Chapter 13
The Pros and Cons of Diffusion, Filters and Fixers in Atmospheric General Circulation Models

Christiane Jablonowski and David L. Williamson

Abstract All atmospheric General Circulation Models (GCMs) need some form of dissipation, either explicitly specified or inherent in the chosen numerical schemes for the spatial and temporal discretizations. This dissipation may serve many purposes, including cleaning up numerical noise generated by dispersion errors or computational modes, and the Gibbs ringing in spectral models. Damping processes might also be used to crudely represent subgrid Reynolds stresses, eliminate undesirable noise due to poor initialization or grid-scale forcing from the physics parameterizations, cover up weak computational stability, damp tracer variance, and prevent the accumulation of potential enstrophy or energy at the smallest grid scales. This chapter critically reviews the wide selection of dissipative processes in GCMs. They are the explicitly added diffusion and hyper-diffusion mechanisms, divergence damping, vorticity damping, external mode damping, sponge layers, spatial and temporal filters, inherent diffusion properties of the numerical schemes, and a posteriori fixers used to restore lost conservation properties. All theoretical considerations are supported by many practical examples from a wide selection of GCMs. The examples utilize idealized test cases to isolate causes and effects, and thereby highlight the pros and cons of the diffusion, filters and fixers in GCMs.

13.1 Introduction

There are many design aspects that need to be considered when building the fluid dynamics component, the so-called dynamical core, for atmospheric General Circulation Models (GCMs). Among them are the choice of the equation set

C. Jablonowski (✉)
Department of Atmospheric, Oceanic and Space Sciences, University of Michigan,
2455 Hayward Street, Ann Arbor, MI 48109, USA
e-mail: cjablono@umich.edu

D.L. Williamson
National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, CO 80305, USA
e-mail: wmson@ucar.edu

P.H. Lauritzen et al. (eds.), *Numerical Techniques for Global Atmospheric Models*,
Lecture Notes in Computational Science and Engineering 80,
DOI 10.1007/978-3-642-11640-7_13,

and prognostic variables, the computational grid and grid staggering options, the spatial and temporal numerical discretizations, built-in conservation properties, and the choice of dissipative processes that (1) might be needed to keep a model simulation numerically stable and (2) might truthfully mimic the cumulative effects of unresolved subgrid-scale processes on the resolved fluid flow. The latter aspect is at least a "hope" in the GCM modeling community. Here, the phrase *subgrid-scale* denotes the dry adiabatic unresolved processes in the dynamical core. This is in contrast to all other unresolved processes that lead to physical parameterizations such as radiation, convection, cloud processes and planetary boundary layer phenomena. These are not considered here, even though they are intimately coupled to the equations of motion. This chapter sheds light on the pros and cons of the most popular processes to handle both "physical" or "unphysical" subgrid-scale flow and mixing, and reviews the use of explicitly added and inherent diffusion, filters and fixers in GCMs. These are rarely documented in the refereed GCM literature but might be detailed in technical model descriptions.

It is common practice in GCMs to include a parameterization of the effects of subgrid-scale motions in the horizontal momentum and thermodynamic equations that is formulated as a local diffusive mixing. In fact, all numerical models need some form of dissipation, either explicitly specified or inherent in the chosen numerical schemes for the spatial and temporal discretizations. This dissipation may serve many purposes, including cleaning up numerical noise generated by dispersion errors, computational modes, or the Gibbs ringing, crudely representing subgrid Reynolds stresses, eliminating undesirable noise due to poor initialization or grid-scale forcing from the physics parameterizations, covering up weak computational stability, damping tracer variance, and preventing the accumulation of potential enstrophy or energy at the grid scale (Wood et al. 2007; Thuburn 2008a). Such an accumulation of energy is due to the physical downscale cascade and can result in excessive small-scale noise. It is known as spectral blocking and leads to an upturn (hook) or flattening in the kinetic energy spectrum at the smallest scales. Furthermore, physical "noise" in GCMs might originate from parameterized grid-scale forcings or from surface boundary conditions such as orography, the land-sea or land-use mask.

An accumulation of energy and enstrophy at the smallest scales may also arise due to a numerical misrepresentation of nonlinear interactions, the so-called aliasing effect. Nonlinear interactions and aliasing mostly originate from the quadratic or higher-order terms in the equations of motion. In essence, products of waves can create new waves that are shorter than $2\Delta x$ where Δx is the physical grid spacing. These waves cannot be represented on a model grid and are aliased into longer waves. Aliasing, if left unchecked, can lead to a blow up of the solution. This phenomenon is characterized as nonlinear computational instability as first discussed by Phillips (1959). Note that almost all GCMs suffer to some degree from aliasing. Exceptions are spectral transform models with quadratic transform grids which eliminate the aliasing of quadratic advection terms, the most problematic form, but do not completely eliminate aliasing from higher-order terms. Nonlinear computational instability does not occur in models that conserve quadratic quantities like

enstrophy and kinetic energy (Arakawa 1966; Arakawa and Lamb 1981). Aliasing errors are not necessarily fatal. Whether an amplification of the waves occurs depends on the phase relation between the misrepresented and original waves in the model. More information on nonlinear computational instability and aliasing is provided in textbooks like Durran (1999, 2010), Kalnay (2003) or Lin (2007).

All mixing processes remove energy and enstrophy from the simulation which would otherwise build up to unrealistic proportions. Frequently, the included dissipation is restricted to be in the horizontal, as there is usually sufficient vertical mixing or diffusion in the physical boundary layer turbulence and convective parameterizations in full GCMs or sufficient inherent numerical diffusion to control noise in the vertical direction. Sometimes, vertical diffusion is also explicitly included in the dynamical core and applied throughout the whole troposphere. An example is the model by Tomita and Satoh (2004) which is discussed later.

A common expectation might be that dissipative formulations based on turbulence theory or observations provide a physical foundation for the subgrid-scale mixing. However, such physical motivation is not guaranteed and each ad hoc mixing process in a GCM must be critically reviewed. As pointed out by Mellor (1985) the horizontal diffusivities in use by GCMs are typically many orders of magnitude larger than those which would be appropriate for turbulence closures. Thus, horizontal diffusion used by most models cannot be considered a representation of turbulence but should be viewed as a substitute mechanism for unresolved horizontal advective processes. Awareness of this might offer some guidance in choosing an adequate subgrid-scale mixing scheme.

Mixing in GCMs generally serves as a numerical filter and neither reflects the mathematical representation of the energy or enstrophy transfer to small scales nor the representation of physical molecular diffusion (Koshyk and Boer 1995). Subgrid-scale processes, although small, can have a profound impact on the large-scale circulation. For example, diffusive mechanisms affect the propagation of waves and thereby the mean flow through wave-mean flow interactions. In addition, both inherent and explicitly added dissipation processes smear out sharp gradients in the tracer fields, and may lead to unphysical and overly strong mixing. Such mixing might provide feedbacks to the physical parameterizations. For example, the precipitation field might be highly influenced by the diffusive characteristics of the moisture transport algorithm in the dynamical core. The notion of overly diffusive GCMs was discussed by Shutts (2005). He argued that numerical advection errors, horizontal diffusion and parameterization schemes like the gravity wave drag or convection, act as energy sinks and lead to excessive energy dissipation in GCMs. However, such a conclusion might be highly model dependent.

In summary, some mixing processes are used for purely numerical reasons to prevent the model from becoming unstable. Others are meant to mimic subgrid-scale turbulence processes that are unsolved on the chosen model grid. In practice, many filters and mixing processes are used at once, which makes it more difficult to evaluate their individual effects. The form of the diffusion processes in atmospheric dynamical cores varies widely and is somewhat arbitrary. There are explicit dissipation processes and filters, inherent numerical dissipation, and fixers in GCMs.

Throughout the chapter we associate the phrase *explicit diffusion* with processes explicity added to the equations of motion. The phrase *implicit diffusion* characterizes the inherent dissipation of numerical schemes. These phrases are intended to make a distinction from explicit and implicit numerical approximations to diffusion operators. Note that the words "diffusion", "dissipation" and "viscosity" are often used interchangeably in the literature. Other characterizations of damping are smoothing, filtering and mixing.

13.1.1 Model Equations and the Representation of Explicit Diffusion

Mixing processes in GCMs can appear in many forms. A very dominant form is based on explicit dissipation mechanisms that get appended to the equations of motion shown in Chap. 15. In the continuous equations this mixing symbolizes molecular diffusion. However, GCMs are not capable of representing molecular diffusion at the nm or mm scale since they are typically applied with horizontal grid spacings between 20 and 300 km. Nonhydrostatic GCMs (Tomita et al. 2005; Fudeyasu et al. 2008) and mesoscale limited-area models like the Weather Research and Forecasting Model WRF (Skamarock et al. 2008) are also run with finer grid spacings of a few kilometers. Other atmospheric models with even finer scales might utilize the Large Eddy Simulation (LES) approach. LES is a mathematical model for turbulence that is based upon the Navier–Stokes equations with built-in low-pass filter. The underlying idea was initially proposed by Smagorinsky (1963) and further developed by Deardorff (1970). LES has been extensively used to study small-scale physical processes and mixing in the atmosphere. But in any case, models truncate the multi-scale spectrum of atmospheric motions well above the molecular diffusion scales. The unresolved part is typically modeled as dissipation and one might hope that it adequately captures the adiabatic subgrid-scale processes in some poorly understood way.

Explicit dissipation can be added to the momentum and thermodynamic equations in the symbolic form

$$\frac{\partial \psi}{\partial t} = Dyn(\psi) + Phys(\psi) + F_\psi \tag{13.1}$$

where $Dyn(\psi)$ denotes the time tendencies of the prognostic variable ψ according to the resolved adiabatic dynamics, $Phys(\psi)$ symbolizes the time tendencies from the subgrid-scale diabatic physical parameterizations, and F_ψ is the dissipation. The actual form of this dissipation is model dependent. For example, models in momentum form, that utilize the zonal and meridional velocities u, v and temperature T, might append the diffusive terms F_u, F_v, F_T. Models in vorticity-divergence (ζ, δ) form add the diffusion F_ζ, F_δ, F_T, or even replace F_ζ with a diffusion of the absolute vorticity $F_{\zeta+f}$ where f symbolizes the Coriolis parameter. Alternatively, if

the potential temperature Θ is selected in the thermodynamic equation a diffusive term F_Θ might be chosen. Dissipation might also be applied to the tracer transport equations, and in case of nonhydrostatic models to the vertical velocity. Whether explicit diffusion is *needed* for computational stability is model dependent. Some models prefer to control the smallest scales via inherent numerical dissipation and select $F_\psi = 0$. However, the form of F_ψ is one of the main foci in Sects. 13.3–13.5, and therefore we introduce the generic form of the forcing here.

13.1.2 Overview of the Chapter

This chapter presents a comprehensive review of dissipative processes and fixers in general circulation models. Many pointers to references are given, and we illustrate the practical implications of the diffusion, filters and fixers on the fluid flow in atmospheric dynamical cores. In particular, we review the principles behind the different dissipative formulations, isolate causes and effects, provide many examples from today's GCMs and utilize idealized dynamical core test cases and so-called aqua-planet simulations to demonstrate the concepts. These test cases are briefly outlined in Sect. 13.2. Overall, we quote or show examples from over 20 different dynamical cores to highlight the broad spectrum of the dissipative processes in GCMs. The models are listed in Sect. 13.2. We characterize fourteen of them in greater detail in the Appendix since they are used as examples throughout the chapter.

The chapter is organized as follows. Sections 13.3 and 13.4 discuss the most popular explicit diffusion and damping mechanisms in the dynamical cores of GCMs. Section 13.3 includes the classical linear and nonlinear horizontal diffusion and hyper-diffusion, their diffusion coefficients and stability constraints, and physical consistency arguments. Section 13.4 discusses the 2D and 3D divergence damping, vorticity damping, Rayleigh friction and diffusive sponges near the model top, and external mode damping. In general, it is debatable whether filters are considered explicit dissipation or just a computational technique to keep a model numerically stable. Here, we choose to present them in their own category in Sect. 13.5 where both temporal and spatial filter are assessed. Section 13.6 captures the basic ideas behind inherent numerical dissipation which is nonlinear and sometimes is interpreted as physically motivated diffusion. Section 13.7 sheds light on the conservation properties of atmospheric GCMs and introduces a posteriori fixers. They include the dry air mass fixer, fixers for tracer masses and total energy fixers. Some final thoughts are presented in Sect. 13.8.

13.2 Selected Dynamical Cores and Test Cases

We illustrate many of the effects of the diffusion, filters and fixers in GCMs with the help of 2D shallow water or 3D hydrostatic model runs to discuss the practical implications of the theoretical concepts. Throughout this chapter, we point to the

specific implementations of the dissipative processes, and quote typical values for the empirical coefficients from a variety of models to shows their spread in the GCMs. The intention is to give hands-on guidance and present the design options.

In particular, this chapter features examples from the dynamical cores CAM Eulerian, CAM semi-Lagrangian, COSMO, ECHAM5, FV, FVcubed, GEOS, GME, HOMME, ICON, IFS, NICAM, UM and WRF. The Appendix explains the acronyms and briefly characterizes the numerical schemes. Most models are global hydrostatic GCMs and use the shallow-atmosphere approximation. The only exceptions are the nonhydrostatic models COSMO, NICAM, UM and WRF that are built upon the deep atmosphere equation set (see White et al. 2005 for a review of the equations). We also briefly refer to other models such as NASA's ModelE by the Goddard Institute for Space Studies, the Atmospheric GCM for the Earth Simulator (AFES) developed by the Center for Climate System Research at the University of Tokyo and the National Institute for Environmental Studies (Japan), the Global Environmental Multiscale (GEM) model from the Canadian Meteorological Centre, the Global Forecast System (GFS) and the Eta model developed by the National Centers for Environmental Prediction (NCEP). The references for these models are given later.

The model simulations utilize a variety of idealized test cases. They include the steady-state and baroclinic wave test cases for dynamical cores suggested by Jablonowski and Williamson (2006a,b), selected shallow water test cases from the Williamson et al. (1992) test suite, the Held–Suarez climate forcing (Held and Suarez 1994), a variant of the Held–Suarez test with modified equilibrium temperatures in the stratosphere (Williamson et al. 1998), and the aqua-planet configuration as proposed by Neale and Hoskins (2000). The adiabatic dynamical core and shallow water test cases generally assess the properties of the numerical schemes in short deterministic model runs of up to 30 days. The idealized Held-Suarez-type simulations utilize a prescribed Newtonian temperature relaxation and boundary layer friction that replace the physical parameterization package. These model runs are typically integrated for multiple years to assess the climate statistics on a dry and spherical earth with no mountains. The aqua-planet assessments are the most complex simulations discussed in this chapter. They represent moist GCM runs that include the full physical parameterization suite but utilize a simplified lower boundary condition. In essence, the lower boundary is replaced by a water covered earth with prescribed sea surface temperatures. In addition, the settings of the physical constants and a symmetric ozone data set are prescribed in aqua-planet simulations.

13.3 Explicit Horizontal Diffusion

This section discusses the ideas behind explicit horizontal diffusion mechanisms in GCMs. In particular, we assess the linear second-order diffusion, the higher-order and thereby more scale-selective hyper-diffusion, reveal the selection criteria for

the diffusion coefficients in both spectral transform and grid point models, discuss the concept of spectral viscosity, and review the stability constraints for the diffusion equation. In addition, we introduce the principles behind nonlinear horizontal diffusion and briefly survey the physical consistency of explicit diffusion schemes.

13.3.1 Generic Form of the Explicit Diffusion Mechanism

The generic form of the explicit linear diffusion is given by

$$F_\psi = (-1)^{q+1} K_{2q} \nabla^{2q} \psi \tag{13.2}$$

where $q \geq 1$ is a positive integer, $2q$ denotes the order of the diffusion, K_{2q} stands for the diffusion coefficient and ∇ is the gradient operator. Both the horizontal 2D gradient operator or an approximated 3D gradient operator have been used for the horizontal diffusion as further explained below. Setting $q = 1$ yields a second-order diffusion that emerges from physical principles such as the heat diffusion, molecular diffusion and Brownian motion. However, molecular diffusion acts on the nanometer to millimeter scale, and is therefore unresolved on a GCM model grid.

In practice, second-order diffusion is often applied as an artificial sponge near the top boundary, and has very little resemblance with its physical counterpart. In general, more scale-selective hyper-diffusion schemes with $q = 2, 3, 4$ are selected in the majority of the model domain. The most popular choice is the fourth-order hyper-diffusion with $q = 2$ that is also called bi-harmonic diffusion or superviscosity. The use of hyper-diffusion is often motivated by the need to maximize the ratio of enstrophy to energy dissipation since 2D turbulence theory predicts a vanishing energy dissipation rate at increasing Reynolds numbers (Sadourny and Maynard 1997). The higher the order of the hyper-diffusion, the higher the ratio of enstrophy to energy dissipation becomes. Farge and Sadourny (1989) even suggested using a 16th-order hyper-diffusion.

13.3.2 Particular Forms of Explicit Diffusion in GCMs

The exact form of the damping varies widely in GCMs. Typically, for convenience the horizontal operators are applied along model levels with the possible exception of the formulation for the scalar temperature diffusion. We now list several examples to illustrate the variety of the diffusion mechanisms. Our first example is taken from the weather forecast model GME which has been developed at the German Weather Service. It applies the fourth-order hyper-diffusion

$$F_u = -K_4 \nabla^4 u \tag{13.3}$$

$$F_v = -K_4 \nabla^4 v \tag{13.4}$$

$$F_T = -K_4 \nabla^4 (T - T_{ref}) \tag{13.5}$$

where ∇ is the horizontal gradient operator and T_{ref} is a reference temperature that depends only on height (Majewski et al. 2002). At upper levels near the model top, a second-order diffusion is applied. GME utilizes local basis functions that are orthogonal and conform perfectly to the spherical surface. They are locally anchored in each triangle of GME's icosahedral grid. Within the local neighborhood of a grid point the coordinate system is therefore nearly Cartesian. Note that Cartesian coordinates simplify the representation of the ∇^{2q} operator since the metric terms are equal to unity.

If the diffusion is expressed in spherical coordinates many metric terms are present. Here, we first show the operators for three spatial dimensions before simplifying them. The scalar 3D Laplacian $\nabla^2_{(3D)}$ operator in spherical coordinates for a prognostic variable ψ has the form

$$\nabla^2_{(3D)}\psi = \frac{1}{r^2 \cos^2\phi}\partial_{\lambda\lambda}\psi + \frac{1}{r^2 \cos\phi}\partial_\phi(\cos\phi\,\partial_\phi\psi) + \frac{1}{r^2}\partial_r(r^2\partial_r\psi) \tag{13.6}$$

where r denotes the radial distance in the local vertical direction from the center of the earth, λ and ϕ are the longitude and latitude, and the notation ∂_x symbolizes a partial derivative in the x direction where x is a placeholder for λ, ϕ and r. In addition, the 3D vector Laplacian for the three-dimensional wind vector $\mathbf{v_3} = (u, v, w)$ is given by

$$\nabla^2_{(3D)}\mathbf{v_3} = \begin{pmatrix} \nabla^2_{(3D)}u - \dfrac{2\sin\phi}{r^2\cos^2\phi}\partial_\lambda v + \dfrac{2}{r^2\cos\phi}\partial_\lambda w - \dfrac{1}{r^2\cos^2\phi}u \\ \nabla^2_{(3D)}v - \dfrac{1}{r^2\cos^2\phi}v + \dfrac{2}{r^2}\partial_\phi w + \dfrac{2\sin\phi}{r^2\cos^2\phi}\partial_\lambda u \\ \nabla^2_{(3D)}w - \dfrac{2}{r^2}w - \dfrac{2}{r^2\cos\phi}\partial_\lambda u - \dfrac{2}{r^2\cos\phi}\partial_\phi(\cos\phi\, v) \end{pmatrix}. \tag{13.7}$$

These expressions are e.g., shown in Appendix A in Satoh (2004). The extra terms besides the scalar $\nabla^2_{(3D)}$ operator arise due to the spatial variation of the unit vectors in spherical coordinates. With the exception of the undifferentiated term in each of the components the extra terms are not necessarily negligible in comparison with those of the scalar diffusion operator. In fact, some of them are crucial in ensuring that the diffusion operator conserves angular momentum as outlined by Staniforth et al. (2006).

Generally, approximated forms of (13.6) and (13.7) are chosen to express the horizontal diffusion. The distance r is often approximated by the constant radius of the earth a, the terms containing the vertical derivative ∂_r and vertical velocity w are dropped, and the vertical component of the vector Laplacian is neglected to create a 2D diffusion operator. The replacement of the distance r by a is in fact a

necessity in hydrostatic models based on the primitive equations (White et al. 2005). The scalar 2D Laplacian then simplifies in the following way

$$\nabla^2\psi = \frac{1}{a^2\cos^2\phi}\partial_{\lambda\lambda}\psi + \frac{1}{a^2\cos\phi}\partial_\phi(\cos\phi\partial_\phi\psi). \tag{13.8}$$

The 2D vector Laplacian for the two-dimensional wind vector $\mathbf{v} = (u, v)$ is given by

$$\nabla^2\mathbf{v} = \begin{pmatrix} \nabla^2 u - \dfrac{2\sin\phi}{a^2\cos^2\phi}\partial_\lambda v - \dfrac{1}{a^2\cos^2\phi}u \\ \nabla^2 v - \dfrac{1}{a^2\cos^2\phi}v + \dfrac{2\sin\phi}{a^2\cos^2\phi}\partial_\lambda u \end{pmatrix}. \tag{13.9}$$

This form of the vector Laplacian leads to the "conventional form" of the horizontal momentum diffusion as characterized by Becker (2001). Unfortunately, this form does not conserve angular momentum as further discussed in Sect. 13.3.7. Some models also drop the extra terms and only apply the scalar 2D Laplacian operator to the vector wind (u, v) in spherical geometry. Such a simplified form is e.g., provided as an optional sponge layer damping mechanism near the model top in the finite-volume (FV) dynamical core in the Community Atmosphere Model CAM (version 5) (Neale et al. 2010). The model CAM FV is used at the National Center for Atmospheric Research (NCAR).

There is another caveat. Both formulations of the Laplacian in (13.7) or (13.9) would lead to an undesired damping of a solid body rotation as thoroughly analyzed by Staniforth et al. (2006) for the Unified Model (UM) developed at the UK Met Office. Therefore in practice, a more complicated form of the momentum diffusion is chosen in the model UM that is applied to the velocity components u, v and w (see Staniforth et al. (2006) for the derivation). The NCAR CAM spectral transform Eulerian (EUL) and semi-Lagrangian (SLD) dynamical cores (Collins et al. 2004) also include such a correction for solid body rotation as explained later.

Concerning the scalar diffusion in the model UM, a form similar to (13.8) is selected for the diffusion of potential temperature. It is applied twice (including a sign reversal) to resemble a fourth-order hyper-diffusion mechanism. The main differences to (13.8) are that (1) the model UM does not utilize a shallow-atmosphere approximation and retains the radial distance r, (2) a slope correction is utilized over steep terrain to lessen the spurious mixing along UM's deformed orography-following vertical coordinate, and (3) the diffusion coefficient is different in the two horizontal directions. The coefficient is constant in the meridional direction, but the strength of the diffusion in longitudinal direction is allowed to vary with latitude. This leads to non-isotropic diffusion and is further explained in Sect. 13.3.5. Note that the model UM does not need horizontal diffusion for computational stability reasons due to the inherent numerical dissipation in the interpolations of its semi-Lagrangian scheme. In practice, the explicit diffusion is therefore optional and not used by default. For example, it is never utilized in short weather prediction simulations (Terry Davies, personal communication).

Scalar diffusion of type (13.8) is also applied in other models such as the spectral transform Integrated Forecasting System (IFS) at the European Centre for Medium-Range Weather Forecast (Ritchie et al. 1995; ECMWF 2010). The model utilizes a second-order ($q = 1$) diffusion scheme close to the model top and a fourth-order ($q = 2$) hyper-diffusion of the prognostic scalar variables relative vorticity ζ, horizontal divergence δ and temperature. It yields the explicit diffusion

$$F_\zeta = (-1)^{q+1} K_{2q} \nabla^{2q} \zeta \tag{13.10}$$

$$F_\delta = (-1)^{q+1} K_{2q} \nabla^{2q} \delta \tag{13.11}$$

$$F_T = (-1)^{q+1} K_{2q} \nabla^{2q} T. \tag{13.12}$$

This form of the diffusion is furthermore utilized by the spectral transform model ECHAM5 developed at the Max–Planck Institute for Meteorology, where even higher-order diffusion operators are chosen below the sponge layer at the model top (Roeckner et al. 2003).

The application of the diffusion along sloping general vertical coordinates, like the hybrid pressure-based η-coordinate (Simmons and Burridge 1981), is straightforward to implement, but as mentioned before can cause spurious mixing over mountains, especially in the neighborhood of steep terrain. This is largely due to the presence of large vertical temperature variations along the sloping surfaces that overlay the horizontal gradients. Such spurious mixing triggered by the vertical variations is undesirable and may grow to significant proportions. Therefore in practice, the diffusion of the temperature in the model IFS is modified to approximate the horizontal diffusion on surfaces of constant pressure rather than on the sloping η-coordinate surfaces. This is further explained in the technical model documentation of the CAM EUL and SLD dynamical cores (Collins et al. 2004). CAM EUL and SLD apply the fourth-order temperature diffusion

$$F_T = -K_4 \left[\nabla^4 T - p_s \frac{\partial T}{\partial p} \frac{\partial p}{\partial p_s} \nabla^4 \ln p_s \right] \tag{13.13}$$

where p is the pressure and p_s symbolizes the surface pressure. The second term in F_T consists of the leading term in the transformation of the ∇^4 operator from η surfaces to pressure surfaces. CAM also applies a second-order sponge-layer diffusion at upper levels. But since the upper levels in CAM coincide with pure pressure levels the correction is not needed there. In general, it is unclear whether diffusion should be applied along constant model levels, constant pressure levels or even along constant height or isentropic levels. If the diffusion primarily counteracts numerical artifacts, arguments can be found that it should be applied along model levels. However, if the primary motivation is to characterize physical mixing, height, pressure or isentropic levels are advantageous as explained in detail by Staniforth et al. (2006).

Note that NCAR's EUL and SLD dynamical cores actually apply a variant of the diffusion shown in (13.10) and (13.11) to the relative vorticity and horizontal divergence fields which generalizes the approach by Bourke et al. (1977). It is given by

$$F_\zeta = (-1)^{q+1} K_{2q} \left[\nabla^{2q} (\zeta + f) + (-1)^{q+1} (\zeta + f) \left(\frac{2}{a^2} \right)^q \right] \quad (13.14)$$

$$F_\delta = (-1)^{q+1} K_{2q} \left[\nabla^{2q} \delta + (-1)^{q+1} \delta \left(\frac{2}{a^2} \right)^q \right] \quad (13.15)$$

which diffuses the absolute vorticity $(\zeta + f)$ instead of ζ in (13.14). The undifferentiated correction terms are added to the vorticity and divergence diffusion to prevent the damping of uniform solid-body rotations. For example, a solid body wind distribution like

$$u = u_0 \cos\phi \quad (13.16)$$

$$v = 0 \quad (13.17)$$

with a constant velocity u_0 does not experience any damping by the diffusion shown in (13.14) and (13.15). The models ECHAM5 and IFS with the diffusion terms (13.10) and (13.11) do not apply this correction and thereby damp such a solid-body rotation.

