The procedure is – for each cell interface – to modify the local fluxes of the diffusive scheme as much as possible toward the fluxes in the high-order scheme without exceeding the magnitude of this flux and without creating new local maxima or minima in the neighboring cells; i.e., the local fluxes are changed differently at all interfaces under the constraint that the change in neighboring cell values do not lead to changed sign of gradients in the neighboring interfaces.

Several different types of flux limiter approaches have been presented in the literature to obtain positive definiteness, e.g., BOTT [1989], or monotonicity, e.g., SMOLARKIEWICZ and GRABOWSKI [1990], BOTT [1992], RASCH [1994], HÓLM [1995], XUE [2000]. For some schemes, such as the schemes by BOTT [1989], BOTT [1992], the flux limiters are inherent parts of the basic flux calculations. HÓLM [1995] was the first to apply flux limiters directly to the fluxes in fully 2D flux-based schemes.

Although one can argue that the specific flux limiters used will be somewhat arbitrary from a physical point of view, such filters can improve the performance of transport schemes significantly at a reasonable cost although there are some logical statements and "max/min" functions involved in the algorithms. As mentioned above, limiters enforcing positive definiteness will generally not ensure conservation of mixing ratios between

different tracers. Furthermore, when flux limiters are applied in operator-split schemes permitting long time-steps, the flux limiters cannot guarantee strict multidimensional shape conservation in flows with strong deformation (Leonard, Lock and Macvean [1996]). An exception is the limiter of Skamarock [2006].

### *2.7. Concluding remarks*

Two fundamentally different FV methods in 2Ds are being used in meteorology: DCISL schemes and the flux-based methods. In DCISL schemes, the forecast for a given Eulerian cell is based on an integral over an isolated area approximating the exact upstream departure area. This means that DCISL schemes are quite direct approximations to the exact forecast that is an integral of the exact subgrid representation over the exact departure area. In flux-based methods, the forecast is obtained as the net flux of mass through each of the faces of the Eulerian cell. For each face of the cell, this flux is shared with a neighboring Eulerian cell and it is determined as an integral over the area swept through the actual face during one time-step. Although less direct than DCISL schemes, the flux-based methods also approximate an integral over the exact departure area. Therefore, the two methods are equivalent and the accuracy of both will depend on the order of the subgrid representation being integrated and the effective approximation to the "true" departure cell.

For DCISL as well as flux-based schemes, the operations related to two directions can be separated or split. For DCISL schemes, this is referred to as cascade integration, and for the flux-based schemes, it is termed operator- or time-splitting. The advantage of the splitting is that only 1D subgrid representations and integrations are required which makes these schemes considerably more efficient.

Although the flux-based schemes are generally quite accurate and conserve mass (or any integral invariant) locally, higher order subgrid representations, i.e., high accuracy, will generally violate conservation of shape, i.e., the schemes become nonmonotonic or nonpositive-definite. A number of constraints to reduce or eliminate such problems can be applied to the subgrid representations entering the upstream integrals. It is also possible to apply a posteriori corrections (e.g., so-called flux limiters) to the forecast that reduce or eliminate these problems.

DCISL schemes are by construction semi-Lagrangian and not subject to any advective CFL criterion that limits the maximum possible time-step apart from the requirement of the departure cells being well defined. In contrast to this, many traditional flux-based schemes are formulated to allow only transport over a maximum distance of one grid cell within one time-step, i.e., the Courant number must be less than unity to obtain stability. For operator-split flux-form schemes, it is, however, possible to extend the integration domain thereby avoiding the CFL criterion.

In the original time-split flux-form schemes, the lake of conservation of constant density fields in non-divergent (but deforming) flow, caused a splitting error, the so-called nudging error, The introduction of combined advective-conservastive flux-form schemes circumvented this problem. In realistic case studies with DCISL schemes no error like the flux-form nudging error has been reported, although the DCISL schemes do not conserve exactly a constant field in non-divergent flow. Such constancy conservation will be

obtained in upstream DCISL schemes only if the departure area $\delta A$ is exactly equal to the arrival area $\Delta A$, or in other words if the discretized Lagrangian divergence $\mathbb{D} = \frac{\partial A - \Delta A}{\Delta t \Delta A}$ is exactly zero. This should be the case in any reasonably smooth divergence free flow if the determination of the departure area was exact. In praxis the departure area is determined only approximately. In present schemes it is dependent on the accuracy of the backward trajectories from the corner points of the arrival area and the strait line approximation to the exact connections between the departure corner points. LAPRISE and PLANTE [1995] found in an idealized solid body rotation experiment an accumulated relative error $(\partial A - \Delta A)/\Delta A$ over 10 time steps of up to 1%. It seems reasonably to assume that this number will be reduced substantially with a trajectory computation like that used by LAURITZEN, LINDBERG, KAAS and MACHENHAUER [2008], that takes accelerations into account. This should be investigated further. Furthermore in realistic flows it must be expected that accumulation of errors will occur less frequent and cancellation of errors will reduce the problem, if any.

For the schemes considered here, the exact departure area in deforming flows is better represented and integrated over in the DCISL schemes than in flux-based schemes. Therefore, with respect to the schemes for effective approximation to the departure area, one may conclude that DCISL schemes are generally more accurate than flux-based schemes.

One may anticipate that further developments of FV methods will include introduction of additional prognostic values and gradients at the cell interfaces as was recently proposed.

## 3. FV models

As stated previously, FV methods are well suited for the numerical simulation of conservation laws. This is demonstrated in Section 3.1 where a complete set of FV prognostic equations, that conserve exactly mass, entropy, total energy, and angular momentum in an adiabatic, friction-free and quasi-hydrostatic atmosphere, is derived. A numerical model based on this set of FV conservation laws, a so-called complete set of conservation laws (CSCL) model, remains to be realized. However, assuming forcing terms and Eulerian vertical discretization as in an existing operational primitive equation model, it is shown how such a prognostic system may be set up. The advection of all invariants is supposed to be calculated by an explicit, absolutely stable DCISL time-stepping scheme borrowed from LAURITZEN, KAAS, MACHENHAUER and LINDBERG [2008]. Here, in each time-step, mass and other invariants are transported conservatively along Lagrangian surfaces determined as in LAURITZEN, KAAS, MACHENHAUER and LINDBERG [2008] by 3D, so-called hybrid trajectories that are horizontally upstream (determined from the horizontal wind field) and vertically downstream (determined indirectly by the condition of hydrostatic balance).

In Sections 3.2 and 3.3, respectively, two recently developed quasi-hydrostatic dynamical cores in spherical coordinates are described, namely, the global NCAR-FFSL (National Center for Atmospheric Research – Flux Form Semi-Lagrangian) dynamical core (COLLINS, RASCH, BOVILLE, HACK, MCCAA, WILLIAMSON, KIEHL, BRIEGLEB, BITZ, LIN, ZHANG and DAI [2004]) and the limited area HIRLAM-DCISL dynamical core (LAURITZEN, KAAS, MACHENHAUER and LINDBERG [2008]). They are pioneering

examples of the two different types of FV dynamical cores developed in the meteorological modeling community. The former is based on the flux-form (Eq. (1.2)) and the latter on the Lagrangian form (Eq. (1.8)) of the continuity equation. These dynamical core examples of FV dynamical cores developed recently in the meteorological modeling community. In both dynamical cores, the continuity equation is solved by absolutely stable FV advection methods, which ensure exact mass conservation. In the HIRLAM DCISL, the remaining primitive equations are solved with finite-difference methods that do not ensure exact conservation of additional integral invariants. In the NCAR-FFSL, additionally potential temperature and absolute potential vorticity are conserved for adiabatic friction-free flow. In the HIRLAM-DCISL dynamical core, the DCISL advection scheme is combined with a semi-implicit time-stepping, thereby allowing large time-steps for all variables at the expense of solutions to elliptic Helmholtz equations each time-step (ROBERT [1969, 1981, 1982]). In the NCAR-flux-form semi-Lagrangian (FFSL) dynamical core, an explicit flux-based advection scheme is used, which means that shorter time-steps must be used for the dynamical variables, while advection of tracers (including water vapor) and physical parameterization can be predicted with long time-steps. Both dynamical cores have been coupled with comprehensive physical parameterization packages. In Section 3.4, the properties of the dynamical cores are discussed.

The main part of this section is dealing almost exclusively with complete quasi-hydrostatic atmospheric models. However, relevant aspects of online and off-line applications are taken up in Section 3.5, and finally in Section 3.6, possibilities of extensions to nonhydrostatic models are briefly discussed.

### *3.1. A complete set of FV conservation laws for a quasi-hydrostatic atmosphere*

As shown by MACHENHAUER [1994], an explicit FV general circulation model, which conserves exactly a maximum number of fundamental integral invariants, may be formulated. Let it as above be called a CSCL model. In this section, the prognostic equations are derived and a possible explicit time-stepping procedure is presented. Finally, the feasibility of such a model is discussed.

#### *3.1.1. The continuous primitive equations*

Consider the continuous equations, the so-called primitive equations, for a general pressure-based terrain-following vertical coordinate $\eta\,(p,\, p_s)$ as formulated, for example, for the European Center for Medium Weather Research (ECMWF) integrated forecast system (IFS) model (SIMMONS and BURRIDGE [1981]) and the HIRLAM (KÄLLÉN [1996]) operational atmospheric models. The prognostic equations are:
*the quasi-horizontal momentum equation*

$$\frac{\mathrm{d}\vec{V}}{\mathrm{d}t} = -\nabla\phi - \alpha\nabla p - f\vec{k}\times\vec{V} + \vec{P}_{\vec{V}} + \vec{K}_{\vec{V}}, \tag{3.1}$$

*the thermodynamic equation*

$$c_p\frac{\mathrm{d}T}{\mathrm{d}t} = \alpha\omega + (P_T + K_T)c_p, \tag{3.2}$$

*the continuity equation for moist air*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial p}{\partial \eta}\right)+\frac{\partial p}{\partial \eta}\nabla\cdot\vec{V}+\frac{\partial p}{\partial \eta}\frac{\partial \dot{\eta}}{\partial \eta}=0, \tag{3.3}$$

and *the moisture equation*

$$\frac{\mathrm{d}q_v}{\mathrm{d}t}=P_{q_v}+K_{q_v}, \tag{3.4}$$

which in combination with Eq. (3.3) may be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(q_v\frac{\partial p}{\partial \eta}\right)+\left(q_v\frac{\partial p}{\partial \eta}\right)\nabla\cdot\vec{V}+\left(q_v\frac{\partial p}{\partial \eta}\right)\frac{\partial \dot{\eta}}{\partial \eta}=\frac{\partial p}{\partial \eta}\left(P_{q_v}+K_{q_v}\right). \tag{3.5}$$

The hybrid vertical coordinate $\eta\,(p,\,p_s)$, introduced by Simmons and Burridge [1981], is a monotonic function of pressure $p$ and surface pressure $p_s$ such that

$$\eta(0,\,p_s)=0 \quad \text{and} \quad \eta(p_s,\,p_s)=1.$$

Here $t$ is time, $\vec{V}$ is the horizontal wind vector, $q_v$ is the specific humidity, $\nabla$ is the horizontal gradient operator along $\eta$-surfaces, $\phi$ is the geopotential, $\vec{k}$ is the vertical upward unit vector, $f$ is the Coriolis parameter ($f=2\Omega\sin\varphi$, where $\Omega$ is the angular velocity of the Earth and $\varphi$ is the latitude), $\omega$ is the $p$-coordinate vertical velocity ($\omega=\mathrm{d}p/\mathrm{d}t$), $\alpha$ is the specific volume, and $\rho$ is the density of moist air determined by *the ideal gas equation*

$$\alpha=\frac{1}{\rho}=\frac{R_dT_v}{p}. \tag{3.6}$$

$R_d$ is the gas constant for dry air and $T_v$ is the virtual temperature defined by $T_v=T\left[1+\left(\frac{1-\varepsilon}{\varepsilon}\right)q_v\right]$, where $T$ is the absolute temperature, $\varepsilon=R_d/R_v$, and $R_v$ is the gas constant for water vapor. $c_p$ is the specific heat of moist air defined by $c_p=c_{pd}(1+(\delta-1)\,q_v)$, where $\delta=c_{pv}/c_{pd}$, $c_{pv}$ and $c_{pd}$ are the specific heat of water vapor and dry air, respectively. The geopotential $\phi$, which appears in Eq. (3.1), is defined by the diagnostic *hydrostatic equation*:

$$\frac{\partial \phi}{\partial \eta}=R_dT_v\frac{\partial \ln\,p}{\partial \eta}. \tag{3.7}$$

The $P$-forcing terms in Eqs. (3.1), (3.2), and (3.4) represent the contributions of the parameterized physical processes while the $k$-forcing terms represent the parameterized horizontal diffusion. The $P$-terms may be specified as for the ECMWF model:

$$\vec{P}_{\vec{V}}=-g\left(\frac{\partial p}{\partial \eta}\right)^{-1}\frac{\partial \vec{J}_{\vec{V}}}{\partial \eta}, \tag{3.8}$$

$$c_PP_T=Q_R+Q_L+Q_D-g\left(\frac{\partial p}{\partial \eta}\right)^{-1}\left[\frac{\partial J_S}{\partial \eta}-c_{pd}T\,(\delta-1)\,\frac{\partial J_{q_v}}{\partial \eta}\right], \tag{3.9}$$

$$P_q=S_{q_v}-g\left(\frac{\partial p}{\partial \eta}\right)^{-1}\frac{\partial J_{q_v}}{\partial \eta}, \tag{3.10}$$

where $\vec{J}_{\vec{V}}$, $J_S$, and $J_{q_v}$ represent net parameterized vertical fluxes of momentum, dry static energy $c_{pd}T + \phi$, and moisture. $Q_R$, $Q_L$, and $Q_D$ represent heating due, respectively, to radiation, internal phase changes, and internal dissipation of kinetic energy associated with the $\vec{P}_{\vec{v}}$-term. $S_{q_v}$ denotes the rate of change of $q_v$ due to rain and snowfall. Comprehensive physical forcing packages have been developed for the calculation of the $P$- and $k$-terms in operational primitive equation models as the IFS and HIRLAM. Assuming that such a package is available, it is convenient to express the forcing of the CSCL model in terms of $P$- and $k$-terms.

*3.1.2. Vertical discretization*

In the formulation of the FV CSCL model, it is convenient to make use of the traditional Eulerian hybrid sigma-pressure vertical discretization which is used widely, e.g., in the ECMWF and the HIRLAM models. This means that the model atmosphere is divided into NLEV layers which are defined by the pressures at the interfaces between them (the "half levels"):

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2}\, p_s \tag{3.11}$$

for $0 \leq k \leq$ NLEV. The pressure thickness of the model layers is denoted as $\Delta_k p = p_{k+1/2} - p_{k-1/2}$. The coefficients $A_{k+1/2}$ and $B_{k+1/2}$ are constants whose values completely define the vertical $\eta$ coordinate.

The finite difference analog, to the hydrostatic Eq. (3.7) for the geopotential thickness of a single layer and for the air mass from surface and up to a half level, is given by

$$\phi_{k+1/2} - \phi_{k-1/2} = -R_d (T_v)_k \, (\Delta \ln p)_k \,, \tag{3.12}$$

$$\phi_{k+1/2} = \phi_s + R_d \sum_{l=k+1}^{\text{NLEV}} (T_v)_l \, (\Delta \ln p)_l \,, \tag{3.13}$$

respectively, where

$$(\Delta \ln p)_k = \ln\left(\frac{p_{k+1/2}}{p_{k-1/2}}\right). \tag{3.14}$$

To obtain the geopotential at a full level, the almost universal approach of Simmons and Burridge [1981] is used:

$$\phi_k = \phi_{k+1/2} + \alpha_k R (T_v)_k \,, \tag{3.15}$$

where

$$\alpha_k = \left\{ \begin{array}{ll} \ln 2 & k = 1 \\ 1 - \dfrac{p_{k-1/2}}{\Delta_k p} \, (\Delta \ln p)_k & k = 2, .., \text{NLEV} \end{array} \right\}. \tag{3.16}$$

Following Lauritzen, Kaas, Machenhauer and Lindberg [2008], the "full level" pressures $p_k$ are computed with

$$p_k = p_{k+1/2} \exp(-\alpha_k) \,. \tag{3.17}$$

Away from the upper boundary, $p_k \cong \frac{1}{2}\left(p_{k+1/2} + p_{k-1/2}\right)$. The various FV conservation laws may be derived directly from the primitive equations listed above (as was done in Machenhauer [1994]). However, the more straightforward procedure used in Section 1 to derive the mass conservation law (Eq. (1.12)) will be used here.

*3.1.3. Conservation of mass of moist air*

The mass in a FV $\delta V$ with vertical walls, horizontal cross-section $\delta A$, and thickness $\delta z = z_1 - z_2$ (see Fig. 3.1) is

$$M_{\delta V} = \iint\limits_{\delta A} \left( \int_{z_2}^{z_1} \rho \, \mathrm{d}z \right) \mathrm{d}x \mathrm{d}y. \tag{3.18}$$

Utilizing the hydrostatic balance $\mathrm{d}p = -g\rho \, \mathrm{d}z$, the inner integral may be written as

$$\int_{z_2}^{z_1} \rho \, \mathrm{d}z = \frac{1}{g} \int_{p_1}^{p_2} \mathrm{d}p = \frac{1}{g} \left(p_2 - p_1\right) = \frac{1}{g} \delta p$$

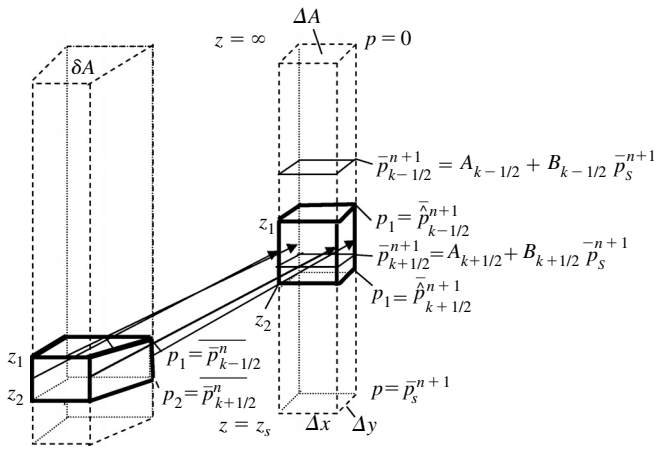


Fig. 3.1 Cell of moist air with vertical walls extending from the height $z_1$ at pressure $p_1$ to the height $z_2$ at pressure $p_2$, situated in the left column at time $t = n\Delta t$ and in the right column at time $t + \Delta t = (n + 1)\,\Delta t$. During a time step, we suppose the cell is moving in an air flow without vertical shear so that its vertical walls remain vertical. Generally, its horizontal cross-section area $\delta A$, its thickness $\delta z = z_1 - z_2$, and the corresponding pressure difference $\delta p = p_2 - p_1$ are changing with time during the time-step. At time $t$ the cell is enclosed in model layer $k$, i.e., $p_1 = \overline{\bar{p}}^n_{k-1/2} = A_{k-1/2} + B_{k-1/2}\,\overline{\bar{p}}^n_s$ and $p_2 = \overline{\bar{p}}^n_{k+1/2} = A_{k+1/2} + B_{k+1/2}\,\overline{\bar{p}}^n_s$, and at $t + \Delta t$, the cell arrives in the regular grid column (the column to the right) with cross-section area $\Delta A = \Delta x \Delta y$ in a layer, which generally do not coincide with a model layer $((p_1, p_2) \neq \left(\overline{\bar{p}}^n_{k-1/2}, \overline{\bar{p}}^n_{k+1/2}\right))$.

so that Eq. (3.18) becomes

$$M_{\delta V} = \frac{1}{g} \iint_{\delta A} \delta p \mathrm{d}x \mathrm{d}y = \frac{1}{g} \overline{\delta p}\, \delta A \tag{3.19}$$

This FV is supposed to move with the flow with vertical walls and without any flux of mass through its boundaries. Thus, the condition for mass conservation is

$$\frac{\mathrm{d}M_{\delta V}}{\mathrm{d}t} = \frac{1}{g} \frac{\mathrm{d}}{\mathrm{d}t} \left( \overline{\delta p}\, \delta A \right) = 0 \tag{3.20}$$

This is similar to Eq. (1.12) except that here the hydrostatic approximation has been applied.

*3.1.3.1. 3D trajectories* In Section 1.1, a traditional FV Lagrangian approach was applied with an integration of Eq. (1.12) in time along (exact) 3D upstream trajectories, starting at time $t$ from the irregular departure cell and ending at $t + \Delta t$ at the regular arrival cell (an Eulerian grid cell with area $\Delta A$ and vertical height difference $\Delta_k h = h_{k+1/2} - h_{k-1/2}$). This resulted in the prognostic Eq. (1.8) that was rewritten finally as Eq. (1.10). A prediction based on Eq. (1.10) would require a 3D integration over the irregular departure FV that would be very complicated and thus inefficient to do in practice. An even more serious objection against using Eq. (1.10) in practice is that it would require the construction of 3D trajectories that would require a priori known vertical velocities. In a quasi-hydrostatic atmosphere, however, the vertical velocity is a diagnostic quantity, which is determined by the diabatic heating and the instantaneous horizontal flow of mass and heat (Richardson [1922]). In pressure coordinates, which is more relevant here, it is even simpler; the pressure vertical velocity $\omega = \mathrm{d}p/\mathrm{d}t$ is diagnostically determined by the continuity equation from just the instantaneous horizontal flow of mass. Thus, the vertical displacements of mass during a time step $\Delta t$ must be determined as those displacements that ensure re-establishment of hydrostatic equilibrium after the given horizontal displacements of mass. Such considerations led Machenhauer and Olk [1997] to suggest a change in the traditional Lagrangian approach used in Section 1.1. They suggested construction and use of combined backward horizontal and forward vertical trajectories as indicated in Fig. 3.1. This idea of introducing quasi-horizontal Lagrangian trajectories and associated hydrostatically determined vertical velocities was concretized by Lauritzen, Kaas, Machenhauer and Lindberg [2008] in the HIRLAM-DCISL (Section 3.2) as described in the following. They called the combined backward horizontal and forward vertical trajectories hybrid trajectories. Also Lin and Rood [1998, 2004] introduced a so-called "floating Lagrangian control volume vertical coordinate" determined from hydrostatic balance, which in its essence is similar to the Lagrangian trajectories introduced by Machenhauer and Olk [1997], although the vertical displacements and thereby the vertical velocity in their scheme are defined from upstream trajectories determined by horizontal winds at the faces of the arrival Eulerian cell only and not from upstream winds as in DCISL.

The DCISL hybrid trajectories depart at $t = n\Delta t$ from the corner points of the irregular area $\delta_k A^n$ (the departure cell in the left column in Fig. 3.1) with a vertical extent

equal to that of a model layer, i.e., with an averaged pressure difference between its top and bottom equal to $\overline{\Delta_k \overline{p}^{\,n}}^{\delta} = \overline{\overline{p}^{\,n}_{k+1/2} - \overline{p}^{\,n}_{k-1/2}}^{\delta} = \frac{1}{\delta A}\iint\limits_{\delta A} \Delta_k \overline{p}^{\,n}\,\mathrm{d}x\mathrm{d}y$, the horizontal mean of $\Delta_k \overline{p}^{\,n}$ over the irregular departure area $\delta_k A^n$. Note that $\overline{(x)}^{\delta}$ denotes a horizontal mean over an irregular departure cell area $\delta_k A^n$, whereas $\overline{(x)}$ denotes a horizontal mean value over a regular arrival cell area $\Delta A$. The area-averaged full level pressure $\overline{\overline{p}^{\,n}_k}^{\delta}$ in the departure area is determined from $\overline{\overline{p}^{\,n}_{k+1/2}}^{\delta}$ and $\overline{\overline{p}^{\,n}_{k-1/2}}^{\delta}$ analogously to Eq. (3.17). The pressure at the trajectory starting point is interpolated from the $\overline{p}^{\,n}_k$ values in the surrounding grid cells. The trajectories are ending at time $t + \Delta t = (n+1)\,\Delta t$ at the corner points of an arrival cell with horizontal area $\Delta A$ located in a regular grid column (the right column in the figure) and in a layer with pressure thickness $\delta_k \overline{\hat{p}}^{\,n+1} = \left(\overline{\hat{p}}_{k+1/2}\right)^{n+1} - \left(\overline{\hat{p}}_{k-1/2}\right)^{n+1}$. The full level pressure in the arrival cell is $\overline{\hat{p}}^{\,n+1}_k = \left(\overline{\hat{p}}_{k+1/2}\right)^{n+1} \exp\left(-\overline{\alpha}^{\,n+1}_k\right)$; and the pressure at the trajectory end point is interpolated from the $\overline{\hat{p}}^{\,n+1}_k$ values in the surrounding grid cells. In general, $\delta_k \overline{\hat{p}}^{\,n+1}$ does not coincide with an Eulerian model layer (as also indicated in the figure).

The Lagrangian FV is supposed to move along the hybrid trajectories so that an integration of Eq. (3.20) from $t = n\Delta t$ to $t + \Delta t = (n+1)\Delta t$ results in the prognostic equation

$$\delta_k \overline{\hat{p}}^{\,n+1} \Delta A = \overline{\Delta_k \overline{p}^{\,n}}^{\delta} \delta_k A^{\,n}, \tag{3.21}$$

where the $n$ and $n+1$ superscripts refer to the time levels.