The derivation of (13.14) and (13.15) can be understood when taking a second look at the scalar 3D Laplacian operator (13.6) as explained by Satoh (2004) (his Chap. 17). Substitute $\mathbf{v}_3$ with $\mathbf{v}_3 = (u, v, 0)$ and assume that the horizontal wind components u and v are proportional to the distance r to help simplify the Laplacian. This could be envisioned by assuming an idealized profile like $u(\lambda, \phi, r) \sim r k_u u_H(\lambda, \phi)$ with the constant k_u and velocity u_H that only varies in the horizontal direction. The scalar 3D Laplacian (13.6) for the velocity $\psi = u$ then yields

$$\nabla^2_{(3D)} u = \frac{1}{r^2 \cos^2\phi} \partial_{\lambda\lambda} u + \frac{1}{r^2 \cos\phi} \partial_\phi (\cos\phi \partial_\phi u) + \frac{2}{r^2} u. \quad (13.18)$$

After replacing the distance r by the radius a we obtain

$$\nabla^2_{(3D)} u = \frac{1}{a^2 \cos^2\phi} \partial_{\lambda\lambda} u + \frac{1}{a^2 \cos\phi} \partial_\phi (\cos\phi \partial_\phi u) + \frac{2}{a^2} u. \quad (13.19)$$

A similar expression holds for v. With these approximations, the horizontal diffusion expressed in (13.14) and (13.15) can be derived as shown by Satoh (2004).

The explicit diffusion mechanisms for the horizontal divergence in spectral models (13.11) or (13.15) readily controls gravity waves as also pointed out by Randall (1994). This suggests an interesting analogy to another type of explicit damping called horizontal divergence damping. The latter is sometimes used in GCMs written in momentum (u, v) form and is explained in detail in Sect. 13.4.1. Although both mechanisms are characterized with different names they accomplish a similar or even identical physical effect, namely they damp the divergent motions with either a second-order or higher order diffusion. We return to this analogy later.

The diffusion discussed so far is linear. It can therefore easily be calculated in spectral space in the spectral models ECHAM5, IFS, EUL or SLD since the spherical harmonics basis functions are eigenfunctions of the Laplacian operator on the sphere. These models utilize a constant diffusion coefficient and apply the diffusion with an implicit temporal discretization to increase the numerical stability of the diffusion mechanism. This is discussed below in Sects. 13.3.4.1 and 13.3.5.

13.3.3 Practical Considerations: Linear High-Order Diffusion Operators

Second-order diffusion schemes are not very scale selective and can therefore impact the well-resolved scales in a negative way. In practice, higher-order hyper-diffusion formulations are generally preferred to improve the scale selectivity. This is despite the fact that higher-order diffusion does not posses a physical foundation. High-order diffusion such as a fourth-order hyper-diffusion based on the ∇^4 operator is most often chosen. Even sixth- or eighth-order hyper-diffusion schemes are applied in GCMs, e.g., in ECHAM5 (Roeckner et al. 2003). The higher order can either be achieved via multiple applications of the ∇^2 operator with sign reversals as in Staniforth et al. (2006) or via the direct discretization of the higher-order operators.

As a sneak preview of the practical aspects of the diffusion discussion in Sects. 13.3, 13.4 and 13.6, we start the assessment by isolating the effects of the fourth-order hyper-diffusion and second-order sponge-layer diffusion (applied in the top three levels < 14 hPa) in an idealized dynamical core simulation. In particular, we choose the CAM (version 4) SLD dynamical core at the triangular (T) truncation T85 (≈ 156 km) with 26 levels. A steady-state test case, described in Jablonowski and Williamson (2006a), is used and run for 30 days with (a) no explicit diffusion and (b) the default second-order and fourth-order diffusion using the default coefficients $K_2 = 2.5 \times 10^5$ m^2 s^{-1} and $K_4 = 1 \times 10^{15}$ m^4 s^{-1}. Note that this K_2 value is the base value at the third level from the top. It is doubled at the second level and doubled again at the top level (Neale et al. 2010). This adds a vertical dependency to the formulation of K_2 that is purely based on the level number without taking the actual pressure or height position into account. The second-order diffusion serves as a sponge near the model top. Figure 13.1 shows the corresponding zonal-mean zonal wind fields at day 30 and the difference plot between the run with default diffusion and no diffusion. The run without diffusion (Fig. 13.1a) keeps an almost perfect steady state that is visually indistinguishable from the initial state. In the default diffusion simulation (Fig. 13.1b) the fourth-order hyper-diffusion acts throughout the entire atmosphere except in the uppermost three levels above 14 hPa. Its influence on this very smooth steady state solution is negligible and the absolute differences in the lower layers are on the order of ± 0.025 m s^{-1}. The difference plot in Fig. 13.1c is clearly dominated by the effects of the second-order diffusion in the sponge layer. The sponge layer diffusion changes the shape of the midlatitudinal

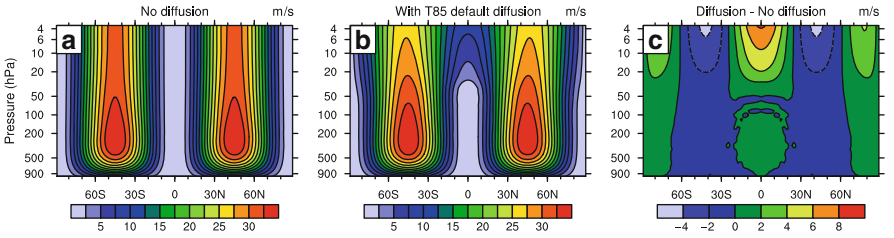

Fig. 13.1 Zonal-mean zonal wind (m/s) at day 30 of the steady-state test case of Jablonowski and Williamson (2006a) in the CAM 4 SLD dynamical core at the resolution T85L26 with (**a**) no diffusion, (**b**) the default second-order and fourth-order diffusion (see text), (**c**) zonal wind difference between the simulation with diffusion and no diffusion. A logarithmic pressure coordinate is used. No decentering is applied. The time step is $\Delta t = 1{,}800$ s

zonal jets considerably which leads to a decrease in the peak wind speeds at the model top by about 4 m s^{-1}. The diffusion also causes significant increases in the wind speeds in the formerly calm equatorial and polar regions by up to 8 m s^{-1}. It thereby smoothes out the sharp gradients in the zonal wind field. Note that the influence of the second-order diffusive sponge is not just limited to the top three layers. It clearly modulates the wind profile in the uppermost six levels which lie above 54 hPa. These sponge layer effects are discussed in greater detail later in Sect. 13.4.5.

As an aside for completeness, no decentering of the trajectories was used ($\epsilon = 0$), as will be explained later in Sect. 13.6.3, and the trajectory calculation utilizes only spherical coordinates to suppress any signal from non-zonal geodesic trajectory calculations in polar regions (typically poleward of 70°). The local geodesic coordinate is essentially a rotated spherical coordinate system whose equator goes through the arrival point of the trajectory (see details in Williamson and Rasch 1989). Of course, omitting decentering is only reasonable in the absence of mountains, and the exclusive use of spherical coordinates is only reasonable in the case of zonal advection as considered in the special case here. These deviations from the default CAM SLD configuration are selected to truly isolate the damping effects from the linear horizontal diffusion. In practice, the damping of all explicit and implicit dissipation mechanisms as well as filters and fixers act in concert, and they are generally difficult to isolate individually.

13.3.4 Choice of the Diffusion Coefficient: Damping Time Scales

The choice of the ∇^2, ∇^4 or even higher-order diffusion coefficient is most often motivated by empirical arguments and chosen in a somewhat arbitrary manner. It is sometimes even considered a model *tuning* parameter. However as seen in Fig. 13.1 and also shown later, the diffusion can have a profound impact on the global circulation, and must be chosen with care. This was also noted by Stephenson (1994) who

points out that relatively few systematic horizontal diffusion studies have been performed with realistic GCMs. Among the few are the studies by Williamson (1978), MacVean (1983) and Laursen and Eliasen (1989).

This raises questions about the scaling of the subgrid-scale horizontal mixing parameterizations with horizontal resolution. For example, Smagorinsky (1963) suggested a second-order diffusion based on turbulence concepts in which the Eddy diffusivity depends on the square of the model horizontal grid spacing and the deformation of the flow field (see Sect. 13.3.6). Higher degree hyper-diffusion operators, such as ∇^4, ∇^6 or ∇^8, are commonly used in spectral transform models. Takahashi et al. (2006) showed with the help of an Eulerian spectral transform model with a truncation limit of about wavenumber $n_0 \leq 100$ that the coefficients for the ∇^4 form may be chosen to yield a straight kinetic energy spectrum with a slope of n^{-3} for spherical wavenumbers n between $[15, n_0]$. For significantly higher truncations, coefficients can be found which yield a slope of $\mathrm{n}^{-5/3}$ beyond $n = 100$. Appendix B of Jakob et al. (1993) provides details about the calculation of such kinetic energy spectra. As a physical motivation, Skamarock (2004) (see also Chap. 14), Takahashi et al. (2006) and Hamilton et al. (2008) discussed the desirability of modeling such slopes based on observational evidence (Nastrom and Gage 1985) and the theoretical reasons why they may or may not be expected. The latter is also addressed in Chen and Wiin-Nielsen (1978) and Boer and Shepherd (1983). Examples of kinetic energy spectra in the resolution range from 224 km down to 3.5 km are shown in Terasaki et al. (2009) for the nonhydrostatic global model NICAM (Satoh et al. 2008).

In practice, coefficients for different resolutions are found experimentally with the model configured for earth-like simulations so that in the mid- to upper-troposphere the kinetic energy spectra have the desired straight tails for each resolution. For example, Boville (1991) empirically determined diffusion coefficients via trial and error with the NCAR Community Climate Model, Version 1 (CCM1). Boville tested coefficients in short model integrations and adjusted them until he obtained kinetic energy spectrum at 250 hPa which did not change shape near the truncation limit. Using the same approach, diffusion coefficients have also been found for CAM 3.1 which provide kinetic energy spectra similar to those of Takahashi et al. (2006). The dynamical component of CAM 3.1 is the Eulerian spectral transform scheme as in the model used by Takahashi et al. (2006), although CAM 3.1 has a different subgrid-scale physics parameterization package and different water vapor transport. The model of Takahashi et al. (2006) used the Eulerian spectral transform method for the water vapor transport and applied the diffusion to water vapor as well as temperature, vorticity and divergence. In contrast, CAM 3.1 uses shape preserving semi-Lagrangian approximations for water vapor transport. Diffusion is applied to temperature, vorticity and divergence, but not to water vapor.

The choice of the diffusion coefficient needs to obey physical and numerical constraints. From a physical viewpoint, the coefficient influences the damping time scales for all waves and should be as small as possible for the resolved large scales to avoid overly strong dissipation of the physically relevant signals while still providing enough damping to prevent the build-up of energy and enstrophy at the smallest

scale. From a numerical viewpoint, the coefficient needs to be sufficiently large to guarantee stable computations if the numerical approximation requires such damping. The stability aspect deserves special attention in explicit time-stepping schemes that place upper stability limits on the strength of the coefficient or impose restrictive time step sizes. An inadequately chosen coefficient can even act as the source of grid-scale noise and numerical instability. Both aspects are further explained below.

Ideally the rate of the energy dissipation near the truncation limit of a model should mimic the true energy transfer rates of the atmosphere at these scales. But unfortunately the knowledge from atmospheric observations of such transfer or dissipation rates is relatively poor as discussed by MacVean (1983). However, it helps to associate the value of the diffusion coefficient with a damping time scale at the smallest spatial scale in the model, since time scales can readily be understood from a physical viewpoint.

13.3.4.1 Diffusion Coefficients in Spectral Transform Models

Recall that $2q$ is the order of the diffusion where $q \geq 1$ is a positive integer. In spectral transform models the diffusion coefficient K_{2q} is typically represented by

$$K_{2q} = \frac{1}{\tau}\left(\frac{a^2}{n_0(n_0+1)}\right)^q \tag{13.20}$$

as for example shown by MacVean (1983), Sardeshmukh and Hoskins (1984) and Roeckner et al. (2003). τ is the e-folding time scale for the diffusion at the smallest wavelength, a denotes the Earth's radius and n_0 symbolizes the maximum wavenumber corresponding to the smallest wavelength. The wavenumber is e.g., specified by a triangular truncation limit like T85 with $n_0 = 85$. Equation (13.20) means that the (n, m)-th spectral component of the diffused quantity in the time-continuous case will be damped by the response function

$$E_n = \exp(-\Delta t\, d_n) = \exp\left\{-\Delta t\left[\frac{1}{\tau}\left(\frac{n(n+1)}{n_0(n_0+1)}\right)^q\right]\right\} \tag{13.21}$$

$$= \exp\left\{-\Delta t\; K_{2q}\left(\frac{n(n+1)}{a^2}\right)^q\right\} \tag{13.22}$$

where Δt symbolizes the length of the time step and d_n determines the strength of the damping (Sardeshmukh and Hoskins 1984; von Storch 2004). Here, n denotes the total (also called spherical) wavenumber and m stands for the zonal wavenumber as discussed in textbooks like Kalnay (2003) or Durran (1999, 2010). Note that the damping is independent of m. The response function is equivalent to an "amplification factor" represented by

$$E_n = \frac{\Psi_n^{t+\Delta t}}{\Psi_n^t} \tag{13.23}$$

that expresses the ratio of the wave amplitudes Ψ_n for each wavenumber n at the discrete future $t + \Delta t$ and current t time levels. The response function provides a damping mechanism for all $E_n < 1$.

The time-discretized form of the response function (13.21) yields

$$E_n \approx 1 - \Delta t \left[\frac{1}{\tau} \left(\frac{n(n+1)}{n_0(n_0+1)} \right)^q \right] \tag{13.24}$$

which can be transformed into the approximate form

$$E_n \approx \left\{ 1 + \Delta t \left[\frac{1}{\tau} \left(\frac{n(n+1)}{n_0(n_0+1)} \right)^q \right] \right\}^{-1} . \tag{13.25}$$

An almost identical response function in comparison to (13.25) is also shown in Collins et al. (2004) for the temperature diffusion in the CAM EUL spectral transform dynamical core. However, this Eulerian dynamical core applies the damping over a duration of $2\Delta t$ due to the chosen leapfrog time-stepping scheme. In fact, CAM EUL uses time splitting and applies the temperature diffusion with an implicit temporal discretization in spectral space after some other temporal updates

$$T_n^m = \tilde{T}_n^m - 2\Delta t \; K_{2q} \left(\frac{n(n+1}{a^2} \right)^q T_n^m \; . \tag{13.26}$$

T_n^m and $\tilde{T}_n^m$ symbolize the spectral coefficients for the temperature at the future and partially updated past time, respectively. Note that (13.26) can also be rewritten as

$$T_n^m = E_n \, \tilde{T}_n^m \tag{13.27}$$

$$E_n = \left\{ 1 + 2\Delta t \; K_{2q} \left(\frac{n(n+1)}{a^2} \right)^q \right\}^{-1} \tag{13.28}$$

which confirms that the response function E_n plays the role of a damping mechanism as stated above in (13.23). The only difference between (13.28) and (13.25) is the duration of the time interval which depends on the time-stepping scheme in the dynamical core. More details on the application of the diffusion in the model CAM EUL are provided in Collins et al. (2004) (their Chap. 3.1.14) that also explains how the correction to pressure levels is computed (13.13). The diffusion mechanism in (13.27) and (13.28) can also be compared to the concept of "spectral viscosity" which is discussed below in Sect. 13.3.4.2.

Equations (13.20) and (13.21) reflect the strength of the damping for the diffusion operators shown in (13.10)–(13.12). If the variant of the diffusion is used that does not diffuse a solid body rotation (13.14)–(13.15) the relationship between the

diffusion coefficient and the time scale becomes

$$K_{2q} = \frac{1}{\tau}\left(\frac{a^{2q}}{[n_0(n_0+1)]^q - 2^q}\right). \tag{13.29}$$

This leads to the time-continuous response function

$$E_n = \exp\left\{-\Delta t\left[\frac{1}{\tau}\left(\frac{[n(n+1)]^q - 2^q}{[n_0(n_0+1)]^q - 2^q}\right)\right]\right\} \tag{13.30}$$

for the vorticity and divergence diffusion. Analogous to (13.25) the time-discretized form of (13.30) can be approximated by

$$E_n \approx \left\{1 + \Delta t\ K_{2q}\left(\frac{[n(n+1)]^q - 2^q}{a^{2q}}\right)\right\}^{-1} \tag{13.31}$$

that reintroduces the corresponding diffusion coefficient K_{2q} (13.29) into the equation.

The strength of the horizontal diffusion, as explicitly stated in (13.25), (13.28) and (13.31), is scale-dependent. This is confirmed for (13.28) in Fig. 13.2. The figure is drawn for a T85 ($n_0 = 85$) triangular truncation with an assumed time step of $2\Delta t = 1{,}200$ s and a damping time scale of $\tau = 8$ h. These settings are close to the default values for the fourth-order hyper-diffusion in CAM EUL (see Sect. 13.3.4.4) that utilizes the leapfrog time-stepping scheme. The response function (13.28) for the second-, fourth-, sixth- and eighth-order horizontal diffusion clearly shows that the higher-order diffusion provides less damping at long spatial scales (low wavenumbers) and rapidly increases in strength towards the high

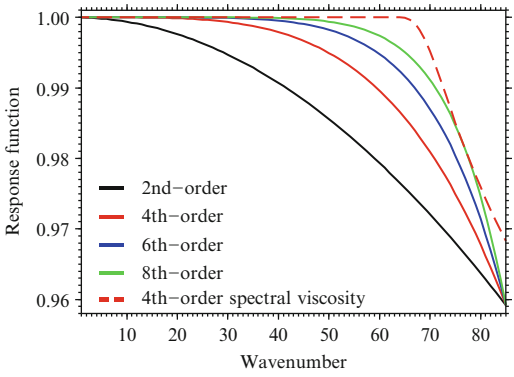

Fig. 13.2 Scale selectivity of the response function E_n shown in (13.28) for a spectral T85 triangular truncation and second-, fourth-, sixth- and eighth-order horizontal diffusion. In addition, the filter function σ_n (13.39) of the spectral viscosity is shown. A time step of $2\Delta t = 1{,}200$ s and damping time scale of $\tau = 8$ h are assumed

wavenumbers. The highest spherical wavenumber $n = 85$ is damped at an identical rate in this example. The figure also shows the damping characteristics of the spectral viscosity method (13.39) that is discussed in the next subsection.

As an aside, Leith (1971) suggested an alternative formulation of the response function. His derivation is based on an analysis of the numerical dissipation function that arises from the energy transfer between wavenumbers in a two-dimensional Cartesian turbulence closure model known as the Eddy-damped Markovian approximation. Gelb and Gleeson (2001) showed the corresponding response function for the Leith (1971) fourth-order hyper-diffusion when utilizing a leapfrog time-stepping scheme. It yields

$$E_n = \left\{ 1 + 2\Delta t \; K_L \; \frac{(n - n_L)^2 (n - n_L + 1)^2}{a^4} \right\}^{-1} \tag{13.32}$$

where the parameter $n_L = 0.55 n_0$ is the cutoff wavenumber and the diffusion coefficient K_L is defined as

$$K_L = \begin{cases} 0 & n \le n_c \;, \\ \dfrac{K_4}{(0.45)^4} & n_c < n \le n_0 \quad . \end{cases} \tag{13.33}$$

K_4 is given in (13.20) for $q = 2$. A second-order version of the Leith (1971) diffusion has been implemented in the weather prediction model "Global Forecast System" (GFS 2003) which is an operational spectral transform model at the National Centers for Environmental Prediction (NCEP). A comparison of the fourth-order traditional diffusion (13.28), the Leith (1971) diffusion (13.32) and spectral viscosity is provided in Gelb and Gleeson (2001).

13.3.4.2 The Concept of Spectral Viscosity

The horizontal diffusion in spectral transform models can also be replaced with a spherical "spectral viscosity" operator as proposed by Gelb and Gleeson (2001). They suggested a spectral viscosity method which is built upon rigorous mathematical principles for nonlinear conservation laws. It yields a viscosity term that depends on the spatial scale of the waves and is equal to zero for low wavenumbers. It is thereby highly scale-selective and does not damp well-resolved wave modes.

Note that the traditional fourth-order hyper-diffusion of the form $F_\Psi = -K_4 \nabla^4 \Psi$ (13.2) has the analytic spectral representation

$$\{F_\Psi\}_n^m = - K_4 \; \frac{n^2 (n+1)^2}{a^4} \; \Psi_n^m . \tag{13.34}$$

The spectral viscosity (SV) approach then translates (13.34) into the new form

$$\{F_\Psi^{SV}\}_n^m = - \epsilon \; \hat{q}_n^2 \; \frac{n^2 (n+1)^2}{a^4} \; \Psi_n^m \tag{13.35}$$

where ϵ is a tuning parameter that, according to the mathematical theory behind spectral viscosity, scales like $\epsilon \sim n_0^{-3}$. $\hat{q}_n$ is given by

$$\hat{q}_n = \begin{cases} 0 & n \le n_c, \\ \exp\left(-\dfrac{(n-n_0)^2}{2(n-n_c)^2} \right) & n_c < n \le n_0. \end{cases} \tag{13.36}$$

As before, n_0 is the maximum spherical wavenumber, a is the earth's radius, and n_c is a cutoff wavenumber. The spectral viscosity is zero for all waves with wavelengths that are larger than the cutoff wavelength. Gelb and Gleeson (2001) recommended the scaling parameters

$$\epsilon = \frac{c\, a^3}{n_0^3} \qquad \text{and} \qquad n_c = 2\, n_0^{\frac{3}{4}} \tag{13.37}$$

with $c = 2\,\mathrm{m\,s^{-1}}$ and tested these in an Eulerian spectral transform shallow water model with a leapfrog time-stepping scheme. The parameter c needs to carry velocity units to match the physical dimensions, which were originally omitted. Gelb and Gleeson (2001) noted that these scalings might not be universal since they were tuned for a single shallow water test case.

The fourth-order spectral viscosity yields

$$\Psi_n^m = \sigma_n\, \tilde{\Psi}_n^m \tag{13.38}$$

$$\sigma_n = \left\{ 1 + 2\Delta t\, [\epsilon\, \hat{q}_n^2]\, \frac{n^2(n+1)^2}{a^4} \right\}^{-1} \tag{13.39}$$

where Ψ_n^m stands for the spectral coefficients of a prognostic variable, such as temperature, at the future time step, and $\tilde{\Psi}_n^m$ represents the spectral coefficients at a partially updated past time level, as discussed earlier in Sect. 13.3.4.1. The spectral viscosity operation specified in (13.38) is equivalent to applying a spectral filter of form σ_n at each time step (Canuto et al. 1987). Equation (13.38) is formally identical to the equation for the fourth-order ($q = 2$) hyper-diffusion (13.27). The most important difference is the definition of the response function σ_n (13.39) in comparison to E_n given in (13.28). The comparison reveals that the wavenumber-dependent viscosity parameter $[\epsilon\, \hat{q}_n^2]$ replaces the constant diffusion coefficient K_4. The different damping characteristics of σ_n and E_n can clearly be seen in Fig. 13.2. The dashed red curve depicts the spectral viscosity response function that can readily be compared to the traditional fourth-order hyper-diffusion (solid red curve). It confirms that the spectral viscosity does not damp the low wavenumbers (e.g., up to spherical wavenumber $n = 56$ in this example) and then quickly increases in strength. However, at this particular resolution and time step the damping effect of the spectral viscosity with parameter set (13.37) is always weaker than the damping of the traditional diffusion.

Gelb and Gleeson (2001) showed via 2D shallow water tests that the spectral viscosity method gives appreciably superior results if the flow field is underresolved.

In a well-resolved flow field the accuracy of SV and hyper-diffusion simulations was comparable. However, they observed that the use of spectral viscosity improved the conservation of invariants by the numerical scheme and led to more accurate energy spectra. Spectral viscosity has not been used in operational 3D spectral transform GCMs so far.

13.3.4.3 Diffusion Coefficients in Grid-Point Models

It is less straightforward to represent the relation between the damping time and the K coefficient in grid point models since the exact relation depends on the type of spatial discretization and the model grid. We therefore pick an example that illustrates the relationship and then generalize it.

The example reflects the UK Met Office's Unified Model on a latitude–longitude grid that utilizes a finite-difference approach in spherical coordinates. Following the arguments by Staniforth et al. (2006) in their Chap. 12 the response function for a centered finite-difference approximation of the second-order diffusion yields

$$E = 1 - \Delta t \left(\frac{K_\lambda \, \sin^2(k_\lambda \Delta\lambda/2)}{a^2 \cos^2\phi (\Delta\lambda/2)^2} + \frac{K_\phi \, \sin^2(k_\phi \Delta\phi/2)}{a^2 (\Delta\phi/2)^2} \right) \tag{13.40}$$

where K_λ and K_ϕ are the second-order diffusion coefficients, k_λ and k_ϕ stand for the longitudinal and latitudinal wavenumbers, and $\Delta\lambda$ and $\Delta\phi$ are the grid spacings (in radians) in the longitudinal and meridional direction, respectively. Selecting an isotropic diffusion with coefficients $K_\lambda = K_\phi$ reveals a stability concern in (13.40), as further discussed in Sect. 13.3.5. In practice, $K_\lambda \cos^2\phi = K_\phi = constant \equiv K$ is therefore chosen (Staniforth et al. 2006) which leads to the damping for each pair of wavenumbers

$$E = 1 - \Delta t \, K \left(\frac{\sin^2(k_\lambda \Delta\lambda/2)}{a^2 (\Delta\lambda/2)^2} + \frac{\sin^2(k_\phi \Delta\phi/2)}{a^2 (\Delta\phi/2)^2} \right). \tag{13.41}$$

Assuming that the highest wavenumbers $k_\lambda = 2\pi/L_x$ and $k_\phi = 2\pi/L_y$ are represented by the smallest resolvable wavelengths $L_x = 2\Delta\lambda$ and $L_y = 2\Delta\phi$ the relation between the diffusion coefficient and the time scale τ for the shortest waves becomes

$$\frac{1}{\tau} = K \left(\frac{1}{a^2 (\Delta\lambda/2)^2} + \frac{1}{a^2 (\Delta\phi/2)^2} \right) \tag{13.42}$$

or equivalently

$$K = \frac{1}{\tau} \left(\frac{1}{(a\Delta\lambda/2)^2} + \frac{1}{(a\Delta\phi/2)^2} \right)^{-1}. \tag{13.43}$$

Note that $\Delta x = a\Delta\lambda$ and $\Delta y = a\Delta\phi$ express the physical grid spacings at the equator.

A generalization of this approach for grid point models with approximately uniform physical grid spacings $\Delta x = \Delta y$ and $2q$-th-order hyper-diffusion yields

$$K_{2q} = \frac{1}{2\tau}\left(\frac{\Delta x}{2}\right)^{2q}. \tag{13.44}$$

Examples of such models are GME on an icosahedral grid (Majewski et al. 2002) or the dynamical core HOMME on a cubed-sphere mesh (see Taylor et al. (2007) and Chap. 12). Other models like the nonhydrostatic icosahedral dynamical core NICAM (Tomita and Satoh 2004) define the horizontal and vertical diffusion coefficients as

$$K_{2q\,H} = \gamma_H \frac{\overline{\Delta x}^{\,2q}}{\Delta t} \tag{13.45}$$

$$K_{2q\,V} = \gamma_V \frac{(\Delta \xi)^{2q}}{\Delta t} \tag{13.46}$$

where γ_H and γ_V are non-dimensional empirical factors and $\Delta\xi$ is the vertical grid spacing in the generalized terrain-following height coordinate ξ. Here $\overline{\Delta x}$ symbolizes the average grid spacing in their quasi-uniform triangular grid given by

$$\overline{\Delta x} = \sqrt{\frac{4\pi a^2}{N}} \tag{13.47}$$

with a total of N grid points per model level. $4\pi a^2$ denotes the total surface area of the sphere. After an empirical factor γ and thereby K_{2q} is chosen, its damping time scale is

$$\tau = \frac{\Delta t}{2^{2q+1}\,\gamma} \tag{13.48}$$

according to (13.44). The symbol γ is used as a placeholder for either γ_H or γ_V. Concrete values for γ are presented below in Sect. 13.3.4.5. As an aside, explicitly added vertical diffusion is often considered part of the physical parameterization suite, and is rarely included in a dynamical core.

The details of the assessments above will differ somewhat based on the numerical scheme and the degree of non-uniformity of the computational grid. Nevertheless, (13.44) gives guidance when selecting the appropriate damping time scales. Note again, that models with explicit time discretizations enforce stability limits on the strength of the coefficient (Jakimow et al. 1992; Staniforth et al. 2006) for a given time step.

13.3.4.4 Examples of Diffusion Coefficients in Spectral Models

As a practical example, we briefly discuss the heterogeneous horizontal diffusion mechanism in the spectral model ECHAM5. First, the order of the hyper-diffusion scheme varies depending on the model level with highly scale-selective sixth or

eighth-order diffusion at low levels and increased fourth- and second-order diffusion in higher regions. This increased diffusion near the model top serves as a sponge to lessen the spurious reflection of planetary and gravity waves at the upper boundary. Second, the strength of the diffusion depends on the type of prognostic variable. Assume that τ_{vor} is the e-folding damping time of the highest resolvable wavenumber for the horizontal vorticity diffusion. In ECHAM5, τ is independent of the vertical position and order of the diffusion but changes with horizontal resolution (Roeckner et al. 2003; Wan et al. 2008). The e-folding times for the horizontal divergence and temperature diffusion are then chosen as

$$\tau_{div} = 0.2\,\tau_{vor} \tag{13.49}$$

$$\tau_T = 2.5\,\tau_{vor}. \tag{13.50}$$

For low resolution T42, T63 and T85 ECHAM5 simulations, the vorticity e-folding times τ_{vor} are typically set to 9, 7 and 5 h, respectively. These resolutions correspond to equatorial grid spacings of about 313, 208 and 156 km.

In contrast, other models such as the semi-Lagrangian and Eulerian spectral transform dynamical cores in CAM utilize the same diffusion coefficient for all prognostic variables. Boville (1991) recommended damping time scales of 14.4 and 7.2 h at the resolutions T42 and T63 for NCAR's Eulerian model with fourth-order hyper-diffusion. Similar damping time scales were suggested by Takahashi et al. (2006) who also investigated the higher-resolution spectral truncations T159, T319 and T639 with the spectral transform Eulerian model AFES (Enomoto et al. 2008). More concretely, Hamilton et al. (2008) reported AFES damping time scales around 9.6, 8.4, 4.8, 2.4 h for the spectral truncations T79, T159, T319 and T639. They yield the K_4 diffusion coefficients $1.19 \times 10^{15}, 8.41 \times 10^{13}, 9.14 \times 10^{12}$ and 1.14×10^{12} m^4 s^{-1}, respectively, according to (13.20). The corresponding grid spacings are approximately 167, 83, 42 and 21 km.

As mentioned before, the diffusivity parameter is generally empirically adjusted in each simulation to produce results in which the ends of the kinetic energy spectra follow a power law and do not change with the model resolution. Both Boville (1991) and Takahashi et al. (2006) found that the diffusivity coefficient needs to be scaled at about the inverse third power of the spectral truncation. As a consequence, the diffusion time scale of the smallest resolved scale drops with increased resolution as also suggested by Williamson (2008a) for CAM 3.1. In particular, Williamson (2008a) proposed damping time scales of 14.0, 8.6, 3.6 and 1.5 h for the T42, T85, T170 and T340 spectral truncations in the Eulerian dynamical core which yield the K_4 coefficients $1 \times 10^{16}, 1 \times 10^{15}, 1.5 \times 10^{14}$ and 2.25×10^{13} m^4 s^{-1}, respectively. In addition, a second-order diffusive sponge was employed in the three uppermost levels. The base diffusion coefficient at the third level from the top was held constant with $K_2 = 2.5 \times 10^5$ m^2 s^{-1} regardless of resolution which corresponds to damping time scales of about 25, 6.2, 1.6 and 0.4 h at the resolutions T42, T85, T170 and T340. The two highest resolutions T170 and T340 correspond to equatorial grid spacings of about 78 and 39 km. Note again that the K_2 coefficient was doubled at the second level and doubled again at the first level from the top. The

second-order sponge layer diffusion was introduced to damp the vertically propagating resolved waves and prevent them from being reflected back down. Therefore, the coefficient is not reduced with resolution. It was chosen by trial and error to yield reasonable stratospheric polar night jet speeds and reduce the polar cold bias in the stratosphere. As resolutions increase further in the future, this choice of the diffusion coefficient K_2 might need to be revisited since a sudden onset of a strong sponge layer with very short time scales can act as a wave reflector by itself. This would defeat the purpose of the sponge layer.

As a concrete example, we now evaluate the effects of horizontal dissipation on the development of baroclinic waves. This was also suggested by MacVean (1983) who conducted a systematic study on the effects of hyper-diffusion and energy transfers in idealized GCM experiments. The effects of different types of hyper-diffusion are also discussed later in Sect. 13.3.8. Here, we start by isolating the effects of a varying fourth-order horizontal diffusion coefficient on growing baroclinic waves, and utilize the idealized dynamical core test case by Jablonowski and Williamson (2006a).

Figure 13.3 shows the surface pressure, 850 hPa temperature and vertical pressure velocity fields at day 9 of two simulations with the CAM Eulerian spectral transform dynamical core at the triangular truncation T85 with 26 levels. Both simulations applied the fourth-order hyper-diffusion with (left column) the default coefficient $K_4 = 1 \times 10^{15}\,\mathrm{m^4\,s^{-1}}$, and (right column) an increased coefficient by the factor

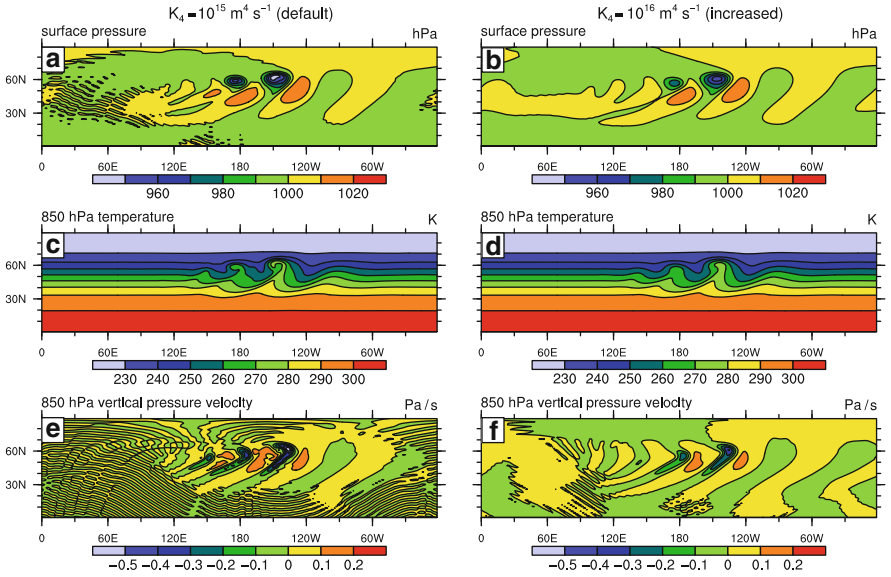

Fig. 13.3 (**a**,**b**) Surface pressure (hPa), (**c**,**d**) 850 hPa temperature (K) and (**e**,**f**) vertical pressure velocity (Pa/s) at day 9 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) in the CAM T85L26 Eulerian spectral dynamical core with ∇^4 diffusion. *Left*: default diffusion coefficient $K_4 = 1 \times 10^{15}\,\mathrm{m^4\,s^{-1}}$, *Right*: increased diffusion coefficient by a factor of 10. A time step of $\Delta t = 600\,\mathrm{s}$ is used

of 10. The corresponding damping time scales are 8.6 and 0.86 h, respectively. A second-order diffusive sponge near the model top is also applied, but is irrelevant for the discussion here. The figure shows that the baroclinic wave grows in both simulations, but that the circulation is relatively damped in the increased diffusion run. The peak magnitudes of the surface pressure and vertical pressure velocity fields are clearly reduced (Fig. 13.3b,f) and the otherwise sharp gradients in the temperature field are smeared out in the highly diffusive run (Fig. 13.3d). However, the increased diffusion diminished (but not eliminated) the numerical noise, the so-called Gibbs ringing, which is very dominant in the default configuration. The Gibbs phenomena are introduced by the need to represent fields with discontinuities or sharp gradients by smooth global basis functions. The noise does not grow unstable in the default EUL run, but such noise in the vertical velocity field can have detrimental effects on other quantities like precipitation in case of full GCMs with physical parameterizations. Spurious rainfall with such a noisy signature is sometimes called "spectral rain", and is undesirable. Therefore, a delicate balance needs to be found between sufficient scale-selective damping and too diffusive simulations. As an aside, except for the spectral ringing the evolution of the baroclinic wave in the default EUL configuration resembles high-resolution reference solutions of other models quite closely as shown in Jablonowski and Williamson (2006a,b) and Lauritzen et al. (2010a).

The impact of explicitly added diffusion on the evolution of baroclinic waves was also investigated by Polvani et al. (2004). They demonstrated that the choice of the diffusion mechanism can fundamentally change the characteristics of the flow field. In particular, their spectral transform simulation with a second-order diffusion scheme had very little resemblance to their nominally identical simulation that utilized a fourth-order hyper-diffusion mechanism. Polvani et al. (2004) also evaluated whether the solutions numerically converge with increasing resolution when keeping a constant diffusion coefficient. For example, they tested a fourth-order hyper-diffusion with the constant coefficient $K_4 = 2.5 \times 10^{16}\,\mathrm{m}^4\,\mathrm{s}^{-1}$ and simulated the evolution of a baroclinic wave at the triangular truncations T21, T42, T85, T170 and T341. This K_4 coefficient corresponds to the damping time scales around 85.8, 5.6, 0.34 h, 78 and 5 s, respectively. The extremely short time scales at the higher resolutions are associated with strong diffusion that dominates the flow and thereby allows the numerical solutions to converge. It means that diffusion can effectively reduce the spatial resolution by suppressing the generation of finer-scale structures that are normally resolved at higher resolutions. As a note of caution, smooth-looking solutions could be caused by overly strong diffusion and are therefore not necessarily accurate. This is also shown later in Sect. 13.6.1 that compares the impact of inherent numerical dissipation.

13.3.4.5 Examples of Diffusion Coefficients in Grid Point Models

All aforementioned diffusion coefficients for spectral models are quite comparable and lie within a factor of 2–3 at a given resolution. But note that the amount of

explicit diffusion needed for a numerical scheme is highly dependent on the inherent diffusive characteristics of the numerical scheme, and all other filters or fixers in the GCM as further explained in this chapter. For example, a second-order grid point model might have different diffusion needs than spectral models. The diffusion coefficients for such a finite-difference-type grid point model on an icosahedral grid are listed in Majewski et al. (2002). They show the grid-size dependencies of the fourth-order linear diffusion coefficients in the weather forecast model GME for various horizontal grid spacings between 10 and 160 km. A comparison with Williamson (2008a) reveals that the K_4 diffusion coefficients in the model GME are higher by factors between 3.5 and 5. A comparison of the damping time scales in GME with the K_4 damping coefficients 5.25×10^{15} and 6.5×10^{14} m^4 s^{-1} at the approximate $\Delta x = \Delta y$ grid spacings 160 and 80 km yields damping time scales of about 1.1 and 0.55 h according to (13.44). The result confirms that GME employs a stronger hyper-diffusion mechanism in comparison to spectral models at similar resolutions.

Lastly, we comment on the empirically tuned diffusions as shown above in (13.45) and (13.46). In particular, Tomita and Satoh (2004) chose a ($q = 2$) fourth-order diffusion in the horizontal direction and a ($q = 3$) sixth-order diffusion in the vertical. At the $\Delta x \approx 240$ km horizontal resolution with a time step of $\Delta t = 1{,}800$ s the empirical factor γ_H was most often set to 6.25×10^{-3}. In Tomita and Satoh (2004) this led to the diffusion coefficient $K_{4H} = 1.152 \times 10^{16}$ m^4 s^{-1} which corresponded to the damping time scale of 2.5 h according to (13.48).

13.3.4.6 Caveats

Unfortunately and as a word of caution, the choice of the diffusion coefficients often remains undocumented in the refereed literature. Even the defaults are difficult to find, and the resolution-dependencies are only rarely mentioned. Official model documentation often lacks the specific information.

13.3.5 Choice of the Diffusion Coefficient: Stability

As pointed out by Mesinger and Arakawa (1976) and Wood et al. (2007) explicit time approximations of diffusion equations, such as those presented in Staniforth et al. (2006)

$$\frac{\psi^{j+1} - \psi^j}{\Delta t} = \frac{1}{a^2}\left[\frac{\partial}{\partial\lambda}\left(\frac{K_\lambda}{\cos^2\phi}\frac{\partial\psi^j}{\partial\lambda}\right) + \frac{\partial}{\partial\phi}\left(K_\phi\frac{\partial\psi^j}{\partial\phi}\right)\right], \tag{13.51}$$

are only conditionally stable. Here, the index j symbolizes a discrete time level. This is especially problematic on latitude–longitude grids since the time step can easily violate the condition for stability close to the poles. This is mainly due to the

convergence of the meridians in polar regions that leads to shrinking longitudinal spacings and higher Courant–Friedrichs–Lewy (CFL) numbers.

If the horizontal diffusion operator is applied with an explicit time discretization, which is for example the case in the UK Met Office model UM, the diffusion coefficient needs to obey strict stability constraints. A comprehensive stability analysis for a finite-difference representation of the diffusion operator on a latitude–longitude grid is shown in Staniforth et al. (2006) (their Chap. 12). In two dimensions, the stability constraint for the second-order diffusion with $q = 1$ in the model UM is given by

$$\frac{\Delta t}{r^2}\left(\frac{K_\lambda}{\cos^2\phi\,\Delta\lambda^2} + \frac{K_\phi}{\Delta\phi^2}\right) \leq \frac{1}{4} \tag{13.52}$$

where the radial distance r from the center of the earth can also be approximated by the radius a, and K_λ and K_ϕ are the diffusion coefficients in the longitudinal and latitudinal directions. This restriction guarantees that the corresponding response function (13.41) lies between [0, 1] and does not change sign on alternate time steps. From a physical viewpoint, an isotropic choice of the diffusion coefficient $K_\lambda = K_\phi$ is advantageous to damp physical scales at the same rate. However, such a choice would enforce very stringent stability conditions on the maximum allowable time step due to the dependence on the $\cos^2\phi$ term. Therefore, $K_\lambda/\cos^2\phi = K_\phi = constant$ is generally selected in the model UM that leads to the less restrictive condition

$$\frac{K_\phi \Delta t}{a^2}\left(\frac{1}{\Delta\lambda^2} + \frac{1}{\Delta\phi^2}\right) \leq \frac{1}{4}. \tag{13.53}$$

An undesirable side effect is that the diffusion becomes highly anisotropic, particularly in polar regions where diffusion is probably most needed, and noise is much less controlled in the east–west direction than in the north–south direction (Staniforth et al. 2006). Note that the second-order response function (13.41) can also be generalized for higher-order diffusion schemes

$$E = 1 - \left\{\frac{K_\phi \Delta t}{a^2}\left[\frac{\sin^2(k_\lambda \Delta\lambda/2)}{(\Delta\lambda/2)^2} + \frac{\sin^2(k_\phi \Delta\phi/2)}{(\Delta\phi/2)^2}\right]\right\}^q \tag{13.54}$$

where $2q$ denotes the order of the hyper-diffusion.

The non-isotropy of the diffusion scheme in the model UM has a caveat. The damping for a particular physical scale decreases polewards because the physical scale represented by any given wavenumber decreases by a factor $\cos^{-1}\phi$ as the poles are approached and the response function of the operator is the same everywhere. As a result, a feature which moves equatorwards from the poles experiences increased damping which has the effect of creating a boundary effect for the propagation of that feature (Staniforth et al. 2006). In addition, the shrinking longitudinal spacing near the poles supports smaller and smaller physical scales. If these small scales are not damped enough it increases the risk of developing noise near the poles. In practice, the application of an explicit diffusion scheme on a latitude–longitude grid is often paired with the application of a polar filter that removes linear and

nonlinear instabilities if any. The applications of an additional filter also allows violations of the CFL stability condition at high latitudes as further explained below in Sect. 13.5.

To avoid the stability limitations of the horizontal diffusion mechanism an implicit temporal approximation of the diffusion is generally desirable as suggested by Jakimow et al. (1992) and Li et al. (1994). On the other hand, this can adversely affect the computational efficiency of the scheme since implicit calculations require the solution of a Helmholtz equation. As an example, an implicit representation of the horizontal diffusion shown in (13.51) yields

$$\frac{\psi^{j+1} - \psi^{j}}{\Delta t} = \frac{1}{a^2}\left[\frac{\partial}{\partial\lambda}\left(\frac{K_\lambda}{\cos^2\phi}\frac{\partial\psi^{j+1}}{\partial\lambda}\right) + \frac{\partial}{\partial\phi}\left(K_\phi\frac{\partial\psi^{j+1}}{\partial\phi}\right)\right] \tag{13.55}$$

which can also be symbolically written in the (Helmholtz) matrix form

$$[I - \Delta t(D_{\lambda\lambda} + D_{\phi\phi})]\psi^{j+1}_{\lambda,\phi} = \psi^{j}_{\lambda,\phi}. \tag{13.56}$$

I stands for the identity matrix, and $D_{\lambda\lambda}$ and $D_{\phi\phi}$ symbolize discretized diffusion operators in matrix form. However, inverting the three-dimensional matrix $[I - \Delta t(D_{\lambda\lambda} + D_{\phi\phi})]$ at every time step would be computationally expensive, although approximations of (13.56) can be formulated to make the implicit formulation more attractive (Staniforth et al. 2006).

As mentioned before and shown in (13.26), (13.27) and (13.28), spectral transform models like ECHAM5, IFS or CAM EUL or SLD always compute the horizontal diffusion implicitly in spectral space. This can be done in a straightforward way since the Laplacian operator has an an analytic spectral representation as e.g., presented for the operator ∇^4 in (13.34). The implicit calculation does not need to obey stability constraints and remains stable even with a high isotropic diffusion coefficient in both directions (Collins et al. 2004; ECMWF 2010). A thorough stability analysis of linear diffusion can also be found in Williamson and Laprise (2000).

13.3.6 Nonlinear Horizontal Diffusion

The choice of the horizontal diffusion mechanism is sometimes linked to turbulence concepts (Boer and Shepherd 1983) which e.g., can be based on nonlinear horizontal mixing coefficients. However, as stated earlier the horizontal diffusivities used in GCMs are typically many orders of magnitude larger than those which would be appropriate for turbulence closures (Mellor 1985). Thus, this association needs to be made with care. It even could offer some guidance in choosing a suitable subgrid-scale mixing scheme. For example, in NCEP's Eta model (Black 1994) this awareness led to the use of only second-order horizontal diffusion since the

choice was not based on the desire to select "scale selective" dissipation. Instead, the intention was that the diffusion scheme should mimic the impact of grid box filamentation due to deformation-dependent stretching (Fedor Mesinger, personal communication, Janjić (1990)).

Nonlinear diffusion typically defines the diffusion coefficient in terms of a nonlinear function of the horizontal wind. Nonlinear second-order diffusion was originally proposed by Smagorinsky (1963) who used a Cartesian coordinate system to derive deformation-based Eddy viscosity coefficients. The basic design of this diffusion mechanism is connected to mixing-length concepts. The latter can be motivated if the ideas of Prandtl are applied to the dissipation of enstrophy in 2D turbulence (Becker and Burkhardt 2007). As shown below Smagorinsky's nonlinear parameterization might appear to be "more physical" than any linear diffusion scheme. However, there is little theoretical basis for such a nonlinear formulation at large geophysical scales.

Nonlinear harmonic (second-order) diffusion depends on the flow field and damps only at times and places of strong horizontal shear. The generic form of this second-order diffusion needs to be written in flux form

$$F_\psi = +\nabla \cdot (K_H \nabla \psi) \tag{13.57}$$

where $\nabla\cdot$ symbolizes the divergence operator. The Eddy viscosity coefficient K_H can be symbolically associated with a length (L) and time (T) scale

$$K_H = \frac{L^2}{T} \tag{13.58}$$

and is proportional to the norm of the strain tensor and the quadratic grid spacing. The inverse time scale $T^{-1} = |D|$ is determined by the deformation rate $|D|$ which is the norm of the strain tensor. Smagorinsky (1963) defines the nonlinear coefficient as

$$K_H = (k_0 \Delta)^2 \sqrt{\left(\partial_x u - \partial_y v\right)^2 + \left(\partial_x v + \partial_y u\right)^2} \tag{13.59}$$

where $\partial_x = (a \cos\phi)^{-1} \partial\lambda$ and $\partial_y = a^{-1} \partial\phi$ denote the partial derivatives in the longitudinal and latitudinal directions in spherical coordinates. k_0 is a unitless empirical constant and Δ is a measure of the physical grid spacing, such as $\Delta = \sqrt{\Delta x \Delta y}$ (Skamarock et al. 2008), $\Delta = \Delta y$ (Koshyk and Hamilton 2001) or $\Delta = \min(\Delta x, \Delta y)$ (Griffies and Hallberg 2000). k_0 is typically set to a value between $[0.1, 0.3]$ as suggested by Smagorinsky (1963, 1993), Andrews et al. (1983), Koshyk and Hamilton (2001) or Skamarock et al. (2008), but both smaller and larger k_0 values have been tried in GCMs. Regardless of the application, the formulation of Smagorinsky's viscosity coefficient utilizes the horizontal tension $D_T = (\partial_x u - \partial_y v)$ and horizontal shearing strain $D_S = (\partial_x v + \partial_y u)$ and is thereby linked to the local deformation rate via $|D| = \sqrt{D_T^2 + D_S^2}$.

Other variants of the nonlinear diffusion coefficient exist as shown by Gordon and Stern (1982) and Andrews et al. (1983). These are summarized in Becker and Burkhardt (2007) who also list the later modification by Smagorinsky (1993)

$$K_H = (k_0 \Delta)^2 \sqrt{\left(\partial_x u - \partial_y v - v \tan\phi/a\right)^2 + \left(\partial_x v + \partial_y u + u \tan\phi/a\right)^2}. \tag{13.60}$$

This revised version can be applied consistently in spherical geometry. Note that nonlinear viscosity is not widely used in atmospheric GCMs today. However, it is commonly used in large-scale ocean models as outlined in Griffies and Hallberg (2000). Griffies and Hallberg (2000) even extended the definition of K_H to incorporate biharmonic (fourth-order) diffusion that enhances the scale selectivity of the Smagorinsky scheme.

A Smagorinsky-type Eddy viscosity is also often chosen as a subgrid-scale (SGS) model in large-eddy simulations to represent the effects of small-scale turbulence in the inertial range. LES has been widely used to study atmospheric boundary layer dynamics and vertical mixing processes. LES explicitly resolves the dynamics of large-scale eddies. They contain most of the energy and are the primary transport mechanism, while small-scale eddies in LES are parameterized by the SGS model as e.g., outlined by Huang et al. (2008). The limited-area mesoscale model WRF also includes a Smagorinsky-type option for its horizontal and vertical diffusion scheme. WRF can base the Eddy viscosities either on a 3D Smagorinsky turbulence closure or on the flow deformation (Skamarock et al. 2008).

13.3.7 Physical Consistency

Explicitly adding linear horizontal diffusion to the equations of motion is popular in GCMs, but most often the implementations are physically inconsistent (Burkhardt and Becker 2006). For example, if the horizontal momentum diffusion is applied in the form of hyper-diffusion the conservation of angular momentum is generally violated. This is e.g., discussed in Becker (2001) who argued that consistent friction in GCMs must be formulated as the divergence of a symmetric Reynolds stress tensor. In contrast, conventional formulations of the horizontal diffusion correspond to a nonsymmetric stress tensor.

In general, viscous dissipation is the conversion of mechanical energy to thermal energy by the flow working against viscous stresses. Diffusion therefore removes kinetic energy from the flow field which could be interpreted as the transfer of kinetic energy from the resolved scales to subgrid scales and finally to a turbulent microscale, where it needs to be converted into heat. A very important aspect is that the conversion is irreversible and always needs to appear as a positive frictional heating on the right-hand side of the thermodynamic equation. However, frictional heating is commonly ignored in GCMs since it is small in comparison to other contributions to the heat budget such as radiative forcing or latent heating (Fiedler 2000; Burkhardt and Becker 2006).

Nevertheless, the kinetic energy loss due to dissipation cannot be neglected over long time periods since it violates the conservation of total energy. This is especially important for climate simulations to prevent artificial climate drifts. According to the estimates by Becker (2003) using a simple GCM the total dissipation averages to approximately 2 W m^{-2} with horizontal dissipation making up about a third of the overall dissipation. Neglecting this heating would cause a spurious thermal forcing of about 0.6 W m^{-2} which is on the order of climate change signals. The conversion of the dissipated energy into heat must therefore be explicitly added to the thermodynamic equation. Similar energy losses around 2 W m^{-2} due to the explicit and inherent damping mechanisms in NCAR's CAM model were also reported by Williamson (2007).

Frictional heating due to horizontal diffusion is explicitly included in the EUL and HOMME dynamical cores of NCAR's CAM 5 model (Neale et al. 2010) and in a variant of the ECHAM4 model (Burkhardt and Becker 2006). But as noted in Becker (2001) CAM does not utilize a symmetric stress tensor formulation. The dissipative heating can therefore only be approximated as outlined in Boville and Bretherton (2003) and thereby allows (at least theoretically) dissipative "cooling" which is unphysical. Furthermore, Boville and Bretherton (2003) and Becker (2003) discussed how to include frictional heating due to vertical diffusion processes in the physical parameterizations. Alternatively or even in addition, the total energy budget can also be simply "fixed" via an a posteriori energy fixer as explained later in Sect. 13.7.3. This is common practice in dynamical cores since other types of damping, such as filters or inherent nonlinear numerical dissipation in semi-Lagrangian or finite-volume schemes, cannot be analytically quantified. Instead, their effects collectively appear in form of a residual in the total energy equation.