As indicated in Fig. 3.2, NLEV-1 FVs arrive at time $t + \Delta t$ in the same grid column in addition to cell $k$ considered above. These FVs originate from all the other model layers, one FV on top of the other. Here, it is assumed that a FV originating from model level $k$ ends up in the arrival column also as number $k$ from the top (without mixing with the one above and the one below). It is described below how the right-hand side of Eq. (3.21) can be estimated for each of the layers. Once the right-hand sides are known for each of the NLEV layers, $\delta_k \overline{\hat{p}}^{\,n+1}$ can be computed from Eq. (3.21) for each level $k$, and finally $\overline{\hat{p}}^{\,n+1}_{k-1/2}$ can be determined by summing up the hydrostatic weight of all the cells above:

$$\overline{\hat{p}}^{\,n+1}_{k-1/2} = \sum_{l=1}^{k-1} \delta_k \overline{\hat{p}}^{\,n+1}. \tag{3.22}$$

Hereby the vertical displacement from $\overline{p}^{\,n}_{k-1/2}$ to $\overline{\hat{p}}^{\,n+1}_{k-1/2}$ is determined in a hydrostatically fully consistent way (see Eq. (3.34) below).

Summing up the hydrostatic weight of all the NLEV cells yields the surface pressure

$$\overline{p}^{\,n+1}_s = \sum_{l=1}^{\mathrm{NLEV}} \delta_k \overline{\hat{p}}^{\,n+1}, \tag{3.23}$$

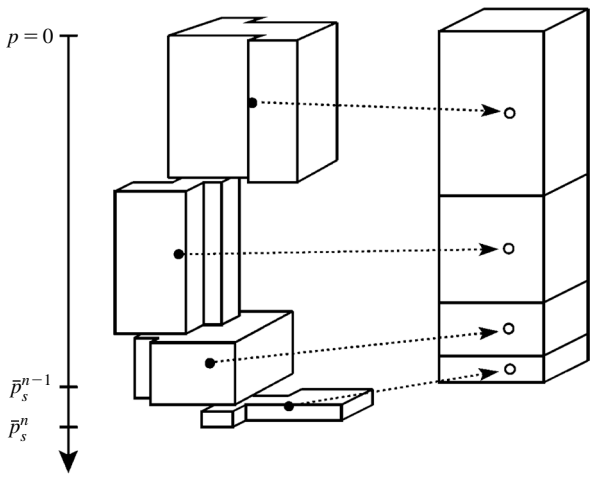


FIG. 3.2 Schematic illustration of the departure and arrival cells which make up the deformed column on the left and the regular column on the right, respectively. The cells move with vertical walls, and the horizontal extension is a polygon. In this figure, the polygon is as in the 2D DCISL scheme of NAIR and MACHENHAUER [2002] but the general idea applies to all DCISL schemes. The filled and unfilled circles indicate the center of mass of the departure and arrival cells, respectively. Note that the vertical levels in the arrival column $\hat{p}_{k+1/2}^{n+1}$ are the ones implied by the advection scheme and not the model levels, $p_{k+1/2}^{n+1}$, based on the hydrostatically determined surface pressure, $p_s^{n+1}$, and the predefined coefficients (Eq. (3.11)).

from which the pressure at the interfaces between the Eulerian model layers can be determined:

$$\overline{p}_{k-1/2}^{n+1} = A_{k-1/2} + B_{k-1/2}\,\overline{p}_s^{n+1}. \tag{3.24}$$

Now Eq. (3.23) we return to the determination of the right-hand side of Eq. (3.21), $\overline{\Delta_k \overline{p}^n}^{\delta} \delta A_k^n$. This is an iterative process where each iteration involves two steps: (I) at first the area $\delta A_k^n$ is determined by constructing hybrid trajectories from the corner points in the irregular departure cell to the corner points of the regular arrival cell. The sides in $\delta A_k^n$ are defined as the straight lines connecting the corner points. (II) Then, $\overline{\Delta_k \overline{p}^n}^{\delta} = \overline{\overline{p}_{k+1/2}^n - \overline{p}_{k-1/2}^n}^{\delta} = \frac{1}{\delta A} \iint\limits_{\delta A} \Delta_k \overline{p}^n \,\mathrm{d}x\mathrm{d}y$, the horizontal mean of $\Delta_k \overline{p}^{\,n}$ over the irregular departure area $\delta A_k^n$ is computed. Steps I and II are iterated.

*3.1.3.2. Trajectory algorithm (I)* Several trajectory algorithms have been developed (see Section 2.3); here we choose the hybrid trajectory scheme developed by LAURITZEN, KAAS, MACHENHAUER and LINDBERG [2008], which is used in the HIRLAM-DCISL dynamical core to be described in Section 3.2.

Since the FV is assumed to move with horizontal winds and vertical walls, the problem is 2D. Thus, we need to consider only the projection of the trajectories on a horizontal plane. The horizontal position vectors for the departure point, the arrival point, and the

trajectory midpoint are denoted $\vec{r}_*^{\,n}$, $\vec{r}^{\,n+1}$, and $\vec{r}_{*/2}^{\,n+1/2}$, respectively. The arrival point is defined as

$$\vec{r}^{\,n+1} = \vec{r}_*^{\,n} + \left(\vec{C}_1\left(\vec{V}_*^{\,n}\right) + \vec{C}_2\left(\tilde{\vec{V}}^{n+1}\right)\right). \tag{3.25}$$

For notational clarity, the level number $(k)$ has been suppressed. The trajectory consists of two parts:

(i) $\vec{C}_1\left(\vec{V}_*^{\,n}\right) = \vec{r}_{*/2}^{\,n+1/2} - \vec{r}_*^{\,n}$ is the vector from the departure point to the trajectory midpoint. It depends on $\vec{V}_*^{\,n}$, the horizontal velocity at the departure point at time $t = n\Delta t$. $\vec{C}_1\left(\vec{V}_*^{\,n}\right)$ is determined by one or more terms in a Taylor series expansion about the departure point:

$$\vec{C}_1 = \frac{\Delta t}{2}\vec{V}_*^{\,n} + \sum_{\upsilon=1}^{N-1} \frac{1}{(\upsilon+1)!}\left(\frac{\Delta t}{2}\right)^{\upsilon+1}\left(\frac{\mathrm{d}^\upsilon \vec{V}}{\mathrm{d}t^\upsilon}\right)_*^{n}, \tag{3.26}$$

where $N$ is the order of the expansion.

(ii) $\vec{C}_2\left(\tilde{\vec{V}}^{n+1}\right) = \vec{r}^{\,n+1} - \vec{r}_{*/2}^{\,n+1/2}$ is the vector from the trajectory midpoint to the arrival point. It depends on $\tilde{\vec{V}}^{n+1}$, a horizontal velocity at the arrival point extrapolated in time to $t + \Delta t = (n+1)\,\Delta t$. The time extrapolation, defined by $\tilde{\vec{V}}^{n+1} = 2\vec{V}^{\,n} - \vec{V}^{\,n-1}$. $\vec{C}_2\left(\tilde{\vec{V}}^{n+1}\right)$, is determined by one or more terms in a Taylor series expansion about the arrival point:

$$\vec{C}_2 = \frac{\Delta t}{2}\tilde{\vec{V}}^{n+1} - \sum_{\upsilon=1}^{N-1} \frac{1}{(\upsilon+1)!}\left(-\frac{\Delta t}{2}\right)^{\upsilon+1}\left(\frac{\mathrm{d}^\upsilon \tilde{\vec{V}}}{\mathrm{d}t^\upsilon}\right)^{n+1}. \tag{3.27}$$

In the HIRLAM-CISL dynamical core, the first two terms in the Taylor series are included ($N = 2$), and thus, estimates of the acceleration $\dfrac{\mathrm{d}\vec{V}}{\mathrm{d}t}$ are taken into account. The acceleration is approximated with $\dfrac{\mathrm{d}\vec{V}}{\mathrm{d}t} \approx \vec{V}\cdot\nabla\vec{V}$ (McGregor [1993]). It follows from Eq. (3.25) that the departure point is given by

$$\vec{r}_*^{\,n} = \vec{r}^{\,n+1} - \left(\vec{C}_1\left(\vec{V}_*^{\,n}\right) + \vec{C}_2\left(\tilde{\vec{V}}^{n+1}\right)\right). \tag{3.28}$$

*3.1.3.3. Upstream integral (II)* An "upstream integration"

$$\overline{\Delta_k \overline{p}^{\,n}}^{\delta} = \overline{\overline{p}_{k+1/2}^{\,n} - \overline{p}_{k-1/2}^{\,n}}^{\delta} = \frac{1}{\delta A}\iint_{\delta A} \Delta_k \overline{p}^{\,n}\,\mathrm{d}x\mathrm{d}y \tag{3.29}$$

determines the horizontal mean of $\Delta_k \overline{p}^{\,n}$ over the irregular departure area $\delta_k A^n$. It may be estimated by one of the DCISL methods described in Section 2. In HIRLAM-DCISL

(Lauritzen, Kaas, Machenhauer and Lindberg [2008]), two alternative methods are available, the method of Nair and Machenhauer [2002] and that of Nair, Scroggs and Semazzi [2002]. For each of the model layers, all the departure areas $\delta_k A^n$ cover the entire integration domain without overlaps or cracks. Consequently, it follows from Eq. (3.21) that mass is conserved both locally and globally.

*3.1.3.4. Iteration* In order to determine the departure point from Eq. (3.28), we need to iterate steps I and II (Sections 3.1.3.2 and 3.1.3.3) since, to start with, the pressure at the end point of the trajectory (at a corner of the regular arrival cell) is not known, and also the horizontal positions of the start point of the trajectory (the corner point of the departure cell) are unknown. Generally, for each grid cell, only the trajectory ending at the south-western corner point needs to be determined since adjacent cells share vertices.

*The first guess* (iteration number $\upsilon = 1$)

1. The winds in model layer $k$ are extrapolated in time to time level $n+1$ and interpolated to the arrival point, the south-western corner of a grid cell. The result is $\left(\tilde{\vec{V}}_k^{n+1}\right)_{SW}$. It is used to determine a first guess $\vec{C}_2$-value: $\left(\vec{C}_2\right)_k^1$.
2. A first guess $\vec{C}_1$-value $(\vec{C}_1)_k^1$ is determined using the time level $n$ winds $\overline{\vec{V}}_k^n$, interpolated also to the arrival point.
3. Thus, the first-guess departure point $\left(\vec{r}_*^n\right)^1$ is determined from Eq. (3.28) using the first guess C's: $\left(\vec{r}_*^n\right)^1 = \vec{r}^{n+1} - \left(\left(\vec{C}_1\right)_k^1 + \left(\vec{C}_2\right)_k^1\right)$.

*Iterations*

1. The upstream integral is made using the departure points $\left(\vec{r}_*^n\right)^\upsilon$ and an updated $(\upsilon+1)^{th}$ guess pressure $\left(\overline{\tilde{p}}_{k-1/2}^{n+1}\right)^{\upsilon+1}$ is determined using Eqs. (3.21) and (3.22). Corresponding full level pressures $\left(\hat{p}_k\right)^{\upsilon+1} = \left(\overline{\tilde{p}}_{k+1/2}\exp(-\alpha_k)\right)^{\upsilon+1}$ are computed and an interpolation of these to the south-western corner point gives $\left(\hat{p}_k\right)_{SW}^{\upsilon+1}$.
2. Vertical interpolation of $\left(\vec{C}_2\right)_k^\upsilon$ is made to this pressure $\left(\left(\hat{p}_k\right)_{SW}^{\upsilon+1}\right)$ giving $\left(\vec{C}_2\right)_k^{\upsilon+1}$.
3. Interpolate $\left(\vec{C}_1\right)_k^\upsilon$ to a preliminary point $\left(\vec{r}_*^n\right)_{prel.}^{\upsilon+1} = \vec{r}^{\,n+1} - \left(\left(\vec{C}_1\right)_k^\upsilon + \left(\vec{C}_2\right)_k^{\upsilon+1}\right)$ giving $\left(\vec{C}_1\right)_k^{\upsilon+1}$.
4. Determine the $(\upsilon+2)^{\text{th}}$ iteration location of the departure point $\left(\vec{r}_*^n\right)_k^{\upsilon+1} = \vec{r}^{n+1} - \left(\left(\vec{C}_1\right)_k^{\upsilon+1} + \left(\vec{C}_2\right)_k^{\upsilon+1}\right)$.
5. If repeat steps 1 to 4.

Note that this departure point algorithm does not require 3D interpolation. Only 1 and 2D interpolations are used.

*3.1.3.5. Values needed for the prediction of other invariants* To be used for the prediction of other invariants, the following values are stored each time-step for every grid cell:

Horizontal position of the final departure corner points, determined from Section 3.1.3.4:

$$\vec{r}_*^{\,n} = \vec{r}^{\,n+1} - \left(\vec{C}_1\left(\vec{V}_*^{\,n}\right) + \vec{C}_2\left(\tilde{\vec{V}}^{n+1}\right)\right). \tag{3.30}$$

These determine the areas $\delta_k A^n$ of the departure cells.

Full level pressure $\overline{\left(\overline{p}_k^{\,n}\right)}_*^{\delta}$ averaged over the departure area $\delta_k A^n$, determined from Eqs. (3.16) and (3.17):

$$\overline{\left(\overline{p}_k^{\,n}\right)}_*^{\delta} = \overline{\overline{p}_{k+1/2}^{\,n} \exp\left(-\overline{\alpha}_k^{\,n}\right)}^{\delta}, \tag{3.31}$$

where

$$\overline{\alpha}_k^{\,n} = \left\{ \begin{array}{ll} \ln 2 & k = 1 \\ 1 - \frac{\overline{p}_{k-1/2}^{\,n}}{\overline{\Delta_k p}^{\,n}} \ln\left(\frac{\overline{p}_{k+1/2}^{\,n}}{\overline{p}_{k-1/2}^{\,n}}\right) & k = 2, .., \mathrm{NLEV} \end{array} \right\}. \tag{3.32}$$

Mean pressure at the top of arrival cells, determined from Eq. (3.22):

$$\overline{\hat{p}}_{k-1/2}^{\,n+1} = \sum_{l=1}^{k-1} \delta_k \overline{\hat{p}}^{\,n+1}. \tag{3.33}$$

They determine together with Eq. (3.31) a mean value of *the vertical pressure velocity* $\omega = \mathrm{d}p/\mathrm{d}t$ of the cell, moving along the trajectories:

$$\overline{(\omega_k)}_{*/2}^{n+1/2} = \frac{1}{\Delta t}\left(\overline{\hat{p}}_k^{\,n+1} - \overline{\left(\overline{p}_k^{\,n}\right)}_*^{\delta}\right). \tag{3.34}$$

This may be used for parameterizations.

The mean surface pressure is determined from Eq. (3.23):

$$\overline{p}_s^{\,n+1} = \sum_{l=1}^{\mathrm{NLEV}} \delta_k \overline{\hat{p}}^{\,n+1}, \tag{3.35}$$

which determine the pressure at the top of the Eulerian cells (3.24):

$$\overline{p}_{k-1/2}^{\,n+1} = A_{k-1/2} + B_{k-1/2}\, \overline{p}_s^{\,n+1}. \tag{3.36}$$

*3.1.4. Conservation of mass of passive tracers*

The mass of a passive tracer with specific concentration $q_i$ in a FV $\delta V$ with vertical walls, horizontal cross-section $\delta A$, and thickness $\delta z = z_1 - z_2$ (see Fig. 3.1) is

$$M_{q_i \delta V} = \iint\limits_{\delta A} \left( \int_{z_2}^{z_1} q_i \rho \, \mathrm{d}z \right) \mathrm{d}x \mathrm{d}y. \tag{3.37}$$

Utilizing again the hydrostatic balance $\mathrm{d}p = -g\rho \, \mathrm{d}z$, the inner integral may be written as

$$\int_{z_2}^{z_1} q_i \rho \, \mathrm{d}z = \frac{1}{g} \int_{p_1}^{p_2} q_i \mathrm{d}p = \frac{1}{g} \widehat{q_i} \, (p_2 - p_1) = \frac{1}{g} \widehat{q_i} \delta p,$$

where $\widehat{q_i}$ is the pressure averaged specific concentration. Hereby, Eq. (3.37) becomes

$$M_{q_i \delta V} = \frac{1}{g} \iint\limits_{\delta A} \widehat{q_i} \delta p \mathrm{d}x \mathrm{d}y = \frac{1}{g} \overline{\widehat{q_i} \delta p} \, \delta A. \tag{3.38}$$

The FV mass conservation law for this passive tracer, which is supposed to move with the flow, with vertical walls, and without any flux of mass through its boundaries, is then

$$\frac{\mathrm{d}M_{q_i \delta V}}{\mathrm{d}t} = \frac{1}{g} \frac{\mathrm{d}}{\mathrm{d}t} \left( \overline{\widehat{q_i} \delta p} \, \delta A \right) = 0. \tag{3.39}$$

The FV is supposed to move along the hybrid trajectories determined for the continuity equation so that an integration of Eq. (3.39) from $t = n\Delta t$ to $t + \Delta t = (n+1)\,\Delta t$ results in the prognostic equation

$$\left( \overline{\widehat{q_i}}^{\delta} \right)_k^{n+1} \delta_k \overline{\hat{p}}^{n+1} \, \Delta A = \overline{\left( \overline{\widehat{q_i}}^{\Delta} \right)_k^n \Delta_k \overline{p}^n}^{\delta} \, \delta A_k^n, \tag{3.40}$$

where $\widehat{x}^{\delta}$ and $\widehat{x}^{\Delta}$ denote vertical mean values over $\delta_k \hat{p}$ and $\Delta_k p$, respectively. Here and in the following, we make the *discretization assumption*: the horizontal mean over the arrival area of a product is equal to the product of the horizontal mean values of the factors.

*3.1.5. Conservation of mass of water vapor*

Apart from forcing terms, the derivation of the discretized prognostic equation for water vapor is identical to the above for passive tracers. The mass of water vapor with specific humidity $q_v$ in a FV $\delta V$ with vertical walls, horizontal cross-section $\delta A$, and pressure thickness $\delta p$ is

$$M_{q_v \delta V} = \frac{1}{g} \iint\limits_{\delta A} \widehat{q_v} \delta p \mathrm{d}x \mathrm{d}y = \frac{1}{g} \overline{\widehat{q_v} \delta p} \, \delta A. \tag{3.41}$$

The FV mass conservation law for water vapor is then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\widehat{\overline{q_v \delta p}}\,\delta A\right) = \overline{\left(\widehat{P}_{q_v} + \widehat{K}_{q_v}\right)\delta p}\,\delta A, \tag{3.42}$$

and the discretized prediction equation becomes

$$\left(\overline{\widehat{q}_v}^{\delta}\right)_k^{n+1} \delta_k \overline{\widehat{p}}^{n+1}\,\Delta A = \overline{\left(\overline{\widehat{q}_v}^{\Delta}\right)_k^{n} \Delta_k \overline{p}^{n}}^{\delta}\,\delta A_k^n + \Delta t\,\overline{\left(\overline{\widehat{P}_{q_v}}^{\Delta} + \overline{\widehat{K}_{q_v}}^{\Delta}\right)_k^{n+1/2} \delta_k \overline{p}^{n+1/2}}^{\delta_{n+1/2}}\,\delta A_k^{n+1/2} \tag{3.43}$$

For simplicity, here and in the following, an instantaneous forcing is assumed to work on the FV at time $t + \Delta t/2$ when it is at the midpoint of the trajectory. Of course, it should ideally be averaged along the trajectory. In reality, it might be convenient to treat the forcing as in an existing semi-Lagrangian model. Thus, in the current HIRLAM-DCISL, for instance, the physics are added at time level $n + 1$ at the arrival cell as in HIRLAM.

*3.1.6. Conservation of total energy*

The total energy $E$ is the sum of the internal energy $E_i$, the potential energy $E_p$, and the kinetic energy $E_k$.

*The internal energy*

$$E_i = \int_{z_2}^{z_1} c_v T\,\rho\,\mathrm{d}z = 1/\mathrm{g}\int_{p_1}^{p_2} c_v T \mathrm{d}p. \tag{3.44}$$

*The potential energy*

$$\begin{aligned} E_p &= \int_{z_2}^{z_1} gz\,\rho\,\mathrm{d}z = \int_{p_1}^{p_2} z\,\mathrm{d}p \\ &= z_2 p_2 - z_1 p_1 - \int_{z_1}^{z_2} p\,\mathrm{d}z \\ &= z_2 p_2 - z_1 p_1 + 1/\mathrm{g}\int_{p_1}^{p_2} RT\,\mathrm{d}p, \end{aligned} \tag{3.45}$$

where integration by parts as well as the equation of state $p = \rho RT$ has been used.

*The "total potential"* energy is then

$$\begin{aligned} E_i + E_p &= z_2 p_2 - z_1 p_1 + 1/\mathrm{g}\int_{p_1}^{p_2} (c_v + R)T\,\mathrm{d}p \\ &= 1/\mathrm{g}(\delta(\phi\, p) + \int_{p_1}^{p_2} c_p T\,\mathrm{d}p), \end{aligned} \tag{3.46}$$

where $\delta(\phi\, p) = \phi_2 p_2 - \phi_1\, p_1$ with $\phi = gz$ and, in addition, the relation $c_p = R + c_v$ has been used. Here, respectively, $c_p$, $c_v$, and $R$ are the specific heat capacity at constant pressure, the specific heat capacity at constant volume, and the individual gas constants, all for moist air.

*The kinetic energy* is defined as

$$E_k = 1/(2\mathrm{g}) \int_{p_1}^{p_2} \left(u^2 + v^2\right) \mathrm{d}p. \tag{3.47}$$

Hence, *the total energy* becomes

$$E = E_i + E_p + E_k = 1/\mathrm{g}(\delta(\phi\, p) + \int_{p_1}^{p_2} c_p T \mathrm{d}p + 1/2 \int_{p_1}^{p_2} \left(u^2 + v^2\right) \mathrm{d}p). \tag{3.48}$$

Introducing vertical mean values and including the level index $k$, Eq. (3.48) may be written as

$$E = 1/\mathrm{g}(\delta_k(\phi\, p) + \left(\widehat{c_p}\widehat{T}\right)_k \delta_k p + (1/2) \left(\widehat{u}_k^2 + \widehat{v}_k^2\right) \delta_k p), \tag{3.49}$$

where we have assumed that $\widehat{c_p}T = \widehat{c_p}\widehat{T}$ and $\widehat{u^2} = \widehat{u}\,\widehat{u}$ and similarly for the northward component v. This is consistent with assuming that the variables are independent of pressure in the layer.

The total energy of a FV $\delta V$ with vertical walls and horizontal cross-section $\delta A$ and thickness $\delta z = z_2 - z_1$ (see Fig. 3.1) is then $\overline{E}\delta A$, where $\overline{(\,)} = \frac{1}{\delta A} \iint_{\delta A} (\,)\, \mathrm{d}x\mathrm{d}y$. Finally, we can construct *the FV total energy conservation law:*

$$\frac{\mathrm{d}\left(\overline{E}\delta A\right)}{\mathrm{d}t} = \iiint_{\delta V} \left(\vec{V} \cdot \left(\vec{P}_{\vec{V}} + \vec{K}_{\vec{V}}\right) + c_p \left(P_T + K_T\right)\right) \mathrm{d}p\, \mathrm{d}x\, \mathrm{d}y, \tag{3.50}$$

where the right-hand side follows from a derivation directly from the primitive equations (see MACHENHAUER [1994])

The FV is supposed to move along the hybrid trajectories determined from the continuity equation so that an integration of Eq. (3.50) from $t = n\Delta t$ to $t + \Delta t = (n+1)\Delta t$ results in the prognostic equation

$$\begin{aligned}
&\delta_k \left(\overline{\hat{\phi}}\overline{\hat{p}}\right)^{n+1} + \left(\left(\overline{\overline{c_p}^{\delta}}\,\overline{\overline{T_k}^{\delta}}\right) + 1/2 \left(\left(\overline{\overline{u_k}^{\delta}}\right)^2 + \left(\overline{\overline{v_k}^{\delta}}\right)^2\right) \delta_k \overline{\hat{p}}\right)^{n+1} \Delta A \\
&\quad = \overline{\Delta_k \left(\overline{\phi}\overline{p}\right)^n + \left(\left(\overline{\overline{c_p}^{\Delta}}\,\overline{\overline{T_k}^{\Delta}}\right) + 1/2 \left(\left(\overline{\overline{u_k}^{\Delta}}\right)^2 + \left(\overline{\overline{v_k}^{\Delta}}\right)^2\right) \Delta_k \overline{p}\right)^n}^{\delta} \delta A_k^n \\
&\qquad + \overline{\Delta t \left(\overline{\vec{V}} \cdot \left(\overline{\vec{P}}^{\Delta}_{\vec{V}} + \overline{\vec{K}}^{\Delta}_{\vec{V}}\right) + \overline{c_p}^{\Delta} \left(\overline{P_T}^{\Delta} + \overline{K_T}^{\Delta}\right)\right)^{n+1/2} \delta_k \overline{p}^{n+1/2}}^{\delta_{n+1/2}} \delta A_k^{n+1/2}.
\end{aligned} \tag{3.51}$$

*3.1.7. Conservation of entropy*

The specific entropy is $S = c_p \ln\theta$, where $\theta$ is the potential temperature defined by $\theta = T\,(p/p_0)^{R/c_p}$ and $p_0 = 1000\,hPa$. Note that here $R$ and $c_p$ are for moist air and not, as usual, for dry air. The total entropy of a FV $\delta V$ with vertical walls and horizontal cross-section $\delta A$ and thickness $\delta p = p_2 - p_1$ (see Fig. 3.1) is then

$$S_{\delta V} = \frac{1}{g}\iiint\limits_{\delta V} c_p \ln\theta \,\mathrm{d}p\,\mathrm{d}x\,\mathrm{d}y. \tag{3.52}$$

Introducing vertical and horizontal mean values and at the same time including the level number $k$, Eq. (3.52) may be written as

$$S_{\delta V} = \frac{1}{g}\overline{\left(\widehat{c}_p \ln\widehat{\theta}\right)_k \delta_k p}\,\delta A_k, \tag{3.53}$$

from which we get *the* FV *entropy conservation law:*

$$\frac{\mathrm{d}S_{\delta V}}{\mathrm{d}t} = \iiint\limits_{\delta V}\left(\frac{c_p\,(P_T + K_T)}{T} + \left\{\left(c_{pv} - c_{pd}\right)\ln T + (R_v - R_d)\ln(p/p_0)\right\}\left(P_{q_v} + K_{q_v}\right)\right)\mathrm{d}p\,\mathrm{d}x\,\mathrm{d}y, \tag{3.54}$$

where again the right-hand side follows from a derivation directly from the primitive equations (see MACHENHAUER [1994]).