Two other aspects need to be considered. As pointed out by Wood et al. (2007) some diffusion schemes do not properly maintain steady-state solutions, but instead wrongly distort them. These errors in steady-state solutions may lead to systematic biases and climate drift as discussed for physical parameterization by Dubal et al. (2004, 2006). In addition, only the second-order diffusion operator guarantees the preservation of the monotonicity of the diffused field. The more scale-selective and thereby desirable higher order operators have the potential to introduce spurious new extrema and violate monotonicity constraints. The amplitude of unphysical over- and undershoots can even increase with increasing order of the diffusion (Sardeshmukh and Hoskins 1984). Over- and undershoots can cause nonlinear interactions between the physics and dynamics and result in undesired side effects such as spurious rainfall. Higher-order diffusion is especially discouraged for positive definite moisture and other tracer fields since it can lead to supersaturation or even negative tracer values. Negative values would then need to be artificially "filled" as discussed later in Sect. 13.7.2. Note that horizontal diffusion operators do not necessarily preserve the global volume integral of the diffused quantity. Most often, they are therefore not conservative as outlined in Staniforth et al. (2006).

In order to avoid over- and undershoots that are triggered by linear high-order diffusion Xue (2000) suggested a monotonic diffusion scheme with simple flux limiters. The basic idea is to interpret the generic form of the diffusion (13.2) as a flux

divergence such as

$$F_\psi = -\nabla \cdot \left[(-1)^q K_{2q} \nabla^{2q-1} \psi\right]. \tag{13.61}$$

The diffusive fluxes inside the outermost bracket can then be limited via flux limiters to prevent numerically triggered over- and undershoots as for example shown by Knievel et al. (2007) for a sixth-order scheme in the model WRF. The form (13.61) of the diffusion was also utilized for the model COSMO by Doms and Schättler (2002) who further improved the monotonic characteristics of the Xue (2000) scheme.

13.3.8 Diffusion Properties in Practice: The Model CAM 3.1

We now illustrate the practical aspects of the linear horizontal hyper-diffusion in GCM simulations with NCAR's CAM Eulerian spectral transform model that is utilized in an aqua-planet mode (Neale and Hoskins 2000) and with the idealized baroclinic wave test case by Jablonowski and Williamson (2006a,b). In addition, we compare the EUL simulations to the CAM Finite Volume (FV) model to gain an appreciation for the subtle differences between explicit hyper-diffusion and inherent numerical dissipation. Selected aspects of the inherent numerical dissipation are discussed below in Sect. 13.6.

The left panel of Fig. 13.4 shows 250 hPa kinetic energy spectra from CAM 3.1 simulations with the Eulerian spectral dynamical core for a variety of spectral truncations. The spectra are calculated from aqua-planet simulations (Williamson 2008a) and except for the highest resolution, are averaged over 100 samples separated by 30 h. The spectrum for the highest resolution is averaged every 6 h for the last 3 days of a 10-day run which started from a lower resolution aqua-planet state. Such sampling is adequate to determine the spectra. As listed above

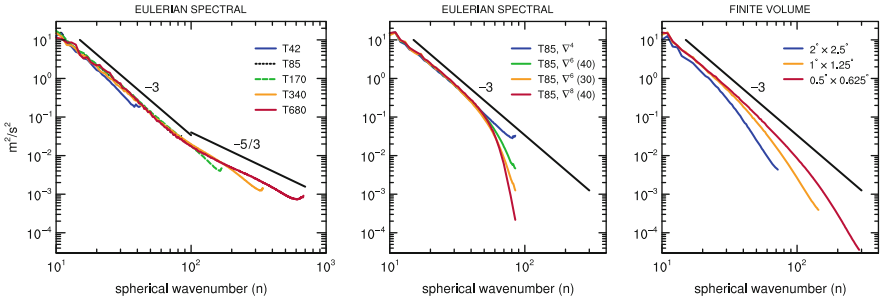

Fig. 13.4 250 hPa kinetic energy spectra as a function of the spherical wavenumber (n) in aqua-planet simulations from (*left*) CAM Eulerian spectral dynamical core with ∇^4 diffusion for different resolutions, (*center*) T85L26 Eulerian spectral dynamical with ∇^4, ∇^6 and ∇^8 diffusion, and (*right*) CAM Finite Volume (FV) dynamical core for different *lat* × *lon* resolutions in degrees and 26 levels

in Sect. 13.3.4.4, the diffusion coefficients used here are 1.0×10^{16}, 1.0×10^{15}, 1.5×10^{14}, 1.5×10^{13} and 1.0×10^{12} m^4 s^{-1} for T42, T85, T170, T340 and T680 truncations, respectively. The corresponding quadratic unaliased transform grids are approximately 2.8°, 1.4°, 0.7°, 0.35° and 0.17°, respectively. For resolutions below T340, the spectra all have a slope close to n^{-3}, and lie very close to each other. There is a small upturn at the truncation limits also found and discussed by Takahashi et al. (2006) and Hamilton et al. (2008). These spectra are also similar to those presented in Williamson (2008a). The T340 spectrum starts to deviate from n^{-3} around $n = 100$. This is the region where the observed atmospheric spectra transition from an n^{-3} slope to an $n^{-5/3}$ slope (Nastrom and Gage 1985). The model starts to make a similar transition but is then overwhelmed at the highest wavenumbers by the diffusion. Perhaps a smaller coefficient would allow the transition to form. The T680 simulation makes the transition and exhibits an $n^{-5/3}$ slope above $n = 100$ in agreement with spectra estimates from observations. The ∇^4 diffusion coefficient for T680 was chosen to yield this $n^{-5/3}$ slope. As an aside, at the 2008 Community Climate System Model (CCSM) Annual Workshop in Breckenridge, Colorado, it was reported that the spectral element model HOMME also exhibits a transition from n^{-3} to $n^{-5/3}$ (Taylor 2008). The transition of the slopes is furthermore evident in high-resolution simulations with the model NICAM when grid spacings below $\approx$ 10 km are employed (Terasaki et al. 2009).

It could be argued that the modeled spectra should follow n^{-3} or $n^{-5/3}$ slope to the truncation limit, depending on resolution as illustrated above, because that is what observations indicate they should do. It could also be argued that that is not a good discrete modeling strategy. For example, the University of Reading spectral model and the model ECHAM5 (Roeckner et al. 2003) have always used ∇^6 or ∇^8 forms of diffusion which lead to steeper kinetic energy spectra approaching the truncation limit. They argue that the scales near the truncation limit are not calculated accurately and cannot be trusted and that a discrete model should treat the end of the spectrum smoothly, including the transition to zero energy at the truncation limit (personal communication, Mike Blackburn). This requires a steeper spectrum which minimizes ringing or noise arising from a sudden transition. The ringing is a non-physical artifact arising from sharp truncation in the discrete system.

MacVean (1983) studied the effect of higher degree hyper-diffusions on baroclinic development in a spectral model truncated at T42. His simplest, but most subjective criterion, was a visual assessment of the degree to which synoptic scale detail was retained and the level of small-scale noise present. He also used other, more objective measures. He concluded that ∇^4 formulation is not scale-selective enough with T42 truncation, which, by the way, is still used today in climate models. The ∇^6 and ∇^8 forms, with appropriate coefficients, appear to be better and both are equally satisfactory.

The concept of applying stronger damping to the smaller scales than needed to maintain the observed spectra to the truncations limit is taken one step further by Lander and Hoskins (1997) with their introduction of the concept of "believable scales". They argue that since the shortest scales are not calculated accurately they should be filtered out before calculating the parameterizations, i.e.,

the parameterizations should be calculated on scales calculated accurately by the dynamics.

Higher degree hyper-diffusions such as ∇^6 or ∇^8 lead to steeper slopes approaching the truncation limit as seen in the center panel of Fig. 13.4 for the CAM EUL. The T85, ∇^4 line duplicates that in the left panel, the ∇^6 diffusion coefficients are chosen to damp wavenumber 40 or 30 at the same rate as the ∇^4 case, and the ∇^8 coefficient is chosen to damp wavenumber 40 the same rate as the ∇^4. The effect of diffusion on spectral ringing or noise in the solution is seen in Fig. 13.5 which shows the surface pressure from day 9 of calculations of the growing baroclinic wave test case of Jablonowski and Williamson (2006a,b). The top panel replots Fig. 13.3a, which was shown earlier for the default T85L26 CAM EUL simulation, but now we choose rather unconventional contour values which accentuate the noise. The second and third panels show T85L26 solutions with ∇^6 diffusion with two different diffusion coefficients. The coefficient for the second panel is chosen so that wavenumber 40 is damped at the same rate as in the ∇^4 calculations. The coefficient for the third panel is chosen so that wavenumber 30 is damped at the same rate as in the ∇^4 calculations. The noise is reduced but still present with ∇^6 diffusion. The larger coefficient (third panel) reduces the noise more than the smaller one (second panel). The noise is largely eliminated with ∇^8 diffusion (fourth panel). The coefficient is chosen so that wave 40 is damped at the same rate as it is in the ∇^4 case. Close examination of the figure shows that minimal noise is still present in the ∇^8 solution as indicated by the 999.98 hPa contour line in the left half of the plot. Quite possibly this noise could be reduced further with other choices for the coefficient values, but the whole process of minimizing the noise via diffusion operators is rather arbitrary and perhaps case dependent.

In contrast to the Eulerian spectral solutions, Fig. 13.5, bottom panel, shows that the solution from the CAM FV dynamical core in CAM 3.1 on a 1° latitude–longitude grid has no indication of noise in the surface pressure field. The contours are smoother than any of the other examples. The FV numerical approximations are shape preserving and thus do not generate small scale noisy structures. This is achieved via monotonicity constraints which are discussed later in Sect. 13.6.

Figure 13.4, right panel, shows the 250 hPa kinetic energy spectra from the CAM FV model for 2°, 1° and 0.5° grids. The spectra tail-off faster than in the Eulerian spectral model with ∇^4 diffusion. They behave more like the Eulerian spectral ∇^6 and ∇^8 diffusion cases, except the departure from an n^{-3} slope begins at lower wavenumbers relative to the truncation limit with the Finite Volume. These FV spectra are dominated by the rotational component of the flow, for which the numerical approximations are shape preserving. The divergent component is not similarly approximated and can become relatively large near the tail of the spectrum if no extra damping of the divergence is included. In these experiments the divergent component is controlled by a divergence damping mechanism (Neale et al. 2010; Whitehead et al. 2011) so that the divergent kinetic energy remains smaller than the rotational kinetic energy. A thorough discussion of horizontal divergence damping is provided in Sect. 13.4.1.

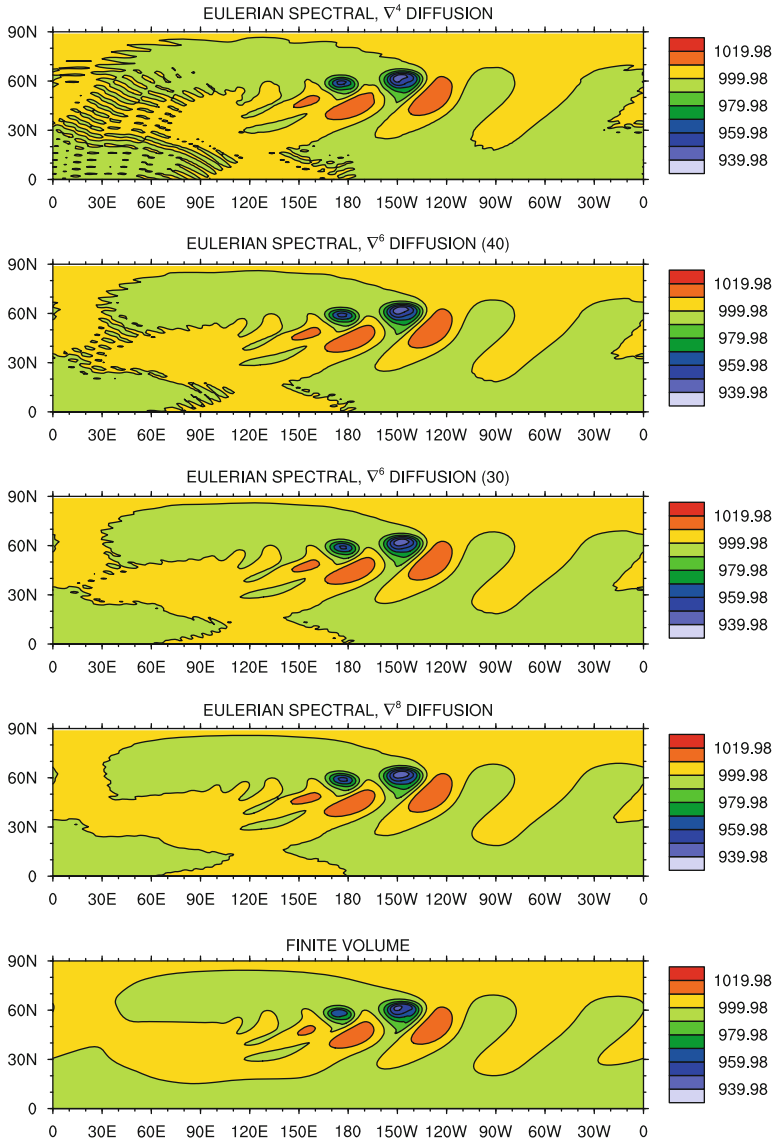

Fig. 13.5 Surface pressure (hPa) at day 9 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) from T85L26 CAM Eulerian spectral dynamical core with ∇^4, ∇^6 and ∇^8 diffusion, and $1^\circ \times 1.25^\circ$ CAM FV model. A time step of $\Delta t = 600$ s is used

The Eulerian spectral model can obtain smooth fields through the application of an arbitrary higher degree hyper-diffusion term. The FV model obtains them through the application of a physical condition, namely shape preservation. Shape preservation provides a more physically based method to obtain smooth solutions.

Perhaps the physical realism in the FV discrete system indicates that the spectrum should fall off faster than n^{-3} in the discrete system, and that it is a better discrete modeling strategy.

In Fig. 13.4 the 1° FV spectrum begins to depart from the n^{-3} slope around $n = 40$, although this evaluation is somewhat subjective. The 2° departs at a much lower wavenumber and generally has lower amplitude at the lower wavenumbers. The T85 spectral model with ∇^6 and ∇^8 diffusions also begins to depart from the n^{-3} slope around $n = 40$. The wavenumber of the departure from n^{-3} might imply something about the resolution of the models. For example, Skamarock (2004) argues that such a departure defines the *effective resolution* of a model (see also Chap. 14). This would imply that the CAM 1° FV and the T85 Eulerian spectral models with ∇^6 and ∇^8 diffusions have the same effective resolution. On the other hand, the T85 spectral model with ∇^4 diffusion maintains the n^{-3} slope to the truncation limit, yet most modelers would not argue that its effective resolution was T85. Even though the spectral method is accurate to the truncation limit for linear problems, the nonlinear interactions cannot be accurate there since they would involve unresolved scales. In addition the arbitrary diffusion is a dominant component in the equations at the smallest scales. Of more practical interest is the question of equivalent resolution of different schemes applied to a problem of interest. By examining a variety of simulated climate statistics in aqua-planet simulations Williamson (2008b) concluded that the T85 Eulerian spectral model with ∇^4 diffusion and 1° FV model reflect equivalent resolutions when applied to the aqua-planet problem. He also mentioned that experiments with the Eulerian spectral model with ∇^6 and ∇^8 diffusions produced results similar to those from the spectral Eulerian model with ∇^4 diffusion.

In summary, it could be argued that the modeled kinetic energy spectra should follow an n^{-3} or $n^{-5/3}$ slope to the truncation limit, depending on resolution, since that is what is observed in the atmosphere. However, this approach can lead to noise in the smallest scales of the solutions as seen in the Eulerian spectral model. It could also be argued that a drop off in the spectra as seen here in the FV model simulations indicates excessive damping by the numerics. Other modelers, however, argue that the modeled kinetic energy spectra should be steeper approaching the truncation limit so that the discrete, truncated spectra transitions to zero energy at the truncation limit more smoothly. In this interpretation shape preserving approximations such as those used in the FV model are not necessarily overly diffusive. They provide a physical condition (smoothness) which determines the shape of the spectra. This is compared to adding arbitrary higher degree hyper-diffusion terms to the equations which require subjective evaluations to determine the diffusion coefficient. The superiority of these two approaches remains a matter of discussion. In general, we recommend using kinetic energy spectra in combination with other quality measures, but not as the sole criterion, to judge the diffusive characteristics of GCMs.

13.4 Divergence and Vorticity Damping, External Mode Damping and Sponge Layers

Besides the very popular explicitly added horizontal diffusion and hyper-diffusion techniques discussed above, GCMs might also apply other explicit damping mechanisms. They include the 2D and 3D divergence damping, vorticity damping, an external mode damping approach or sponge layers near the model top. These are discussed in the next subsections.

13.4.1 2D Divergence Damping

Adding a horizontal divergence damping term to the horizontal momentum equations is a simple way of reducing high-frequency gravity wave noise. In this approach the time rates of change of the zonal and meridional velocities u and v are forced by a damping term. Recall the generic prognostic equation (13.1) for variable ψ where ψ now stands for the horizontal velocity vector $\mathbf{v} = (u, v)$ and $\mathbf{F}_\mathbf{v}$ symbolizes the vector of the horizontal divergence damping. The divergence damping mechanism of order $2q$ is then given by

$$\mathbf{F}_\mathbf{v} = (-1)^{q+1} \nabla \left(\nu_{2q} \, \nabla^{2q-1} \cdot \mathbf{v} \right) \tag{13.62}$$

where $q \geq 1$ is a positive integer value analogous to the discussion of the horizontal diffusion in Sect. 13.3. ν_{2q} stands for the divergence damping coefficient. Applying the horizontal divergence operator $\nabla \cdot$ to (13.1) and utilizing (13.62) yields an evolution equation for the divergence

$$\frac{\partial D}{\partial t} = \cdots + (-1)^{q+1} \nu_{2q} \nabla^{2q} \, D \tag{13.63}$$

where $D = \nabla \cdot \mathbf{v}$ denotes the horizontal divergence defined by

$$D = \frac{1}{a \cos\phi} \left(\frac{\partial u}{\partial \lambda} + \frac{\partial (v \cos\phi)}{\partial \phi} \right) \tag{13.64}$$

in spherical coordinates. As before, a symbolizes the Earth's radius and λ,ϕ are the longitude and latitude, respectively. Equation (13.63) assumes that the coefficient is constant in the horizontal direction. The equation demonstrates that the divergence damping represents a horizontal diffusion of the divergent part of the flow that is generally closely associated with inertia-gravity waves. Divergence damping can easily be explicitly added to models written in (u, v) form that do not utilize a prognostic equation for D.

As a concrete example, the second-order divergence damping mechanism (with $q = 1$) in component form yields

$$\frac{\partial u}{\partial t} = \cdots + \frac{1}{a\cos\phi}\frac{\partial}{\partial\lambda}\left(\nu_2 D\right) \tag{13.65}$$

$$\frac{\partial v}{\partial t} = \cdots + \frac{1}{a}\frac{\partial}{\partial\phi}\left(\nu_2 D\right). \tag{13.66}$$

This type of divergence damping was e.g., suggested by Shuman and Stackpole (1969), Sadourny (1975), Dey (1978), Haltiner and Williams (1980), Bates et al. (1990) and Lin (2004), mainly for numerical stability reasons. Note that Dey (1978) called it *divergence diffusion*. It leaves the rotational motion unaffected, selectively damps inertia-gravity waves, controls numerical noise and prevents the spurious build-up of energy near the cut-off grid scale. Sadourny and Maynard (1997) argued that 2D divergence damping can be viewed as a model of nonlinear interactions between inertia-gravity waves and rotational motion. The use of divergence damping was also explored by McDonald and Haugen (1992) and Gravel et al. (1993) to control gravity wave noise in their two-time-level semi-Lagrangian schemes.

The horizontal Laplacian-type ∇^{2q} hyper-diffusive term shown in (13.63) damps all scales, but is strongest at higher wavenumbers. Generally, divergence damping becomes more scale-selective at higher orders. This is e.g., shown by Whitehead et al. (2011) who also provide a linear von Neumann stability analysis of the divergence damping technique for explicit time-stepping schemes. In particular, the fourth-order divergence damping (with $q = 2$) can be expressed as

$$\frac{\partial u}{\partial t} = \cdots - \frac{1}{a\cos\phi}\frac{\partial}{\partial\lambda}\left(\nu_4 \nabla^2 D\right) \tag{13.67}$$

$$\frac{\partial v}{\partial t} = \cdots - \frac{1}{a}\frac{\partial}{\partial\phi}\left(\nu_4 \nabla^2 D\right) \tag{13.68}$$

where ν_4 is the fourth-order damping coefficient. This leads to the evolution equation for the divergence

$$\frac{\partial D}{\partial t} = \cdots - \nu_4 \nabla^4 D \tag{13.69}$$

in case of a horizontally constant coefficient. Fourth-order damping is an option in NCAR's model CAM 5 (Neale et al. 2010) which utilizes the FV dynamical core on a latitude–longitude grid. Even high-order divergence damping mechanisms like a sixth or eighth-order scheme are under investigation in the cubed-sphere version of the FV model (FVcubed) at the NOAA Geophysical Fluid Dynamics Laboratory (S.-J. Lin and William Putman, personal communication). FVcubed (Putman and Lin 2007, 2009) is also part of the most recent internal version of NASA's GEOS model which will be officially named GEOS6 upon its public release.

As pointed out earlier in Sect. 13.3 the explicit diffusion mechanisms for the horizontal divergence in spectral models (13.11) and (13.15) resemble the form of the divergence damping in (13.63) and (13.69). Again, both mechanisms are characterized with different names, but they accomplish a similar or even identical physical effect, namely they damp the divergent motions with either a second-order or higher order diffusion.

13.4.1.1 Selection of the 2D Divergence Damping Coefficient

As an example, we present the formulation of the damping coefficients in NCAR's CAM 5 model with the FV dynamical core (Lin 2004). The default second-order diffusion coefficient ν_2 in CAM is given by

$$\nu_2 = C_2(i) \frac{a^2 \Delta\lambda \Delta\phi}{\Delta t} \tag{13.70}$$

where the parameter $C_2(i)$ is a function of pressure to provide a sponge at the model top. Let p_{top} be the pressure at the model top and $[p_{ref}(i) = (a(i) + b(i))p_0]$ the reference pressure at a given model level with index i. Here, $a(i)$ and $b(i)$ are the hybrid coefficients of the vertical η coordinate (Simmons and Burridge 1981; Jablonowski and Williamson 2006b) and the surface pressure is assumed to be $p_0 =$ 1000 hPa. Then, as implemented in CAM, the parameter C_2 is given by

$$C_2(i) = \frac{1}{128} \max \left\{ 1, 8 \left[1 + \tanh \left(\ln \left(\frac{p_{top}}{p_{ref}(i)} \right) \right) \right] \right\} \tag{13.71}$$

which lets the divergence damping coefficient ν_2 increase by up to a factor of 8 close to the model top. Generally, the model top in CAM lies around $p_{top} \sim 3$ hPa or even at a lower pressure so that the troposphere and lower stratosphere are unaffected by this increase in the strength. C_2 is then constant with $C_2 = 1/128 \approx 0.0078$ at almost all levels except the uppermost two or three.

The formulation of ν_2 (13.70) is proportional to the area of a grid cell at the equator, and inversely proportional to the time step. Dimensionally this is an appropriate choice of the damping coefficient, but reliance on the area of the grid cell at the equator, and not the true area of the grid cell (with appropriate latitudinal dependence) places the same damping effect on a given physical wavelength, regardless of discretization or latitudinal location. Alternatively, a latitude-dependent coefficient could also be selected that takes the variation of the grid cell areas on a latitude–longitude grid into account

$$\nu_2 = C_2(i) \frac{a^2 \cos\phi \Delta\lambda \Delta\phi}{\Delta t}. \tag{13.72}$$

Such an area-dependent coefficient is selected for the optional fourth-order divergence damping mechanism in CAM 5 (Neale et al. 2010). It is given by

$$\nu_4 = C_4 \frac{a^4 \cos^2\phi \, (\Delta\lambda)^2 \, (\Delta\phi)^2}{\Delta t} \tag{13.73}$$

where the parameter C_4 is set to a default value of 0.01. The alternative formulation without the variation of the grid cell area yields

$$\nu_4 = C_4 \frac{a^4 (\Delta\lambda)^2 (\Delta\phi)^2}{\Delta t}. \tag{13.74}$$

Again, the stability aspects of both coefficients ν_2 and ν_4 are discussed in Whitehead et al. (2011) who determine upper limits for the parameters C_2 and C_4 in case of explicit time-stepping schemes.

Alternatively, NASA's FVcubed model on the almost uniform-resolution cubed-sphere grid defines the second-order and fourth-order divergence damping coefficients as

$$\nu_2 = C_2 \frac{A_{min}}{\Delta t} \tag{13.75}$$

$$\nu_4 = C_4 \frac{A_{min}^2}{\Delta t} \tag{13.76}$$

where $A_{min} \sim \Delta x_{min} \Delta y_{min}$ symbolizes the minimum grid area of the cubed sphere grid cells at a given resolution. C_2 and C_4 are often set to 0.005 and 0.0025, respectively, as documented in Lauritzen et al. (2010a). At coarser resolutions $\geq 1°$ these coefficients are typically lowered (William Putman, personal communication). Second- and fourth-order divergence damping can be used concurrently in the model FVcubed.

As the last point of comparison Bates et al. (1990) used the constant second-order damping coefficient $\nu_2 = 9 \times 10^7$ m^2 s^{-1} for their semi-implicit semi-Lagrangian shallow water simulations at the grid resolution $\Delta\lambda = \Delta\phi = 3.75°$ with both a $\Delta t = 600$ s and $\Delta t = 3{,}600$ s time step. According to (13.70) these correspond to the parameters $C_2 \approx 0.31$ and $C_2 \approx 1.86$, respectively, and are thereby considerably higher than the values in CAM FV. Such high values are unstable for explicit time-stepping schemes on latitude–longitude grids and necessitate an implicit treatment.

13.4.1.2 Example: The Effects of Divergence Damping

As a concrete example, we now illustrate the effects of the second-order and fourth-order divergence damping mechanisms in the Finite Volume dynamical core on the latitude–longitude grid. Both techniques can be selected in NCAR's model version CAM 5 at run time, with the second-order divergence damping technique being the default. In addition, we present FV simulations without divergence damping. We again utilize the growing baroclinic wave dynamical core test case by Jablonowski and Williamson (2006a) as depicted before in Sect. 13.3 (Figs. 13.3 and 13.5).

Figure 13.6 shows the surface pressure and 850 hPa meridional velocity field at day 9 for three CAM 5 FV simulations at the resolution $1° \times 1°$ with 26 levels. The only difference between the simulations is the selected divergence damping technique with their specific default coefficient. The upper row displays the simulation with the default second-order damping using the coefficient shown in (13.70)

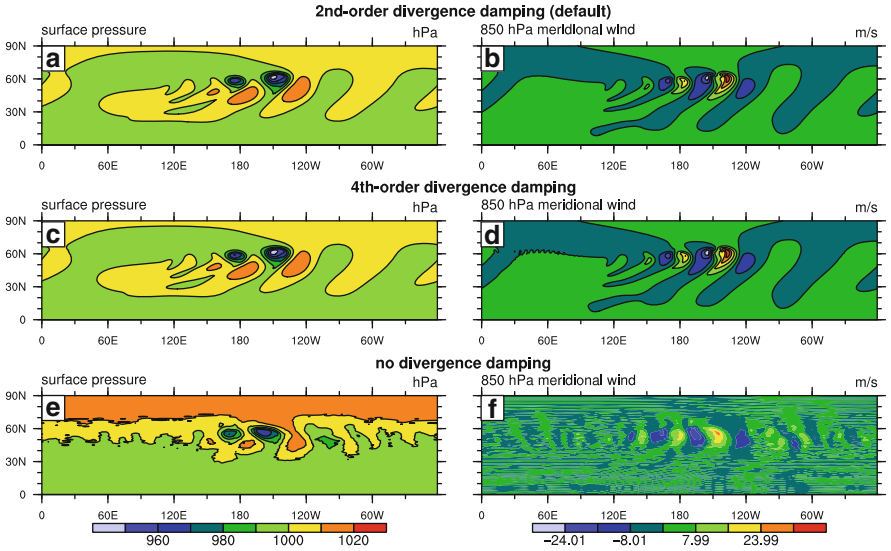

Fig. 13.6 Surface pressure (hPa) and 850 hPa meridional velocity (m/s) at day 9 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) from the CAM FV dynamical core at the resolution $1^\circ \times 1^\circ$ L26. (**a**,**b**): default second-order divergence damping, (**c**,**d**) fourth-order divergence damping, (**e**,**f**) no divergence damping. The unusual contour interval for the v wind emphasizes the very weak oscillations in (**d**). A dynamics time step of $\Delta t = 180$ s is used

with the default base value $C_2 = 1/128$, the middle row depicts the simulation with the optional fourth-order damping and the coefficient from (13.73) with the default $C_4 = 0.01$, and the bottom row reflects the simulation without divergence damping. Slightly unusual contour values for the meridional velocity are chosen to highlight the differences between the second-order and fourth-order divergence damping mechanisms. The figure shows that the simulation without divergence damping is corrupted by small-scale noise which suppresses the evolution of the baroclinic wave. This solution has little resemblance with Fig. 13.6a–d or additional high-resolution reference solutions from other models (Jablonowski and Williamson 2006a).

The differences between the simulation with second-order and fourth-order divergence damping are more subtle. As argued in Whitehead et al. (2011) fourth-order divergence damping is more scale-selective and introduces very strong damping near the grid scale (2-4 Δx) whereas longer scales are damped slightly less in comparison to the second-order damping scheme. The exact break-even point of the damping and the corresponding wavelength depends on the grid aspect ratio $\Delta\lambda/\Delta\phi$, the damping coefficients and the latitudinal position ϕ. However, it lies around 4-5 Δx at 60° in the current simulation. Presumably, this is the reason why there are still some very minor oscillations in the otherwise smooth 850 hPa meridional wind field in Fig. 13.6d. These oscillations have an approximate wavelength of 4-5 Δx. They are not obvious though in the surface pressure field or if the

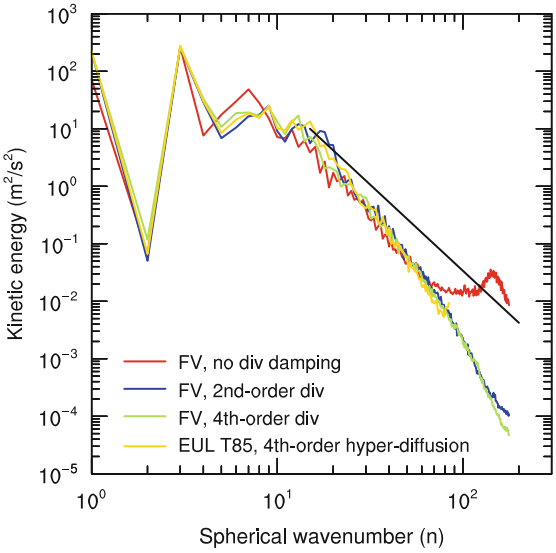

Fig. 13.7 700 hPa kinetic energy spectra at day 30 as a function of the spherical wavenumber n using the baroclinic wave test case of Jablonowski and Williamson (2006a). The spectra of the CAM FV model at the resolution $1^\circ \times 1^\circ$ L26 without divergence damping, the default second-order divergence damping and the fourth-order divergence damping are depicted. The dynamics time step is $\Delta t = 180$ s. A CAM Eulerian T85L26 run (with $\Delta t = 600$ s) with fourth-order hyper-diffusion is shown for comparison. The *black line* indicates an n^{-3} slope

contour values are integer multiples of 8 m s^{-1} (not shown). These oscillations do not grow over time. The visual comparison of the surface pressure fields suggests that the fourth-order divergence damping provides indeed less damping since the low pressure systems have deepened slightly more as indicated by the contour lines.

A more quantitative comparison of the divergence damping is provided in Fig. 13.7. The figure shows the 700 hPa kinetic energy spectra at day 30 of the simulations. The slightly rugged tails of the spectra could be smoothed via time-averaging, but is not of importance here. The kinetic energy spectra present a single snapshot in time. The figure displays that there is insufficient damping near the tail of the kinetic energy spectrum in the simulation without divergence damping. The upturn in the spectrum is a sign of numerical grid-scale noise and small-scale gravity waves which are connected to unbalanced (divergent) ageostrophic motions (O'Sullivan and Dunkerton 1995). The presence of too much energy in the divergent part of the spectrum is confirmed in the right figure of Fig. 14.6 in Chap. 14. The hook in the spectrum in Fig. 14.6 is solely triggered by the divergent motions (dotted red line) in this CAM FV aqua-planet run without divergence damping. In our example in Fig. 13.7 the divergent part of the spectrum causes a similar build-up of energy at the smallest spatial scales which is diffused by the divergence damping in the two additional FV simulations. The two FV simulations with divergence damping almost overlay each other, but show small differences in the steepness of the

kinetic energy spectra at the very end of the tail. It again suggests that fourth-order divergence damping damps the shortest scales more aggressively than second-order damping since the tail falls off more quickly. As a point of comparison, the figure also depicts the kinetic energy spectrum of the CAM Eulerian T85L26 simulation with default fourth-order hyper-diffusion (see Sect. 13.3.4.4). The EUL curve is very similar to the Finite Volume simulations. Recall that CAM EUL applies fourth-order hyper-diffusion to both the divergence and vorticity fields.

Similar results were also found by Becker and Burkhardt (2007) who observed a hook at the end of the kinetic energy spectrum in a simple general circulation model. The hook was eliminated via hyper-diffusion of the horizontal divergence in their spectral transform model. Note that the spectra in Fig. 13.7 fall off faster than the spectra shown earlier in Fig. 13.4. This is mainly due to the nature of the circulations in the different test cases and the fact that they are plotted at different pressure levels. The low 700 hPa level is selected here since the idealized baroclinic wave simulation is most pronounced at lower levels in this deterministic dynamical core assessment.

Divergence damping provides the major source of the diffusion for the divergent part of the motion in CAM FV. The rotational motions are damped via inherent numerical diffusion as shown later (see also Lin (2004)). The divergence damping stabilizes the FV dynamical core by smoothing the small-scale noise and preventing the hook in the tail of the kinetic energy spectrum. However, there are other physical effects that need to be considered. For example, it has been observed that divergence damping impacts the precipitation field in aqua-planet simulations at very coarse resolutions (Peter H. Lauritzen, personal communication). This is depicted in Fig. 13.8 that shows the annual average of the frequency distribution for tropical precipitation between 10°S and 10°N. The yellow curve represents a CAM 3.5 FV simulation with standard second-order divergence damping (13.70) at the coarse $2.7° \times 3.3°$ L26 resolution, the blue curve denotes an identical FV simulation, but with a doubled divergence damping coefficient. The red curve depicts a CAM EUL aqua-planet simulation at the resolution T31L26. The variation of the divergence damping coefficient has profound impact on the likelihood of heavy precipitation events in FV. The increase in the diffusion (blue curve) leads to a sharp decrease in the likelihood of heavy precipitation events and lets the FV simulation resemble the CAM EUL run. However, there is no "true solution" so the physical realism cannot be judged from these experiments alone. The figure only demonstrates the high sensitivity of the precipitation to the choice of the coefficient. The plots can also be compared to similar figures shown in Williamson (2008a,b).

13.4.1.3 2D Divergence Damping: Avoiding Confusion

As a word of caution, the spectral Eulerian dynamical core in NCAR's CAM model is formulated in vorticity-divergence form (Collins et al. 2004) and defines a Rayleigh friction technique (see Sect. 13.4.5.1 below) with the term *divergence damping*. However, the two mechanisms are very different. CAM EUL's definition is

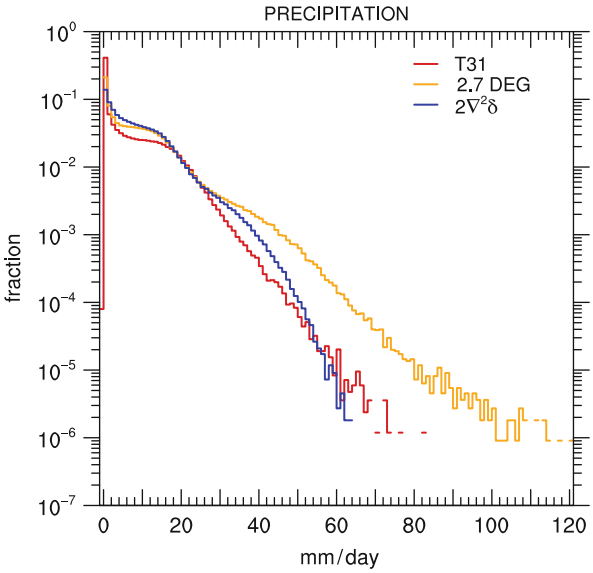

Fig. 13.8 Fraction of the time the tropical precipitation is in 1 mm day^{-1} bins ranging from 0 to 120 mm day^{-1}, calculated from 6-h averages for all grid points between $\pm 10°$. This frequency distribution is an annual average. The aqua-planet simulations are (*blue*, *yellow*) CAM FV at the coarse *lat* × *lon* resolution 2.7° × 3.3° L26 and (*red*) CAM EUL at the resolution T31L26 (with time step $\Delta t = 1{,}800$ s). *Yellow FV curve*: standard second-order divergence damping (13.70). *Blue curve*: FV simulation with a doubled coefficient. The figure is courtesy of Peter H. Lauritzen, NCAR

$$\frac{\partial D}{\partial t} = -rD \tag{13.77}$$

where r symbolizes an inverse damping time scale like $1/T$. The damping time scale determines the strength of the friction and is user-defined. In particular, the damping has an initial e-folding time of T and linearly decreases to zero over a time period of T_d, usually set to 2 days. It yields

$$r = \max\left[\frac{1}{T}\frac{T_d - t}{T_d}, 0\right] \tag{13.78}$$

where t stands for the elapsed time after the start of the model. In the CAM Eulerian or semi-Lagrangian dynamical core the damping is computed implicitly in spectral space via time splitting after the horizontal diffusion. If activated by the user it is only applied at the beginning of a model climate simulation to damp the gravity wave propagation arising from poorly balanced initial states. They usually result from interpolating a model simulated state to a different resolution with no attempt to maintain geostrophic balance. The initial behavior of a climate simulation is generally of no interest. This damping should never be used for short-term forecasts

when the initial behavior is of interest. After day T_d this damping mechanism is no longer active. Dey (1978) uses the phrase *divergence dissipation* for this type of damping.

13.4.2 3D Divergence Damping (or Acoustic Mode Filtering)

3D divergence damping is a smoothing mechanism for nonhydrostatic models and is, from a design perspective, very similar to the 2D divergence damping presented above in Sect. 13.4.1. However, there is a principal difference concerning its impact on the circulation. 2D divergence damping mainly affects internal gravity waves, whereas gravity waves are not noticeably impacted by 3D divergence damping. This is due to the fact that their velocity fields are almost non-divergent in three dimensions.

3D divergence damping serves two purposes. First, it is an effective damping mechanism for acoustic modes in nonhydrostatic models. Second, it eliminates spurious high-wavenumber modes caused by the instabilities in a partially-split (split-explicit) time-stepping scheme (Tatsumi 1983; Skamarock and Klemp 1992). Time-split schemes separate the terms in the equations of motion into slow and fast processes, and integrate them with large and multiple small timesteps, respectively. This technique is sometimes used to increase the computational efficiency of mesoscale models since fast, but meteorologically less important, sound waves can then be treated with a lower order and cheaper numerical scheme. Partially split numerical schemes are for example used in the models WRF (Skamarock et al. 2008) and COSMO (Doms and Schättler 2002; Gassmann and Herzog 2007; Baldauf 2010) and both models apply a 3D divergence damper to stabilize the schemes. Skamarock and Klemp (1992) showed that 3D divergence damping only caused very minor reductions in the amplitudes of gravity waves whereas it effectively damped both acoustic waves and the spurious noise associated with the time-split discretization. They also used the phrase *acoustic mode filtering* to describe 3D divergence damping. As an aside, acoustic mode filtering can also be accomplished by forward biasing an implicit time-stepping scheme of Crank-Nicolson or trapezoidal type (see textbooks like Durran (1999, 2010) or Kalnay (2003). This is briefly discussed in Sect. 13.6.3 that characterizes such an off-centering approach as inherent numerical diffusion.

The second-order 3D divergence damping formulation in vector form is given by

$$\mathbf{F_v} = \nabla_{2/3}\left(\nu\, \nabla_3 \cdot \mathbf{v}\right) \tag{13.79}$$

where ν is the divergence damping coefficient, $\mathbf{v} = (u, v, w)$ is the 3D velocity vector, $\nabla_{2/3}$ symbolizes either the two- or three-dimensional gradient operator and $(\nabla_3 \cdot)$ denotes the 3D divergence operator along constant coordinate surfaces. The term $\mathbf{F_v}$ can be appended exclusively to the 2D horizontal momentum equations as in Dudhia (1993), Doms and Schättler (2002) and Skamarock et al. (2008), or

to the 3D momentum equations as in Skamarock and Klemp (1992), Xue et al. (2000), Tomita and Satoh (2004) and Gassmann and Herzog (2007). Adding (13.79) to the momentum equations effectively introduces a second-order diffusion of the 3D divergence. This is highly specific to filtering acoustic modes.

The analysis in Gassmann and Herzog (2007) showed that an isotropic (constant ν) application to all three momentum equations should be the preferred choice. However, if the horizontal grid spacing is much larger than the vertical spacing as in typical GCMs, it is not possible to use the same value of ν in all directions. This was argued in Tomita and Satoh (2004) who also give guidance on the choice of the divergence damping coefficient for the model NICAM. In particular, Tomita and Satoh (2004) selected

$$\nu_x = \alpha'_x \, c_{s0}^2 \, \Delta\tau \tag{13.80}$$

where x serves a placeholder for either the horizontal (H) or vertical (V) direction, c_{s0} is the speed of sound at a temperature of $T = 273$ K, and $\Delta\tau$ is the length of the small time step in their split-explicit time-stepping scheme ($\sim \Delta t/4$). Typically, horizontal values around $\alpha'_H \in [0.05, 0.2]$ were chosen for horizontal grid spacings between 120 km and 240 km, and α'_V was set to zero (nonisotropic case) in their selected dynamical core experiments.

13.4.3 Vorticity Damping

A second-order vorticity damping formulation was suggested by Shuman (1969) and McPherson and Stackpole (1973) for models written in momentum (u, v) form. It is represented by

$$\frac{\partial u}{\partial t} = \cdots - \nu_\zeta \, \frac{1}{a} \frac{\partial \zeta}{\partial \phi} \tag{13.81}$$

$$\frac{\partial v}{\partial t} = \cdots + \nu_\zeta \, \frac{1}{a \cos\phi} \frac{\partial \zeta}{\partial \lambda} \tag{13.82}$$

where ν_ζ symbolizes the vorticity damping coefficient. In spherical coordinates the relative vorticity ζ is defined as

$$\zeta = \frac{1}{a \cos\phi} \left(\frac{\partial v}{\partial \lambda} - \frac{\partial (u \cos\phi)}{\partial \phi} \right) \tag{13.83}$$

and expresses the vertical component of the 3D vorticity vector. Formulating the evolution equation for the vorticity based on (13.81) and (13.82) yields

$$\frac{\partial \zeta}{\partial t} = \cdots + \nu_\zeta \, \nabla^2 \zeta. \tag{13.84}$$

It shows that the vorticity damping selectively diffuses the rotational part of the motion. Divergent motions remain unaffected.

An interesting analogy can again be drawn between the second-order vorticity damping presented here and the vorticity diffusion described earlier in the context of spectral transform models (13.10) and (13.14). Despite the different names, the physical effects of both damping mechanisms are similar. In practice though, none of today's GCMs in (u, v) form use such a vorticity damping. They employ alternatives such as the hyper-diffusion of u and v in the model GME (Majewski et al. 2002) that damps both the rotational and divergent motions. Other alternatives include the use of inherent numerical dissipation. For example, the model FV (Lin 2004) damps rotational motions via monotonicity constraints that are built into in its finite volume scheme (see also Sect. 13.6.2).

13.4.4 External Mode Damping

Noise in a numerical model can also manifest itself in form of pressure oscillations that are almost independent of the vertical level. These can be identified as the Lamb wave that is also called the "external inertia-gravity wave" mode. Lamb waves are fast moving horizontal acoustic modes with imaginary vertical wavenumbers that do not propagate in the vertical (and are therefore described as *external*). As shown in Kalnay (2003) Lamb waves are equivalent to gravity waves in a shallow water model.

Lamb waves are associated with fluctuations of the mean divergence in an atmospheric column. Recall that the change of pressure at a point is determined by the vertical integral of the mass divergence at this location. Therefore, damping the mass-weighted vertical integral of the divergence controls spurious pressure oscillations. Washington and Baumhefner (1975) explored this type of damping mechanism for model initialization purposes. In particular, they *Lamb filtered* the initial velocity data by modifying the divergence and successfully suppressed the external mode and high-speed oscillations. This connection was also pointed out in Dey (1978).

The so-called *external mode damping* for models in the native (u, v) momentum form is given by

$$\frac{\partial \mathbf{v}}{\partial t} = \cdots + K_e \nabla \left[\frac{1}{p_s - p_{top}} \int_{p_{top}}^{p_s} D \; dp \right] \tag{13.85}$$

where the term in the bracket is the mass-weighted vertical integral of the horizontal divergence D, and K_e is the damping coefficient. p_s stands for the surface pressure, and p_{top} is the pressure at the model top. Applying the horizontal divergence operator to (13.85) and integrating this equation again in the vertical then yields

$$\frac{\partial \overline{D}}{\partial t} = \cdots + K_e \nabla^2 \overline{D} \tag{13.86}$$

with the vertically integrated divergence $\overline{D}$. It shows that external mode damping indeed acts as a second-order diffusion mechanism and affects the vertically averaged divergence of the column.

All that is left is the definition of the external mode damping coefficient. Analogous to the discussion of the horizontal second-order diffusion it carries the physical dimensions $m^2\ s^{-1}$. Typically, it is defined by

$$K_e = \beta \frac{A}{\Delta t} = \beta \frac{\Delta x \Delta y}{\Delta t} \tag{13.87}$$

where the area A of a grid cell can be expressed by the physical grid spacings Δx and Δy in the two horizontal directions. A similar form of the external mode filter was also shown in Klemp et al. (2007) who discussed its definition in the weather forecast model WFR (note that there is a minus sign missing in Klemp et al.'s definition of D_h after their (46)).

In the model WRF the dimensionless and positive coefficient β is typically set to $\beta = 0.01$ (Klemp et al. 2007). In NASA's FVcubed dynamical core on a cubed sphere grid (Putman and Lin 2007, 2009) β is often set to 0.02 (William Putman, personal communication). In addition, the area measure A for the individual grid cells is replaced by the minimum cell area A_{min} within the FVcubed model. Equation (13.87) therefore defines a constant coefficient for all grid points regardless of their actual size. For latitude–longitude grids this choice might need to be reconsidered. External mode damping is only rarely used in GCMs today.

13.4.5 Sponge Layer Mechanisms at the Model Top

Setting appropriate upper boundary conditions in atmospheric models has been proven difficult for many years. The choices include radiation boundary conditions that allow energy to radiate outward at some finite height, the choice of a zero vertical velocity at the model top, or absorbing boundary conditions that absorb most or all incoming energy (Rasch 1986). Radiation boundary condition are popular in research models as e.g., suggested by Klemp and Durran (1983) but they cannot easily be implemented in GCMs. Therefore, operational models generally do not apply a radiation boundary condition but impose the condition that the vertical velocity is zero at the top. However, the presence of such a rigid top can lead to spurious wave reflections and even trigger instabilities at the top. Extra diffusion is then often utilized near the model top to absorb the reflections and slow down the wind speeds. This is common practice in almost all GCMs.

The type of extra dissipation in these sponge layers varies widely though. For example, the models CAM EUL, CAM SLD, ECHAM5 and GME switch from a

linear hyper-diffusion to a second-order diffusion which is applied in a few (around three) levels near the top. Sometimes the diffusion coefficient also increases upward. The model FV increases the strength of the divergence damping (see Sect. 13.4.1) and furthermore, utilizes a lower order numerical scheme to provide inherent diffusion as explained later in Sect. 13.6.1. The ECMWF model IFS (until cycle Cy35r3, September 2009), the spectral transform model AFES (Enomoto et al. 2008), NASA's ModelE (Schmidt et al. 2006) developed at the Goddard Institute for Space Studies and the model COSMO generally apply Rayleigh friction to the horizontal wind field. Rayleigh friction is also optional in the model WRF (Skamarock et al. 2008). These sponge layer mechanisms are outlined in more detail below.

All sponge layer mechanisms have one feature in common. As pointed out by Rasch (1986) sponge layers need to be thick enough and have adequate resolution to capture the waves reasonably well that they are supposed to damp. Sponge layers should ideally resolve the vertical wavelength with 8–10 vertical levels in order to damp waves effectively. They also need to guarantee a smooth transition region since a sudden onset of a strong sponge can cause wave reflections by itself. Sponge layers are simple to use, but can become computationally expensive in case extra vertical levels are needed for the sole purpose of providing a damping layer. An analysis of the properties of some dissipative sponge layer mechanisms can be found in Klemp and Lilly (1978).

13.4.5.1 Rayleigh Friction

A Rayleigh friction sponge is based on a linear relaxation which can be appended to the prognostic equations in the generic form

$$\frac{\partial \psi}{\partial t} = \cdots - k_R \, (\psi - \psi_r). \tag{13.88}$$

Here, k_R symbolizes a possibly spatially varying Rayleigh friction coefficient that expresses the inverse of a friction time scale, ψ is a placeholder for the velocities u or v, and ψ_r is a prescribed reference profile that might vary in space and time. Most often, the reference profile for the wind components u_r and v_r is set to zero. If "Rayleigh friction" is applied to the temperature field, it is called a *temperature relaxation* or *Newtonian heating or cooling*. Such a temperature relaxation always utilizes the reference profile since a temperature relaxation towards zero would be overly strong and unrealistic. This type of Rayleigh forcing was for example suggested by Held and Suarez (1994) and Boer and Denis (1997) for idealized dynamical core assessments.

In practice, Rayleigh friction might be used for two reasons at the upper levels in GCMs. First, it is considered a very crude approximation for gravity wave breaking in models with a very high model top in the upper stratosphere or mesosphere (Boville 1986; Boville and Randel 1986). Basically, it then replaces missing physics

mechanisms that would otherwise capture the momentum deposition by breaking gravity waves that travel upwards from the troposphere. These gravity waves induce a force on the large-scale flow, but note that Rayleigh friction is too simplistic to drive the observed reversal of the jets between the upper mesosphere and the lower thermosphere, or the tropical zonal wind oscillations in the stratosphere. Second, Rayleigh friction might be applied as a pure numerical sponge to prevent wave reflections at the model top. However, this requires some care in the formulation of the Rayleigh damping coefficient since as argued above, a sudden onset of a strong Rayleigh friction can also act as a reflector. Most vertical profiles of the Rayleigh damping coefficient therefore provide a smooth transition with increasing strength towards the model top. Typical damping time scales are 50 days or longer in the middle stratosphere and 1–2 days at the model top in the mesosphere.

We now present several damping profiles which have been suggested in the literature. A commonly used profile for the damping coefficient k_R is

$$k_R = \frac{1}{\tau}\left[1 + \tanh\left(\frac{z - z_1}{h}\right)\right] \tag{13.89}$$

where the height z is given by

$$z = H \ln\left(\frac{p_0}{p}\right). \tag{13.90}$$

$H = R_d \overline{T} g^{-1}$ symbolizes a constant scale height of the atmosphere with the gas constant for dry air R_d, a constant temperature $\overline{T}$ and the gravity g. p_0 is a reference pressure set to 1,000 hPa, p denotes the pressure at the model level, τ is a damping time scale, z_1 presents the approximate height of the model top, and h is a scaling parameter with height units. Such a profile has for example been defined in Boville and Randel (1986) and was the default in the ECMWF model IFS until September 2009 (Orr and Wedi 2009; Orr et al. 2010). Boville and Randel (1986) suggested the parameters $\tau = 3$ days, $z_1 = 63$ km and $h = 7.5$ km for a middle atmosphere GCM with a model top around 63 km. The model IFS (cycle 18R3, November 1997) set the parameters to $\tau = 3$ days, $z_1 = 61$ km and $h = 7.7$ km for a 50-level version with a model top at 0.1 hPa (about 61 km). IFS only applied the Rayleigh friction to the zonal wind field at model levels above 10 hPa. We refer to these IFS settings as "strong Rayleigh friction" in the examples below. Alternatively, Boville and Baumhefner (1990) used $\tau = 3$ days, $z_1 = 75$ km and $h = 7.5$ km which we characterize as "weak Rayleigh friction" (note that there is a sign error in their original definition). These latter two profiles of the Rayleigh damping coefficient are shown in Fig. 13.9. The figure also depicts the additional profile

$$k_R = \begin{cases} 0 & \text{if } p \geq p_{sponge} \\ k_{max}\left[(p_{sponge} - p)/p_{sponge}\right]^2 & \text{if } p < p_{sponge} \end{cases} \tag{13.91}$$

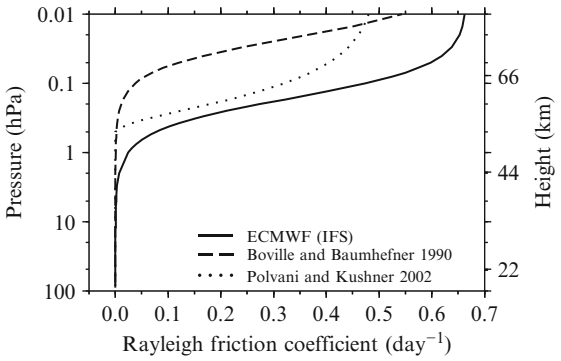

Fig. 13.9 Vertical profiles of three Rayleigh friction coefficients k_R in units day^{-1}

as suggested by Polvani and Kushner (2002) with the parameters $k_{max} = 0.5$ day^{-1} and $p_{sponge} = 0.5$ hPa. The onset of this Rayleigh friction at 0.5 hPa is more sudden. Rayleigh friction is sometimes only added to the zonal momentum equation. If the numerical stability of this damping process is a concern, it can easily be applied implicitly via a time-splitting approach.