The FV is supposed to move along the hybrid trajectories determined for the continuity equation so that an integration of Eq. (3.54) from $t = n\Delta t$ to $t + \Delta t = (n+1)\,\Delta t$ results in the prediction equation

$$\left(\overline{\widehat{c}_p}^{\,\delta} \ln\overline{\widehat{\theta}}^{\,\delta}\right)_k^{n+1} \delta_k \overline{\hat{p}}^{\,n+1}\,\Delta A = \overline{\left(\overline{\widehat{c}_p}^{\,\Delta} \ln\overline{\widehat{\theta}}^{\,\Delta}\right)_k^{n} \Delta_k \overline{p}^{\,n}\;\delta A_k^n}^{\,\delta} + \Delta t\overline{\left(\frac{\widehat{c}_p^{\,\Delta}\left(P_T^{\Delta} + \widehat{K}_T^{\Delta}\right)}{\widehat{T}} + \left\{\left(\widehat{c}_{pv}^{\,\Delta} - \widehat{c}_{pd}^{\,\Delta}\right)\ln\widehat{T}^{\Delta} + \left(\widehat{R}_v^{\Delta} - R_d\right)\ln\left(p_k/p_0\right)\right\} \times \left(\widehat{P}_{q_v}^{\Delta} + \widehat{K}_{q_v}^{\Delta}\right)\right)^{n+1/2} \delta_k p^{n+1/2}}^{\,\delta_{n+1/2}} \delta A_k^{\,n+1/2}. \tag{3.55}$$

Here, a further "*discretization assumption*" is made: the mean value of $\ln x$ over the arrival area is set equal to the logarithm of the mean value.

*3.1.8. Conservation of angular momentum*

The absolute angular momentum per unit mass of air is

$$m = (\Omega a\cos\varphi + u)a\cos\varphi \tag{3.56}$$

or

$$m = \Omega a^2 \cos^2 \varphi + ua \cos \varphi, \tag{3.57}$$

where $u$ is the eastward component of velocity, $a$ is the radius of the Earth (for simplicity and as usual assumed constant), and $\varphi$ is the latitude.

The absolute angular momentum of the mass in a FV $\delta V$ with vertical walls and horizontal cross-section $\delta A$ and thickness $\delta p = p_1 - p_2$ (see Fig. 3.1) is then

$$m_{\delta V} = \frac{1}{g} \iint_{\delta A} \int_{p_1}^{p_2} (\Omega a \cos \varphi + u) a \cos \varphi \, \mathrm{d}p \, \mathrm{d}x \, \mathrm{d}y. \tag{3.58}$$

Introducing vertical and horizontal mean values and at the same time adding the level number $k$, Eq. (3.58) may be written as

$$m_{\delta V} = \frac{1}{g} \left( \overline{\widehat{m}_k \, \delta_k p} \, \delta A_k \right) = \frac{1}{g} \left( \overline{\left( \Omega a^2 \cos^2 \varphi_k + \widehat{u}_k a \cos \varphi_k \right) \delta_k p} \, \delta A_k \right) \tag{3.59}$$

So, *the absolute angular momentum conservation law* becomes

$$\frac{\mathrm{d}(m_{\delta V})}{\mathrm{d}t} = \frac{1}{g} \iint_{\delta A} \int_{p_1}^{p_2} \left( -\left( \frac{\partial \phi}{\partial \lambda} + R_d T_v \frac{\partial}{\partial \lambda} (\ln p) \right) + (P_u + K_u) a \cos \varphi \right) \mathrm{d}p \, \mathrm{d}x \, \mathrm{d}y. \tag{3.60}$$

The FV is supposed to move along the hybrid trajectories determined for the continuity equation so that an integration of Eq. (3.60) from $t = n\Delta t$ to $t + \Delta t = (n+1)\Delta t$ results in the prognostic equation

$$\begin{aligned} &\left( \Omega a^2 \cos^2 \varphi + \left( \overline{\widetilde{u}_k}^{\delta} \right)^{n+1} a \cos \varphi \right) \delta_k \overline{\widehat{p}}^{n+1} \Delta_k A \\ &= \overline{\left( \Omega a^2 \cos^2 \varphi + \left( \overline{\widetilde{u}_k}^{\Delta} \right)^{n} a \cos \varphi \right) \Delta_k \overline{p}^{n} \;\; \delta A_k^n}^{\delta} \\ &+ \Delta t \overline{\left\{ -\left( \frac{\partial \widehat{\phi}^{\Delta}}{\partial \lambda} + R_d \widehat{T}_v^{\Delta} \frac{\partial}{\partial \lambda} \left( \ln \widehat{p}^{\Delta} \right) \right) + \left( \widehat{P}_u^{\Delta} + \widehat{K}_u^{\Delta} \right) \right\} a \cos \varphi \, \delta_k p \qquad \delta A_k^{n+1/2}}^{\delta_{n+1/2}}. \end{aligned} \tag{3.61}$$

### *3.1.9. Choice of invariants*

As mentioned by THUBURN [2006], the continuous adiabatic frictionless equations have an infinite number of invariants. In CSCL, we have chosen to fulfill those conservation laws which are fundamental for the dynamics and thermodynamics of the atmosphere, namely the basic conservation laws from which the primitive equations are derived. It should be mentioned that other invariants might substitute for some of those selected above. One obvious example is to replace the conservation of angular momentum by conservation of Ertel potential vorticity. In this case, we would still have a complete set of prognostic equations. An advantage of using angular momentum is that it leads to a direct separation of the $u$ and $v$ contributions to kinetic energy.

*3.1.10. Explicit integration procedure*

Together with the diagnostic hydrostatic equation (Eq. (3.13)), the discretized FV versions of the conservation laws for (1) mass of moist air (Eq. (3.21)), (2) mass of passive tracers (Eq. (3.40)), (3) mass of water vapor (Eq. (3.43)), (4) total energy (Eq. (3.51)), (5) entropy (Eq. (3.55)), and (6) angular momentum (Eq. (3.61)) constitute a complete prognostic system equivalent to the primitive equations. Initially, the following quantities should be given: Eulerian grid cell area averaged surface pressure $\overline{p}_s$ for each vertical grid column, grid cell averaged values of temperature $\overline{\widehat{T}}_k^{\Delta}$, specific humidity $\left(\overline{\widehat{q}_v}^{\Delta}\right)_k$, specific concentration of passive tracers $\left(\overline{\widehat{q}_i}^{\Delta}\right)_k$, and eastward and northward horizontal velocity components, $\overline{\widehat{u}}_k^{\Delta}$ and $\overline{\widehat{v}}_k^{\Delta}$. These would be the history carrying variables. Explicit time-stepping with such a system would be relatively easy. At first the continuity equation (Eq. (3.21)) is solved as described in Section 3.1.3. The outcome, summarized in Section 3.2.1, is the grid cell averaged surface pressure $\overline{p}_s^{n+1}$. In addition, the hybrid trajectories needed for the transport of all the other invariants and diagnostic values of $\omega$ (which might be needed in the physical parameterization package) are determined. Next step is to solve the continuity equations for water vapor (Eq. (3.43)) and passive tracers (Eq. (3.40)) giving the updated cell averaged prognostic variables $\left(\overline{\widehat{q}_v}^{\delta}\right)_k^{n+1}$ and $\left(\overline{\widehat{q}_i}^{\delta}\right)_k^{n+1}$. The cell averaged values, $\delta_k \overline{\hat{p}}$, $\left(\overline{\widehat{q}_v}^{\delta}\right)_k^{n+1}$, and $\left(\overline{\widehat{q}_i}^{\delta}\right)_k^{n+1}$, over the Lagrangian cells, which originally are transported into a vertical Eulerian column, must be remapped into the Eulerian cells, giving $\Delta_k \overline{p}$, $\left(\overline{\widehat{q}_v}^{\Delta}\right)_k^{n+1}$, and $\left(\overline{\widehat{q}_i}^{\Delta}\right)_k^{n+1}$. Next, the conservation law for entropy (Eq. (3.55)) is solved giving (after some algebra) the updated cell averaged prognostic variables $\overline{\widehat{T}}_k^{\delta}$. Again, the cell-averaged Lagrangian values $\overline{\widehat{T}}_k^{\delta}$ of the cells must be remapped into the Eulerian cells giving $\overline{\widehat{T}}_k^{\Delta}$. Next, the conservation law of angular momentum (Eq. (3.61)) is solved giving (after some algebra) the updated cell averaged prognostic variables $\overline{\widehat{u}}_k^{\delta}$. Again, they must be remapped into the Eulerian cells giving $\overline{\widehat{u}}_k^{\Delta}$. Finally, the conservation law of total energy (Eq. (3.51)) is solved giving (after some algebra and vertical remapping) the updated cell averaged prognostic variables $\overline{\widehat{v}}_k^{\Delta}$.

*3.1.11. Feasibility of a CSCL model*

To the author's knowledge, a dynamical core that includes FV versions of all the conservation laws considered here has not yet been realized in spite of the "fact" that (as mentioned in the introduction) it may be expected that a simultaneous exact conservation of all the fundamental physical invariants valid for the atmosphere will result in a particular fast convergence to any "true" solution. The reason for not realizing such a system seems to be difficulties with the application of any of the popular fast-wave-stabilizing techniques, i.e., the semi-implicit or the split-explicit technique, which would eliminate fast wave CFL restrictions on the time-step. MACHENHAUER and OLK [1997] succeeded in the construction of two 1D shallow water semi-implicit semi-Lagrangian dynamical cores, one that conserves mass and total energy and another that conserves

mass and angular moment (see Section 3.2.2). A 1D shallow water system has just two prognostic variables. Therefore, just two invariants can be conserved exactly. In both cases, the fast wave CFL restriction on the time-step was eliminated and large time steps could be used. However, it seems difficult to extend this to 2 and 3Ds (see Section 3.2.3). Another possibility would be the application of the split-explicit technique. However, as already noted in Section 1, this possibility was abandoned by MACHENHAUER and OLK [1997] for FV models because when splitting the system of continuous equations into an advective part (which should use large time steps) and an adjustment gravity wave part (which should use short time steps), it was found that neither of the sub systems was conserving momentum or total energy. Consequently, it seems unlikely that these invariants could be conserved exactly for the full system in any FV version. In the two examples of dynamical cores with FV techniques, which are described in the following two sections, the continuity equations are solved with the FV technique so that mass is conserved exactly. In the NCAR-FFSL also potential temperature is conserved for adiabatic and friction-free flow. In the HIRLAM-DCISL, a semi-implicit time-stepping is implemented, thereby allowing large time-steps for all variables at the expense of solutions to elliptic Helmholtz equations. This has been feasible because just the continuity equation is solved with the FV technique, while the other primitive equations are kept in their original form, i.e., Eq. (3.1) with $u$ and $v$ and Eq. (3.2) with $T$ as prognostic variables. Furthermore, a special "predictor-corrector" approach (see Section 3.2.3.2) has been used successfully in the semi-implicit continuity equation. In the other system, the NCAR-FFSL, an explicit time-stepping scheme, is used. Consequently, shorter time-steps have to be used for the dynamical core. Tracers (including water vapor) and physical parameterization can, however, be updated with long time-steps. Such a time-stepping procedure would, of course, be possible also in a CSCL model, which then, most likely, would be comparable to the NCAR-FFSL in efficiency.

### *3.2. The HIRLAM-DCISL with a departure cell-integrated semi-implicit semi-Lagrangian dynamical core*

The HIRLAM-DCISL, described in details by LAURITZEN, KAAS, MACHENHAUER and LINDBERG [2008], is a pioneering example of a FV model based on the Lagrangian form of the continuity equation (Eq. (1.8)). The continuity equations for moist air, water vapor, cloud water, and miscellaneous passive tracers are updated each time-step using a DCISL FV scheme while the remaining prognostic equations are in finite difference form and solved using a traditional upstream semi-Lagrangian scheme. It has been developed from the HIRLAM system, (KÄLLÉN [1996] and UNDÉN et al. [2002]). The new HIRLAM-DCISL uses the same horizontal C-grid (ARAKAWA and LAMB [1977]) and vertical LORENZ [1960] staggering of variables as the HIRLAM model (see Section 3.1.2). Also the lateral boundary relaxation scheme is the same. HIRLAM-DCISL is the first model that combines a FV semi-Lagrangian integration scheme with a semi-implicit treatment of gravity wave terms. Thus, this semi-implicit version is absolutely stable as long as the trajectories do not cross (Lipschitz criterion), which in practice means that it runs stably with relatively long time-steps, similar to those used by HIRLAM, and still sufficiently small compared to the time scale of weather system developments. In

Section 3.2.1, we shall introduce at first an explicit version of HIRLAM-DCISL and then, in Sections 3.2.2 and 3.2.3, the changes needed to make it semi-implicit are discussed.

*3.2.1. Explicit HIRLAM-DCISL*

The explicit continuity equation for moist air is solved for each model layer as described in Section 3.1.3. (see Eq. (3.21)). Hybrid trajectories determine the irregular upstream departure area $\delta_k A^n$, and an "upstream integration" determines the horizontal mean of $\Delta_k \overline{p}^n$ over the departure area $\delta_k A^n$ (3.29). Here $\Delta_k \overline{p}^n$ is defined as

$$\Delta_k \overline{p}^n = \overline{p}^n_{k+1/2} - \overline{p}^n_{k-1/2}. \tag{3.62}$$

The departure cells are the same for all tracers, including water vapor, and Lagrange interpolations between the hybrid trajectory departure points determine the departure points for temperature $T$ and the velocity components $u$ and $v$. In HIRLAM-DCISL, two alternative upstream integration methods are available, the method of Nair and Machenhauer [2002] and that of Nair, Scroggs and Semazzi [2002]. The mean top pressures of the arrival cells $\overline{\hat{p}}^{n+1}_{k-1/2}$ are determined hydrostatically from Eq. (3.22), i.e., from the Lagrangian pressure thicknesses $\delta_k \overline{\hat{p}}^{n+1}$ in Eq. (3.33). Together with Eq. (3.62), these values determine a mean value of the vertical pressure velocity $\omega = \mathrm{d}p/\mathrm{d}t$ along the trajectory (Eq. (3.34)). This $\omega$ is consistent with the hydrostatic assumption and the horizontal flow, contrary to the inconsistent vertical velocities, based on partly Eulerian solutions to the continuity equation, which are applied in traditional semi-Lagrangian models such as HIRLAM. $\omega$ is used in the thermodynamic equation (Eq. (3.2)) in the energy conversion term $\frac{\alpha\omega}{c_p} = \frac{R_d T_v}{c_p}\frac{\omega}{p}$, which is approximated with

$$\Delta t \left[\left(\frac{R_d T_v}{c_p}\frac{\omega}{p}\right)_k\right]^{n+1} = \frac{R_d}{c_p}\left[T^n_v + \tilde{T}^{n+1}_v\right]_k \left[\frac{\overline{\hat{p}}^{n+1}_k - \overline{\left(\overline{p}^n_k\right)_*}^{\delta}}{\overline{\hat{p}}^{n+1}_k + \overline{\left(\overline{p}^n_k\right)_*}^{\delta}}\right]. \tag{3.63}$$

The hydrostatic mean surface pressure (Eq. (3.23)) is the weight of all NLEV model layers above the surface:

$$\overline{p}^{n+1}_s = \sum_{l=1}^{\mathrm{NLEV}} \delta_k \overline{\hat{p}}^{n+1}, \tag{3.64}$$

determining the top pressure of Eulerian cells (Eq. (3.24))

$$\overline{p}^{n+1}_{k-1/2} = A_{k-1/2} + B_{k-1/2}\,\overline{p}^{n+1}_s. \tag{3.65}$$

The explicit continuity equations for passive tracers (Eq. (3.40)) and water vapor (Eq. (3.43)) are

$$\left(\overline{\overline{q}_i}^{\delta}\right)^{n+1}_k \delta_k \overline{\hat{p}}^{n+1} \Delta A = \overline{\left(\overline{\overline{q}_i}^{\Delta}\right)^n_k \Delta_k \overline{p}^n}^{\delta} \delta A^n_k \tag{3.66}$$

and

$$\left(\overline{\widetilde{q}_v}^{\delta}\right)_k^{n+1}\delta_k\overline{\tilde{p}}^{n+1}\Delta A=\overline{\left(\overline{\widetilde{q}_v}^{\Delta}\right)_k^{n}\Delta_k\overline{p}^{n}}^{\delta}\ \delta A_k^n + \Delta t\overline{\left(\overline{\widetilde{P}_{q_v}}^{\Delta}+\overline{\widetilde{K}_{q_v}}^{\Delta}\right)_k^{n+1/2}\delta_k\overline{p}^{n+1/2}}^{\delta_{n+1/2}}\ \delta A_k^{n+1/2}, \quad (3.67)$$

respectively, determine updated specific concentrations, $\left(\overline{\widetilde{q}_i}^{\delta}\right)_k^{n+1}$ and $\left(\overline{\widetilde{q}_v}^{\delta}\right)_k^{n+1}$, in Lagrangian arrival cells $\left(\delta V=\delta\hat{p}\Delta V\right)$ from $\left(\overline{\widetilde{q}_i}^{\Delta}\right)_k^{n}$ and $\left(\overline{\widetilde{q}_v}^{\Delta}\right)_k^{n}$ plus Eq. (3.62). Finally, the updated specific concentrations, $\left(\overline{\widetilde{q}_i}^{\Delta}\right)_k^{n+1}$ and $\left(\overline{\widetilde{q}_v}^{\Delta}\right)_k^{n+1}$, in the Eulerian cells $(\Delta V=\Delta p\Delta A)$ are determined from $\left(\overline{\widetilde{q}_i}^{\delta}\right)_k^{n+1}$ and $\left(\overline{\widetilde{q}_v}^{\delta}\right)_k^{n+1}$ by 1D vertical remappings.

The discretized explicit momentum and thermodynamic equations are straightforward grid-point semi-Lagrangian and finite difference approximations to Eqs. (3.1) and (3.2), respectively (see KÄLLÉN [1996] and UNDÉN et al. [2002]), except that in the thermodynamic equation the consistent energy conversion term is approximated consistently with (3.63). Regarding the addition of the physics in Eq. (3.67): since DMI-HIRLAM adds the physics at the arrival level (no averaging along the trajectory), that procedure was also adopted in HIRLAM-DCISL. Of course, it should ideally be done as indicated in Eq. (3.67).

### *3.2.2. 1D semi-implicit CSCL shallow water models*

MACHENHAUER and OLK [1997] made a preliminary study, in which a successful implementation of a semi-implicit scheme was made in two different cell-integrated versions of the simple 1D shallow water model. One version conserves mass and momentum and another version conserves mass and total energy. The momentum and continuity equations for the 1D shallow water model are, respectively,

$$\frac{\mathrm{d}u}{\mathrm{d}t}+g\frac{\partial h}{\partial x}=0 \quad \text{and} \quad (3.68)$$

$$\frac{\mathrm{d}h}{\mathrm{d}t}+h\frac{\partial u}{\partial x}=0, \quad (3.69)$$

where $u$ is velocity (constant with height), $h$ height (of the fluid surface), and $x$ distance. A periodic domain is assumed $0\le x\le L$. The implementation of a semi-implicit scheme in the cell-integrated model versions will be compared with the traditional approach in a traditional finite difference grid-point model based on Eqs. (3.68) and (3.69). The traditional explicit semi-Lagrangian prediction equations are

$$u_{\exp}^{n+1}=u^n-\frac{\Delta tg}{\Delta x}(\delta(h))^{n+1/2}, \quad (3.70)$$

$$h_{\text{exp}}^{n+1} = h^n - \frac{\Delta t}{\Delta x}(h\delta u)^{n+1/2}, \tag{3.71}$$

where $\delta$ indicates a finite difference operator. Like in HIRLAM, we choose a second-order centered finite difference. A semi-implicit system corresponding to this system is obtained simply by averaging $(n+1)\Delta t$ and $n\Delta t$ values of the pressure gradient term and the linear part of the divergence term ($H\delta u$) along the trajectories instead of taking them at $(n+1/2)\Delta t$ at the midpoint of the trajectory as in Eqs. (3.70) and (3.71). The resulting equations may be written as

$$u^{n+1} = u_{\text{exp}}^{n+1} - \frac{\Delta tg}{2\Delta x}\Big[\delta h^{n+1} + \delta h^n - 2(\delta(h))^{n+1/2}\Big], \tag{3.72}$$

$$h^{n+1} = h_{\text{exp}}^{n+1} - \frac{\Delta tH}{2\Delta x}\Big[\delta u^{n+1} + \delta u^n - 2(\delta(u))^{n+1/2}\Big]. \tag{3.73}$$

We note that a linear version of Eqs. (3.70) and (3.71), linearized around a state at rest with a mean fluid height $H$, has gravity wave solutions. These solutions are characterized by purely divergent velocity fields. The height field and the divergence field in these solutions oscillate with a frequency $v \approx 2\pi/\sqrt{gH}$ driven by an oscillating pressure gradient force $-g\partial h/\partial x$ and divergence $-H\partial u/\partial x$, respectively. This explains intuitively why the implicit system, obtained by averaging these terms, can be expected to have stable gravity wave solutions. The system (Eqs. (3.72) and (3.73)) is absolutely stable (as long as the trajectories do not cross).

We may write the system as

$$u^{n+1} = q_1 - \frac{\Delta tg}{2\Delta x}\delta h^{n+1}, \tag{3.74}$$

$$h^{n+1} = q_2 - \frac{\Delta tH}{2\Delta x}\delta u^{n+1}. \tag{3.75}$$

In $q_1$ and $q_2$, the terms which do not depend on values at $(n+1)\Delta t$ have been collected. Applying the $\delta$ operator on Eq. (3.75) and substituting in Eq. (3.74) gives

$$u^{n+1} - \frac{gH\Delta t^2}{4\Delta x^2}\delta^2 u^{n+1} = q_1 - \frac{g\Delta t}{2\Delta x}\delta q_2. \tag{3.76}$$

This is an elliptic equation which can be solved to give $u^{n+1}$ and then (3.75) can be used to determine $h^{n+1}$.

The fact that the elliptic equation is with constant coefficients, a so-called Helmholtz equation, means that it is relatively easy and fast to solve. In operational semi-implicit multi-level models as HIRLAM, a series of elliptic equations must be solved. This reduces the advantage of large time-steps. It is therefore important that the elliptic equations in any new implementation of the semi-implicit scheme are kept as simple and fast to solve as possible. The strategy of Machenhauer and Olk [1997] for the present

models and later that of Lauritzen, Kaas, Machenhauer, Lindberg, [2006, 2008] for the HIRLAM-DCISL has been to do the semi-implicit implementation in the FV model in such a way that the resulting elliptic equation becomes similar to that of the traditional model it replaces.

Now let us derive the FV models corresponding to Eqs. (3.70) and (3.71). The mass, momentum, and total energy in a volume of length $\delta x$ are, respectively,

$$M_{\delta x} = \rho \int_x^{x+\delta x} \left( \int_0^h z \mathrm{d}z \right) \mathrm{d}x = \rho \overline{h}\ \delta x, \tag{3.77}$$

$$m_{\delta x} = \rho \int_x^{x+\delta x} \left( u \int_0^h \mathrm{d}z \right) \mathrm{d}x = \rho \overline{uh}\, \delta x, \text{ and} \tag{3.78}$$

$$E = \rho \int_x^{x+\delta x} \int_0^h (gz + (1/2)(u^2)) \mathrm{d}z \mathrm{d}x = (1/2)\rho \left( g\overline{h^2} + \overline{hu^2} \right) \delta x, \tag{3.79}$$

where $\rho$ is a constant density (mass per unit length). Then the conservation laws for mass, momentum, and total energy are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\overline{h}\delta x\right) = 0, \tag{3.80}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\overline{uh}\delta x\right) = -g\delta x \overline{\frac{\partial}{\partial x}\left(\frac{1}{2}h^2\right)} = -\frac{g}{2}\delta\left(h^2\right), \text{ and} \tag{3.81}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(g\overline{h^2} + \overline{hu^2}\right)\delta x\right) = -g\delta x \overline{\frac{\partial}{\partial x}\left(\frac{1}{2}h^2 u\right)} = -\frac{g}{2}\delta\left(h^2 u\right), \tag{3.82}$$

respectively, and after integration over a time-step from $t = n\Delta t$ to $t + \Delta t = (n+1)\Delta t$, we get the corresponding discretized conservation laws

$$\overline{h}^{n+1}\Delta x = \overline{h}^{n}\delta x, \tag{3.83}$$

$$\overline{uh}^{n+1}\Delta x = \overline{uh}^{n}\delta x - g\frac{\Delta t}{2}\delta\left(h^2\right)^{n+1/2}, \quad \text{and} \tag{3.84}$$

$$\left(g\overline{h^2} + \overline{hu^2}\right)^{n+1}\Delta x = \left(g\overline{h^2} + \overline{hu^2}\right)^{n}\delta x - g\frac{\Delta t}{2}\delta\left(h^2 u\right)^{n+1/2} \tag{3.85}$$

A summation over all grid cells of each of these equations and application of the periodic boundary condition show that mass, momentum, and total energy are globally conserved.