As an aside, some meso-scale models such as WRF (Skamarock et al. 2008) offer optional ultra-strong Rayleigh friction sponges with profiles like

$$k_R = \begin{cases} 0 & \text{if } z < (z_{top} - z_d) \\ k_{max} \sin^2\left[\frac{\pi}{2}\left(1 - \frac{z_{top}-z}{z_d}\right)\right] & \text{if } z \geq (z_{top} - z_d) \end{cases} \tag{13.92}$$

where z_{top} denotes the height of the model top, z_d stands for the thickness of the damping layer as measured from the model top, and k_{max} is set to 0.2 s^{-1}. The latter corresponds to the damping time scale of $\tau = 5$ s at the model top which is very short in comparison to typical GCM settings of 1–2 days at upper levels. This damping is not used operationally in WRF. However, in idealized mountain wave test cases it has been found that about a third of the vertical domain must be classified as a sponge layer to suppress gravity wave reflections. Doms and Schättler (2002) also suggested using one third of the total domain height or at least one vertical wavelength as a sponge in the regional nonhydrostatic weather forecast model COSMO. They picked the formulation

$$k_R = \begin{cases} 0 & \text{if } z < z_{damp} \\ k_{max}\left[1 - \cos\left(\pi\left(\frac{z-z_{damp}}{z_{top}-z_{damp}}\right)\right)\right] & \text{if } z \geq z_{damp} \end{cases} \tag{13.93}$$

where z_{damp} symbolizes the starting position of the damping layer. The default values in COSMO are $z_{damp} = 11$ km and a coefficient of $k_{max} = (20\Delta t)^{-1}$ where Δt denotes the model time step. Similar Rayleigh friction profiles for meso-scale mountain wave simulations are also presented in Durran and Klemp (1983).

We now illustrate the effects of "weak" and "strong" Rayleigh friction in dynamical core simulations with CAM SLD at the triangular truncation T63 (≈ 210 km)

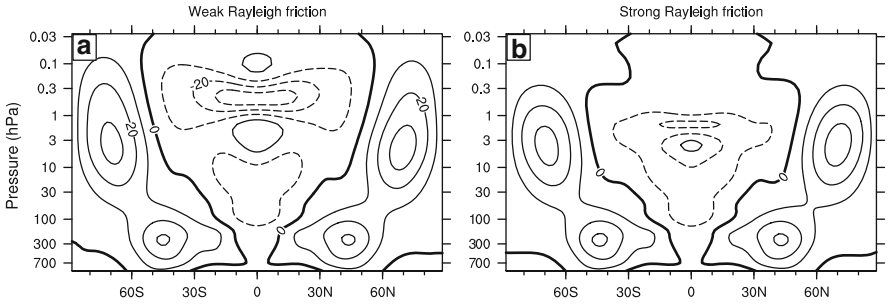

Fig. 13.10 Zonal-mean 1110-day-mean zonal wind in the CAM semi-Lagrangian dynamical core run with the Williamson et al. (1998) forcing at the resolution T63L55 with a model top at 0.015 hPa with (**a**) weak Rayleigh friction corresponding to the Boville and Baumhefner (1990) profile, (**b**) strong Rayleigh friction (ECMWF IFS). The friction is applied above 1 hPa in both simulations. The contour interval is 10 m/s. Negative contours are *dashed*. The time step is $\Delta t = 1{,}800$ s

with 55 vertical levels. The model top is placed at 0.015 hPa which lies around 75 km. The sponge-layer Rayleigh friction is only appended to the zonal momentum equation above 1 hPa and applied at every time step with $\Delta t = 1{,}800$ s. The experiments utilize the Held and Suarez (1994) forcing with the modification by Williamson et al. (1998) that provides strong meridional gradients in the stratospheric reference temperature field. This reference field is used for a temperature relaxation and causes zonal jets in the middle atmosphere according to the thermal wind relationship.

Figure 13.10 depicts the zonal-mean 1100-day-mean zonal wind fields of the two SLD simulations with (a) the weak Rayleigh friction according to Boville and Baumhefner (1990) and (b) the strong Rayleigh friction used in previous versions of IFS. Both model simulations were spun up with identical initial conditions and run for 1,440 days. The time average is over days 360–1,440. The figure shows that the flow field at the upper levels is very different in the two simulations. The "strong" Rayleigh friction almost damps out all motion above 1 hPa which is quite drastic. In contrast, the effect in the "weak" Rayleigh friction simulation is less pronounced in the region between 1 and 0.1 hPa. However, it is clearly visible near the model top where the wind speed has slowed down considerably. These figures can also be directly compared to a gentler diffusive sponge discussed in the following subsection (Fig. 13.11).

As a word of caution, Rayleigh drag changes the upper atmospheric angular momentum which should be balanced by a correction in the troposphere in order to conserve angular momentum. From a physical viewpoint, such a compensation mimics in some way the transfer of momentum by the unresolved gravity waves and is included in the Rayleigh friction implementation in NASA's ModelE (Schmidt et al. 2006). However, angular momentum conservation is not generally considered in GCMs. In case Rayleigh drag violates the angular momentum conservation it leads to a physically-spurious large-scale meridional circulation extending to the

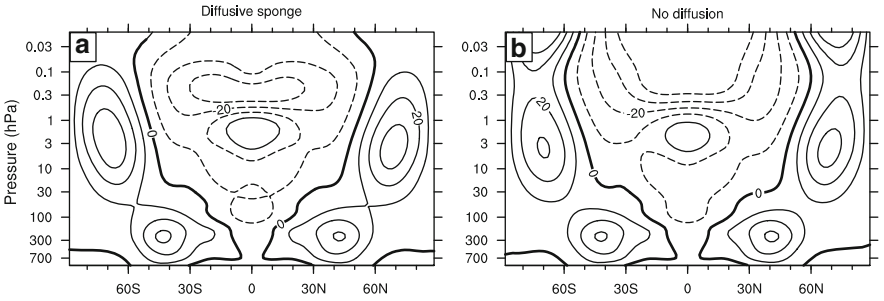

Fig. 13.11 Zonal-mean 1110-day-mean zonal wind in the CAM semi-Lagrangian dynamical core run with the Williamson et al. (1998) forcing at the resolution T63L55 with a model top at 0.015 hPa with (**a**) second-order diffusive sponge near the model top, (**b**) no ∇^2 and ∇^4 diffusion. The contour interval is 10 m/s. Negative contours are *dashed*. The time step is $\Delta t = 1{,}800$ s

surface of the earth as shown in Shepherd et al. (1996) and Shepherd and Shaw (2004). Therefore, Shepherd and Shaw (2004) recommended using Rayleigh friction very selectively, if at all, and only applying it to the wind perturbations from the mean state. A Rayleigh friction sponge is rarely used in operational GCMs today. Newtonian heating or cooling for tropospheric models is also often frowned upon because it forces the simulation toward a prescribed state rather than letting the state evolve freely.

13.4.5.2 Diffusive Sponges

The most popular sponge layer mechanism in GCMs is an increase in the horizontal diffusion, either via an increase in the diffusion coefficient, as e.g., discussed in Klemp and Lilly (1978), or a decrease in the order of the diffusion. The latter or even a combination of the two are most often chosen. Then a second-order diffusion replaces the usual hyper-diffusion in a few layers near the model top. For example, the model ECHAM5 decreases the order of the diffusion from a sixth-order to a fourth-order and finally to a second-order diffusion at the model top while using a constant time scale for all K_6, K_4 and K_2 diffusion coefficients (Roeckner et al. 2003). The effects of this diffusive sponge in ECHAM5 are clearly visible in the idealized dynamical core simulations by Wan et al. (2008) who utilized the Held and Suarez (1994) forcing. The sponge leads to a strong damping of the zonal wind field at upper levels in the equatorial region which can be compared to the simulations with the model CAM in Sect. 13.7.1 (Fig. 13.29).

A diffusive sponge is quite effective in reducing reflections and slowing down the wind speeds near the model top. The latter is illustrated in dynamical core simulations with CAM SLD at the triangular truncation T63 (≈ 210 km) with 55 vertical levels. As explained above in Sect. 13.4.5.1 the experiments utilize the Held-Suarez forcing with the modifications of the equilibrium temperature in the stratosphere according to Williamson et al. (1998). The model top is set to 0.015 hPa and no

Rayleigh friction is applied. Instead, we compare a simulation with a linear second-order sponge in the uppermost three layers (between 0.015 and 0.05 hPa) and very weak hyper-diffusion in the rest of the domain to a simulation without any (neither a K_2 nor K_4) explicit diffusion. In particular, the base diffusion coefficient for the sponge layer starting at the third level from the top is $K_2 = 2.5 \times 10^5\ \mathrm{m^2\ s^{-1}}$ which corresponds to a damping time scale of 11.2 h, and the hyper-diffusion parameter K_4 is set to $1 \times 10^{15}\ \mathrm{m^4\ s^{-1}}$ which corresponds to the long damping time scale of 28.1 h according to (13.20).

Figure 13.11 shows the zonal-mean 1100-day-mean zonal wind fields of the two SLD simulations with (a) the diffusive sponge and (b) neither ∇^2 nor ∇^4 diffusion. As before, both model simulations were spun up with identical initial conditions and run for 1440 days. The time average is over days 360–1440. The wind speeds near the model top are strongly reduced in the simulation with the sponge (Fig. 13.11a) whereas the winds almost reach their maxima at the model top in the model run without diffusion (Fig. 13.11b). This emphasizes the very strong influence of the additional sponge-layer dissipation on the circulation in the upper atmosphere. The impact of the sponge is not just confined to the top three layers but extends further down into the domain to about 1 hPa which incorporates 14 vertical levels. A similar effect was also shown earlier for a 26-level setup in Fig. 13.1. The impact of the ∇^4 diffusion at lower levels below the sponge is harder to evaluate since the stratospheric polar jets show some variability even in these long time averages. However, the impact of the ∇^4 diffusion seems to be minor in Fig. 13.11a. As a word of caution, if sponge layers are applied to model simulations the upper layers cannot be used for scientific evaluations. But unfortunately, it remains unclear whether the simulation without diffusion is any more reliable at upper levels due to the potential impact of artificial wave reflections. As seen before in the Rayleigh friction case, the sponge-layer dissipation dominates the upper level flow field. Recall that is has no physical foundation.

13.4.5.3 Vertical Velocity Damping

If Rayleigh damping is explicitly applied to the prognostic vertical velocity field in a nonhydrostatic model it is sometimes called *vertical velocity damping*. Rayleigh friction can also be implicitly applied within the implicit solution technique for vertically propagating acoustic modes. The latter has been found to be a very effective and robust absorbing sponge layer mechanism in nonhydrostatic models as suggested by Klemp et al. (2008).

As an example, the model WRF offers both an implicit Rayleigh damping of the vertical velocity and optional explicit vertical velocity damping to foster the robustness of the numerical scheme (Skamarock et al. 2008). Here, we briefly describe the explicit vertical velocity damping which is not just a sponge-layer technique. It generally damps the vertical motion whenever a violation of the CFL condition in the vertical direction is imminent, and thereby prevents the model from becoming unstable. The damping coefficient is locally determined and utilizes a critical

Courant number C_r defined by

$$C_r = \left| \frac{\dot{\eta} \Delta t}{\Delta \eta} \right| \tag{13.94}$$

where $\dot{\eta}$ is the vertical velocity in the generalized vertical coordinate η with grid spacing $\Delta \eta$. If C_r exceeds an activation Courant number C_a a Rayleigh friction is switched on. Assuming a forecast equation for the vertical velocity w, it yields

$$\frac{\partial w}{\partial t} = \cdots - sign(w)\ \gamma_w\ (C_r - C_a) \tag{13.95}$$

where γ_w is a damping coefficient and $sign(w)$ symbolizes the sign of w. The typical coefficients in WRF are $C_a = 1$ and $\gamma = 0.3\ \mathrm{m\ s^{-2}}$. This process does not possess a physical foundation and needs to be used with care. The regional nonhydrostatic weather forecast model COSMO (Doms and Schättler 2002) even includes a similar CFL-dependent Rayleigh friction as a Courant number limiter in the forecast equations for the horizontal wind speeds.

13.4.5.4 Courant Number Limiter

Occasionally in global GCM simulations the polar night jet becomes very strong at the top of the model and then the CFL stability condition is violated for the shortest longitudinal waves. Without further action the model would blow up. Rather than damp the jet speed further, some models simply remove the short waves that are unstable. This is generally only done at the top few levels of the model and only while the jet speed remains overly strong. The elimination is particularly simple in spectral transform models. If the maximum wind speed is sufficiently large, then the amplitudes of waves with wavenumber $n > n_c$ are set to zero, where the cutoff wave length is $n_c = a \Delta t / \max|V|$. a symbolizes the radius of the earth. This condition is applied whenever the maximum wind speed $\max|V|$ is large enough that n_c is less than the truncation limit and temporarily reduces the effective resolution of the model at the affected levels, but it does not affect the remaining scales.

To avoid adding code to sweep through spectral space, the dynamical core CAM EUL includes this process in the solution of the horizontal diffusion. Recall (13.12) that expresses the $2q$-th order temperature diffusion and for simplicity, let us focus the discussion on the diffusion that is applied along model levels. The time-discretized response function (13.28) for CAM EUL then becomes

$$E_n = \left\{ 1 + 2\Delta t\ D_n\ K_{2q} \left(\frac{n(n+1)}{a^2} \right)^q \right\}^{-1} \tag{13.96}$$

where the so-called "Courant number limiter" D_n has been added. The response function (13.96) can also be viewed as a wavenumber-dependent damping function

that is denoted by $K_n^{(2)}$ and $K_n^{(4)}$ in the technical CAM documentation (Collins et al. 2004).

Generally, the factor D_n is set to 1. However, the diffusion coefficient K_{2q} is increased by a factor of $D_n = 1000$ for those waves which are to be eliminated. This is possible because the diffusion is approximated implicitly in spectral space in the model. Since the diffusion is linear, the solution for each wavenumber is independent of all other wavenumbers. Note that this should not be thought of as increased overall diffusion since that is a process in physical space and affects all scales. The diffusion code is simply a convenient place to eliminate the selected waves. This Courant number limiter should be thought of as simply removing the shortest waves that would otherwise be unstable and thereby temporarily reducing the horizontal resolution. Therefore it is generally only applied near the model top where the solution is already contaminated by the sponge layer. Here, we only list the principle design of this limiter and refer to Collins et al. (2004) for the exact application in spectral space. Note this type of limiter can only be safely used if the diffusion is approximated implicitly as it is the case in CAM EUL. Otherwise the diffusion process will likely be unstable. The stability limitations for explicit time stepping schemes are discussed above in Sect. 13.3.5.

13.5 Explicit Filtering Techniques

Filtering is a fairly common smoothing technique in GCMs. Two types of filters need to be distinguished as explained below. The first category encompasses the temporal filters. The second type includes spatial filters such as digital grid point filters, or spectral filters like the Fast Fourier Transform (FFT) or Boyd-Vandeven filter. Spatial filters are popular in grid point models, especially on latitude–longitude grids at high latitudes, where they are often called *polar filters*. In addition, temporal digital filters are sometimes used as model initialization schemes to damp out high-frequency noise in analyzed data as discussed in Lynch and Huang (1992).

Spatial filters damp both linear and nonlinear computational instabilities. Linear instability arises if the CFL stability condition is violated by e.g., fast moving inertia-gravity waves which can easily occur on longitude-latitude grids with converging meridians near the poles. Filtering then allows the use of longer CFL-violating time steps which would otherwise grow unstable. Nonlinear computational instability is associated with quadratic terms in the equations of motion (Phillips 1959; Orszag 1971). In particular, products of waves can create new waves with wavelengths shorter than $2\Delta x$ as discussed in Durran (1999, 2010). Since these waves cannot be resolved on a model grid, they are aliased into longer wavelengths. This tends to accumulate energy at the smallest scales in a nonlinear model which can be removed by a filter to maintain computational stability. High-frequency noise might also be introduced by truncation or observational errors in the initial data. Whether filtering is necessary for stable computations depends on the characteristics of the numerical scheme and the choice of the model grid. In general, no attempt

is made to associate filters with physical processes. Filtering is an ad-hoc smoothing process in atmospheric models.

13.5.1 Time Filters

13.5.1.1 Robert-Asselin Filter

Time filters are commonly applied to multi-step time-stepping schemes, such as the three-time-level Adam-Bashforth scheme or the popular leapfrog method. These time-stepping schemes are e.g., described in Haltiner and Williams (1980), Durran (1999, 2010) or Kalnay (2003). In particular, the leapfrog scheme for the variable ψ is given by

$$\psi^{j+1} = \psi^{j-1} + 2\Delta t \; F(\psi^j) \tag{13.97}$$

where $j+1$, j and $j-1$ represent the future, current and previous time level of the three-time-level scheme, and $F(\psi^j)$ is the forcing by the dynamical and physical processes. A major problem with the leapfrog scheme is that every wavenumber is associated with two frequencies (Durran 1999, 2010). One is a physical mode, the second is an undamped computational mode that is associated with the handling of the even and odd time steps. It manifests itself as a spurious oscillation between even and odd time steps that amplifies during nonlinear simulations. It ultimately grows explosively and causes the model to blow up.

The decoupling of the solutions at odd and even time steps can be avoided when applying a recursive time filter at each time step as suggested by Robert (1966) and Asselin (1972). This filter is referred to as the Robert–Asselin, Robert or Asselin filter. It is defined by

$$\overline{\psi}^j = \psi^j + \alpha \left(\overline{\psi}^{j-1} - 2\psi^j + \psi^{j+1} \right) \tag{13.98}$$

where the overbar denotes the time-filtered variable, and the unitless positive coefficient α determines the strength of the filter. It leads to the Asselin-leapfrog scheme

$$\psi^{j+1} = \overline{\psi}^{j-1} + 2\Delta t \; F(\psi^j). \tag{13.99}$$

In general, the Asselin filter can be viewed as a second-order diffusion mechanism in time. It effectively damps the computational mode of the leapfrog time-stepping scheme, but unfortunately also affects the physical mode by slightly changing the phase and amplitude of the resolved waves. This reduces the formal order of accuracy of the leapfrog scheme to first order and can degrade the accuracy of model simulations. For example, physical quantities such as energy that are conserved by the time-continuous equations are not necessarily conserved by the time-discretized equations which is even true with or without the Asselin filter. The damping and non-conservation may be acceptable for short model simulations, but could become more prevalent and questionable in long GCM runs. The Asselin filter is generally

applied at each time step to all prognostic variables along model levels. The accuracy and stability aspects of the Asselin filter were investigated by Tandon (1987), Robert and Lépine (1997), Cordero and Staniforth (2004) and Williams (2009). In particular, Williams (2009) proposed a modification of the Asselin-filter that greatly reduces its negative impact on the physical mode and increases the numerical accuracy, but does so at the expense of a very slight instability.

A comprehensive list of atmospheric models with the Asselin filter is documented in Williams (2009). However, the strength of the damping coefficient α is highly model-dependent and rarely mentioned in the refereed literature. The typical values in GCMs vary between 0.02 and 0.2. More specifically, the spectral element dynamical core HOMME utilizes a default coefficient of $\alpha = 0.05$ as listed in Lauritzen et al. (2010a), the spectral element model by Giraldo and Rosmond (2004) uses $\alpha = 0.02$, the NCAR Eulerian spectral dynamical core in CAM 3.1 applies $\alpha = 0.06$ (Collins et al. 2004), Skamarock and Klemp (1992) and McDonald and Haugen (1992) set α to 0.1, the model ICON defines $\alpha = 0.1$, and the model GME specifies $\alpha = 0.15$ (Majewski et al. 2008). Schlesinger et al. (1983) even suggested using values in the range 0.25–0.3 for certain advection-diffusion problems. A systematic sensitivity analysis to the Asselin filter coefficient has been conducted by Rípodas et al. (2009) with the shallow-water version of the ICON model. They found that $\alpha \geq 0.05$ was required to keep their model stable. As an aside, there are other methods for controlling the computational mode of the leapfrog scheme, which do not involve the application of a time filter. These methods are briefly outlined in Williams (2009).

Figure 13.12 assesses the effect of the Asselin filter on a short dynamical core simulation with the CAM Eulerian spectral transform model at the triangular truncation T85 and 26 levels (L26). The figure shows the surface pressure and 850 hPa temperature field at day 9 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) with and without strong Asselin time filtering. The quantitative comparison suggests that the simulation of the baroclinic wave without Asselin filtering develops slightly stronger high and low pressure systems by day 9. There might also be a very minor shift in the position of the wave as indicated by the difference plots (Fig. 13.12e,f). However, the difference fields are mostly dominated by amplitude errors provided the unfiltered simulation is considered more accurate and closer to the high-resolution reference solutions shown in Jablonowski and Williamson (2006a,b). The computational mode in the unfiltered simulation is not obvious at day 9, but grows unstable by day 14. The application of the Asselin filter is therefore paramount. Recall that the default in CAM EUL is generally set to $\alpha = 0.06$ which provides significantly weaker filtering.

As a word of caution, Déqué and Cariolle (1986) stated that despite the demonstrated ability of the Asselin filter to stabilize numerical solutions to the equations of motion for certain combinations of temporal differencing and physical forcings, even a very weak Asselin filter may have the potential to trigger an instability. Déqué and Cariolle (1986) suppressed this instability by a severe reduction of the time step in their model runs. Some unexpected anomalous behavior of the time filter was also highlighted by Robert and Lépine (1997).

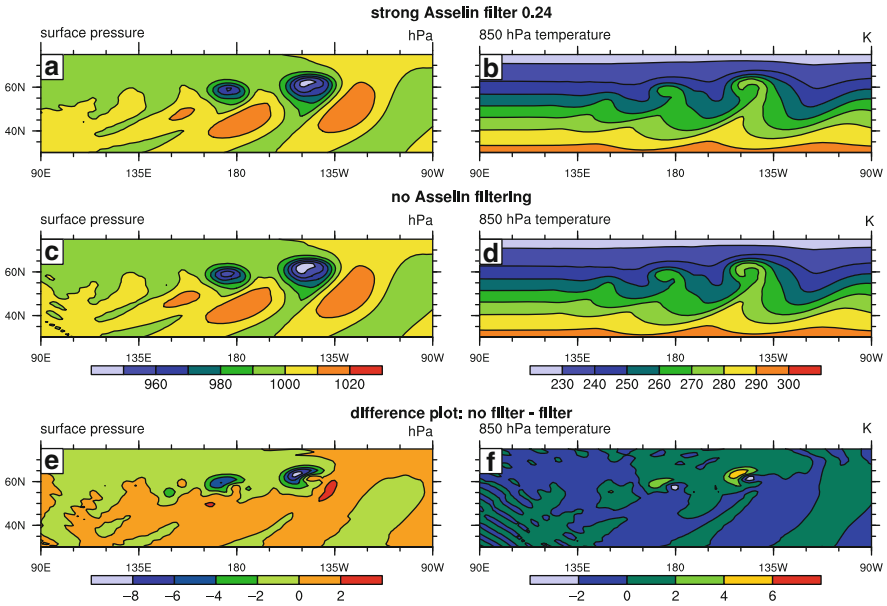

Fig. 13.12 Surface pressure (hPa) and 850 hPa temperature (K) at day 9 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) in the CAM T85L26 Eulerian spectral dynamical core with and without Asselin time filtering. (**a**,**b**) strong Asselin filter with $\alpha = 0.24$, (**c**,**d**): no Asselin filter, (**e**,**f**) difference between the unfiltered and filtered simulation. A time step of $\Delta t = 600$ s is used

13.5.1.2 Time Filter for Extrapolated Values

Semi-implicit semi-Lagrangian models need information about future parcel trajectories and therefore require information about the wind velocities at the future half-time level $t^{j+1/2}$ (Staniforth and Côté 1991). A popular method for estimating these time-centered wind speeds is time extrapolation. Two different time extrapolators are commonly used that either utilize two or three time levels. For the wind vector $\mathbf{v}$ the two-term extrapolator with times t^j and t^{j-1} yields

$$\mathbf{v}^{j+1/2} = \frac{3\mathbf{v}^j - \mathbf{v}^{j-1}}{2}. \tag{13.100}$$

which has been widely used in two-time level semi-implicit semi-Lagrangian schemes by e.g., Temperton and Staniforth (1987), McDonald and Haugen (1992, 1993) or ECMWF (2010). The three-term extrapolator is defined as

$$\mathbf{v}^{j+1/2} = \frac{15\mathbf{v}^j - 10\mathbf{v}^{j-1} + 3\mathbf{v}^{j-2}}{8} \tag{13.101}$$

which includes the additional time level t^{j-2}.

However, time extrapolation is a potential source of instability. For example, McDonald and Haugen (1992) showed that the noise introduced by the extrapolations in a σ-level model could be efficiently controlled by an *implicit divergence damper* that damps gravity waves. Here, the phrase *divergence damper* referred to an inherent numerical dissipation via a decentering scheme as explained later (Sect. 13.6.3). But the noise, that mainly originates from discretized nonlinear terms, was no longer controlled by decentering when switching to hybrid vertical coordinates (McDonald and Haugen 1993). Therefore, they introduced the following time filtered equations for both the extrapolated wind velocities and nonlinear terms

$$\psi^{j+1/2} = \frac{3\psi^j - \overline{\psi}^{j-1}}{2} \tag{13.102}$$

$$\psi^{j+1/2} = \frac{15\psi^j - 10\overline{\psi}^{j-1} + 3\overline{\psi}^{j-2}}{8}. \tag{13.103}$$

The time-filtered quantity is defined by

$$\overline{\psi}^j = \psi^j + \epsilon\left(\overline{\psi}^{j-1} - 2\psi^j + \psi^{j+1}\right) \tag{13.104}$$

where ψ serves as a placeholder variable. This time filter with the unitless and positive filter coefficient ϵ is formally the same as the Asselin filter (13.98). However, it is only selectively applied to nonlinear terms and the centering of the trajectory departure points in the semi-Lagrangian method. The linear terms are untouched which causes minimal decreases in accuracy according to McDonald and Haugen (1993). Their recommended ϵ values, in combination with a decentering scheme, ranged from $[0.05, 0.3]$.

13.5.2 Spatial Filters

Spatial filtering techniques have long been used for global atmospheric modeling. As mentioned above, spatial filtering suppresses linear and nonlinear instabilities, but conservation properties can get lost and might necessitate the use of a-posteriori restoration mechanisms (Takacs 1988). For example the conservation of mass gets lost if the mass variable needs to be filtered for numerical stability reasons. Similar difficulties emerge with respect to the conservation of total energy which is most often lost even without filtering. These aspects are assessed in Sect. 13.7. Here, we focus on the discussion of digital and spectral filters.

13.5.2.1 Digital Grid Point Filters

Digital filters are local grid point filters that only take neighboring grid points in the horizontal direction into account. They can be applied in one or two dimensions.

There are many digital filtering techniques that have been proposed in the literature. Examples are the filters by Shuman (1957), Shapiro (1970, 1971, 1975), Nelson and Weible (1980), Raymond and Garder (1988), Raymond (1988), Purser (1987) and Staniforth et al. (2006).

The most popular digital filter still used in some models today is the Shapiro filter which is based on constant-coefficient grid point operators of order n. The order determines the width of the numerical stencil. In general, the higher the order the higher the computational cost due to the wider stencil, which leads to more scale-selective filters. These minimize the damping of long and thereby flow-relevant waves.

In this section we focus on the one-dimensional Shapiro operator of order n to illustrate the general characteristics of digital filtering and show examples from shallow water simulations with a finite-volume model (Jablonowski 2004). The smoothing operation of the so-called *optimal* or *ideal* second-order ($n = 2$) Shapiro filter (Shapiro 1975) is defined by

$$\bar{\psi}_i = \frac{1}{16}\big(-\psi_{i-2} + 4\psi_{i-1} + 10\psi_i + 4\psi_{i+1} - \psi_{i+2}\big) \tag{13.105}$$

where the overbar symbolizes the smoothed variable ψ at grid point index i. The width of the stencil in one dimension is $2n + 1$. The coefficients for higher-order filters are listed in Shapiro (1975). As shown by Purser (1987) the Shapiro filter can also be generalized to describe a family of symmetric-stencil filters. This was first discussed in Hamming (1977) but has received very little attention in the atmospheric sciences.

The $n = 2$ Shapiro filter completely eliminates the unwanted $2\Delta x$ waves and significantly reduces the amplitudes of other poorly-resolved short waves, especially the $3, 4\Delta x$ waves that also tend to accumulate energy during model integrations. Each application of the Shapiro filter reduces the amplitude of a Fourier wave component $\exp(\imath k_x \Delta x)$ by the factor

$$\begin{aligned} \rho_n(k_x) &= 1 - sin^{2n}\Big(\frac{k_x\,\Delta x}{2}\Big) \\ &= 1 - sin^{2n}\Big(\frac{\pi\,\Delta x}{L}\Big) \end{aligned} \tag{13.106}$$

where $\imath = \sqrt{-1}$ is the imaginary unit, $k_x = 2\pi/L_x$ is the wavenumber in the x-direction and L_x is the wavelength, expressed as multiples of the grid spacing Δx. $\rho_n(k_x)$ is the response or damping function of the Shapiro filter (Shapiro 1971). In two dimensions the response function becomes

$$\rho_n(k_x, k_y) = \left[1 - sin^{2n}\Big(\frac{k_x\,\Delta x}{2}\Big)\right]\left[1 - sin^{2n}\Big(\frac{k_y\,\Delta y}{2}\Big)\right] \tag{13.107}$$

where k_y symbolizes the wavenumber in the y-direction with grid spacing Δy. This form expresses the response function for subsequent applications of the 1D filter operators.

Note that there is an inconsistency in the literature how to denote the order of the Shapiro filter. Here, we choose to follow the notation n as suggested in Shapiro (1971). Other modelers like Fox-Rabinovitz et al. (1997) or Ruge et al. (1995) use the notation $2n$ which corresponds to the order 4 in (13.105). The differences in the notations might have been motivated by Shapiro's observation that the one-dimensional ideal operator of order n is equivalent to the incorporation of a one-dimensional linear diffusion of order $2n$ with a coefficient $K = (-1)^{n-1}(\Delta x/2)^{2n}/\Delta t$. This draws an interesting analogy to the 1D linear diffusion mechanism. However, this result does not entirely generalize in two dimensions as discussed in Shapiro (1971). In 2D, the ideal n-th-order Shapiro operators only resemble the $2n$-th-order linear horizontal diffusion with the addition of a $4n$-th-order mixed damping term. This renders the Shapiro filter more scale selective than linear horizontal diffusion and fully eliminates $2\Delta x$ waves after each application.

The response of different 1D Shapiro operators (13.106) with respect to the wave spectrum is illustrated in Fig. 13.13. The figure shows the filter responses after one and 1,000 applications, and clearly depicts the cumulative character of the smoothing operation, especially for low-order filters. In practice, this is not a concern for overresolved waves close to the pole points on latitude–longitude grids, but must be taken into consideration in case filtering is to be applied at lower latitudes. For example, if filtering in midlatitudes or even in the tropics becomes necessary due to stability reasons, a low-order filter like the second-order Shapiro filter should be avoided and replaced by either a highly scale-selective FFT or higher order digital filters. A commonly used higher order filter is the eighth- or sixteenth-order Shapiro filter. An $n = 8$ Shapiro operator effectively eliminates all components

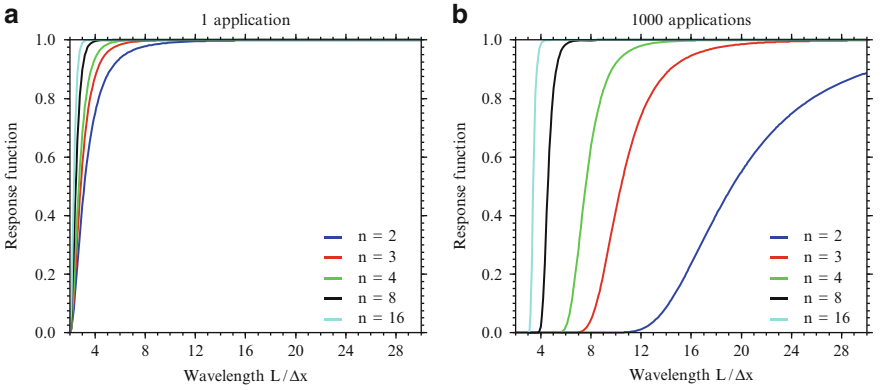

Fig. 13.13 Response function of different 1D Shapiro filters after (**a**) one application and (**b**) 1,000 applications. n indicates the order of the ideal Shapiro operator

with wavelengths less than four grid intervals but does not damp the waves with wavelengths greater than six grid intervals (Shapiro 1975). This filter has been successfully utilized by Ruge et al. (1995) who applied the $n = 8$ Shapiro filter every twelve time steps to eliminate the nonlinear computational instability in a shallow water model.

13.5.2.2 Examples: Applications of the Shapiro Filter

In the following shallow water examples, the Shapiro filter is solely applied in the longitudinal direction poleward of 60° N/S. The shallow water model is described in Jablonowski (2004) and applied at the constant resolution $2.5° \times 2.5°$. It utilizes the Lin and Rood (1997) finite-volume approach and is built upon a latitude–longitude grid with converging meridians near the pole points. The examples are chosen to briefly outline the advantages and disadvantages of the Shapiro filtering approach.

Filtering techniques must be carefully chosen. As mentioned before, strong low-order filters, like the second-, third- or fourth-order Shapiro filters, applied at low- or midlatitudes can significantly diffuse and degenerate the flow pattern and should therefore only be rarely used. An example that illustrates the detrimental effects of rather aggressive and unnecessary filtering is presented in Fig. 13.14. The example depicts a Rossby-Haurwitz wave which corresponds to test case 6 of the standard shallow water test suite proposed by Williamson et al. (1992). This flow field comprises a wavenumber four pattern that moves from west to east with only minor changes in shape. We assess two model runs that both apply an FFT filter to the horizontal wind components poleward of 75°N/S. Additionally, the strong second-order Shapiro filter is applied to the wind speeds at every time step in (a) the limited model domain 60°–75°N/S and (b) the whole remaining 75°S–75°N model area. Figure 13.14 shows the geopotential height field at model day 14. It can clearly be seen in Fig. 13.14b that the additional digital filtering in midlatitudes and the tropics leads to a very diffusive and inaccurate solution in comparison to the (a) less-filtered simulation which resembles reference solutions. The filter effects accumulate significantly during the course of the 3,360 time step simulation which confirms the cumulative effect in Fig. 13.13b. The errors are pure amplitude errors without changes in the phase speed of the wave.

The crucial need for a filtering mechanism in the polar 60°–75°N/S range in the finite-volume shallow water model is further depicted in Fig. 13.15 which displays a geostrophically balanced flow field in the Northern Hemisphere (test case 3 in Williamson et al. (1992)). The test results are shown for the geopotential height field after a 23-h simulation with and without the digital second-order Shapiro filtering technique. Again, an FFT filter is applied poleward of 75°N/S. Here, the chosen time step $\Delta t = 600$ s purposely violates the CFL condition for gravity waves in the polar region and as a consequence, a numerical instability equatorwards of 75°N develops in Fig. 13.15a. This position clearly marks the edge of the FFT filtering mechanism. In Fig. 13.15b the instability is removed by an $n = 2$ Shapiro filter. Of

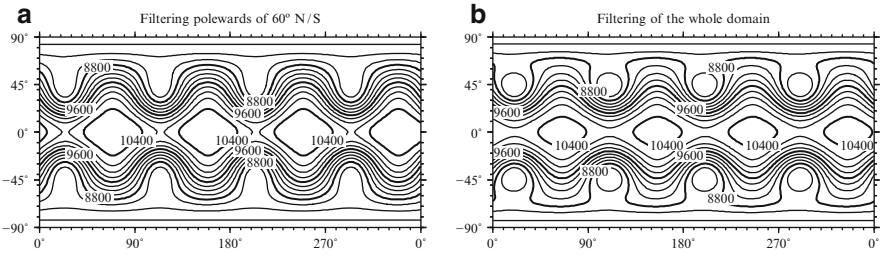

Fig. 13.14 Latitude–longitude plot of the geopotential height at day 14 for the Rossby–Haurwitz wave (test case 6) in the shallow water version of the FV model (Jablonowski 2004). (**a**) Second-order Shapiro filter is applied to the wind fields u and v between 60°–75°N/S. (**b**) Same filter is applied in the whole 75°S–75°N domain. Contour interval is 200 m. Resolution is 2.5° × 2.5° with $\Delta t = 360$ s

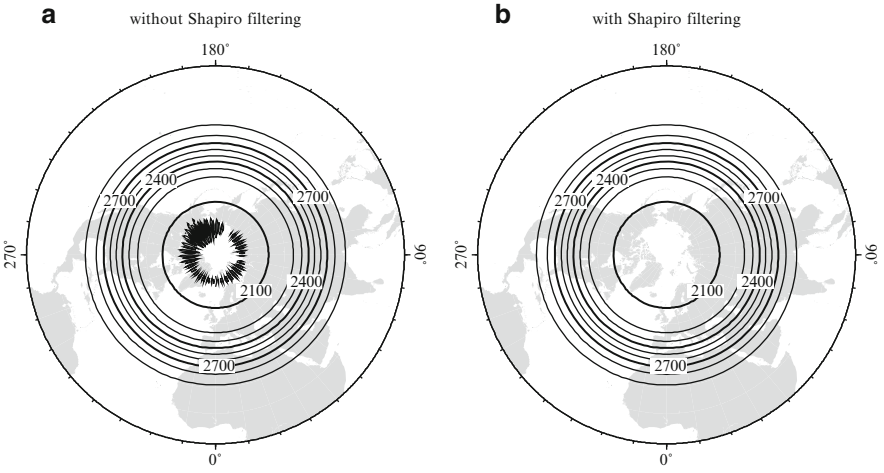

Fig. 13.15 Geopotential height field after 23 model hours (test case 3) in the shallow water version of the FV model (Jablonowski 2004). A north-polar stereographic projection is shown (*outer circle* is the equator). (**a**) No Shapiro filtering poleward of 60°, (**b**) Second order Shapiro filter is applied between 60°–75°. Contour interval is 100 m. Resolution is 2.5° × 2.5° with $\Delta t = 600$ s

course, the Fourier filter could have also been used to eliminate this linear instability at lower latitudes.

In summary, digital filtering promotes computational stability by eliminating or severely dampening the CFL unstable waves, especially at high latitudes. The filtering must be selectively applied in order to avoid a detrimental damping effect in the midlatitudes or tropical regions. In the finite-volume shallow water example, digital filtering techniques complement the even more effective FFT filtering technique used near the pole points. However, attempts to entirely replace the FFT filter with a digital Shapiro (1975) or Purser (1987) algorithm did not prevent numerical instabilities close to the poles. As an aside, an FFT-Shapiro filter mix was also promoted

by Kalnay-Rivas et al. (1977), Ruge et al. (1995), Fox-Rabinovitz et al. (1997) and Takacs et al. (1999).

The main motivation behind such a mix is that local digital filtering has computationally advantages on parallel computing architectures. Shapiro filtering only requires the use of $2n + 1$ neighboring points whereas Fourier filtering incorporates all grid cells along a latitude circle. The latter becomes problematic on parallel machines that distribute the entire GCM domain over many processors. It necessitates parallel communication and triggers computational overhead. A comprehensive discussion of such computational aspects in GCMs is also provided in Chap. 16.

Many choices are still left open. Digital filters can be applied at each time step, sporadically or only with a fraction of their full strength. The latter was suggested by Fox-Rabinovitz et al. (1997). In addition, decisions need to be made about the filtered variables. For example, Fox-Rabinovitz et al. (1997) do not apply their 2D high-order Shapiro filter to the surface pressure field so that the conservation of mass is not affected. Instead, the time tendencies of all other forecast variables are filtered.

13.5.2.3 Spectral Filters: Fourier Filtering

Fast Fourier Transform (FFT) filters are spectral filters that are popular in grid point models with latitude–longitude grids, especially if explicit time-stepping is used. In general, a Fourier filter is only applied in the zonal direction to promote computational stability at mid- and high latitudes, and to allow a violation of the CFL conditions for gravity waves in the filtered region by eliminating short unstable waves. Examples of models with FFT filters include Williamson and Browning (1973), Williamson (1976), Purser (1987, 1988), Fox-Rabinovitz et al. (1997) and Lin (2004). In particular, Purser (1988) examined different filtering strategies and highlighted their advantageous and detrimental effects. Takacs and Balgovind (1983) compared the spectral filtering of tendencies to the spectral filtering of the prognostic variables and assessed the side effects of polar filtering. This fuels the discussion of conservation properties in Sect. 13.7. Generally, Fourier filtering is solely applied along coordinate surfaces without adjustments to constant pressure or height levels as was discussed in Sect. 13.3.2 for the divergence and vorticity diffusion.

The polar FFT filtering of a variable ψ is accomplished by first applying a 1D Fourier (forward) transformation along an entire latitude circle with constant longitudinal grid spacing $\Delta\lambda = 2\pi/n_x$. Here, n_x symbolizes the total number of grid points in the zonal direction. The Fourier coefficients for wavenumbers that exceed a prescribed threshold are then modified which corresponds to a damping mechanism in spectral space. The filtered coefficients are finally transformed back to physical space which completes the filter step. This filter application can be written for all dimensionless wavenumbers k as

$$\hat{\psi}(k)_{filtered} = f(k)\,\hat{\psi}(k) \tag{13.108}$$

where $\hat{\psi}(k)$ and $\hat{\psi}(k)_{filtered}$ are the Fourier coefficients before and after the filtering step. The filter coefficients $f(k)$ are generally defined as a function of k and the latitudinal position ϕ. Most commonly, they are defined as

$$f(k) = \min\left[1., \left(\frac{\cos\phi}{\cos\phi_c}\right)^{2q} \frac{1}{\sin^2(k\,\Delta\lambda/2)}\right] \tag{13.109}$$

with the cutoff latitude ϕ_c. The polar filter is then only active polewards of ϕ_c where $|\phi| > \phi_c$. Its strength gradually increases toward the pole by increasing the number of affected wavenumbers and decreasing the response function $f(k)$ by which they are damped. The positive integer coefficient q can be used to adjust the filter strength. Models like WRF (Skamarock et al. 2008), CAM FV (Collins et al. 2004) or the stretched-grid finite-difference model by Fox-Rabinovitz et al. (1997) set the default to $q = 1$. Older versions of NASA's GEOS model (version 2, Suarez and Takacs (1995)) and its stretched-grid variant (Takacs et al. 1999) utilized a weaker Fourier filter with the coefficients

$$f(k) = \min\left[1., \left(\frac{\cos\phi}{\cos\phi_c}\right) \frac{1}{\sin(k\,\Delta\lambda/2)}\right]. \tag{13.110}$$

For an even number of grid points n_x the mesh supports the dimensionless wavenumbers $k = 1, 2, 3, \cdots, n_x/2$ that can travel in both the west or east direction. Note that the last entry $k = n_x/2$ corresponds to a $2\Delta\lambda$ wave which is the shortest resolvable wavelength. The $2\Delta\lambda$ mode is stationary and its phase cannot be resolved. For odd n_x, the unitless wavenumber range is $k = 1, 2, 3, \cdots, (n_x - 1)/2$.

The cutoff latitude is model-dependent. In CAM FV and NASA's GEOS5 model with the FV dynamical core on a latitude–longitude grid the cutoff ϕ_c is determined by

$$\phi_c = \arccos\Big[\min\left(0.81, \Delta\phi/\Delta\lambda\right)\Big]. \tag{13.111}$$

For equal grid spacings in both directions this cutoff lies around $\phi_c \sim 36°$ which is the minimum latitude. For default CAM FV grid spacings with $\Delta\phi/\Delta\lambda \approx 0.754$ though, the cutoff is placed near $\phi_c \sim 41°$. Fox-Rabinovitz et al. (1997) set their cutoff latitude to $\phi_c = 45°$. In general, the cutoff is most often placed in the midlatitudes which leaves the tropical region unfiltered. The midlatitudes are often empirically chosen since the grid spacing does not change much equatorwards. An example of the response function $f(x)$ for the cutoff latitude $\phi_c = 40°$ is presented in Fig. 13.16. The figure shows the wavenumber dependency of the filter for both the strong ((13.109) with $q = 1$) and weaker (13.110) Fourier filtering at the latitudes $\phi = 45°, 60°$ and $85°$. The strong filter response is depicted by the solid lines. The figure confirms that the filter is scale-selective and primarily damps the higher wavenumbers (shorter scales) which depend on the latitudinal position. The filtering

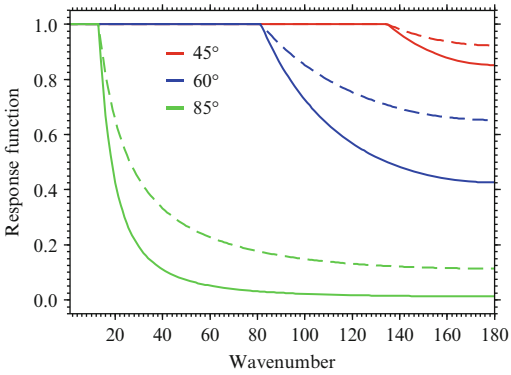

Fig. 13.16 Response function $f(x)$ for the Fourier filter at the latitudes $\phi = 45°, 60°$ and $85°$. The *solid line* denotes the strong filter (13.109), the *dashed line* shows the weaker filter (13.110). The cutoff latitude is $\phi_c = 40°$

gets stronger at higher latitudes. Close to the poles almost all wavenumbers are heavily damped.

The Fourier filtering can be made mass-conservative as discussed in Skamarock et al. (2008), but it is neither monotonic nor positive definite. The latter becomes especially important if tracer variables are filtered since the filter can create negative tracer mass.

13.5.2.4 Local Spectral Filters

Local spectral filters are popular in GCMs that utilize local spectral methods like the spectral element method or the discontinous Galerkin (DG) approach. Despite their high accuracy at high orders the spectral element or DG methods are susceptible to nonlinear aliasing errors which introduce nonphysical high-frequency oscillations. A local spectral filter is then often employed as an alternative to hyper-diffusion to prevent this noise from contaminating the solution.

The most common local spectral filter is the Boyd-Vandeven filter which is a variant of the Vandeven (1991) filter that was developed by Boyd (1996, 1998). The Boyd-Vandeven filter of order p is given by

$$\sigma(x) = \frac{1}{2}\mathrm{erfc}\left(2\sqrt{p}\Omega\sqrt{-\frac{\log(1-4\Omega^2)}{4\Omega^2}}\right) \qquad \text{with } \Omega = |x| - \frac{1}{2} \tag{13.112}$$

where $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$ symbolizes the complementary error function and $\mathrm{erf}(x)$ is the error function. Taylor et al. (1997) explain how this filter is applied to a 1D function $f(x)$ in the x-direction that can be written as the sum of the first N Legendre polynomials $P_k(x)$ with coefficients f_k

$$f(x) = \sum_{k=0}^{N-1} f_k P_k(x). \tag{13.113}$$

The filtered function $f'(x)$ is then given by

$$f'(x) = (1-\mu)f(x) + \mu \sum_{k=0}^{N-1} w_k(x) f_k P_k(x) \tag{13.114}$$

where the weights w_k are defined as

$$w_k(x) = \begin{cases} 1 & \text{if } k < s \\ \sigma(x)\left(\frac{k-s}{N-s}\right) & \text{if } s \le k \le N. \end{cases} \tag{13.115}$$

The parameter μ represents the filter viscosity which can range from 0 (no filtering) to 1 (full filtering). The coefficient s specifies the filter lag, e.g., setting $s = 2N/3$ determines that the filter is only applied to the last one third of the spectrum of f. Taylor et al. (1997) showed that this filter is very scale-selective, especially at high order. It can be applied without sacrificing the spectral accuracy of the spectral element or DG scheme.

The Boyd-Vandeven filter has for example been used by Taylor et al. (1997), Giraldo et al. (2002), Giraldo and Rosmond (2004), Thomas and Loft (2005) and St-Cyr et al. (2008) for idealized dynamical core assessments. These references also discuss the specifics of the 2D filter implementation in the particular model. A common choice is to apply a weak twelfth-order ($p = 12$) filter every few time steps with the parameters $\mu = 0.2$ and $s = 2N/3$. These parameters are generally chosen through experimentation. Note that Thomas and Loft (2005) observed that a much larger $\mu = 0.4$ was often required for stable integrations when switching from a pure vertical σ-coordinate to hybrid η-coordinates. They also needed to apply the filter at every time step.

In practice though, the filter might not be strong enough for full GCMs simulations with parameterized physics (Mark Taylor, personal communication). For example, it has been found in the spectral element model HOMME that spectral filtering alone did not prevent the so-called grid imprinting of a cubed-sphere computational mesh in the numerical solution. The cubed-sphere structure of the mesh was mirrored e.g., in the precipitation field. Such artificial effects were avoided when switching to a fourth-order hyper-diffusion in HOMME.

13.6 Inherent Numerical Damping

Inherent numerical dissipation comes in many forms and is a source of nonlinear flow-dependent damping in model simulations. For example, it is embedded in semi-Lagrangian advection schemes due to the necessary interpolations at every time step. In addition, dissipation is inherent in finite volume methods that are upwind-biased or utilize monotonicity constraints to avoid unphyscial over- and undershoots in the solution. The inherent damping is not necessarily a weakness

of the numerical scheme. It can be turned into a useful property as e.g., demonstrated by Váňa et al. (2008) who used the damping abilities of interpolations in a semi-Lagrangian scheme as targeted diffusive filtering. They named their technique "semi-Lagrangian horizontal diffusion" which brought beneficial new skills to their forecast model. Skamarock and Klemp (2008) pointed out that their upwind-biased advection schemes in the model WRF provides significant filtering of the small scales. They estimated that the effective hyperviscosity coefficient is proportional to the Courant number, and thereby most active at higher Courant numbers where phase errors are most likely to produce noise. Skamarock and Klemp (2008) also found that the horizontal mixing provided by the fifth-order upwind-biased advection scheme in WRF is sufficient to control small-scale noise in weather prediction applications for grid spacings larger than 10 km.

The topic of inherent numerical dissipation is rather broad and cannot be exhaustively covered in this chapter. Therefore, we only present selected aspects to highlight the principal design considerations and characteristics of this type of nonlinear dissipation. In particular, we discuss the inherent dissipation that is embedded in the nominal order of a finite volume scheme, assess the diffusive properties of monotonicity constraints, briefly review the use of the decentering technique as e.g., used in semi-implicit semi-Lagrangian schemes, and shed light on the damping characteristics of semi-Lagrangian methods.

13.6.1 Order of the Numerical Scheme

As a specific example of the inherent dissipation in a finite volume scheme we discuss the properties of a first-, second- and third-order approximation in the model FV. The advection method implemented in the FV dynamical core can be viewed as a multi-dimensional extension of higher-order Godunov-type schemes like the van Leer scheme (van Leer 1974, 1977) or the Piecewise Parabolic Method (PPM, Colella and Woodward (1984)). Finite volume schemes are based on the reconstruct-evolve-average approach as e.g., explained by LeVeque (2002) and in Chap. 8, and use constant, piecewise linear (van Leer), piecewise parabolic (PPM) or even piecewise cubic subgrid distributions for the piecewise continuous reconstruction of the flow field. The transport problem is then solved exactly and new initial data at the future time step are obtained by averaging the transported quantity over each control volume.

The first-order upwind method is based on a constant subgrid distribution and is thereby very diffusive by design. It is for example explained by Lin and Rood (1996) who introduced the FV advection scheme. The second-order van Leer advection scheme (van Leer 1977) is based on the reconstruction of linear subgrid distributions in each finite volume cell. We now briefly review the design of such subgrid distributions to motivate the subsequent discussion. The linear subgrid distribution $h(x)$ of a model variable h is given by

$$h(x, y) = \bar{h} + \Delta a^x \, x + \Delta a^y \, y \tag{13.116}$$

where $\bar{h} = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} h(x, y) \, dx \, dy$ is the volume-averaged value with normalized local coordinates $x, y \in [\frac{-1}{2}, \frac{1}{2}]$. Δa^x and Δa^y denote the slopes in the x and y direction at a grid point (i, j) which in the model FV are defined via van Leer's scheme I

$$\Delta a^x = \frac{1}{2} \left(h_{i+1,j} - h_{i-1,j} \right) \tag{13.117}$$

$$\Delta a^y = \frac{1}{2} \left(h_{i,j+1} - h_{i,j-1} \right). \tag{13.118}$$

This assessment uses centered finite differences. The slopes can be further manipulated if monotonicity constraints are required. Then the monotonized central-difference (MC) slope limiter (van Leer 1977) can e.g., be used

$$\begin{aligned} \Delta a^x &= min\big(|\Delta a^x|, 2|h_{i+1,j} - h_{i,j}|, 2|h_{i,j} - h_{i-1,j}|\big) \; sgn(\Delta a^x) \\ &\qquad \text{if } (h_{i+1,j} - h_{i,j})(h_{i,j} - h_{i-1,j}) > 0 \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{13.119}$$

that picks out the smallest magnitude among three slopes which are the centered difference and the two one-sided differences. However, if the discrete value of $h_{i,j}$ represents an extreme value, the slope is set to zero. The same principle applies to the MC slope limiter in the y direction. The $sgn(\Delta a^x)$ function extracts the sign of the argument.

Alternatively, the third-order PPM scheme can be applied. The corresponding 2D biparabolic subgrid distribution is then given by

$$h(x, y) = \bar{h} + \delta a^x \, x + b^x \left(\frac{1}{12} - x^2 \right) + \delta a^y \, y + b^y \left(\frac{1}{12} - y^2 \right) \tag{13.120}$$

where the coefficients of the parabola δa^x, b^x, δa^y and b^y are defined by Colella and Woodward (1984) or Carpenter et al. (1990). The coefficients can again be modified in order to enforce monotonicity constraints as explained in Lin and Rood (1996) and Lin (2004). Both the order and the choice of the monotonicity constraint (see Sect. 13.6.2) determine the inherent diffusion in the FV advection scheme.