We have used a slightly simplified notation compared to that introduced in Section 3.1. On the right-hand side, we have, for example, simplified $\overline{\overline{u}^n \overline{h}^n}^{\delta}$ to $\overline{uh}^n$. Also, the left-hand sides are assumed to be discretized like in Section 3.1 so that here $\overline{uh}^{n+1}\Delta x$ stands for $\overline{u}^{n+1}\overline{h}^{n+1}\Delta x$ and $\left(g\overline{h^2} + \overline{hu^2}\right)^{n+1}\Delta x$ stands for $\left(g\left(\overline{h}^{n+1}\right)^2 + \overline{h}^{n+1}\left(\overline{u}^{n+1}\right)^2\right)\Delta x$. Eqs. (3.83) and (3.84) constitute a complete set of prognostic equations, which we may call "the momentum set" and Eqs. (3.83) and (3.85) constitute another complete set of prognostic equations, which we may call "the energy set." Explicit time-stepping can be performed with both set of equations as was the case for the complete system of 3D FV conservation laws considered in the Section 3.1. The explicit time-stepping scheme is absolutely stable with regard to advection but only conditional stable with regard to gravity waves. The strategy of MACHENHAUER and OLK [1997] was to duplicate as far as possible the implementation of the semi-implicit scheme done above in the corresponding traditional model. Several problems are encountered when trying to do this.

The first problem is that the divergence, which we want to average over the two time levels $n\Delta t$ and $(n+1)\Delta t$, is not explicit in the cell-integrated continuity equation, Eq. (3.83), as it is in the traditional one, Eq. (3.70). This can be dealt with by using the Lagrangian expression for divergence.

$$\mathbb{D} = \frac{1}{\delta_i x}\frac{\mathrm{d}\delta_i x}{\mathrm{d}t} \tag{3.86}$$

where here $\delta_i x$ is an infinitesimal small length. A finite difference approximation to this expression is

$$\mathbb{D} = \frac{1}{\Delta x}\frac{\Delta x - \delta x}{\Delta t}. \tag{3.87}$$

Isolating $\delta x$ in Eq. (3.87) and inserting it in Eq. (3.83) gives

$$\overline{h}^{n+1} = \overline{h}^n - \Delta t \overline{h}^n \mathbb{D} = \overline{h}^n - \overline{h}^n \frac{\Delta x - \delta x}{\Delta x}. \tag{3.88}$$

Noting that approximately $\Delta x - \delta x = \Delta t \delta u^{n+1/2}$ where $\delta u^{n+1/2}$ is the velocity increment over the cell at time $(n+1/2)\Delta t$, we get finally

$$\overline{h}^{n+1}_{\mathrm{exp}} = \overline{h}^n - \Delta t \overline{h}^n \frac{\delta u^{n+1/2}}{\Delta x}. \tag{3.89}$$

This expression is similar to Eq. (3.71) and can be used in the same way as Eq. (3.71) in the implementation of the semi-implicit scheme. Doing so, the semi-implicit equation becomes

$$\overline{h}^{n+1} = \overline{h}^n - \frac{\Delta t}{\Delta x}\left(\overline{h}^n\right)' \delta u^{n+1/2} - \frac{\Delta t H}{2\Delta x}\left(\delta u^{n+1} + \delta u^n\right). \tag{3.90}$$

The second problem is that both the momentum and total energy equation are nonlinear quantities in the basic variables $\overline{u}^{n+1}$ and $\overline{h}^{n+1}$. We shall see how that becomes a problem

and how MACHENHAUER and OLK [1997] dealt with it. The explicit momentum equation (Eq. (3.84)) may be written as

$$\overline{uh}^{n+1}_{\text{exp}}\,\Delta x = \overline{uh}^{n}\,\Delta x - \Delta t\Big(UH + \overline{(uh)'}^{n}\Big)\delta u^{n+1/2} - g\frac{\Delta t}{2}\delta\Big(h'^2 + 2h'H\Big)^{n+1/2}, \tag{3.91}$$

where we have used $\overline{(uh)}^{n} = \overline{(uh)'}^{n} + UH$ and $\delta x = \Delta x - \Delta t\delta u^{n+1/2}$. The corresponding semi-implicit equation becomes

$$\begin{aligned}\overline{uh}^{n+1} = \overline{uh}^{n}\,\Delta x - \frac{\Delta t}{\Delta x}\overline{(uh)'}^{n}\delta u^{n+1/2} - g\frac{\Delta t}{2\Delta x}\delta\Big(h'^2\Big)^{n+1/2} \\ - \frac{\Delta t H}{2\Delta x}\Big(\Big(U\Big(\delta u^{n+1} + \delta u^{n}\Big) + g\Big(\delta h^{n+1} + \delta h^{n}\Big)\Big)\Big).\end{aligned} \tag{3.92}$$

The two semi-implicit equations (Eqs. (3.90) and (3.92)) may now be written in the form

$$\overline{h}^{n+1} = \overline{h}^{n+1}_{\text{exp}} - \frac{\Delta t H}{\Delta x}\Big(\delta u^{n+1} + \delta u^{n} - 2\delta u^{n+1/2}\Big) \tag{3.93}$$

and

$$\overline{(uh)}^{n+1} = \overline{(uh)}^{n+1}_{\text{exp}} - \frac{\Delta t H}{\Delta x}\Big(U\Big(\delta u^{n+1} + \delta u^{n} - 2\delta u^{n+1/2}\Big) + g\Big(\delta h^{n+1} + \delta h^{n} - 2\delta h^{n+1/2}\Big)\Big). \tag{3.94}$$

As for the explicit momentum system (Eqs. (3.83) and (3.84)), a summation over all grid cells of each of these equations and application of the periodic boundary condition show that also the semi-implicit system conserves mass and momentum globally. This follows from the fact that the sum of the semi-implicit corrections to the explicit updated values is zero.

To proceed as for the traditional system in Eqs. (3.72) and (3.73), $(uh)^{n+1}$ in Eq. (3.94) must be linearized. At first it is expanded as

$$(uh)^{n+1} = \Big(\big(U + u'\big)\big(H + h'\big)\Big)^{n+1} = UH + \big(u'\big)^{n+1}H + \big(h'\big)^{n+1}U + \big(u'h'\big)^{n+1}. \tag{3.95}$$

In order to make it linear in $\big(u'\big)^{n+1}$ and $\big(h'\big)^{n+1}$, we approximate $\big(u'h'\big)^{n+1}$ with $\big(u'h'\big)^{n}$ getting

$$(uh)^{n+1} \cong UH + u'^{n+1}H + h'^{n+1}U + \big(u'h'\big)^{n}. \tag{3.96}$$

When Eq. (3.96) is inserted in Eq. (3.94), it becomes (after using Eq. (3.93) and some algebra)

$$u'^{n+1} = \frac{1}{H}\left((uh)^{n+1}_{\text{exp}} - Uh^{n+1}_{\text{exp}} - \big(u'h'\big)^{n} - \frac{\Delta t g}{\Delta x}\Big(\delta h^{n+1} + \delta h^{n} - 2\delta h^{n+1/2}\Big)\right) - U. \tag{3.97}$$

This equation is in the same form as the traditional velocity equation (Eq. (3.72)). So, using it together with Eq. (3.93), we can establish a system, similar to the traditional system (Eqs. (3.74) and (3.75)):

$$u'^{n+1} = \tilde{q}_1 - \frac{\Delta t g}{2\Delta x}\delta h^{n+1}, \tag{3.98}$$

$$h^{n+1} = \tilde{q}_2 - \frac{\Delta t H}{2\Delta x}\delta u^{n+1}, \tag{3.99}$$

where terms not depending on values at time level $(n+1)\Delta t$ are collected in $\tilde{q}_1$ and $\tilde{q}_2$, respectively. Applying the operator $\delta(.)$ on Eq. (3.99) and substituting in Eq. (3.98), a Helmholtz equation like Eq. (3.76) is obtained:

$$u'^{n+1} - \frac{gH\Delta t^2}{4\Delta x^2}\delta^2 u^{n+1} = \tilde{q}_1 - \frac{g\Delta t}{2\Delta x}\delta\tilde{q}_2. \tag{3.100}$$

The solution to this equation determines $u'^{n+1}$ and Eq. (3.101) determines $h^{n+1}$, which may then be used to determine the semi-implicit correction terms in Eqs. (3.93) and (3.94).

Machenhauer and Olk [1997] derived a similar semi-implicit model for the cell-integrated energy system. As for the momentum system, a summation over all grid cells of each of these equations and application of the periodic boundary condition showed that the semi-implicit corrections to the explicit updated values become zero so that also the semi-implicit energy system conserves mass and total energy globally.

In this case, the question is if the prognostic variable in Eq. (3.76), the total energy $g\overline{h^2} + \overline{hu^2}$, can be linearized as we did with the momentum. To show that, we first expand it and then approximate nonlinear terms in perturbations with their values at time $n\Delta t$.

$$\begin{aligned} g\overline{h^2}^{n+1} + \overline{hu^2}^{n+1} &= \overline{gh^2 + hu^2}^{n+1} \\ &= \overline{g(H+h')^2 + (H+h')(U+u')^2}^{n+1} \\ &= 2HU\overline{u'}^{n+1} + (UU + 2gH)\,\overline{h'}^{n+1} + \overline{\left(gh'^2 + Hu'^2 + 2Uh'u' + h'u'^2\right)}^{n} \\ &\quad + gH^2 + HU^2. \end{aligned} \tag{3.101}$$

Apparently, one gets something that might work as it did for the momentum equation. The main question is if the coefficient $2HU$ in front of $\overline{u'}^{n+1}$ is sufficiently large. Of course, it will not work if $U = 0$, i.e., if the mean zonal flow is zero. In a realistic flow, however, both $H$ and $U$ are relatively large compared to the perturbations $\overline{u'}^{n+1}$. In test runs with such a flow, the resulting semi-implicit energy system did work satisfactorily, even with a time-step 50 times larger than the CFL maximum.

However, for the full multilevel CSCL system considered in Section 3.1, a corresponding linearization of the total energy $E = 1/g(\delta_k(\phi\, p) + (\widehat{c}_p\widehat{T})_k\delta_k p + 1/2$

$\left(\widehat{u}_k^2 + \widehat{v}_k^2\right)\delta_k p$), defined in Eq. (3.48), cannot be expected to work. Therefore, the present approach cannot be expected to lead to a semi-implicit system that works satisfactory. The reason is that, in the full CSCL system, the explicit angular momentum equation determines the zonal velocity $u$ and the explicit energy equation determines the meridional velocity $v$. Thus, the question is how big the coefficient in front of $v'$ will be in a linearized expression for $E$. It is easily seen that the term in question is $\left(\Delta_k p^{ref} V\right) v_k^{'}$, where we have used $\delta_k p = \left(\Delta_k p^{ref} + \delta_k p\right)$ and $v_k = V + v_k^{'}$. Thus, the question is whether the coefficient $\left(\Delta_k p^{ref} V\right)$ is large enough. Generally, the answer is no as a time independent $V$ in most places will be close to zero. Thus, most likely the approach of Machenhauer and Olk [1997] cannot be extended to the full CSCL system – at least not in a system including the total energy conservation law. An efficient full semi-implicit CSCL system may be possible and may be developed, eventually. However, it will require the invention of a new way to transform the explicit system to a semi-implicit system. The classical approach, described above, goes back to Robert [1969]. It was developed for a system consisting of the momentum equations (or the vorticity and the divergence equations), the continuity equation, and the thermodynamic equation. A new approach, if possible, should be based directly on modifications of the explicit CSCL system.

### *3.2.3. The semi-implicit version of HIRLAM-DCISL*

As a consequence of the conclusions in the last paragraph of the preceding section, Lauritzen, Kaas, Machenhauer and Lindberg [2006, 2008] decided to develop a semi-implicit version of the HIRLAM model in which $u$ and $v$ are kept as prognostic variables, and just the continuity equation is implemented in FV form. In this case, no linearization of the prognostic variables is needed. In the present section, we present the derivation of a semi-implicit system of prognostic equations based on the system of explicit equations in Section 3.2.1. The derivation is described in more details in Lauritzen, Kaas, Machenhauer and Lindberg [2006, 2008]. Here we concentrate on deviations from the traditional derivation procedure used, e.g., for HIRLAM. The traditional procedure for deriving the elliptic equations associated with the baroclinic HIRLAM model is for the central parts the same as for the 1D shallow water models considered in Section 3.2.2.

First, the explicit system is made semi-implicit by time averaging certain right-hand side terms in the discretized primitive equations between time levels $n\Delta t$ and $(n+1)\,\Delta t$.

These linearized terms are:

1. the linearized pressure gradient force in the momentum equation (Eq. (3.1)), i.e., $-\nabla G_k$, which depends on temperature and surface pressure
2. two linearized divergence terms: (1) one in the FV continuity equation (Eq. (3.21)) and (2) one in the energy conversion term $\alpha\omega$ in the thermodynamic equation (Eq. (3.2)), i.e., $-\frac{R_d}{c_{pd}}\left(\frac{T}{p_{k+1/2}}\right)^{ref} \sum_{l=1}^{k} (\Delta_l p)^{ref}\, \mathbb{D}_l^{n+1/2}$ in Eq. (3.120).

Secondly, the formula for the surface pressure and temperatures at time level $(n+1)$ from the semi-implicit continuity and thermodynamic equations, respectively, are inserted in

the formula for the linearized pressure gradient force, $-\nabla G_k$, in the momentum equations. Finally, the divergence operator, $\nabla \cdot (\ )$, is applied to the momentum equation resulting in a set of coupled elliptic equations with updated divergence as an independent variable. The vertically coupled equations are separated into a set of vertically decoupled shallow water Helmholtz equations via diagonalization. The final solution to the elliptic system determines the semi-implicit corrections to the explicit solutions for all the prognostic variables.

*3.2.3.1. The linearized pressure gradient force* The explicit semi-Lagrangian momentum equation at model level $k$ is

$$\frac{\left(\vec{V}_k^{n+1}\right)_{\exp} - \vec{V}_k^n}{\Delta t} = -\nabla\phi_k - R_d(T_v)_k \nabla \ln p_k - f\vec{k} \times \vec{V}_k + \left(\vec{P}_{\vec{V}} + \vec{K}_{\vec{V}}\right)_k . \tag{3.102}$$

The pressure gradient force

$$\vec{F}_k = -\nabla\phi_k - \frac{R_d(T_v)_k}{p_k}\nabla p_k = -\nabla\left(\phi_s + R_d \sum_{l=k+1}^{\mathrm{NLEV}} (T_v)_l (\Delta \ln p)_l + R_d \alpha_k (T_v)_k\right) - R_d(T_v)_k \nabla \ln p_k \tag{3.103}$$

is linearized as in HIRLAM:

$$-\nabla G_k = -\nabla\left(\phi_s + R_d \sum_{l=k+1}^{\mathrm{NLEV}} (T_v)_l \left(\Delta \ln p^{ref}\right)_l + R_d \alpha_k^{ref} (T_v)_k\right) - \frac{R_d T^{ref}}{p_s^{ref}} \nabla p_s, \tag{3.104}$$

where $T^{ref}$ and $p_s^{ref}$ are constant reference temperature and constant surface pressure, respectively. $\left(\Delta \ln p^{ref}\right)_k$ and $\alpha_k^{ref}$ are defined by Eqs. (3.14) and (3.16) with the "half-level" pressures obtained from Eq. (3.11) by choosing $p_s = p_s^{ref}$.

After temporal averaging of $-\nabla G_k$, the semi-implicit momentum equation may be written as in HIRLAM:

$$\left[\vec{V} + \frac{\Delta t}{2}\nabla G - f_0 \vec{k} \times \vec{V}\right]_k^{n+1} = \left(\vec{R}_{\vec{V}}\right)_k , \tag{3.105}$$

where $\left(\vec{R}_{\vec{V}}\right)_k$ represents explicit terms. In the traditional HIRLAM derivation of the elliptic system, a substitution in the linearized pressure gradient force, $-\nabla G_k^{n+1}$, of the updated surface pressure from the continuity equation and the updated temperatures from the thermodynamic equations is performed. Thereby the linearized geopotential, $G_k^{n+1}$, is expressed in terms of the divergence, $\mathcal{D}_k^{n+1}$. The HIRLAM-DCISL derivation proceeds similarly. Here, just those parts involving the FV continuity equation will be dealt with. These are the parts which deviate from the traditional derivation.

*3.2.3.2. The semi-implicit DCISL continuity equations* The discretization of the explicit continuity equation for moist air was discussed in Section 3.2.1. The derivation of the semi-implicit continuity equation considered here is a direct extension of the derivation for the 1D shallow water models in Section 3.2.2. Defining the discretized Lagrangian divergence,

$$\mathbb{D}_k^{n+1/2} = \frac{1}{\Delta A}\frac{\Delta A - \delta A_k^n}{\Delta t} = \frac{1}{\Delta t}\left(1 - \frac{\delta A_k^n}{\Delta A}\right) \tag{3.106}$$

and substituting $\frac{\delta_k A^n}{\Delta A}$ from the explicit continuity equation (Eq. (3.21)) in the form

$$\overline{\left(\delta_k \hat{p}\right)}_{\exp}^{n+1} = \overline{\Delta_k p}^n \frac{\delta A_k^n}{\Delta A}, \tag{3.107}$$

it may be written as

$$\begin{aligned}\overline{\left(\delta_k \hat{p}\right)}_{\exp}^{n+1} &= \overline{\Delta_k p}^n - \Delta t \overline{\Delta_k p}^n \mathbb{D}_k^{n+1/2} \\ &= \overline{\Delta_k p}^n - \Delta t \left(\overline{\Delta_k p}^n\right)' \mathbb{D}_k^{n+1/2} - \Delta t (\Delta_k p)^{ref} \mathbb{D}_k^{n+1/2}.\end{aligned} \tag{3.108}$$

Treating the linear term as a temporal average, the ("ideal") semi-implicit continuity equation results

$$\overline{\delta_k \hat{p}}^{n+1} = \overline{\left(\delta_k \hat{p}\right)}_{\exp}^{n+1} - \frac{\Delta t}{2}(\Delta_k p)^{ref}\left(\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right) + \mathbb{D}_k^n\left(\vec{V}_k^n\right) - 2\mathbb{D}_k^{n+1/2}\right) \tag{3.109}$$

or

$$\overline{\delta_k \hat{p}}^{n+1} = \overline{\left(\delta_k \hat{p}\right)}_{\exp}^{n+1} - \frac{\Delta t}{2}(\Delta_k p)^{ref}\left(\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right) - \mathbb{D}_k^{n+1}\left(\tilde{\vec{V}}_k^{n+1}\right)\right), \tag{3.110}$$

where $\mathbb{D}_k^{n+1}\left(\tilde{\vec{V}}_k^{n+1}\right)$ is defined as the Lagrangian divergence for the last part of the hybrid trajectory

$$\mathbb{D}_k^{n+1}\left(\tilde{\vec{V}}_k^{n+1}\right) = \frac{1}{\Delta A}\frac{\Delta A - \delta A_k^{n+1/2}}{\Delta t/2} \tag{3.111}$$

and $\mathbb{D}_k^n\left(\vec{V}_k^n\right)$ is defined as the Lagrangian divergence for the first part of the hybrid trajectory

$$\mathbb{D}_k^n\left(\vec{V}_k^n\right) = \frac{1}{\Delta A}\frac{\delta A_k^{n+1/2} - \delta A_k^n}{\Delta t/2}, \tag{3.112}$$

(see Fig. 3.3). Thus,

$$\mathbb{D}_k^n\left(\vec{V}_k^n\right) + \mathbb{D}_k^{n+1}\left(\tilde{\vec{V}}_k^{n+1}\right) = 2\mathbb{D}_k^{n+1/2}. \tag{3.113}$$

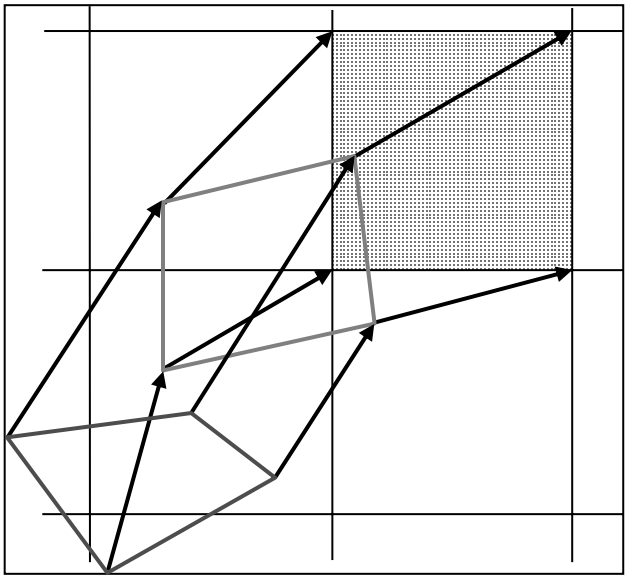

FIG. 3.3 Illustrating the different areas in Eqs. (3.113) and (3.114): $\delta A_k^n$ (*red*), the departure area at time $n\Delta t$, $\delta A_k^{n+1/2}$ (*green*), the "mid-way" area at time $(n+1/2)\,\Delta t$, and $\delta A_k^{n+1}$ (*blue*), the arrival area at time $(n+1)\,\Delta t$. (See also color insert).

This was used to derive Eq. (3.110) from Eq. (3.109). Note that corresponding to Eq. (3.112), $\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$ is defined as

$$\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right) = \frac{1}{\Delta A}\frac{\delta A_k^{n+1/2} - \delta A_k^{n+1}}{\Delta t/2} = \frac{1}{\Delta A}\frac{\delta A_k^{n+3/2} - \delta A_k^{n+1/2}}{\Delta t}, \tag{3.114}$$

where the last expression is centered about time level $n+1$.

In order to proceed with the derivation of the semi-implicit system, the Lagrangian divergence (Eq. (3.114)) should be expressed as a function of the velocity components. In 1D, it was straightforward (see Eq. (3.89)), but in 2Ds, it is complicated although not impossible. However, LAURITZEN, KAAS and MACHENHAUER [2006], LAURITZEN, KAAS, MACHENHAUER and LINDBERG [2008], found that the resulting elliptic equations would be much more complicated than the elliptic equations associated with the traditional HIRLAM system, and therefore, it would be more time consuming to solve. This would significantly reduce the efficiency of the semi-implicit model version. Therefore, they decided to use instead a predictor-corrector approach, which results in elliptic equations in the same form as in HIRLAM. The predictor-corrector approach applied to Eq. (3.110) gives finally the semi-implicit continuity equation for moist air:

$$\begin{aligned}\overline{\delta_k \hat{p}}^{n+1} &= \overline{\left(\delta_k \hat{p}\right)}_{\exp}^{n+1} - \frac{\Delta t}{2}(\Delta_k p)^{ref}\left(\mathcal{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right) - \mathbb{D}\left(\tilde{\vec{V}}_k^{n+1}\right)\right) \\ &\quad + \frac{\Delta t}{2}(\Delta_k p)^{ref}\overline{\left[\mathcal{D}_k^n\left(\vec{V}_k^n\right) - \mathbb{D}_k^n\left(\vec{V}_k^n\right)\right]_*}^{\delta}\frac{\delta A_k^n}{\Delta A},\end{aligned} \tag{3.115}$$

where the discretized Eulerian divergence $\mathcal{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$ is defined in the C-grid in spherical coordinates as

$$\mathcal{D}_k^n\left(\vec{V}_k^n\right) = \frac{1}{a\cos\varphi}\left[\frac{\delta_\lambda u_k^n}{\Delta\lambda} + \frac{\delta_\varphi\left(v_k^n\cos\varphi\right)}{\Delta\varphi}\right] = \frac{\delta_\lambda u_k^n}{\Delta x} + \frac{1}{\cos\varphi}\frac{\delta_\varphi\left(v_k^n\cos\varphi\right)}{\Delta y}, \tag{3.116}$$

$\lambda$ is the longitude, $\varphi$ is the latitude, $\Delta x = a\cos\varphi\Delta\lambda$, and $\Delta y = a\Delta\varphi$.

By replacing $\mathbb{D}_k^{n+1}(\vec{V}_k^{n+1})$ in Eq. (3.110) with the discretized Eulerian divergence $\mathcal{D}_k^{n+1}(\vec{V}_k^{n+1})$, as done in Eq. (3.115), an elliptic equation in the same form as in the traditional HIRLAM system results. However, if just this replacement was done, the scheme would be inconsistent since $\mathcal{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$ is different from $\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$ (in fact small-scale noise would develop and it would result in an instability). Therefore, a correction term is added, the last term in Eq. (3.115), which corrects for the error introduced in the first term of Eq. (3.115). The correction term is equal to the error $\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right) - \mathcal{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$ introduced at time level $n+1$, but it is computed in the subsequent time-step when $\mathbb{D}_k^n\left(\vec{V}_k^n\right)$ is known. (The current $\mathbb{D}_k^n\left(\vec{V}_k^n\right)$ is then equal to $\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$ from the previous time step.) Note that the bar over the last term in Eq. (3.115) indicates a spatial average over the departure area $\delta A_k^n$, i.e., the same departure area as the one used to calculate $\overline{\left(\delta_k\hat{p}\right)}_{\text{exp}}^{n+1}$. In practice, $\frac{\Delta t}{2}\left(\Delta_k p\right)^{ref}\left[\mathcal{D}_k^n\left(\vec{V}_k^n\right) - \mathbb{D}_k^n\left(\vec{V}_k^n\right)\right]$ can be added to the term $\Delta_k p^n$ as the first operation in each time-step. This means that only one upstream integration is needed to evaluate the first and the last term on the right-hand side of Eq. (3.115). The correction term is necessary because the discretized Eulerian divergence $\mathcal{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$, defined in Eq. (3.116), corresponds to a discretized Lagrangian divergence (see Fig. 3.4a), which is different from $\mathbb{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)$, defined in Eq. (3.114). This difference is illustrated in Fig. 3.4.

As demonstrated in Section 3.1.3.3, the explicit continuity equation (Eq. (3.21)) conserves mass both locally and globally. Since the correction terms in Eq. (3.115), which correct the explicit prediction, consist of linear divergence terms, integration over the entire integration area become zero if the Lagrangian and the Eulerian divergence both are zero at the boundaries or if the integration area is global. Consequently, with these assumptions fulfilled, the semi-implicit continuity equation also conserves global mass. It is the impression from preliminary tests that the semi-implicit correction terms generally are small compared to the explicit local mass changes, so it is our impression that the local mass conservation is only slightly modified by the semi-implicit corrections.

The explicit continuity equation for a passive tracer was derived in Section 3.1.4. The result was Eq. (3.40) which may be written as

$$\left(\overline{\left(\widehat{q}_i^{\delta}\right)_k^{n+1}\delta_k\hat{p}^{n+1}}\right)_{\text{exp}} = \overline{\left(\left(\overline{\widehat{q}_i}^{\Delta}\right)_k^n\Delta_k\overline{p}^n\right)_*}^{\delta}\frac{\left(\delta A_k^n\right)}{\Delta A}.$$

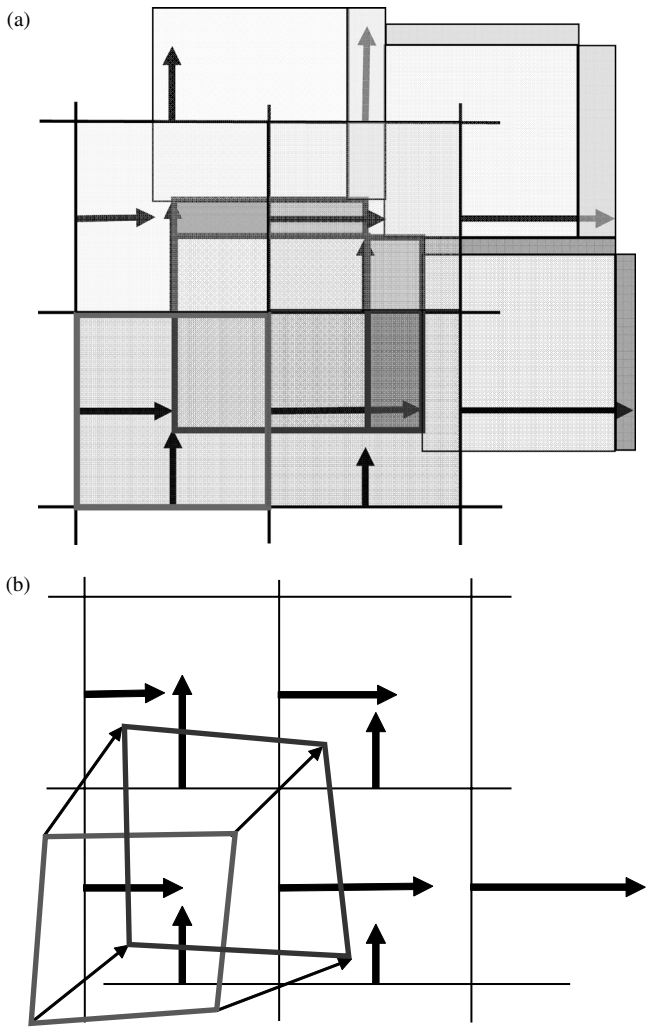


FIG. 3.4 Panel A: Illustrating the Lagrangian divergence $\mathcal{D}_k^n\left(\vec{V}_k^n\right) = \frac{1}{\Delta A}\frac{\delta A_k^{n+1}-\Delta A}{\Delta t}$, which corresponds to the Eulerian divergence (Eq. (3.116)). The periphery of a regular departure area, $\Delta A$, is marked red and the periphery around its arrival area, $\delta A_k^{n+1}$, is marked blue. Additional departure and arrival areas for three neighbor cells are shown. Note that the areas $\delta A_k^{n+1}$ do not cover the whole domain; there are cracks between them. Panel B: Illustrating the Lagrangian divergence $\mathbb{D}_k^n\left(\vec{V}_k^n\right) = \frac{1}{\Delta A}\frac{\delta A_k^{n+1/2}-\delta A_k^{n-1/2}}{\Delta t}$, where the departure area $\delta A_k^{n-1/2}$ is marked red and the arrival area $\delta A_k^{n+1/2}$ is marked blue. Obviously, $\mathbb{D}_k^n\left(\vec{V}_k^n\right)$ is generally different from $\mathcal{D}_k^n\left(\vec{V}_k^n\right)$ in A. In both panels: Black arrows are velocity components in the C-grid. (See also color insert).

The variable on the left-hand side is the weight of the cell-integrated tracer mass per unit horizontal area (see Section 3.1.4). For a nonpassive tracer, source and sink terms should be added on the right-hand side of the equation. In a semi-implicit model, semi-implicit correction terms must be added to the explicit predicted tracer weights in order to make them consistent with the predicted moist air weights. The corrected semi-implicit tracer prediction equation must be identical to that for moist air (Eq. (3.115)) for $\left(\overline{\widehat{q}_i}\right)_k^n \equiv 1$. This means that the semi-implicit tracer continuity equation must be

$$\begin{aligned}\overline{\left(\widehat{q}_i^{\delta}\right)_k^{n+1}\delta\hat{p}_k^{n+1}} &= \left(\overline{\left(\overline{\widehat{q}_i}^{\delta}\right)_k^{n+1}\delta_k\overline{\hat{p}}^{n+1}}\right)_{\exp} \\ &\quad - \frac{\Delta t}{2}\left(\left(\overline{\widehat{q}_i}^{\Delta}\right)_k\Delta_k\overline{p}\right)^{ref}\left(\mathcal{D}_k^{n+1}\left(\vec{V}_k^{n+1}\right)-\mathbb{D}\left(\tilde{\vec{V}}_k^{n+1}\right)\right) \\ &\quad + \frac{\Delta t}{2}\left(\left(\overline{\widehat{q}_i}^{\Delta}\right)_k\Delta_k\overline{p}\right)^{ref}\overline{\left[\mathcal{D}_k^n\left(\vec{V}_k^n\right)-\mathbb{D}_k^n\left(\vec{V}_k^n\right)\right]_*}^{\delta}\frac{\delta A_k^n}{\Delta A}.\end{aligned}$$

*3.2.3.3. The semi-implicit energy conversion term* Also a dependence on divergence in the thermodynamic equation (Eq. (3.2)) needs to be temporally averaged in the semi-implicit model. Specifically, it is the energy converting term $\frac{\alpha\omega}{c_p} = \frac{R_d T_v}{c_p}\frac{\omega}{p}$, approximated in the explicit model with Eq. (3.63), which is divergence dependent. To isolate this dependence $\overline{\omega_k}^{n+1/2}$, given by Eq. (3.34), is expanded as follows:

$$\overline{[\omega_k]}_{\exp}^{n+1/2} = \frac{1}{\Delta t}\left(\overline{\hat{p}}_k^{n+1} - \overline{\left(\overline{p}_k^n\right)_*}^{\delta}\right) = \frac{1}{\Delta t}\left(\exp\left(-\alpha_k^{n+1}\right)\sum_{l=1}^{k}\overline{\delta_l\hat{p}}^{n+1} - \overline{\left(\overline{p}_k^n\right)_*}^{\delta}\right), \tag{3.117}$$

where Eqs. (3.17) and (3.22) have been used. When then $\overline{\delta_l\hat{p}}^{n+1}$ is substituted from Eq. (3.108), the result is

$$\begin{aligned}\overline{[\omega_k]}_{\exp}^{n+1/2} &= \frac{1}{\Delta t}\left(\exp\left(-\alpha_k^{n+1}\right)\sum_{l=1}^{k}\left(\overline{\Delta_l p}^n - \Delta t\,\overline{\Delta_l p}^n\mathbb{D}_l^{n+1/2}\right) - \overline{\left(\overline{p}_k^n\right)_*}^{\delta}\right) \\ &= \frac{1}{\Delta t}\left(\exp\left(-\alpha_k^{n+1}\right)\sum_{l=1}^{k}\overline{\Delta_l p}^n - \overline{\left(\overline{p}_k^n\right)_*}^{\delta}\right) - \exp\left(-\alpha_k^{n+1}\right)\sum_{l=1}^{k}\overline{\Delta_l p}^n\mathbb{D}_l^{n+1/2}.\end{aligned} \tag{3.118}$$

Thus, the explicit energy converting term may be written as

$$\left[\left(\frac{R_d T_v\omega}{c_p p}\right)_k\right]_{\exp}^{n+1}$$

$$= \frac{R_d (T_v)_k^{n+1/2}}{(c_p)_k^{n+1/2} \overline{p}_k^{n+1/2}} \overline{[\omega_k]}_{\exp}^{n+1/2} = \frac{R_d\,(T_v)_k^{n+1/2}}{(c_p)_k^{n+1/2} \exp\left(-\alpha_k^{n+1/2}\right) \overline{p}_{k+1/2}^{n+1/2}} \overline{[\omega_k]}_{\exp}^{n+1/2}$$

$$= \frac{R_d (T_v)_k^{n+1/2}}{(c_p)_k^{n+1/2} \exp\left(-\alpha_k^{n+1/2}\right) \overline{p}_{k+1/2}^{n+1/2}} \frac{1}{\Delta t} \left( \exp\left(-\alpha_k^{n+1}\right) \sum_{l=1}^{k} \overline{\Delta_l p}^n - \overline{\left(\overline{p}_k^n\right)}_*^{\delta} \right)$$

$$- \frac{R_d (T_v)_k^{n+1/2} \exp\left(-\alpha_k^{n+1}\right)}{(c_p)_k^{n+1/2} \exp\left(-\alpha_k^{n+1/2}\right) \overline{p}_{k+1/2}^{n+1/2}} \sum_{l=1}^{k} \overline{\Delta_l p}^n \mathbb{D}_l^{n+1/2}, \tag{3.119}$$

where again Eqs. (3.17) and (3.22) have been used. When the last term is linearized about a reference temperature $T^{ref}$ and a reference surface pressure $p_s^{ref}$, the result is

$$\left[ \left( \frac{R_d T_v \omega}{c_p p} \right)_k \right]_{\exp}^{n+1}$$

$$= \frac{R_d\,(T_v)_k^{n+1/2}}{(c_p)_k^{n+1/2} \exp\left(-\alpha_k^{n+1/2}\right) \overline{p}_{k+1/2}^{n+1/2}} \frac{1}{\Delta t} \left( \exp\left(-\alpha_k^{n+1}\right) \sum_{l=1}^{k} \overline{\Delta_l p}^n - \overline{\left(\overline{p}_k^n\right)}_*^{\delta} \right)$$

$$- \sum_{l=1}^{k} \left( \overline{\frac{R_d\,(T_v)_k^{n+1/2} \exp\left(-\alpha_k^{n+1}\right)}{(c_p)_k^{n+1/2} \exp\left(-\alpha_k^{n+1/2}\right) \overline{p}_{k+1/2}^{n+1/2}} \Delta_l p}^{\,n} \right)' \mathbb{D}_l^{n+1/2}$$

$$- \frac{R_d}{c_{pd}} \left( \frac{T}{p_{k+1/2}} \right)^{ref} \sum_{l=1}^{k} (\Delta_l p)^{ref}\, \mathbb{D}_l^{n+1/2}. \tag{3.120}$$

Treating the linear term as a temporal average, we finally get the ("ideal") semi-implicit energy conversion term

$$\left( \frac{R_d T_v \omega}{c_p p} \right)_k^{n+1} = \left[ \left( \frac{R_d T_v \omega}{c_p p} \right)_k \right]_{\exp}^{n+1}$$

$$- \frac{R_d}{c_{pd}} \left( \frac{T}{p_{k+1/2}} \right)^{ref} \sum_{l=1}^{k} (\Delta_l p)^{ref} \left( \mathbb{D}_l^{n+1} \left( \vec{V}_l^{n+1} \right) + \mathbb{D}_l^n \left( \vec{V}_k^n \right) - 2\mathbb{D}_l^{n+1/2} \right) \tag{3.121}$$

or

$$\left( \frac{R_d T_v \omega}{c_p p} \right)_k^{n+1} = \left[ \left( \frac{R_d T_v \omega}{c_p p} \right)_k \right]_{\exp}^{n+1}$$

$$- \frac{R_d}{c_{pd}} \left( \frac{T}{p_{k+1/2}} \right)^{ref} \sum_{l=1}^{k} (\Delta_l p)^{ref} \left( \mathbb{D}_l^{n+1} \left( \vec{V}_k^{n+1} \right) - \mathbb{D} \left( \tilde{\vec{V}}_l^{n+1} \right) \right), \tag{3.122}$$

where again $\mathbb{D}\left(\tilde{\vec{V}}_l^{n+1}\right)$ is defined by Eq. (3.111).

In order to obtain the same uncomplicated elliptic equations as in HIRLAM, the predictor-corrector approach is again utilized. This changes Eq. (3.122) to

$$\left(\frac{R_d T_v \omega}{c_p p}\right)_k^{n+1} = \left[\left(\frac{R_d T_v \omega}{c_p p}\right)_k\right]_{\exp}^{n+1}$$

$$- \frac{R_d}{c_{pd}}\left(\frac{T}{p_{k+1/2}}\right)^{ref}\left[\sum_{l=1}^{k}(\Delta_l p)^{ref}\left(\mathcal{D}_l^{n+1}\left(\vec{V}_l^{n+1}\right) - \mathbb{D}\left(\tilde{\vec{V}}_l^{n+1}\right)\right)\right.$$

$$\left. - \sum_{l=1}^{k}(\Delta_l p)^{ref}\overline{\left[\mathcal{D}_l^n\left(\vec{V}_k^n\right) - \mathbb{D}_l^n\left(\vec{V}_k^n\right)\right]_*}^{\delta}\frac{\delta A_l^n}{\Delta A}\right]. \tag{3.123}$$

### *3.3. The NCAR-FFSL dynamical core*

This section describes main features of the FV dynamical core, included in the NCAR Community Atmospheric Model (CAM 3.0) description, Collins, Rasch, Boville, Hack, Mccaa, Williamson, Kiehl, Briegleb, Bitz, Lin, Zhang and Dai [2004], Chapter 3.3. It is the pioneering example of a meteorological FV model based on the flux form (Eq. (1.2)) of the continuity equation. It was initially developed and used at the NASA Data Assimilation Office for data assimilation, numerical weather prediction, and climate simulations. The dynamical core is quasi-hydrostatic, global, and formulated for traditional latitude-longitude coordinates.

#### *3.3.1. The 3D transport scheme*

The quasi-horizontal transport of air mass, tracer mass, and potential temperature is based on a 2D FV FFSL scheme developed by Lin and Rood [1996] and Lin and Rood [1997]. This Eulerian scheme of the operator splitting or time splitting type, described in Section 2.3.2, is among the most modern flux-based FV schemes. For the sake of the following description of the dynamical core, it is convenient to rewrite the prediction equation, Eq. (2.52), with the notations of Lin and Rood [1997].

At first, the following standard finite difference $\delta$ and average $\overline{(\,)}$ operators are defined

$$\begin{aligned}\delta_\sigma q &= q\left(\sigma + \frac{\Delta\sigma}{2}\right) - q\left(\sigma - \frac{\Delta\sigma}{2}\right),\\ \overline{q}^{\sigma} &= \frac{1}{2}\left[q\left(\sigma + \frac{\Delta\sigma}{2}\right) + q\left(\sigma - \frac{\Delta\sigma}{2}\right)\right].\end{aligned} \tag{3.124}$$

The conservation law for a density-like variable $Q$ is

$$\frac{\partial Q}{\partial t} + \nabla\cdot\left(Q\vec{V}\right) = 0. \tag{3.125}$$

As explained in Section 2.3.2, the FFSL scheme involves the application of 1D flux convergence operators and advective operators, successively applied along the two horizontal coordinate directions, in such a way that the scheme becomes both *conservative* and *constancy preserving*.

The 1D flux convergence operators $F$ and $G$ are defined as

$$\begin{aligned} F\big(u^*, \Delta\tau; Q^n\big) &= -\frac{\Delta\tau}{a\Delta\lambda\cos\varphi}\delta_\lambda\big[\mathbb{X}\big(u^*, \Delta\tau; Q^n\big)\big], \\ G\big(v^*, \Delta\tau; Q^n\big) &= -\frac{\Delta\tau}{a\Delta\varphi\cos\varphi}\delta_\varphi\big[\cos\varphi\mathbb{Y}\big(v^*, \Delta\tau; Q^n\big)\big]. \end{aligned} \tag{3.126}$$

$F$ and $G$ updates $Q$ for one time-step in the zonal ($\lambda$) and meridional ($\varphi$) directions, respectively. Here, $\mathbb{X}$ and $\mathbb{Y}$, the time-averaged fluxes of $Q$ in the zonal ($\lambda$) and meridional ($\varphi$) directions, respectively, are defined as

$$\begin{aligned} \mathbb{X}\big(u^*, \Delta\tau; Q^n\big) &\cong \frac{1}{\Delta\tau}\int_t^{t+\Delta\tau} uQ\,dt - hot \cong u^*\big(Q^n\big)^*_\lambda, \\ \mathbb{Y}\big(v^*, \Delta\tau; Q^n\big) &\cong \frac{1}{\Delta\tau}\int_t^{t+\Delta\tau} vQ\,dt - hot \cong v^*\big(Q^n\big)^*_\varphi, \end{aligned} \tag{3.127}$$

where $u^*$ and $v^*$ are predicted time-centered velocity components at $t + \Delta\tau/2$ in C-grid positions at the east and south face of the cell, respectively. $(Q^n)^*_\lambda$ is determined by an upstream integral

$$\big(Q^n\big)^*_\lambda = \frac{1}{a\cos\varphi\Delta\lambda}\int_0^{u^*\Delta\tau} Q^n a\cos\varphi d\lambda, \tag{3.128}$$

with a corresponding expression for $(Q^n)^*_\varphi$.

Thus, to approximate the time-averaged fluxes across the cell faces the time-centered winds, $u^*$ and $v^*$, and the cell averaged field, $Q^n$, at time level $n$ are required. Furthermore, for modeling cross-stream advection in the zonal and meridional directions, respectively, the *advective* flux operators $\tilde{f}$ and $\tilde{g}$ are introduced. They are defined in Collins, Rasch, Boville, Hack, Mccaa, Williamson, Kiehl, Briegleb, Bitz, Lin, Zhang and Dai [2004] in terms of the corresponding F and G operators. Here, $\tilde{f}$ is defined as

$$\tilde{f}\big(u^*, \Delta\tau; Q^n\big) = F\big(u^*, \Delta\tau; Q^n\big) + \frac{\Delta\tau}{a\Delta\lambda\cos\varphi}\delta_\lambda u^*, \tag{3.129}$$

with a corresponding definition for $\tilde{g}$. With these definitions, the following prognostic equation (corresponding to Eq. (2.52)) results

$$\begin{aligned} Q^{n+1} = Q^n &+ F\left[u^*, \Delta\tau; Q^n + \frac{1}{2}\tilde{g}\big(v^*, \Delta\tau; Q^n\big)\right] \\ &+ G\left[v^*, \Delta\tau; Q^n + \frac{1}{2}\tilde{f}\big(u^*, \Delta\tau; Q^n\big)\right] \end{aligned} \tag{3.130}$$

or

$$Q^{n+1} = Q^n + F\left[u^*, \Delta\tau; Q^{\varphi}\right] + G\left[v^*, \Delta\tau; Q^{\lambda}\right], \tag{3.131}$$

where

$$Q^{\varphi} = Q^n + \frac{1}{2}\tilde{g}\left(v^*, \Delta\tau; Q^n\right) \quad \textit{and} \quad Q^{\lambda} = Q^n + \frac{1}{2}\tilde{f}\left(u^*, \Delta\tau; Q^n\right). \tag{3.132}$$

As in the HIRLAM-DCISL, the fluxes are assumed to be along 3D trajectories; however, here the trajectories are line-segment parallel to the coordinate axes. The final *vertical displacements* after each time-step are determined so that hydrostatic balance is maintained. A *Lagrangian vertical coordinate* $\xi$, is introduced (see Section 3.1.3.1), which per definition is constant along the 3D trajectories. The quasi-horizontal flow along such coordinate surfaces is 2D and relative to the coordinate surfaces, where the vertical velocity is zero as expressed in Eq. (1.1). The governing Eulerian equations, presented below in Section 3.3.2, are therefore without vertical advection terms.

In the present setup, the Eulerian vertical discretization defining the vertical extend of the Eulerian grid cells is similar to the hybrid sigma-pressure discretization (Simmons and Burridge [1981]) described in Section 3.1.2, except that here the top of the model atmosphere is at a constant pressure $p_\infty$. Thus, the pressure at a $\eta$-model-surface is $p^n_{k+1/2} = p_\infty + A_{k+1/2} + B_{k+1/2}p^n_s$ and the vertical "pressure thickness" of an Eulerian grid cell is $\delta_k p^n = \Delta_k A + \Delta_k B\, p^n_s$. As illustrated in Fig. 3.1, the transport of air during a time-step, ending up in an Eulerian grid column, is effectuated by a Lagrangian cell, with initial pressure thickness $\delta_k p^n = p^n_{k+1/2} - p^n_{k-1/2}$. The Lagrangian cell is moving with the 3D flow and is ending up with a pressure thickness $\delta_k \hat{p}^n = \hat{p}^n_{k+1/2} - \hat{p}^n_{k-1/2}$. Its pressure ($\hat{p}^{n+1}_{k+1/2}$) is determined hydrostatically from the weight of the cells arriving above in the same Eulerian grid column. The scheme is globally conservative; however, it is less locally conservative than the DCISL schemes. Thus, as illustrated in Fig. 2.15(d), the FFSL scheme uses information from an area that is somewhat dispersed compared to the exact departure area.

### *3.3.2. The governing equations*

Neglecting physical forcing terms, the governing continuous quasi-hydrostatic equations in spherical latitude-longitude coordinates are:

*The hydrostatic balance equation* on the form

$$\delta_k p = -g\rho\delta_k z, \tag{3.133}$$

which shows that the variable in the continuity equation, $\delta_k p$, is the weight per unit horizontal area in the layer of air with thickness $\delta_k z$.

The continuity equation or in other words the conservation law for mass is written as

$$\frac{\partial}{\partial t}\delta_k p + \frac{1}{a\cos\varphi}\left[\frac{\partial}{\partial\lambda}(u_k\delta_k p) + \frac{\partial}{\partial\varphi}(v_k\delta_k p\cos\varphi)\right] = 0. \tag{3.134}$$

Similarly, the mass conservation law for tracer species (including water vapor) is

$$\frac{\partial}{\partial t}\left((q_i)_k\,\delta_k p\right) + \frac{1}{a\cos\varphi}\left[\frac{\partial}{\partial\lambda}\left(u_k(q_i)_k\,\delta_k p\right) + \frac{\partial}{\partial\varphi}\left(v_k\,(q_i)_k\,\delta_k p\cos\varphi\right)\right] = 0. \tag{3.135}$$

The thermodynamic equation or the conservation law for potential temperature is

$$\frac{\partial}{\partial t}(\theta_k\delta_k p) + \frac{1}{a\cos\varphi}\left[\frac{\partial}{\partial\lambda}(u_k\theta_k\delta_k p) + \frac{\partial}{\partial\varphi}(v_k\theta_k\delta_k p\cos\varphi)\right] = 0, \tag{3.136}$$

where $\theta_k = (T_v)_k\left(\frac{p_k}{1000\,hPa}\right)^{\frac{R_d}{c_{pd}}}$ is the (virtual) potential temperature.
The momentum equations are used on the so-called "vector invariant form:"

$$\frac{\partial}{\partial t}u_k = \eta_k v_k - \frac{1}{a\cos\varphi}\left[\frac{\partial}{\partial\lambda}((e_{kin})_k + \phi_k - \nu\mathcal{D}_k) + \frac{1}{\rho}\frac{\partial p_k}{\partial\lambda}\right], \tag{3.137}$$

$$\frac{\partial}{\partial t}v_k = -\eta_k u_k - \frac{1}{a}\left[\frac{\partial}{\partial\varphi}((e_{kin})_k + \phi_k - \nu\mathcal{D}_k) + \frac{1}{\rho}\frac{\partial p_k}{\partial\varphi}\right], \tag{3.138}$$

where $\mathcal{D}_k$ is the divergence, defined as in Eq. (3.146), $\nu$ is the coefficient for an optional divergence damping, and the absolute vorticity $\eta_k$ is

$$\eta_k = 2\Omega\sin\varphi + \frac{1}{a\cos\varphi}\left[\frac{\partial v_k}{\partial\lambda} - \frac{\partial(u_k\cos\varphi)}{\partial\varphi}\right]. \tag{3.139}$$

Here, $\Omega$ is the angular velocity of the earth. Finally, the kinetic energy $(e_{kin})_k$ is defined as

$$(e_{kin})_k = \frac{1}{2}\left((u_k)^2 + (v_k)^2\right). \tag{3.140}$$

*3.3.3. Time-stepping*

In the model, the dynamics and the NCAR CAM physics are time-split as in HIRLAM in the sense that all prognostic variables are updated sequentially, at first by the dynamics and then by the physics. The time-stepping is fully explicit with subcycling over small time-steps, $\Delta\tau = \Delta t/m$, within the 2D dynamics. The number of subcycles needed to stabilize the fast gravity waves is $m$. To avoid excessive small time-steps due to the convergence of the meridians near the poles, a polar Fourier filter, which filters out the shortest zonal waves, is applied to $u^*$ and $v^*$ and certain tendency terms in the prognostic equations. The transport for tracers, however, can take a much larger time-step $\Delta t$ equal to the interval between the physics updates. In the present setup, the cells are transported along the Lagrangian surfaces during the long tracer time-steps, starting initially at the beginning of the first small time-step as a model layer with "pressure thickness" $\Delta_k p^n = \Delta_k A + \Delta_k B\,p_s^n$, without any remapping of the variables to the Eulerian model levels. So during these $m$ short time-steps, the transport is fully Lagrangian. Only at the end of each tracer time-step, a remapping takes place. This is done to avoid excessive smoothing caused by too frequent vertical remapping after each small time-step.

*3.3.3.1. Conservative predictions* The *prognostic variables*, the cell averaged values $\delta_k p$, $(q_i)_k$ , $\theta_k$, $u_k$, and $v_k$, are updated by the use of the prognostic equations (Eqs. (3.134) – (3.138)), but only the Eqs. (3.134)–(3.136) for the density-related variables are in the proper flux-form (Eq. (3.125)) and are integrated directly by the FFSL prediction equation (Eq. (3.130)). Thus, the dynamical core conserves exactly mass of air, tracer mass, including water vapor (apart from evaporation and condensation), and potential temperature (in adiabatic friction free flow). The integration of the momentum equations, (Eqs. (3.137) and (3.138)) is discussed in Section 3.3.3.2.

At the start of time-step $n$ (time $n\Delta\tau$), the prognostic variables, $\delta_k p$, $(q_i)_k$ , $\theta_k$, $u_k$, and $v_k$, are given in the D-grid as indicated in Fig. 3.5. In addition, the advective winds $u^*$ and $v^*$are needed for the update of $\delta_k p$, $(q_i)_k$, and $\theta_k$. So at first, these are updated to time $(n+1/2)\Delta\tau$on the C-grid using the momentum prediction equation (Eq. (3.144)). When they are available, Eqs. (3.134)–(3.136) can be advanced one time-step using the FFSL prediction form (Eq. (3.130)). $u^*$ and $v^*$ are not history carrying variables. They are overwritten after being used. When the continuity equation has been solved, the updated pressure thickness of each Lagrangian layer, $\overline{(\delta_k \hat{p})}^{n+1}$, determines the pressure of the Lagrangian surfaces by summing up the hydrostatic weight of all the cells above:

$$\overline{\hat{p}}^{n+1}_{k-1/2} = p_\infty + \sum_{l=1}^{k-1} \overline{\delta_l \hat{p}}^{n+1}. \tag{3.141}$$

Summing up the hydrostatic weight of all the NLEV Lagrangian layers yields the surface pressure:

$$\overline{p}^{n+1}_s = p_\infty + \sum_{l=1}^{\mathrm{NLEV}} \overline{\delta_l \hat{p}}^{n+1}. \tag{3.142}$$

This is needed for determination of the pressure at the interfaces between the Eulerian model layers:

$$\overline{p}^{n+1}_{k-1/2} = p_\infty + A_{k-1/2} + B_{k-1/2}\, \overline{p}^{n+1}_s, \tag{3.143}$$

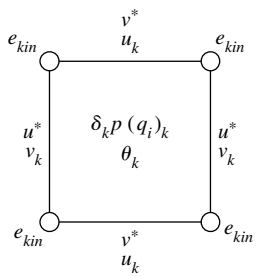


FIG. 3.5 Schematic stencil of the location of variables in finite volume schemes for the Arakawa C- (indicated with superscript *) and D-grids.

which are needed for the physical parameterization and the vertical remapping of the prognostic variables. The density-like prognostic variables are given after each large time-step, $\Delta t$, as mean values over the Lagrangian layers, $\overline{\delta_k \hat{p}}^{n+1}$. To be used for physical parameterization, they must be remapped on the Eulerian model layers $\overline{\Delta_k p}^{n+1}$ (determined from Eq. (3.143)).

*3.3.3.2. Integration of the momentum equation* Inspired by the papers of SADOURNY [1972] and ARAKAWA and LAMB [1981], a discretization of the momentum equations has been achieved that results in conservation of the absolute vorticity. Since also mass and potential energy are conserved and all three invariants are consistently transported, it might be expected that approximately the same will be the case for potential vorticity.

The resulting prognostic equations are

$$u^{n+1} = u^n + \Delta\tau \left\{ \mathbb{Y}\left(v^*, \Delta\tau; \eta^\lambda\right) - \frac{1}{a\cos\varphi\,\Delta\lambda}\delta_\lambda[(e_{kin})^* - \nu\mathcal{D}^*] + \widehat{P}_\lambda \right\},$$
$$v^{n+1} = v^n - \Delta\tau \left\{ \mathbb{X}\left(u^*, \Delta\tau; \eta^\varphi\right) - \frac{1}{a\Delta\varphi}\delta_\varphi\,[(e_{kin})^* - \nu\mathcal{D}^*] + \widehat{P}_\varphi \right\}, \tag{3.144}$$

where $(e_{kin})^*$, the upstream-biased kinetic energy (defined in the four corners of the D-grid (Fig. 3.5)), is formulated as

$$(e_{kin})^* = \frac{1}{2}\left\{ \mathbb{X}\left(\overline{u^*}^\varphi, \Delta\tau; u^n\right) \text{`} + \mathbb{Y}\left(\overline{v^*}^\lambda, \Delta\tau; v^n\right) \right\} \tag{3.145}$$

and

$$\mathcal{D}^* = \frac{1}{a\cos\varphi}\left[ \frac{\delta_\lambda u^{n+1}}{\Delta\lambda} + \frac{\delta_\varphi\left(\cos\varphi\, v^{n+1}\right)}{\Delta\varphi} \right]. \tag{3.146}$$

The FV mean pressure gradient terms $\widehat{P}_\lambda$ and $\widehat{P}_\varphi$ in Eq. (3.144) are computed by the method presented in LIN [1997], which eliminates a long-standing problem in terrain-following coordinates, i.e., the inaccuracy caused by different truncation errors in the two terms that the pressure gradient force traditionally are split into.

The velocity components $u_k$ and $v_k$ are given after each large time at Lagrangian levels, $\hat{p}_k$. Like the remaining prognostic variables, they need to be remapped to Eulerian levels. An accurate and conservative remapping procedure has been developed. The current remapping version, described in detail in COLLINS, RASCH, BOVILLE, HACK, MCCAA, WILLIAMSON, KIEHL, BRIEGLEB, BITZ, LIN, ZHANG and DAI [2004], conserves exactly mass, momentum, and total energy.

## 3.4. Properties of the dynamical cores

### 3.4.1. Conservation properties

With proper boundary conditions, the HIRLAM-DCISL dynamical core conserves exactly global mass of moist air and tracers, including water vapor, liquid water, and

solid water, if included (apart from evaporation and condensation). Also the NCAR-FFSL dynamical core conserves exactly these global masses. However, in addition, it conserves globally potential temperature (in adiabatic friction-free flow) and absolute vorticity (in adiabatic friction-free flow). Thus, the NCAR-FFSL dynamical core comes closer than the HIRLAM-DCISL to the ideal CSCL model considered in Section 3.1. However, as we have seen in the idealized tests presented in Section 2, the local conservation is more accurate in the HIRLAM-DCISL transport schemes than in the NCAR-FFSL scheme (compare Figs. 2.10(b) and 2.10(c) with Fig. 2.11(d)). Although the vertical remapping is designed to conserve the total energy, it is not globally conserved in NCAR-FFSL. The horizontal discretization and the use of a "diffusive" transport scheme with monotonicity constraint tend to decrease the kinetic energy and thereby the total energy. Whether this is realistic is difficult to assess. A total energy "fixer" is applied to effectively add the loss in kinetic energy due to "diffusion" back to the model as total potential energy so that the total energy is globally conserved. However, even without the fixer, the loss is found to be very small, less than 2 W/m$^2$ with a 2-degree resolution and it is found to decrease with increasing resolution. It is stated in Collins, Rasch, Boville, Hack, Mccaa, Williamson, Kiehl, Briegleb, Bitz, Lin, Zhang and Dai [2004] that in the future it may considered to use the total energy as a transported prognostic so that the total energy could be automatically conserved. In the HIRLAM-DCISL dynamical core, no total energy "fixer" is applied.

*3.4.2. FV Lagrangian pressure gradient force*

As mentioned above, a particular feature of the NCAR-FFSL model is the Eulerian expression for the FV mean pressure-gradient force (Lin [1997]), which is used in the model. It eliminates a long-standing problem: the inaccuracy caused by different truncation errors in the two terms that the pressure gradient force traditionally is split into in terrain-following coordinates. A similar finite volume Lagrangian expression for the mean pressure gradient force along the trajectories during a time-step is suggested by Lauritzen, Kaas, Machenhauer and Lindberg [2008]. It may be used in any DCISL model. It has not yet been implemented in HIRLAM-DCISL; however, it is expected that it may lead to increased accuracy, even though the linearized pressure gradient force (Eq. (3.104)), used to derive the semi-implicit correction terms, is based on the two-term expression of the pressure gradient force. The proposed Lagrangian mean pressure gradient force $PGF_{\Delta s'}$ along the sloping trajectory $\Delta s'$is easily computed from the pressure of the arrival cell, at the end of the trajectory, $\left(\overline{\hat{p}}_k^{n+1}\right)_{\exp}$ and the pressure of the departure cell, at the start of the trajectory, $\overline{\left(\overline{p}_k^n\right)}_*^{\delta}$. Note that these are the same pressures that are used to define $\omega$ in Eq. (3.117). The proposed expression is

$$PGF_{\Delta s'} = \left[\left(-\frac{1}{\rho}\frac{\partial p}{\partial s'}\right)_k\right]_{\exp}^{n+1/2} = -\frac{1}{\overline{\overline{\rho}}}\frac{\left(\overline{\hat{p}}_k^{n+1}\right)_{\exp} - \overline{\left(\overline{p}_k^n\right)}_*^{\delta}}{\Delta s'}$$

$$= -\frac{1}{\overline{\overline{\rho}}}\frac{\left(\overline{\hat{p}}_k^{n+1}\right)_{\exp} - \overline{\left(\overline{p}_k^n\right)}_*^{\delta}}{\Delta s}\cos(\vartheta). \tag{3.147}$$

$\overline{\overline{\rho}}$ is an approximation of the mean density along the trajectory. $\Delta s$ is the horizontal distance between the midpoints of the departure and the arrival cells, and $\Delta s'$ is the corresponding distance along the sloping trajectory. $\cos(\vartheta) = \frac{\Delta s}{\Delta s'}$ defines the slope of the trajectory. In order to determine the horizontal component of the pressure gradient force $PGF_h$, the vertical component $PGF_v$, which is balanced by gravity $g$, must be subtracted:

$$PGF_h = \sqrt{(PGF_{\Delta s'})^2 - g^2}. \tag{3.148}$$

*3.4.3. Tests of performance*

THUBURN [2006] has made an attempt to estimate the relative importance of different conservation laws. It is argued that satisfactory model performance requires spurious sources of a conservable quantity to be much weaker than any true physical sources; for several conservable quantities, the magnitudes of the physical sources are estimated in order to provide benchmarks against which any spurious sources may be measured. A model with weak spurious sources of a conservable quantity compared with the physical sources may in practice produce as accurate forecasts, especially long simulations, as a model which conserves the quantity exactly if the spurious sources are not systematic. However, even if the spurious sources are weak but the spurious sources are systematic, long simulations may be very inaccurate. Of course, if possible, the spurious sources should be estimated relative to the physical sources, but that may be very difficult and it is not enough; it is also necessary to know if the spurious sources are systematic. So, in practice, as when other potential model improvements are considered, it is necessary to carry out a series of real cases that is validated against observations in competition with any model it is supposed to substitute. To the authors' knowledge, such a series of real case tests have not yet been carried out for any of the two FV dynamical cores considered here.

*3.4.4. Idealized baroclinic wave test*

However, both have been preliminary tested in the idealized baroclinic wave test case of JABLONOWSKI and WILLIAMSON [2006a], in the following called JW06a (a more detailed technical note JABLONOWSKI and WILLIAMSON [2006b] is available). These tests have, as we shall see, confirmed that both dynamical cores work properly, producing in general as realistic baroclinic developments as present day's nonconservative state-of-the-art dynamical cores. It would have been of considerable interest to validate the importance of the mass conservation property of the FV dynamical cores considered here in real case tracer transport simulations. There seems to be no doubt that this property is essential for tracer transport however to the authors' knowledge, it has not yet been verified.

The idealized baroclinic wave test case of JW06a consists of an analytic steady-state zonal solution to the global primitive equations. The steady-state is unstable so that an overlaid perturbation in global reference integrations triggers the development of an idealized baroclinic wave in the northern hemisphere. By day 4, a well-defined wave train is established, and by days 7–9, a significant deepening of the highs and lows takes place before a breakdown by days 20–30 leads to a full circulation in both hemispheres.

*3.4.5. Idealized test of NCAR-FFLS*

JW06a applied the test to four different global dynamical cores at varying horizontal and vertical resolutions. Namely, the NCAR Eulerian three-time-level semi-implicit spectral transform dynamical core (EUL), the NCAR two-time-level semi-Lagrangian semi-implicit spectral transform dynamical core (SLD), the German Weather Service icosahedral finite-difference three-time-level semi-implicit dynamical core (GME), and finally the NCAR-FFSL FV dynamical core.

*3.4.6. Diffusion processes*

Before summarizing the performance of FV in the test, it is relevant to list the diffusion processes which had to be included in all the dynamical cores considered in order to ensure stable integrations.

- EUL includes a $\nabla^4$horizontal diffusion on temperature, divergence, and vorticity to control the energy on the smallest resolved scales and a $\nabla^2$ horizontal diffusion on the top three levels to control upward propagating waves. The thermodynamic equation includes a frictional heating term corresponding to the momentum diffusion. It includes also *a posterior* mass fixer applied at every time-step. The three-time-level core includes a time filter to control the $2\Delta t$ time computational mode.
- SLD do not include the $\nabla^4$ and $\nabla^2$ horizontal diffusions; the interpolants control the energy at the smallest scales. Every time-step *posterior* mass and energy fixers are applied. A standard decentering parameter $\varepsilon = 0.2$ is used in the semi-implicit scheme.
- GME includes the $\nabla^4$ and $\nabla^2$ horizontal diffusions as EUL. Neither a mass fixer nor an energy fixer is applied.
- FV do not include explicit $\nabla^4$and $\nabla^2$ horizontal diffusion operators; the horizontal remapping, using a monotonic PPM sub-grid representation, is supposed to control the energy at the smallest scales. An explicit divergence damping is, however, applied. The monotonic and conservative vertical remapping is performed every 10 explicit time-steps. FV employs in addition both a 3-point digital filter in mid-latitudes and an FFT filter in polar regions to control unstable waves in the zonal direction. A *posterior* energy fixer is applied at every time step.

*3.4.7. Resolution, time step, run times*

After the addition of small perturbation to the unstable steady-state zonal flow in the four dynamical cores, they are run for 30 model days with different horizontal resolutions. The five horizontal resolutions and the corresponding time-steps used in the FV integrations are shown in Table 3.1. The other models were run with five approximately equivalent resolutions. The time-steps used in the integrations with the different dynamical cores are also included in the table. All these integrations were run with 26 standard vertical levels (L26).

JW06a also publish the runtimes for the four dynamical cores at their midrange and second highest resolutions. They are added in Table 3.1. They are meant to serve as a general guide for the computational costs of each model acknowledging that the cost are

TABLE 3.1
Horizontal resolutions (FV), time-step $\Delta t$ (s) and wall clock time WT (s) for one model day

| Resolution (FV) $\Delta\varphi \times \Delta\lambda$ | EUL $\Delta t$/WT | SLD $\Delta t$/WT | GME $\Delta t$/WT | FV $\Delta\tau(=\Delta t/10)$/WT |
|---|---|---|---|---|
| $4^\circ \times 5^\circ$ | 2400 | 7200 | 1600 | 720 |
| $2^\circ \times 2.5^\circ$ | 1200 | 3600 | 800 | 360 |
| $1^\circ \times 1.25^\circ$ | 600/44 | 1800/24 | 400/48 | 180/66 |
| $0.5^\circ \times 0.625^\circ$ | 300/483 | 900/271 | 200/325 | 90/625 |
| $0.25^\circ \times 0.3125^\circ$ | 150 | 450 | 100 | 45 |

hardware-dependent and vary with the ingenuity of the programmer. The runtime data represent the wall clock time (WT) needed to complete one model day on a 32-processer node of an IBM (International Business Machines) Power 4 architecture when using a pure message passing interface parallelization approach. Identical compiler optimization flags were used for all models. No efforts was made to optimize the numerical schemes or to configure the models in their optimal setup, such as selecting an optimal time-step or switching from quadratic to a linear truncation technique in case of SLD. The dynamical cores represent the standard versions in CAM3 (The NCAR Community Atmospheric Model system, version3) and GME.

The FV model is seen to be the most expensive in computational costs. Thus, the conservative property is achieved on the expense of efficiency. At the medium resolution, it is 2.75 times more expensive than the semi-Lagrangian semi-implicit SLD dynamical core. However, it is also more expensive than the two Eulerian semi-implicit dynamical cores at medium resolution: 50% more expensive than EUL and 38% more expensive than GME. At the second highest resolution, it is still the most expensive dynamical core. Its WT has increased by a factor 9.5, whereas those of EUL and SLD increased by a factor of 11.0 and 11.3, respectively. However, GME increased much less, by a factor of only 6.8. So at this resolution, FV is 92% more expensive than GME. The advantage of EUL on the other hand has been reduced, but FV is still 29% more expensive than EUL.

It should be noted that here FV is compared with operational dynamical cores that have been carefully optimized. It must be noted furthermore that with massive parallel computers with a high number of nodes, expected to be more common in the future, the FV explicit code is supposed to gain in relative efficiency due to better parallelization than Eulerian as well as semi-Lagrangian semi-implicit codes with elliptic solvers that involve more data exchange between nodes.

### *3.4.8. Results of the global simulations*

As already mentioned, the JW06a tests indicated that the synoptic performance of the FV (or NCAR-FFSL) dynamical core generally is satisfactory. It is producing as realistic idealized baroclinic developments as the state-of-the-art dynamical cores it is compared with. All four dynamical cores compared are found to converge toward a common

solution. Thus, the second-highest and highest resolution FV L26 surface pressure solutions at day 9 are visually almost indistinguishable (see Fig. 6 in JW06a). Up to day 10, differences between the solutions from different dynamical cores can only be seen at the smallest scales that are most influenced by the diffusive characteristics of the numerical schemes, summarized above. An example is the closed cells in the low-pressure center of the surface pressure fields at day 9 (Fig. 7 in JW06a) of the EUL and GME solutions. They are slightly deeper than those from the FV dynamical core. Such small-scale differences are seen more clearly in the 850 hPa relative vorticity fields (shown in Fig. 8 in JW06a). At day 7, the high-resolution FV dynamical core exhibits a slightly weaker vorticity pattern in comparison with EUL, SLD, and GME at high resolutions. According to JW06a, the slightly weaker vorticity fields are caused by the frequent remappings with monotonicity constraint for every short dynamic time-step in the FV dynamical core. This constraint adds nonlinear intrinsic diffusion in the regions where the monotonicity principle is locally violated. Note that JW06a has shown that the slightly more diffusive solution of the FV dynamical core can be matched very closely by EUL and SLD when increasing their diffusive coefficients. So there is no doubt that the excessive smoothing in FV is caused by the intrinsic diffusion caused by the frequent remappings at the end of every short dynamic time-step. Since the vertical remappings are performed only every 10 dynamic time-step, their smoothing effects are less pronounced.

The small-scale differences between the solutions of the different dynamical core were interpreted as an uncertainty of their individual estimates of the true reference solution. JW06a defined the uncertainty as the maximum root mean square deviation $l_2$ between a highest and a second highest horizontal resolution surface pressure simulation among all model versions (see JW06a for details). The uncertainty is increasing with the number of days simulated, becoming more and more large scale, until saturation between day 25 and 30, when the $l_2$ difference is as big as the $l_2$ difference between two randomly picked global surface pressure fields. Using this uncertainty measure, JW06a found that both two highest resolutions of the FV, the SLD and the EUL dynamical cores, converge within the estimated uncertainty to the true solution, whereas only the highest resolution of the GME dynamical core was found to converge.

### *3.4.9. Idealized test of HIRLAM-DCISL*

Also the HIRLAM-DCISL dynamical core has been tested with the Jablonowski-Williamson test case. As it would be difficult to extend the limited area of the HIRLAM-DCISL dynamical core to a global domain, its domain was made as global as possible and an effort was made to minimize the effects of its boundaries. The active domain was extended meridionally to 80°S – 80°N and zonally to 80°W – 280°E, without changing the zero divergence boundary condition, used in the elliptic system solver, to a periodic boundary condition at the zonal boundaries. The zonal extension was chosen so that the initial perturbation, centered at (20°E, 40°N), which triggers a *main wave* is separated (by exactly 100°) from the western domain boundary where the boundary scheme initially trigger a weak *boundary wave*. Both waves develop into wave trains which become less and less separated, although they move with approximately the same speed toward the east. The usual HIRLAM boundary relaxation scheme is applied in a 6°

wide zone along the boundary inside the active domain. Within this zone, the updated prognostic variables are relaxed toward the initial values with a weight that decrease from 1 at the boundary to zero approximately 6° inside it. To accommodate the DCISL upstream integrations, there is also a halo zone around the active domain in which the prognostic variables are held fixed at the initial values. In order to facilitate a comparison with the global FV dynamical core reference solution, the boundary wave was effectively eliminated from both the HIRLAM-DCISL and the HIRLAM solutions. This was done by utilizing that a completely similar boundary wave is created in simulations without an initial perturbation. The simulations with the boundary wave removed were then compared with the global reference solution. This was done only up to day 8 after which the main wave reached the eastern boundary zone.

*3.4.10. Resolution, time-step and runtimes*

Two horizontal resolutions were used. The lower resolution, corresponding to the middle global resolution (see Table 3.1), is $\Delta\varphi \times \Delta\lambda = 1.15° \times 1.45°$ and the higher resolution, corresponding to the second highest global resolution, is $\Delta\varphi \times \Delta\lambda = 0.59° \times 0.74°$. In the vertical, 27 levels are placed as in JW06b but with one more level at the top of the model atmosphere to accommodate the zero top pressure of HIRLAM. Like the global SLD dynamical core, the time-steps used for the two horizontal resolutions were 30 and 15 minutes, respectively. On a single NEC SX6 processor using ad hoc coding with almost no optimization, the lowest resolution HIRLAM-DCISL dynamical core is approximately twice as expensive as the corresponding highly optimized reference HIRLAM dynamical core. There is no doubt, however, that the efficiency of the HIRLAM-DCISL dynamical core can be increased considerably by a dedicated optimization.

*3.4.11. Diffusion processes*

- HIRLAM use decentering with a decentering parameter $\varepsilon = 0.1$. The nonlinear terms in continuity equation, the thermodynamic equation, and the momentum equations are needed at time level $n + 1/2$. As they are potential sources of instability, they are extrapolated from filtered values at time level $n - 1$ as follows: $\psi^{n+1/2} = (3\psi^n - \psi_f^{n-1})/2$, where $\psi$ is any of the nonlinear terms and $\psi_f^{n-1} = \psi^{n-1} + \varepsilon_M[\psi^n - 2\psi^{n-1} + \psi_f^{n-2}]$ with $\varepsilon_N = 0.1$. At the end of each time-step, all prognostic variable, except liquid water, are diffused using an approximate implicit $\nabla^4$ horizontal diffusion with the diffusion coefficient $K = 3.5 \times 10^{14}$ for $\Delta x = 0.5°$ and $\Delta t = 300$ s (see p. 12–13 in Undén 2002). The coefficients are scaled for resolution so that the e-folding time of the $2\Delta x$ wave is the same regardless of resolution (McDonald, 1998). There has been no attempt to tune the diffusion coefficient for the present idealized dry adiabatic simulations. The horizontal diffusion was increased at the uppermost 4 model layers. In addition, the horizontal and vertical interpolations, using cubic Lagrange interpolation, are supposed to control the energy at the smallest scales. No mass and energy fixers are applied.
- HIRLAM-DCISL does not use decentering and filtering of the nonlinear terms, but it was necessary to retain a weak implicit $\nabla^6$ horizontal diffusion on $T$, $u$, and $v$. The horizontal diffusion was increased at the uppermost 4 model layers. In addition, the horizontal and vertical remapping and interpolations, using, respectively, a

positive-definite PPM subgrid representation and cubic Lagrange interpolation, are supposed to control the energy at the smallest scales. No energy fixer is applied

*3.4.12. Results of simulations*

For both resolutions of HIRLAM and HIRLAM-DCISL, the $l_2$ difference between the simulation and the simulation of the global highest resolution FV dynamical core ($\Delta\varphi \times \Delta\lambda = 0.25° \times 0.3125°$) was computed. The results showed that up to day 8 both the highest resolution ($\Delta\varphi \times \Delta\lambda = 0.59° \times 0.74°$) simulation of HIRLAM and the highest resolution simulation of HIRLAM-DCISL had converted within the uncertainty of the reference solution. For the lower resolution simulations, the simulation of the HIRLAM-DCISL version had not converted, whereas that of the HIRLAM version had; so the FV version needs higher resolution than the grid-point version for the same level of accuracy (Fig. 3(a) in Lauritzen, Kaas, Machenhauer and Lindberg [2008]). The explanation seems, as for the global FV dynamical core, to be too heavy smoothing due to the repeated remappings and interpolations. Regarding phase error, the HIRLAM-DCISL simulation is slightly better than the HIRLAM simulation. When using the cascade scheme of Nair, Scroggs and Semazzi [2002] instead of the fully 2D CISL scheme of Nair and Machenhauer [2002], the accuracy in terms of the $l_2$ difference is not altered (Fig. 3(b) in Lauritzen, Kaas, Machenhauer and Lindberg [2008]). An important result of the idealized baroclinic wave tests is that the consistent Lagrangian discretization of the energy conversion term, introduced in Section 3.2.1, is seen clearly to be better with both a smaller $l_2$ difference and a smaller phase error, than when using the traditional Eulerian discretization (Fig. 3(b) in Lauritzen, Kaas, Machenhauer and Lindberg [2008]).

HIRLAM-DCISL has also been coupled with the HIRLAM physics package and initial test runs from the initial conditions of a strongly developing extratropical storm have been performed. The mass conserving version ran stably and produced simulations that were quite similar to the reference HIRLAM simulations, except again for slightly more smoothing in the DCISL version. Also for the full-physics run, the Lagrangian discretization of the energy conversion term leads to a more accurate simulation than the traditional discretization. The results of these tests are mentioned in Lauritzen, Kaas, Machenhauer and Lindberg [2008]. In this paper also a possible cure for the slightly excessive smoothing is suggested, although it was not tested in practice. It is suggested to keep the Lagrangian cells in the Lagrangian model layers for a number of consecutive large semi-implicit time-steps before performing the vertical remapping and interpolation to the Eulerian model layers and levels, just as it is done in the NCAR-FFSL over 10 consecutive small time-steps. In HIRLAM-DCISL, an additional vertical remapping and interpolation must be performed after each of the long semi-implicit time-steps as it is needed for the physical parameterization. Thus, it will not affect the computational efficiency.

*3.5. Online and offline applications – The problem of mass-wind inconsistency*

An obvious application of FV models such as the NCAR-FFSL and the HIRLAM-DCISL is tracer transport since tracer-mass conservation is important. These quasi-hydrostatic FV models use a pressure-based vertical coordinate and the prognostic variable for tracer

mass is the cell averaged value $\overline{\overline{q_i}}^{\Delta} \Delta\overline{p}$, where $\overline{p}$ is the horizontal mean pressure over the area $\Delta A$ of an Eulerian grid cell and $\Delta\overline{p}$ is the pressure thickness of the cell. $\overline{\overline{q_i}}^{\Delta}$ is the Eulerian cell average specific concentration[6] of the tracer in question. Thus, $\overline{\overline{q_i}}^{\Delta} \Delta\overline{p}$ is the weight of tracer mass in the cell per unit horizontal area.[7] At each time-step, the FV model solves at first the continuity equation for air mass. The input to the continuity equation for air mass is the horizontal wind field which together with the hydrostatic balance determines the 3D trajectories along which the air is transported. The output is the updated values of $\Delta\overline{p}$. Hereafter, the continuity equation for each specific tracer mass is solved, using the same trajectories, giving the updated values of $\overline{\overline{q_i}}^{\Delta} \Delta\overline{p}$. Thus, both the predicted air mass and the predicted tracer mass fields are consistent with the "driving" horizontal velocity field. If, the specific concentration $q_i$ is needed (e.g., for a compability check as illustrated in Fig. 2.1) $\overline{\overline{q_i}}^{\Delta} \Delta\overline{p}$ must be divided by $\Delta\overline{p}$. In a model setup where the tracer continuity equation is an integrated part of the dynamical core, this can be done without loss of mass because $\overline{\overline{q_i}}^{\Delta} \Delta\overline{p}$ and $\Delta\overline{p}$ are internally consistent, i.e., they are both computed by the same mass conserving transport and remapping operations.

Often a FV transport scheme is imported into a GCM, which dynamical core does not conserve the mass of air locally. For example, in the Eulerian ECHAM5 (ECmwf/HAMburg, version 5) model, where the vertically integrated mass variable is log $p_s$ the continuity equation for air is solved using the spectral transform method. This model is neither globally nor locally mass conserving, while the tracer transport is performed using the inherently mass-conserving advection scheme of Lin and Rood [1996]. This set up is called an *online coupling*. Here, the FV transport scheme is solved on the same grid as used by the GCM and the horizontal GCM winds $\vec{V}_{GCM}$ needed by the transport scheme are provided by the GCM at every GCM time-step. A problem of such an online coupling is that the GCM predicts its own *air mass field* $\Delta p_{GCM}$, which is generally different from the $\Delta\overline{p}$ predicted from *the GCM wind field* by the mass conservative tracer transport scheme with $q_i = 1$ (see Fig. 3.6). This is a manifestation of the so-called *mass-wind inconsistency* discussed in detail by Jöckel, Von Kuhlmann, Lawrence, Steil, Brenninkmeijer, Crutzen, Rasch and Eaton [2001].

The consequences of the mass-wind inconsistency in long online coupled simulations can be severe. Jöckel, Von Kuhlmann, Lawrence, Steil, Brenninkmeijer, Crutzen, Rasch and Eaton [2001] ran a low-resolution FV transport scheme online coupled to a nonmass conserving GCM. For passive tracers initialized at different locations in the atmosphere, the variations in the total mass were up to 70% in a one-year simulation. The amount of artificially (spuriously) created and destroyed mass due to the mass-wind inconsistency is strongly dependent on the vertical gradient of the tracer. Since tracer gradients are usually steepest around the tropopause the problem is large in the tropopause region.

[6] *The specific tracer concentration* $q_i$ is the ratio between the mass of the tracer $m_t$ and the mass of the moist air it is mixed into $m_v$.

[7] $\overline{\overline{q_i}}^{\Delta} \Delta\overline{p} = \frac{m_t}{m_v}\overline{\overline{\rho}}_v g\,\Delta z = \frac{m_t}{m_v}\frac{m_v}{\Delta A \Delta z} g\,\Delta z = g\frac{m_t}{\Delta A}$ = the weight of tracer mass per unit horizontal area.

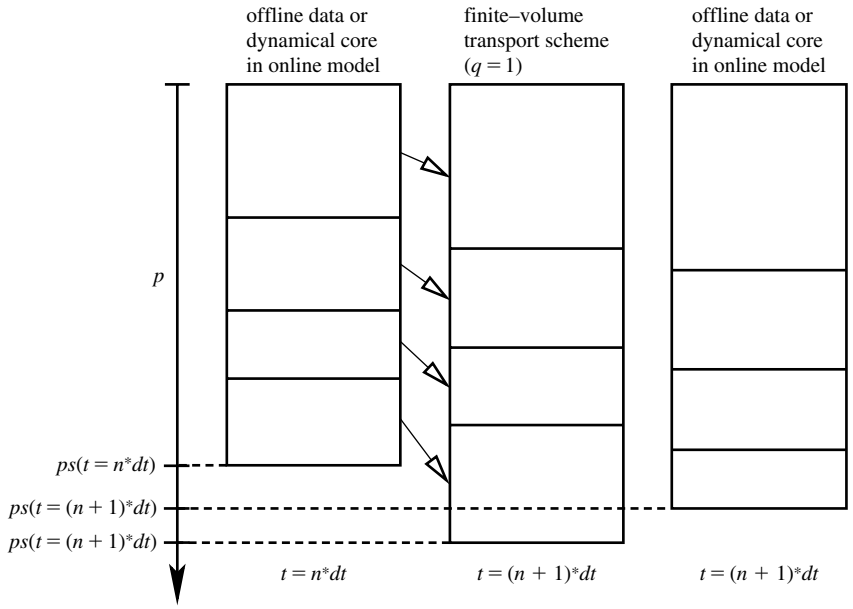


FIG. 3.6 Graphical illustration of the mass-wind inconsistency. The figure shows the location of pressure levels at the beginning of a time-step $t = n\Delta t$ (left), after one time-step $t = (n + 1)\,\Delta t$ using a FV transport scheme (middle) and given by offline data or predicted by the continuity equation of the dynamical core (right), respectively. If the vertical levels implied by the transport scheme and the dynamical core or offline data do not coincide, an inconsistency between the mass and wind fields exists and affects the mass of the tracer advection.

If the wind and pressure data driving a FV transport model are not given at every time-step and may be specified on another grid than used in the FV transport model (typically an archived meteorological data-set such as ECMWF reanalysis (REA), ERA40, or the National Centers for Environmental Prediction (NCEP)/NCAR REA) so that both interpolation in time and space is needed, then the coupling is called *offline*. This is typically the situation in a *chemical transport model*. (It should be noted that this does not apply to the NCAR-FFSL model although this model for the tracer transport uses a large time-step that is equal to an integer number $m$ of the small time-steps, which are used for determining the air mass transport. This is because the fluxes used in the tracer transport during a large time-step is obtained by accumulating fluxes over the $m$ small time-steps.) In an *offline* setting, assimilated analysis or REA data is often provided to the FV transport scheme. Both the driving horizontal wind field $V_{REA}$ and the corresponding mass field $\Delta p_{REA}$ must be spatially and temporally interpolated to accommodate the grid and time-step used by the transport scheme. In such a situation, the consistency between the mass and wind fields cannot be achieved unless a posteriori consistency correction methods are applied to the *offline* data. This inconsistency with respect to the ECMWF analyses is discussed in TRENBERTH [1991].

One can attempt to restore the mass conservation by altering the specific tracer concentrations a posteriori. JÖCKEL, VON KUHLMANN, LAWRENCE, STEIL, BRENNINKMEIJER,

CRUTZEN, RASCH and EATON [2001] investigated the effects of various a posteriori mass-fixing algorithms. But all these fixers have severe disadvantages such as violation of shape preservation or introduction of non physical transport components.

Instead of altering the specific tracer concentrations a posteriori, one can as well adjust the horizontal velocity field such that the tracer advection equation is consistent with the mass field, i.e., the winds are corrected so that the vertical integrated divergence of mass matches the surface pressure tendency of the meteorological data (for details see PRATHER, MCELROY, WOFSY, RUSSEL and RIND [1987]; ROTMAN, TANNAHILL, KINNISON, CONNELL, BERGMANN, PROCTOR, RODRIGUEZ, LIN, ROOD, PRATHER, RASCH, CONSIDINE, RAMAROSON and KAWA [2001]; CAMERON-SMITH, CONNELL and PRATHER [2002])[8]. This type of restoration algorithm is referred to as a *pressure fixer*. Contrary to the algorithms described in the preceding paragraph, it ensures that constant specific concentrations and mass conservation are retained with this pressure fixer. However, the approach is not completely satisfactory either since "true" wind data are enforced to provide mass-wind consistency. Introduction of a pressure fixer in a semi-Lagrangian FV model would be somewhat different from that used by Eulerian type FV models since divergence is defined directly by the trajectories. Therefore modifications of trajectories are needed in such models to achieve an analogy to the traditional pressure fixer.

The pressure fixer method can be used to indicate the severity of the mass-wind inconsistency problem, i.e., by running a model with and without a pressure fixer and assuming that the pressure fixer does not have a significant effect on the wind field. HOROWITZ, WALTERS, MAUZERALL, EMMONS, RASCH, GRANIER, TIE, LAMARQUE, SCHULTZ, TYNDALL, ORLANDO and BRASSEURET [2003] have run the global Model of Ozone Research version 2 (MOZART-2) with and without a pressure fixer. Near the tropopause (where the vertical gradient of the ozone specific concentration is large) the difference between the two runs was approximately 187 Tg/yr. Assuming that the pressure fixer is perfect, it can be estimated that a spurious source of ozone of 187 Tg/yr is caused by the mass-wind inconsistency problem which is not a negligible amount. For example, the spurious source of ozone is similar in magnitude to the estimated amount of influx of ozone per year to the troposphere from the stratosphere. This is, of course, only an indication of the magnitude of the problem. In order to estimate the systematic spurious sources and sinks a fully consistent model must be run. But the estimates provided by HOROWITZ, WALTERS, MAUZERALL, EMMONS, RASCH, GRANIER, TIE, LAMARQUE, SCHULTZ, TYNDALL, ORLANDO and BRASSEURET [2003] suggest that the mass-wind inconsistency can introduce significant errors.

As discussed above, the problem of performing accurate offline or online tracer transport in a model using a pressure-based vertical coordinate is not limited to the use of an accurate FV tracer transport scheme, but it is also a question of mass-wind consistency. That is consistency between on one hand the mass field $\Delta\bar{p}$ determined by the FV transport scheme and the driving wind field and on the other hand the associated mass field $\Delta p_{REA}$ or $\Delta p_{GCM}$. In offline applications using existing REA data, there is little choice but to use some kind of correction method. However, it is hoped that

[8]Alternatively one may adjust the surface pressure field instead of the horizontal velocity field to achieve consistency (P. JÖCKEL personal communication).

in the future analysis data set with a better inherent mass conservation, produced by mass conserving data-assimilation models, may become available. In an online transport coupling, the consistency can be guaranteed only if the same numerical method is used for the continuity equation of the driving model as for the tracer transport. That is the case if a consistent FV model like the NCAR-FFSL or the HIRHAM-DCISL is used. The problem is that the majority of the GCM models available are based on nonlocally conservative schemes for the air mass continuity equation (e.g., traditional semi-Lagrangian models such as the IFS at ECMWF, the HIRLAM, and the Max-Planck Institute model (ECHAM)). The problem is, as described above, that the changes needed to convert a non-conserving model to a locally mass conserving one, like the change of HIRLAM to the HIRLAM-DCISL, are rather extensive. All discretizations in the model as a whole must be carefully rethought in order to obtain inherent local mass conservation. It involves changes to almost all parts of the model and all the prognostic equations and not only changes to the continuity equations. However, it is necessary in order to guarantee consistency and thereby accurate tracer transport.

### *3.6. Extensions to nonhydrostatic models*

In the quasi-hydrostatic FV dynamical cores considered in Sections 3.1–3.5, the continuity equation was solved in so-called Lagrangian ($\xi$) vertical coordinates, with, per definition, $\xi =$ constant along 3D trajectories. That is, all transport during a time-step was assumed to be effectuated by Lagrangian finite control volumes moving with the 3D flow, usually in a semi-Lagrangian sense, starting or ending as Eulerian grid cells. The advantage is that in such Lagrangian coordinates, the vertical velocity is zero as expressed in Eq. (1.1), so the transport problem becomes 2D. However, the vertical components of the 3D trajectories still need to be determined. In the quasi-hydrostatic models considered so far, this is done hydrostatically, i.e., the vertical displacement of the Lagrangian control-volume during a time-step was determined by requiring that the arrival cell is in hydrostatic balance. Of course, this approach cannot be used in a nonhydrostatic model, but instead the vertical displacement can be determined directly from the vertical velocity, which in a nonhydrostatic model is an independent prognostic variable. Realizing this one can formulate a DCISL solution to the continuity equation also for the nonhydrostatic case.

We may start from the continuity equation on the form (Eq. (1.8)), which was derived without assuming hydrostatic balance. Using the notations in Section 1.1 Eq. (1.8) is

$$\overline{\left(\tilde{\rho}\ \delta_k h\right)}^{+} \Delta A = \overline{\left(\tilde{\rho}\ \delta_k h\right)}\, \delta_k A. \tag{3.149}$$

This is a prognostic equation when $\left(\tilde{\rho}\ \delta_k h\right)$ is known in the departure area at time $t$. Unlike what we did in the "exact" case, we now set the height of the Lagrangian surfaces $\xi_{k-1/2}$ and $\xi_{k+1/2}$ equal to the Eulerian surfaces $h_{k-1/2}$ and $h_{k+1/2}$, respectively, in the departure area so that here $\delta_k h = \Delta_k h$. Thus, Eq. (3.149) becomes

$$\overline{\left(\tilde{\rho}\ \delta_k h\right)}^{+} \Delta A = \Delta_k h\, \overline{\left(\tilde{\rho}\right)}\, \delta_k A, \tag{3.150}$$

Now in the nonhydrostatic case, we utilize the prognostic variable $w$, the vertical velocity, to determine $\delta_k h$. So it is a known constant in Eq. (3.150). Therefore, it may be written as

$$\overline{(\tilde{\rho})}^{+}\delta_k h \Delta A = \Delta_k h\, \overline{(\tilde{\rho})}\, \delta_k A, \tag{3.151}$$

or

$$\overline{(\tilde{\rho}\,)}^{+} = \frac{\Delta_k h}{\delta_k h\ \Delta A} \iint\limits_{\delta_k A} \tilde{\rho}\, \mathrm{d}x\, \mathrm{d}y. \tag{3.152}$$

$\overline{(\tilde{\rho})}^{+}\delta_k h\ \Delta A$ is the updated mass in the arrival grid cell at time $t + \Delta t$. According to Eq. (3.152), it is equal to the mass in the upstream departure cell at time $t$, which can be computed simply as a 2D integral over the departure area of the vertical mean density in the Eulerian model layer considered.

This demonstrates that FV methods of the DCISL type, like the one used in HIRLAM-DCISL, can be used also for nonhydrostatic models. It also means that the Lagrangian ($\xi$) vertical coordinate approach can be used in nonhydrostatic Eulerian flux-type FV models, just as the corresponding quasi-hydrostatic approach was used in the NCAR-FFSL dynamical core.

Of course, alternatively, a nonhydrostatic model may be based on a traditional 3D operator-split flux-form method, with fluxes entering an Eulerian grid cell at both horizontal and vertical faces. However, it becomes rather complicated if one uses the most accurate schemes of the symmetric FFSL type. Thus, Leonard, Lock and Macvean [1996] presents a symmetric 3D scheme. Using a notation similar to that used in Section 2.3.2, it becomes

$$\begin{aligned}
\overline{\psi}^{n+1} = \overline{\psi}^{n} &+ X_C \left\{ \frac{'1}{6}\left[ \left( \overline{\psi}^{n} + \overline{\psi}_{AY} + \overline{\psi}_{AYZ} \right) + \left( \overline{\psi}^{n} + \overline{\psi}_{AZ} + \overline{\psi}_{AZY} \right) \right] \right\} \\
&+ Y_C \left\{ \frac{'1}{6}\left[ \left( \overline{\psi}^{n} + \overline{\psi}_{AZ} + \overline{\psi}_{AZX} \right) + \left( \overline{\psi}^{n} + \overline{\psi}_{AX} + \overline{\psi}_{AXZ} \right) \right] \right\} \\
&+ Z_C \left\{ \frac{'1}{6}\left[ \left( \overline{\psi}^{n} + \overline{\psi}_{AX} + \overline{\psi}_{AXY} \right) + \left( \overline{\psi}^{n} + \overline{\psi}_{AY} + \overline{\psi}_{AYX} \right) \right] \right\},
\end{aligned} \tag{3.153}$$

where, for example,

$$\overline{\psi}_{AYZ} = \overline{\psi}_{AZ}\left(\overline{\psi}_{AY}\right) = \overline{\psi}_{AY} + Z_A\left(\overline{\psi}_{AY}\right). \tag{3.154}$$

Such schemes have been used extensively; a recent example is the MIT-GCM (Adcroft, Campin, Hill and Marshall [2004]).

## 4. Summary

Recent developments in FV methods have provided the basis for new meteorological dynamical cores that conserve integral invariants exactly, globally as well as locally. In particular, these new FV methods have been the basis for design of exact mass conserving

tracer transport models. The new technologies are reviewed and the perspectives for the future are discussed.

During about two decades, the traditional semi-implicit and semi-Lagrangian spectral or grid-point dynamical cores have been dominating worldwide in meteorological models applied for weather prediction and climate simulations. They are efficient and otherwise accurate but lack exact mass conservation, which is considered a serious drawback for the hydrometeorological variables as well as an increasing number of chemical variables included in the models. In Section 3, we presented two recently developed pioneering meteorological dynamical cores which potentially solve these problems. These are the semi-implicit cell-integrated semi-Lagrangian limited area dynamical core HIRLAM-DCISL and the global flux-form NCAR-FFSL dynamical core. Each of them extends newly developed 2D FV semi-Lagrangian schemes, described among others in Section 2, to 3D utilizing a common newly developed Lagrangian time-stepping technique building on horizontally upstream and vertically downstream time-steps. Values of certain quantities integrated horizontally over an upstream departure area in an Eulerian model layer are assumed to be transported with vertical walls along 3D trajectories into a Lagrangian layer in a column of Eulerian grid cells. The vertical coordinates of this Lagrangian layer are then determined hydrostatically. As a result, the quantities in question are conserved exactly globally and with high, slightly different accuracy, also locally in both dynamical cores. Idealized tests presented in Section 2 showed that the local conservation is slightly more accurate in the HIRLAM-DCISL transport schemes than in the NCAR-FFSL scheme. With proper boundary conditions, the HIRLAM-DCISL dynamical core conserves the global mass of moist air and tracers exactly, including water vapor, liquid water, and solid water, if included (apart from evaporation and condensation). Also the NCAR-FFSL dynamical core conserves exactly these masses. In addition, it conserves, except for time truncation errors, potential temperature and absolute vorticity (in adiabatic friction-free flow). Thus, the NCAR-FFSL dynamical core comes closer than the HIRLAM-DCISL to the ideal CSCL model considered in Section 3.1. Both FV dynamical cores have been tested and compared with nonconservative dynamical cores in an idealized baroclinic wave test. All dynamical cores considered were found to converge toward a common solution. However, in their present formulation, the FV dynamical cores needed higher resolution than the nonconservative dynamical cores they were compared with for the same level of accuracy. The explanation seems to be a slight smoothing due to the repeated remappings and interpolations. As a possible cure to HIRLAM-DCISL, it is suggested to keep the vertical Lagrangian cells for a number of consecutive large semi-implicit time-steps before performing the vertical remapping to the Eulerian model layers. The idealized as well as other tests with HIRLAM DCISL showed that a consistent Lagrangian discretization of the energy conversion term in the thermodynamic equation leads to a more accurate simulation than the traditional discretization. A further increase in accuracy is expected from a corresponding Lagrangian discretization of the horizontal pressure gradient term in the momentum equation. The idealized intercomparison tests showed that among the dynamical cores the FV ones are the most expensive in computational costs. Thus, the conservative property is achieved at the expense of efficiency. There is no doubt, however, that the efficiency of the present experimental ad hoc coded FV dynamical cores can be increased considerably by a dedicated optimization.

An obvious application of FV models such as the NCAR-FFSL and the HIRLAM-DCISL is tracer transport since tracer-mass conservation is essential. As described in Section 3.5, a FV transport scheme has often been imported into a GCM with a dynamical core that does not conserve mass locally. A problem of such an online coupling is that it leads to the so-called mass-wind inconsistency, which in long simulations can lead to severe errors with large amounts of artificially (spuriously) created or destroyed tracer mass. The only way to avoid completely such errors is to use a complete FV model with exactly the same locally mass conserving algorithms for all tracers and for the moist air.

To facilitate a wider application, the limited area HIRLAM-DCISL may be extended to a global domain. This may be done by using the extension of the horizontal FV schemes used in this dynamical core that have already been developed in spherical latitude-longitude coordinates. Another possibility which may be relevant also for the NCAR-FFSL is to change to a new grid, as the icosahedral-hexagonal grid, which is almost uniform on the sphere. Finally, the possibility of an extension to nonhydrostatic dynamical cores is discussed in Section 3.6. The same extension, as used in the hydrostatic dynamical cores, of the available 2D FV semi-Lagrangian schemes to 3Ds may be used in nonhydrostatic dynamical cores. That is, utilizing horizontally upstream and vertically downstream time-stepping. The only difference is that the vertical displacement of the Lagrangian cells must be determined directly by predicted vertical velocities and not from hydrostatic balance.

## 5. Acknowledgments

The authors wish to express their gratitude to Dr. Ramachandran D. Nair and Dr. Phil Rasch for helpful suggestions and useful discussions on parts of Section 2. Thanks to Dr. Patrick Jöckel and Rune Graversen for their comments on the ‘mass-wind inconsistency’ section. The third author is grateful to NCAR’s Advanced Study Program and Climate Modeling Section for providing necessary support for this research.

# References


Adcroft, A., Campin, J-M., Hill, C., Marshall, J. (2004). Implementation of an atmosphere-ocean general circulation model on the expanded spherical cube. *Mon. Wea. Rev.* **132** (12), 2845–2863.

Arakawa, A. (2000). A personal perspective on the early years of general circulation modeling at UCLA. In: Randall, D.A. (ed.), *General Circulation Model Development. Past, Present, and Future* (Academic Press), pp. 1–65.

Arakawa, A., Lamb, V. (1977). Computational design and the basic dynamical processes of the UCLA general circulation model. *Methods Comput. Phys.* **17**, 173–265.

Arakawa, A., Lamb, V. (1981). A potential enstrophy and energy conserving scheme for the shallow water equations. *Mon. wea. Rev.* **109**, 18–36.

Bates, J.R., Mcdonald, A. (1982). Multiply-upstream, semi-Lagrangian advective schemes: analysis and application to a multi-level primitive equation model. *Mon. Wea. Rev.* **110** (12), 1831–1842.

Bell, J.B., Dawson, C.N., Shubin, G.N. (1988). An unsplit, higher order Godunov method for scalar conservation laws in multiple dimensions. *J. Comput. Phys.* **74**, 1–24.

Bermejo, R. Staniforth, A. (1992). The conversion of semi-Lagrangian advection schemes to quasi-monotone schemes. *Mon. Wea. Rev.* **120** (11), 2622–2632.

Boris, J.P., Book, D.L. (1973). Flux corrected transport. I. SHASTA: a fluid transport algorithm that works. *J. Comput. Phys.* **11**, 38–69.

Bott, A. (1989). A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes. *Mon. Wea. Rev.* **117** (5), 1006–1016.

Bott, A. (1992). Monotone flux limitation in the area-preserving flux-form advection algorithm. *Mon. Wea. Rev.* **120** (11), 2592–2602.

Bott, A. (1993). The monotone area-preserving flux-form advection algorithm: reducing the time-splitting error in two-dimensional flow fields. *Mon. Wea. Rev.* **121** (9), 2637–2641.

Burridge, D.M., Hasler, J. (1977). A model for medium range forecasting. Adiabatic formulation. ECMWF Tech. Rep. 4, Reading, United Kingdom, pp. 46.

Byun, B.W. (1999). Dynamically consistent formulations in meteorological and air quality models for multiscale atmospheric studies, Part II: mass conservation issues. *J. Atmos. Sci.* **56** (21), 3808–3820.

Cameron-Smith, P., Connell, P., Prather, M. (2002). Pressure fixers: conserving mass in chemical transport models by adjusting the winds. Available at 2002; http://eed.llnl.gov/pfix/.

Carpenter, J.R., Droegemeier, K.K., Woodward, P.R., Hane, C.E. (1990). Application of the piecewise parabolic method (PPM) to meteorological modeling. *Mon. Wea. Rev.* **118** (3), 586–612.

Colella, P. (1990). Multidimensional upwind methods for hyperbolic conservation laws. *J. Comput. Phys.* **87**, 171–200.

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

Collins, W.D., Rasch, P.J., Boville, B.A., Hack, J.J., Mccaa, J.R., Williamson, D.L., Kiehl, J.T., Briegleb, B., Bitz, C., Lin, S.-J., Zhang, M., Dai, Y. (2004). Description of the NCAR Community Atmosphere Model (CAM3). Technical Report, NCAR/TN- 464+STR, National Center for Atmospheric Research, Boulder, Colorado 80307-3000, pp. 226.

Cotter, C.J., Frank, J., Reich, S. (2007). The remapped particle-mesh semi-Lagrangian advection scheme. *Q. J. R. Meteorol. Soc.* **133**, 251–260.

Crowley, W.P. (1968). Numerical advection experiments. *Mon. Wea. Rev.* **96** (1), 1–11.

Doswell III, C.A. (1984). A kinematic analysis of frontogenesis associated with a nondivergent vortex. *J. Atmos. Sci.* **41** (7), 1242–1248.

Dukowicz, J.K., Baumgardner, J.R. (2000). Incremental remapping as a transport/advection algorithm. *J. Comput. Phys.* **160**, 318–335.

Dukowicz, J.K., Kodis, J.W. (1987). Accurate conservative remapping (rezoning) for arbitrary Lagrangian-Eulerian computations. *SIAM J. Sci. Stat. Comput.* **8**, 305–321.

Dukowicz, J.K., Ramshaw, J.D. (1979). Tensor viscosity method for convection in numerical fluid dynamics. *J. Comput. Phys.* **32**, 71–79.

Durran, D.R. (1999). *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics* (Springer-Verlag), pp. 465.

Easter, R.E. (1993). Two modified versions of Bott's positive-definite numerical advection scheme. *Mon. Wea. Rev.* **121** (1), 297–304.

Eumard, R., Gallouët, T., Herbin, R. (2000). Finite volume methods. In: Ciarlet, P.G., Lions, J.L. (eds.), *Handbook of Numerical Analysis, vol. VII* (Elsevier Science B.V.), pp. 715–1020.

Godunov, S.K. (1959). Finite difference method for numerical computation of discontinous solution of the equations of fluid dynamics. *Math. Sb.* **47**, 271, translated from Russian by I. Bohachevsky.

Hirsch, C. (1990). *Numerical Computation of Internal and External Flows, vol. 2*, Wiley Series in Numerical Methods (John Wiley and Sons), pp. 714.

Hirt, C.W., Amsden, A.A., Cook, J.L. (1974). An arbitrary Lagrangian-Eulerian computing method for all flow speeds. *J. Comput. Phys.* **14**, 227.

Hólm, E.V. (1995). A fully two-dimensional, nonoscillatory advection scheme for momentum and scalar transport equations. *Mon. Wea. Rev.* **123** (2), 536–552.

Horowitz, L.W., Walters, S., Mauzerall, D.L., Emmons, L.K., Rasch, P.J., Granier, C., Tie, X., Lamarque, J., Schultz, M.G., Tyndall, G.S., Orlando, J.J., Brasseuret, G.P. (2003). A global simulation of tropospheric ozone and related tracers: description and evaluation of MOZART, version 2. *J. Geophys. Res.* **108** (D24), 4784.

Hortal, M. (2002). The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model. *Q. J. R. Meteorol. Soc.* **128**, 1671–1687.

Hourdin, F., Armengaud, A. (1999). The use of finite-volume methods for atmospheric advection of trace species. Part I: test of various formulation in a general circulation model. *Mon. Wea. Rev.* **127** (5), 822–837.

Jablonowski, C. and Williamson, D.L. (2006a). A baroclinic instability test case for atmospheric model dynamical cores. *Q. J. R. Meteorol. Soc.* **132**, 2943–2975.

Jablonowski, C. and Williamson, D.L. (2006b). A baroclinic wave test case for dynamical cores of general circulation models: Model intercomparisons. NCAR Technical Note, NCAR/TN-469+STR, Boulder CO, pp. 89.

Jöckel, P., Von Kuhlmann, R., Lawrence, M.G., Steil, B., Brenninkmeijer, C.A.M., Crutzen, P.J., Rasch, P.J., Eaton, B. (2001). On a fundamental problem in implementing flux-form advection schemes for tracer transport in 3-dimensional general circulation and chemistry transport models. *Q. J. R. Meteorol. Soc.* **127**, 1035–1052.

Kaas, E. (2008). A simple and efficient locally mass conserving semi-Lagrangian transport scheme. *Tellus A*, in press.

Källén, E. (1996). HIRLAM documentation manual, System 2.5. Available from SMHI, Norrköping, Sweden.

Kuo, H., Williams, R.T. (1990). Semi-Lagrangian solutions to the inviscid burgers equation. *Mon. Wea. Rev.* **118** (6), 1278–1288.

Laprise, L.P.R., Plante, A. (1995). A class of semi-Lagrangian integrated-mass (SLIM) numerical transport algorithms. *Mon. Wea. Rev.* **123** (2), 553–565.

Lauritzen, P.H. (2007). A stability analysis of finite-volume advection schemes permitting long time steps. *Mon. Wea. Rev.* **135** (7), 2658–2673.

Lauritzen, P.H., Kaas, E., Machenhauer, B. (2006). A mass-conservative semi-implicit, semi-Lagrangian limited-area shallow water model on the sphere. *Mon. Wea. Rev.* **134** (4), 1196–1212.

Lauritzen, P.H., Kaas, E., Machenhauer, B., Lindberg, K. (2008). A mass-conservative version of the semi-implicit semi-Lagrangian HIRLAM. *Q. J. R. Meteorol. Soc.*, **134** (635), 1583–1595.

Lauritzen, P.H., Nair, R.D. (2008). Monotone and conservative cascade remapping between spherical grids (CaRS): regular latitude-longitude and cubed-sphere grids. *Mon. Wea. Rev.*, **136** (4), 1416–1432.

Leith, C.E. (1965). Numerical simulation of the earth's atmosphere. *Methods Comput. Phys.* **4**, 1–28.

Leonard, B.P. (1994). Note on the von Neumann stability of explicit one-dimensional advection schemes. *Comp. Meth. Appl. Mech. Ens.* **118**, 29–46.

Leonard, B.P., Lock, A.P., Macvean, M.K. (1996). Conservative explicit unrestricted time step multidimensional constancy-preserving advection schemes. *Mon. Wea. Rev.* **124** (11), 2588–2606.

Leslie, M.L., Purser, R.J. (1995). Three-dimensional mass-conserving semi-Lagrangian scheme employing forward trajectories. *Mon. Wea. Rev.* **123** (8), 2551–2566.

Leveque, R.J. (2002). *Finite Volume Methods for Hyperbolic Problems* (Cambridge University Press).

Li, Y., Chang, J.S. (1996). A mass-conservative, positive-definite, and efficient Eulerian advection scheme in spherical geometry and on a non-uniform grid system. *J. Appl. Meteorol.* **35**, 1897–1913.

Lin, S.-J. (1997). A finite-volume integration method for computing pressure gradient forces in general vertical coordinates. *Q. J. R. Meteorol. Soc.* **123**, 1749–1762.

Lin, S.-J. (2004). A "Vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.* **132** (10), 2293–2307.

Lin, S.-J., Rood, R.B. (1996). Multidimensional flux-form semi-Lagrangian transport schemes. *Mon. Wea. Rev.* **124** (9), 2064–2070.

Lin, S.-J., Rood, R.B. (1997). An explicit flux-form semi-Lagrangian shallow water model on the sphere. *Q. J. R. Meteorol. Soc.* **123**, 2477–2498.

Lin, S.-J., Rood, R.B. (1998). A flux-form semi-Lagrangian general circulation model with a Lagrangian control-volume vertical coordinate. In: *Proc. The Rossby-100 Symp.* (University of Stockholm, Stockholm, Sweden), pp. 220–222.

Lorenz, E.N. (1960). Energy and numerical weather prediction. *Tellus* **12**, 364–373.

Machenhauer, B. (1979). The spectral method. In: Kasahara, A. (ed.), *Numerical Methods Used in Atmospheric Models, vol. 2. GARP Publications Series No.17* (WMO and IGSU, Geneva, Switzerland), pp. 121–275.

Machenhauer, B. (1994). A Note on a Mass-, Energy- and Entropy Conserving Semi-Lagrangian and Explicit Integration Scheme for the Primitive Equations. MPI Workshop on Semi-Lagrangian Methods, Hamburg 8-9 Oct. 1992. **146** (*Max Planck Institute for Meteorology Tech. Rep*, Hamburg, Germany), pp. 73–102.

Machenhauer, B., Olk, M. (1996). On the development of a cell-integrated semi-Lagrangian shallow water model on the sphere. ECMWF Workshop on semi-Lagrangian methods, 6-8 November, Workshop Proceedings, pp. 213–228.

Machenhauer, B., Olk, M. (1997). The implementation of the semi-implicit scheme in cell-integrated semi-Lagrangian models. *Atmos. Ocean.* **35**, 103–126.

Machenhauer, B., Olk, M. (1998). Design of a semi-implicit cell-integrated semi-Lagrangian model. Max Planck Institute for Meteorology Tech. Rep., 265, Hamburg, Germany, pp. 76–85.

Majewski, D., Liermann, D., Prohl, P., Ritter, B., Buchhold, M., Hanisch, T., Paul, G., Wergen, W., Baumgardner, J. (2002). The operational global icosahedral-hexagonal gridpoint model GME: description and high-resolution tests. *Mon. Wea. Rev.* **130** (2), 319–338.

McDonald A. (1998). Default horizontal diffusion coefficients in HIRLAM-4.1. HIRLAM Newsletter No. 31, electronic version available at http://hirlam.knmi.nl/.

Mcgregor, J.L. (1993). Economical determination of departure points for semi-Lagrangian models. *Mon. Wea. Rev.* **121** (1), 221–230.

Mcgregor, J.L. (1996). Semi-Lagrangian advection on a conformal cubic grid. *Mon. Wea. Rev.* **124** (6), 1311–1322.

Moorthi, S., Higgins, R.W., Bates, J.R. (1995). A global multilevel atmospheric model using a vector semi-Lagrangian finite-difference scheme. Part II: version with physics. *Mon. Wea. Rev.* **123** (5), 1523–1541.

Nair, R.D., Côté, J., Staniforth, A. (1999a). Monotonic cascade interpolation for semi-Lagrangian advection. *Q. J. R. Meteorol. Soc.* **125**, 197–212.

Nair, R.D., Côté, J., Staniforth, A. (1999b). Cascade interpolation for semi-Lagrangian advection over the sphere. *Q. J. R. Meteorol. Soc.* **125**, 1445–1468.

Nair, R.D., Thomas, S.J., Loft, R.D. (2005). A discontinuous galerkin transport scheme on the cubed sphere. *Mon. Wea. Rev.* **133** (4), 814–828.

Nair, R.D. (2004). Extension of a conservative cascade scheme on the sphere to large courant numbers. *Mon. Wea. Rev.* **132** (1), 390–395.

Nair, R.D., Jablonowski, C. (2007). Moving vortices on the sphere: a test-case for horizontal advection problems. *Mon. Wea. Rev.*, to appear.

Nair, R.D., Scroggs, J.S., Semazzi, F.H.M. (2003). A forward-trajectory global semi-Lagrangian transport scheme. *J. Comput. Phys.* **190**, 275–294.

Nair, R.D., Scroggs, J.S., Semazzi, F.H.M. (2002). Efficient conservative global transport schemes for climate and atmospheric chemistry models. *Mon. Wea. Rev.* **130** (8), 2059–2073.

Nair, R.D., Machenhauer, B. (2002). The mass-conservative cell-integrated semi-Lagrangian advection scheme on the sphere. *Mon. Wea. Rev.* **130** (3), 649–667.

Peng, X., Xiao, F., Ohfuchi, W., Fuchigami, H. (2005). Conservative semi-Lagrangian transport on a sphere and the impact on vapor advection in an atmospheric general circulation model. *Mon. Wea. Rev.* **133** (3), 504–520.

Petschek, A.G., Libersky, L.D. (1975). Stability, accuracy, and improvement of Crowley advection schemes. *Mon. Wea. Rev.* **103** (12), 1104–1109.

Prather, M., Mcelroy, M., Wofsy, S., Russel, G., Rind, D. (1987). Chemistry of the global troposphere: fluorocarbons as tracers of air motion. *J. Geophys. Res.* **92** (D6), 6579–6613.

Prather, M.J. (1986). Numerical advection by conservation of second order moments. *J. Geophys. Res.* **91**, 6671–6681.

Pudykiewicz, J., Benoit, R., Staniforth, A. (1985). Preliminary results from a partial LRTAP model based on an existing meteorological forecast model. *Atmos. Ocean.* **22**, 283–308.

Purser, R.J., Leslie, L.M. (1991). An efficient interpolation procedure for high-order three-dimensional semi-Lagrangian models. *Mon. Wea. Rev.* **119** (10), 2492–2498.

Rančić, M. (1992). Semi-Lagrangian piecewise biparabolic scheme for two-dimensional horizontal advection of a passive scalar. *Mon. Wea. Rev.* **120** (7), 1394–1406.

Rančić, M. (1995). An efficient, conservative, monotone remapping for semi-Lagrangian transport algorithms. *Mon. Wea. Rev.* **123** (4), 1213–1217.

Rančić, M.R., Purser, J., Mesinger, F. (1996). A global-shallow water model using an expanded spherical cube. *Q. J. R. Meteorol. Soc.* **122**, 959–982.

Rasch, P.J. (1994). Conservative shape-preserving two-dimensional transport on a spherical reduced grid. *Mon. Wea. Rev.* **122** (6), 1337–1350.

Rasch, P.J., Lawrence, M. (1998). Recent development in transport models at NCAR. In: Mchenhauer, B. (ed.), MPI Workshop on Conservative Transport Schemes Hamburg 2-3 June 1997. **265**, pp. 65–75, (*Max Planck Institute for Meteorology Tech. Rep*, Hamburg, Germany).

Rasch, P.J., Williamson, D.L. (1990). Computational aspects of moisture transport in global models of the atmosphere. *Q. J. R. Meteorol. Soc.* **116**, 1071–1090.

Reich, S. (2007). An explicit and conservative remapping strategy for semi-Lagrangian advection. *Atmos. Sci. Let.*, in press.

Richardson, L.F. (1922). *Weather Prediction by Numerical Process* (Cambridge University Press).

Robert, A.J. (1969). The integration of a spectral model of the atmosphere by the implicit method. In: *Proc. WMO/IUGG Symp. on Numerical Weather Prediction* (Japan Meteorological Agency, Tokyo, Japan), pp. 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. Meteorol. Soc. Jpn.* **60**, 319–325.

Ronchi, C., Iacono, R., Paolucci, P.S. (1996). The "cubed sphere": A new method for the solution of partial differential equations in spherical geometry. *J. Comput. Phys.* **124**, 93–114.

Rotman, D.A., Tannahill, J.R., Kinnison, D.E., Connell, P.S., Bergmann, D., Proctor, D., Rodriguez, J.M., Lin, S.J., Rood, R.B., Prather, M.J., Rasch, P.J., Considine, D.B., Ramaroson, R., Kawa, S.R. (2001). Global modeling initiative assessment model: model description, integration, and testing of the transport shell. *J. Geophys. Res.* **106** (D2), 1669–1692.

Sadourny, R., Arakawa, A., Mintz, Y. (1968). Integration of the non-divergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere. *Mon. Wea. Rev.* **96** (6), 351–356.

Sadourny, R. (1972). Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids. *Mon. Wea. Rev.* **100** (2), 136–144.

Schär, C., Smolarkiewicz, P.K. (1996). A synchronous and iterative flux-correction formalism for coupled transport. *J. Comput. Phys.* **128**, 101–120.

Scroggs, J.S., Semazzi, F.H.M. (1995). A conservative semi-Lagrangian method for multidimensional fluid dynamics applications. *Numer. Methods Partial Differ. Equ.* **11**, 445–452.

Simmons, A.J., Burridge, D.M. (1981). An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.* **109** (4), 758–766.

Skamarock, W.C. (2006). Positive-definite and monotonic limiters for unrestricted-time-step transport schemes. *Mon. Wea. Rev.* **134** (8), 2241–2250.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W., Powers, J.G. (2007). A description of the advanced research WRF Version 2. NCAR Tech Note, NCAR/TN–468+STR, Jun. 2005, Upd. Jan. 2007.

Smolarkiewicz, P.K. (1982). The multi-dimensional Crowley advection transport algorithm with small implicit diffusion. *Mon. Wea. Rev.* **110** (12), 1968–1983.

Smolarkiewicz, P.K. (1984). A fully multidimensional positive definite advection transport algorithm with small implicit diffusion. *J. Comput. Phys.* **54**, 325–362.

Smolarkiewicz, P.K., Pudykiewicz, J.A. (1992). A class of semi-Lagrangian approximations for fluids. *J. Atmos. Sci.* **49** (22), 2082–2096.

Smolarkiewicz, P.K., Grabowski, W.W. (1990). The multidimensional positive definite advection transport algorithm: non oscillatory option. *J. Comput. Phys.* **86**, 355–375.

Staniforth, A., Côté, J. (1991). Semi-Lagrangian integration schemes for atmospheric models—a review. *Mon. Wea. Rev.* **119** (9), 2206–2223.

Staniforth, A., Coté, J., Pudykjewicz, J. (1987). Comments on "Swolarkiewicz's Deformational Flow." *Mon. Wea. Rev.* **115** (4), 894–900.

Starr, V.P. (1945). A quasi-Lagrangian system of hydrodynamical equations. *J. Atmos. Sci.* **2** (4), 227–237.

Sun, W.Y., Yeh, K.S., Sun, R.Y. (1996). A simple semi-Lagrangian scheme for advection equations. *Q. J. R. Meteorol. Soc.* **122**, 1211–1226.

Taylor, M., Tribbia, J., Iskandrani, M. (1997). The spectral element method for the shallow water equations on the sphere. *J. Comput. Phys.* **130**, 92–108.

Thuburn, J. (1997). A PV-based shallow-water model on a hexagonal-icosahedral grid. *Mon. Wea. Rev.* **125** (9), 2328–2347.

Thuburn, J. (2006). Some conservation issues for dynamical cores of NWP and climate models. *J. Comput. Phys.*, in press.

Tremback, C.J., Powell, J., Cotton, W.R., Pielke, R.A. (1987). The forward in time upstream advection scheme: extension to higher orders. *Mon. Wea. Rev.* **115** (2), 540–555.

Trenberth, K.E. (1991). Climate diagnostics from global analyses: conservation of mass in ECMWF analyses. *J. Clim.* **4** (7), 707–722.

Per Undén, Laura Rontu, Heikki Järvinen, Peter Lynch, Javier Calvo, Gerard Cats, Joan Cuxart, Kalle Eerola, Carl Fortelius, Jose Antonio Garcia-Moya, Colin Jones, Geert Lenderlink, Aidan McDonald, Ray McGrath, Beatriz Navascues, Niels Woetman Nielsen, Viel Ødegaard, Ernesto Rodriguez, Markku Rummukainen, Rein Rõõm, Kai Sattler, Bent Hansen Sass, Hannu Savijärvi, Ben Wichers Schreur, Robert Sigg, Han The, Aleksander Tijm (2002). Hirlam-5 scientific documentation. SMHI Hirlam-5 Project Tech. Rep., pp. 146 [Available online at http://www-hirlam.knmi.nl.].

Van Leer, B. (1974). Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second-order scheme. *J. Comput. Phys.* **14**, 361–370.

Van Leer, B. (1977). Towards the ultimate conservative difference scheme. Part IV: a new approach to numerical convection. *J. Comput. Phys.* **23**, 276–299.

Williamson, D.L. (1968). Integration of the barotropic vorticity equation on a spherical geodesic grid. *Tellus* **20**, 642–653.

Williamson, D.L., Drake, J.B., Hack, J.J., Jakob, R., Swarztrauber, R.N. (1992). A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comput. Phys.* **102**, 211–224.

Wooward, P.R., Colella, P. (1984). The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys.* **54**, 115–173.

Xue, M. (2000). High-order monotonic numerical diffusion and smoothing. *Mon. Wea. Rev.* **128**, 2853–2864.

Xiao, F., Yabe, T., Peng, X., Kobayashi, H. (2002). Conservative and oscillation-less atmospheric transport schemes based on rational functions. *J. Geophys. Res.* **107** (D22), 4609, doi: 10.1029/2001JD001532.

Xiao, F., Yabe, T. (2001). Completely conservative and oscillation-less semi-Lagrangian schemes for advection transport. *J. Comput. Phys.* **170**, 498–522.

Yabe, T., Ishikawa, T., Wang, P.Y., Aoki, T., Kodota, Y., Ikeda, F. (1991). A universal solver for hyperbolic-equations by cubic-polynomial interpolation. II. Two- and three-dimensional solvers. *Comput. Phys. Commun.* **66**, 233–242.

Zalesak, S.T. (1979). Fully multidimensional flux-corrected transport algorithm for fluids. *J. Comput. Phys.* **31**, 335–362.

Zerroukat, M., Wood, N., Staniforth, A. (2002). SLICE: a semi-Lagrangian inherently conservative and efficient scheme for transport problems. *Q. J. R. Meteorol. Soc.* **128** (586), 2801–2820.

Zerroukat, M., Wood, N., Staniforth, A. (2004). SLICE-S: A semi-Lagrangian inherently conserving and efficient scheme for transport problems on the sphere. *Q. J. R. Meteorol. Soc.* **130** (602), 2649–2664.

Zerroukat, M., Wood, N., Staniforth, A. (2005). A monotonic and positive-definite filter for a semi-Lagrangian inherently conserving and efficient (SLICE). *Q. J. R. Meteorol. Soc.* **131** (611), 2923–2936.

Zerroukat, M., Wood, N., Staniforth, A. (2006). The parabolic spline method (PSM) for conservative transport. *Int. J. Numer. Methods Fluids* **11**, 1297–1318.

Zerroukat, M., Wood, N., Staniforth, A. (2007). Application of the parabolic spline method (PSM) to a multi-dimensional conservative semi-Lagrangian transport scheme (SLICE). *J. Comput. Phys.* **225** (1), 935–948.