Figure 13.17 visualizes the effects of the inherent diffusion in the dynamical core CAM FV with the grid spacing $1° \times 1°$ and 26 vertical levels. CAM FV has runtime options to run with the aforementioned first-order upwind advection scheme, the second-order van Leer algorithm or with the nominally third-order PPM method. Note however, that the 2D implementation of the PPM algorithm in the model FV does not exhibit a third-order convergence and is in practice closer to a second-order scheme (Lin and Rood 1996; Jablonowski et al. 2006). Both the van Leer scheme and PPM apply monotonicity constraints which are the MC limiter in case of van Leer, and the "relaxed" monotonicity constraint in case of PPM (Lin 2004).

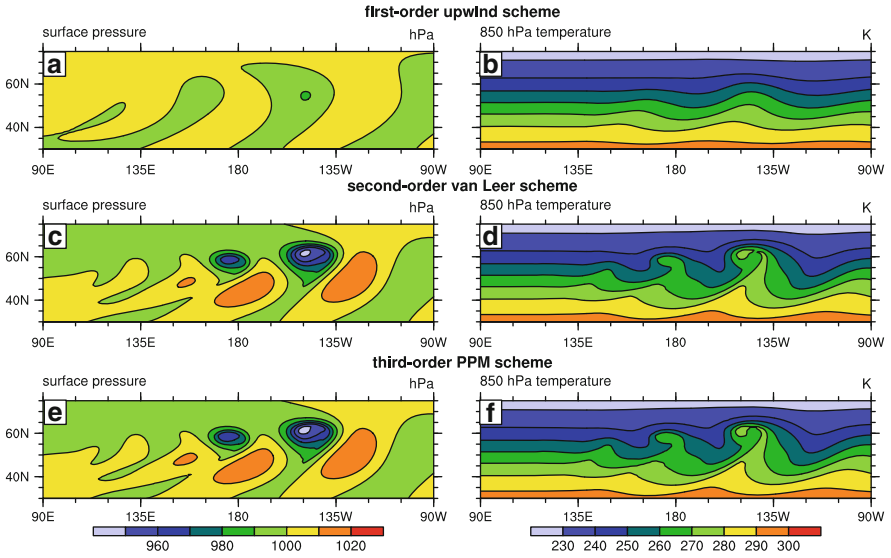

Fig. 13.17 Surface pressure (hPa) and 850 hPa temperature (K) at day 9 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) in the CAM FV dynamical core at the resolution $1^\circ \times 1^\circ$ L26. (**a**,**b**) first-order upwind scheme, (**c**,**d**) second-order van Leer scheme, (**e**,**f**) default third-order PPM scheme with the relaxed monotonicity constraint (Lin 2004). The dynamics time step is $\Delta t = 180$ s

The figure shows the surface pressure and 850 hPa temperature fields at day 9 of the growing baroclinic wave described in Jablonowski and Williamson (2006a). The impact of the inherent numerical dissipation can clearly be seen in all fields. The first-order method (Fig. 13.17a,b) hardly captures the evolving baroclinic instability and only shows hints of a wave. The second-order van Leer scheme shows a clear evolution of the baroclinic wave and exhibits a slightly early wave breaking event in the temperature field (Fig. 13.17d). The peak amplitudes of the surface pressure field at day 9 are $(p_s)_{min} = 948.19$ hPa and $(p_s)_{max} = 1018.78$ hPa with the second-order van Leer (Fig. 13.17c) algorithm. The evolution of the wave simulated with the PPM scheme is very similar to the van Leer simulation. However, the peak surface pressure amplitudes are slightly intensified and the values read $(p_s)_{min} = 947.04$ hPa and $(p_s)_{max} = 1018.74$ hPa in Fig. 13.17e. In addition, the temperature field in the PPM simulation (Fig. 13.17f) shows slightly sharper frontal zones without wave breaking.

A more quantitative comparison of the baroclinic wave simulation is provided in Fig. 13.18. The figure depicts the 700 hPa kinetic energy spectra at day 30 of the baroclinic wave simulation for the first-order upwind method, the second-order van Leer scheme and two PPM simulations. The only difference between the PPM simulations is the selection of the monotonicity constraint. Here, we compare the "relaxed" constraint by Lin (2004) and "positive definite" constraint described in Lin and Rood (1996). The latter only prevents negative undershoots and has

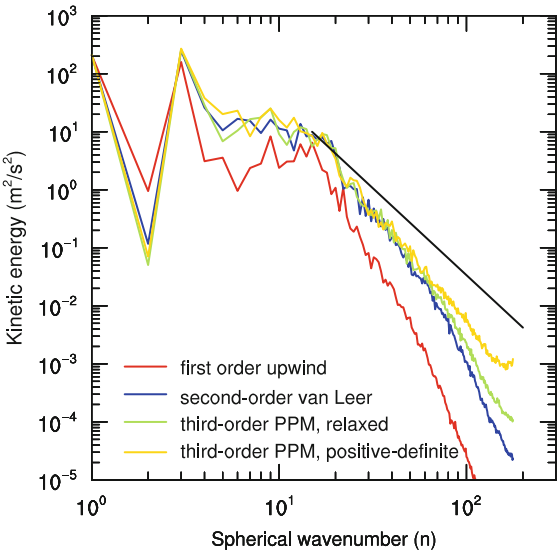

Fig. 13.18 700 hPa kinetic energy spectra at day 30 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) in the CAM FV dynamical core at the resolution $1° \times 1°$ L26: first-order upwind, second-order van Leer, third-order PPM scheme with the *relaxed* and *positive definite* monotonicity constraints (Lin 2004). The *black line* indicates an n^{-3} slope. The dynamics time step is $\Delta t = 180$ s

originally been designed for pure tracer advection. Therefore, its application in the dynamical core is generally not recommended, but we only use it here to demonstrate different diffusion properties. The curves confirm that the first-order method is the most diffusive as indicated by the sharp drop off of the spectrum and the significant damping of the longer wavenumbers 4–10. The differences between the van Leer and PPM simulations are more subtle. The van Leer and relaxed PPM curve almost overlay each other until about wavenumber 60 or so before the van Leer curve exhibits a slightly faster drop off than the relaxed PPM run. In contrast, the positive definite PPM simulation is less diffusive than the relaxed PPM scheme and almost runs parallel to the n^{-3} slope. More analysis on the monotonicity constraints in provided in the next Sect. 13.6.2.

Increased inherent dissipation can also be used as a sponge at the model top. This is for example accomplished by lowering of the order of the numerical scheme in the uppermost levels. Such a technique is used in the CAM FV dynamical core that switches the numerical scheme from the PPM algorithm to the more diffusive van-Leer scheme in the uppermost $nlev/8$ layers where $nlev$ is the number of total levels.

Inherently dissipative schemes can often be run without explicit diffusion which is for example true for the semi-Lagrangian models CAM SLD and UM. However, additional explicit diffusion might still be applied in long climate simulations. Inherent numerical dissipation can also be viewed as an application of a symmetric low-pass sine filter as suggested by Raymond and Garder (1991). These similarities between inherent dissipation in finite-difference models and numerical filters were also pointed out in Purser and Leslie (1994).

13.6.2 Monotonicity Constraints and Shape Preservation

The advection algorithm in the CAM Finite Volume model is upstream-biased and monotonic if limiters are applied to the subgrid distributions. As mentioned in the previous subsection such monotonicity constraints are used in case of the default PPM scheme which leads to a total variation diminishing (TVD) method. A short review of limiters for finite volume schemes is given in Chap. 8 or in the textbook by Durran (2010). Note that Chap. 9 introduces limiters for discontinuous Galerkin methods which are an active research topic (see also Nair (2009)). In general, limiters can be grouped into slope/curvature limiters or flux limiters. Here we only briefly assess the impact of slope/curvature limiters that a-priori limit the bi-parabolic subgrid distribution used in the PPM scheme.

The limiting of the subgrid distribution clips extreme values and thereby introduces inherent nonlinear dissipation into the finite volume scheme. From a design perspective the clipping suppresses over- and undershoots in the advection step. It should be as strict as necessary to prevent unphysical oscillations but as nonintrusive as possible to minimize the associated dissipation. Figure 13.19 schematically illustrates how the clipping by a monotonicity constraint affects a flow field (Jablonowski 2004). In this particular example, an interpolation of a zonal wind field from a 2.5° × 2.5° to a 1.25° × 1.25° latitude–longitude grid is performed. The analytically prescribed zonal wind field at the 2.5° × 2.5° resolution is depicted in

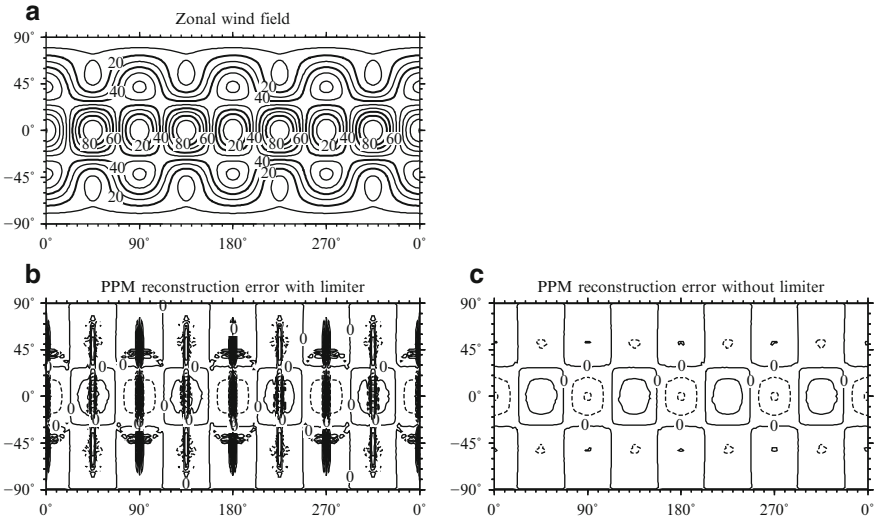

Fig. 13.19 Latitude–longitude plot of the zonal wind field in the Rossby–Haurwitz wave test case. (**a**) Zonal wind field at a 2.5° × 2.5° resolution, (**b**) absolute error of the zonal wind field after a PPM-based interpolation with limiters to a 1.25° × 1.25° grid, (**c**) absolute error of the zonal wind field after a PPM-based interpolation without limiters to a 1.25° × 1.25° grid. The contour intervals are (**a**) 10 m s^{-1} and (**b**,**c**) 0.025 m s^{-1}. Negative contours are *dashed*

Fig. 13.19a. The field corresponds to the Rossby–Haurwitz wave with wavenumber 4 that is part of the standard shallow-water test suite by Williamson et al. (1992). The interpolation uses the bi-parabolic subgrid-scale reconstruction defined by the PPM algorithm (Colella and Woodward 1984) which are then integrated over the new $1.25° \times 1.25°$ domain.

Figure 13.19b,c assess the absolute errors of these interpolations with and without monotonicity constraints. In particular, the original PPM monotonicity constraint without steepening (Colella and Woodward 1984) was selected. The distribution of the errors in Fig. 13.19b confirms that the monotonicity constraint mostly affects the areas with extreme values. This can be seen when comparing the error pattern to the original zonal wind field (Fig. 13.19a). In contrast, the interpolation errors are clearly diminished without the limiter (Fig. 13.19c). In fact, small-magnitude over- and undershoots are present in Fig. 13.19c with peak values around ± 0.05 m s^{-1}. The overshoots appear in the regions of the wind maxima, the undershoots are concentrated near the wind minima. These are eliminated by the monotonicity constraint (Fig. 13.19) that, on the downside, increases the overall errors to ± 0.22 m s^{-1}. Note that these errors are assessed with the help of the analytic solution at the $1.25° \times 1.25°$ resolution.

Limiting can enforce a strictly monotonic advection algorithm in the 1D case as also discussed by Thuburn (1993, 1997) and Mesinger and Jovic (2002). However, very minor violations of the monotonicity constraint in two-dimensional flows are possible and have indeed been observed for the horizontal FV advection algorithm by Lin and Rood (1996). Limiters can also be designed to only avoid negative undershoots for tracer transport application, and allow an overestimation of the transported quantity. The limiters in finite volume schemes or shape-preservation constraints in semi-Lagrangian models are physically motivated and provide a smoothness constraint. They are a form of nonlinear inherent numerical dissipation that is guided by the flow field. Shape preservation constraints for semi-Lagrangian dynamical cores and advection schemes are for example discussed in Williamson and Rasch (1989), Williamson (1990), Rasch and Williamson (1990b) and Bermejo and Staniforth (1992).

A detailed documentation of the many limiter options in the model FV is beyond the scope of this section and we refer to the associated literature for the exact explanations of the algorithms (Colella and Woodward 1984; Carpenter et al. 1990; Lin and Rood 1996; Huynh 1996; Lin 2004; Putman and Lin 2007). The main focus here is to qualitatively demonstrate that limiters determine the amount of inherent dissipation in finite volume scheme and need to be used with care. Figure 13.20 visualizes the influence of different PPM monotonicity constraints on the evolution of the baroclinic wave in the FV dynamical core. Here, two models are depicted which both utilize almost identical versions of the FV dynamics described in Lin (2004). They are the CAM 5 FV implementation with the grid spacing $1° \times 1°$ and the GFDL FV model with the *lat* $\times$ *lon* grid spacing $1° \times 1.25°$. The figure can also be directly compared to the surface pressure fields in Fig. 13.17.

The figure demonstrates the effects of the positive-definite (Fig. 13.20a) and quasi- (or semi-) monotone constraints (Fig. 13.20b) described in Lin and Rood

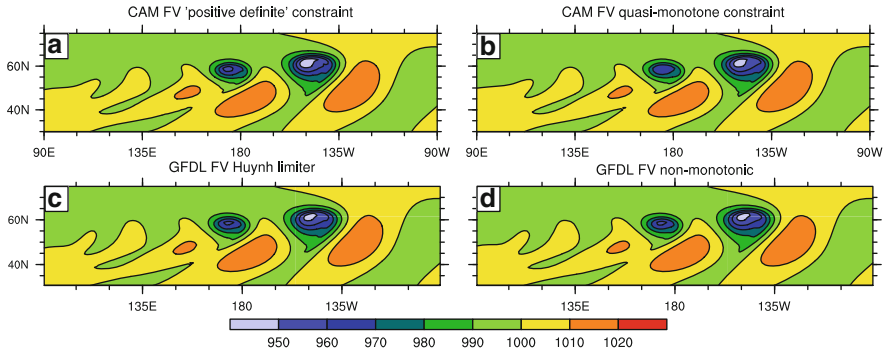

Fig. 13.20 Surface pressure (hPa) at day 9 of the growing baroclinic wave test case of Jablonowski and Williamson (2006a) in the (**a**,**b**) CAM FV dynamical core at the resolution $1^\circ \times 1^\circ$ L26 and (**c**,**d**) GFDL FV model at the resolution M90 $1^\circ \times 1.25^\circ$ with 26 levels. (**a**) positive definite constraint that only eliminates negative undershoots, (**b**) quasi-monotone constraint (Lin and Rood 1996) that is the default in CAM FV, (**c**) second constraint discussed in Huynh (1996), (**d**) non-monotonic simulation with slope- but no curvature-limiter in the PPM scheme

(1996), the second monotonicity constraint by Huynh (1996) (Fig. 13.20c) and a non-monotonic simulation (Fig. 13.20d) that only utilized the MC slope limiter but no curvature constraint. Option (b) is the default in CAM FV. Overall, the strengths of the pressure systems in all four simulations are comparable and the differences are subtle. The quasi-monotone simulation appears to be slightly more diffusive than the other three model runs as indicated by the fewer contour lines in the low pressure system located at about 175°E. The "positive definite" simulation, that only prevents negative undershoots, is the least diffusive and has developed the deepest low pressure systems. This has also been demonstrated by the kinetic energy spectra in Fig. 13.18. However as mentioned before, this monotonicity option has been specifically designed for positive definite tracer transports and should not be used as the basis for the dynamical core. Other even less stringent limiters are available for PPM-type algorithms as e.g., documented by Colella and Sekora (2008) and McCorquodale and Colella (2010).

The choice of the limiter should be motivated by the design criteria of the model as argued in Chap. 15. Note that the cumulative damping effect of the limiters cannot be quantified analytically. Therefore, the model CAM 5 FV applies a total energy fixer that provides dissipative heating (Neale et al. 2010).

13.6.3 Decentering Mechanisms

Decentering adds inherent dissipation to the numerical scheme and is tightly linked to semi-implicit time discretizations. Here we briefly review the decentering used in semi-implit semi-Lagrangian models and in other semi-implicit approaches.

13.6.3.1 Decentering in Semi-implicit Semi-Lagrangian Models

The decentering mechanism, sometimes also called uncentering or off-centering technique, is usually applied in semi-implicit semi-Lagrangian (SISL) models. Its primary purpose is to suppress computational noise and orographic resonance in regions of steep orography and high Courant numbers, and maintain stability, especially at high resolutions (Bates et al. 1993; Rivest et al. 1994). A comprehensive discussion of the orographic resonance problem in semi-Lagrangian models is provided in Rivest et al. (1994), Côté et al. (1995) and Lindberg and Alexeev (2000), and is not repeated here. Tanguay et al. (1992) suggested a first-order decentering of the semi-implicit terms along the trajectory. Rivest et al. (1994) discussed both first and second-order decentering schemes in a 1D SISL shallow water model. A thorough stability analysis of the decentering method is presented in Tanguay et al. (1992) and Gravel et al. (1993).

To illustrate the basic idea behind the first-order decentering technique consider the prognostic equation

$$\frac{D\psi}{Dt} = S \tag{13.121}$$

where D/Dt is the total time derivative, ψ is a scalar variable and S is a source term which may incorporate ψ. In a two-time-level semi-Lagrangian scheme a conventional discretization of the trajectory calculation leads to

$$\psi^{j+1} - \psi_d^j = \int_{t^j}^{t^{j+\Delta t}} S\, dt = \bar{S}\, \Delta t \tag{13.122}$$

where ψ^{j+1} is the value of the prognostic variable at the arrival point at time t^{j+1} and ψ_d^j is the value at the departure point of the trajectory at time t^j. $\bar{S}$ denotes the time-averaged source term along the trajectory that can be replaced by

$$\psi^{j+1} - \psi_d^j = \Delta t \left(\frac{1+\epsilon}{2} S^{j+1} + \frac{1-\epsilon}{2} S_d^j \right) \tag{13.123}$$

with the decentering (time-weighting) parameter ϵ. The averaged source term represents both a temporal and spatial average. A centered two-time-level scheme with $\epsilon = 0$ is second-order accurate. For a decentered scheme with $0 < \epsilon < 1$ the truncation error is first-order. A decentered SISL scheme is generally more accurate and less damping the closer ϵ is to 0, and less accurate and more damping the closer ϵ is to unity. In practice though, some decentering is desirable or even necessary in SISL schemes to suppress the spurious orographic resonance.

Decentering is for example used in the operational Global Environmental Multiscale (GEM) model developed at the Canadian Meteorological Centre (Côté et al. 1998a,b), in the spectral transform model CAM SLD (Collins et al. 2004) and the grid point model UM (Staniforth et al. 2006). Typical decentering parameters in GCMs are $\epsilon = 0.1$ in the model GEM, $\epsilon = 0.2$ in CAM SLD and $\epsilon \in [0.2, 0.4]$ in the model UM (Davies et al. 2005). Davies et al. (2005) reported that the smaller

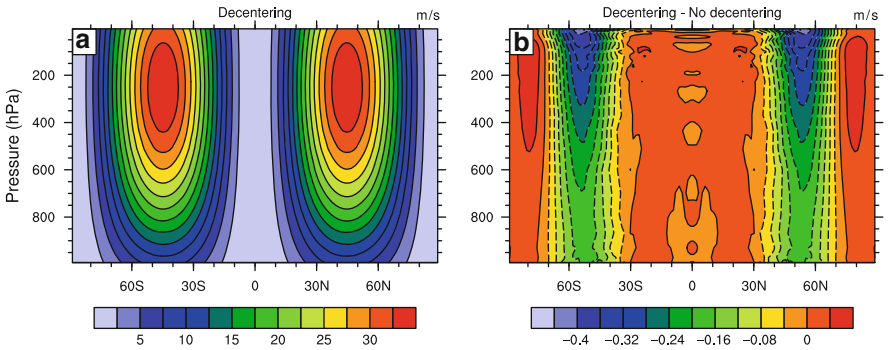

Fig. 13.21 (**a**) Zonal-mean zonal wind (m/s) at day 30 of the steady-state test case of Jablonowski and Williamson (2006a) in the CAM SLD dynamical core at the resolution T85L26. The default decentering parameter $\epsilon = 0.2$ is used. (**b**) Difference of the zonal-mean zonal wind at day 30 between the run with decentering and a run without decentering ($\epsilon = 0$). No explicit diffusion is used, the time step is $\Delta t = 1{,}800$ s

value $\epsilon = 0.2$ is sufficient at low resolutions, but it needed to be replaced with $\epsilon = 0.4$ at high weather forecast resolutions to suppress noise near strong jets. A discussion of the impact of decentering and its stability properties can also be found in Chap. 14.

As a practical example, we now isolate the effect of the inherent dissipation from the decentering mechanism in idealized dynamical core simulations. As before in Sect. 13.3.3, we choose the CAM 4 semi-Lagrangian dynamical core at the triangular truncation T85 with 26 levels. A steady-state test case, described in Jablonowski and Williamson (2006a), is used and run for 30 days with varying decentering parameter ϵ. Figure 13.21 shows the zonal-mean zonal wind field at day 30 with the default decentering parameter $\epsilon = 0.2$ and the zonal-mean zonal wind difference between the run with decentering and no decentering. No explicitly added diffusion was used. The influence of the decentering can clearly be seen in the difference plot (Fig. 13.21b) throughout the entire atmosphere but the impact is strongest in the midlatitudes in this test case, especially near the model top. The decentering damps the zonal wind speed with magnitudes of up to 0.5 m s^{-1} during this 30-day simulation. Note that Fig. 13.21 can also be readily compared to Fig. 13.1 that isolates the effects of the fourth-order hyper-diffusion and second-order sponge layer diffusion (without decentering) with the help of the same test case.

A quantitative comparison of the damping due to decentering is depicted in Fig. 13.22 that shows the time evolution of the global root-mean square l_2 zonal wind error during the 30-day steady-state simulation. For this analysis the zonal-mean zonal wind field $\overline{u}$ is compared to the analytic solution at time $t = 0$ (see Jablonowski and Williamson (2006a) for the definition of the error norm). The decentering parameter ϵ is set to 0, 0.1, 0.2 and 0.3, respectively. Again, no explicitly added diffusion was used and, as briefly discussed before for Fig. 13.1, the semi-Lagrangian trajectory calculation utilized only spherical coordinates to suppress

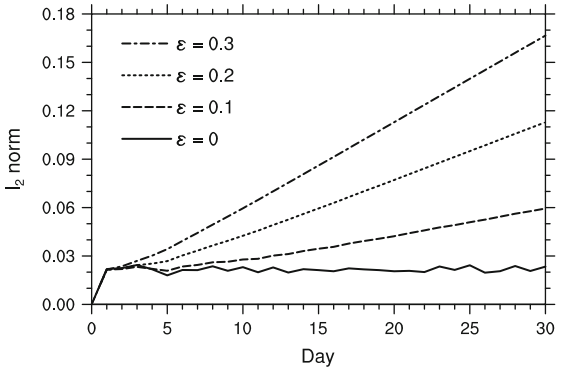

Fig. 13.22 Time evolution of the $l_2(\overline{u}(t) - \overline{u}(t=0))$ error norms (in m/s) of the zonal-mean zonal wind field $\overline{u}$ in the steady-state test case of Jablonowski and Williamson (2006a). The CAM SLD dynamical core simulations at the resolution T85L26 with decentering parameters between $\epsilon = 0$ and $\epsilon = 0.3$ are shown. No explicit diffusion is used, the time step is $\Delta t = 1{,}800$ s

any additional damping from non-zonal geodesic trajectory calculations in polar regions (Williamson and Rasch 1989). The latter is only reasonable in the case of zonal advection as considered here. These deviations from the default CAM SLD configuration are selected to truly isolate the damping effects from the decentering parameter. Of course, in practice the damping of all explicit and inherent dissipation mechanisms as well as filters and fixers act in concert, and they are generally difficult to isolate individually. Figure 13.22 confirms that the inherent dissipation in these steady-state simulations strongly depends on the decentering parameter. The l_2 zonal wind errors grow steadily over time, and there is an almost linear relationship between the magnitude of the error at day 30 and the magnitude of the decentering parameter. Recall that CAM SLD selects $\epsilon = 0.2$ by default which poses a compromise between accuracy and the suppression of orographic noise in practice.

13.6.3.2 Forward-Biasing of Trapezoidal Time Integrations

Of similar spirit as the SISL decentering approach is the forward-biasing technique for the implicit trapezoidal time integration method that provides damping of high-frequency modes. A short discussion can be found in Durran (1999) (his Chap. 7.3). The principal difference between the SISL decentering and the forward-biasing is that the SISL decentering represents a mix of a spatial and temporal average since both the departure and arrival point information are involved in the estimate of the decentered trajectory. In contrast, the forward-biasing technique only represents a temporal average at a single location. However, forward-biasing is sometimes called off-centering, uncentering or decentering, but despite the same nominal names the differences should be kept in mind.

Forward biasing is accomplished by replacing time-centered source terms of the form $(S^{j+1} + S^j)/2$ with the off-centered expression

$$\left(\frac{1+\epsilon}{2}\right)S^{j+1} + \left(\frac{1-\epsilon}{2}\right)S^j \tag{13.124}$$

where $0 \leq \epsilon \leq 1$. Choosing $\epsilon = 0$ recovers a second-order in time centered discretization that e.g., represents the implicit Crank–Nicolson scheme

$$\frac{\psi^{j+1} - \psi^j}{\Delta t} = \frac{1}{2}(S^{j+1} + S^j). \tag{13.125}$$

Any $\epsilon > 0$ formally reduces the order of accuracy of the temporal discretization. Off-centering the time discretization with $\epsilon > 0$ adds inherent numerical dissipation and numerically stabilizes the solution. Values in the range of $[0.2, 0.4]$ are quite common for models with two-time-level semi-implicit schemes. Note that ϵ is also often called *implicitness parameter*.

As an example, Durran and Klemp (1983) used forward-biasing of the trapezoidal time-differencing scheme for vertical derivatives. They found that a value of $\epsilon = 0.2$ provided sufficient damping that did not noticeably modify the gravity waves. Bonaventura and Ringler (2005) also used $\epsilon = 0.2$ and argued that such an inherently dissipative scheme can often be used without adding further explicit diffusion. As discussed by Skamarock et al. (2008) forward-in-time weighting of the vertically-implicit acoustic-time-step terms damps instabilities associated with vertically-propagating sound waves and the partially-split temporal discretization. The forward weighting also damps instabilities associated with sloping model levels and horizontally propagating sound waves as shown in Durran and Klemp (1983). A value of $\epsilon = 0.1$ is used as the default in the nonhydrostatic limited-area model WRF. The regional model COSMO sets the default parameter to $\epsilon = 0.4$ (Gassmann and Herzog 2007). Recently, Baldauf (2010) assessed suitable limits for the off-centering parameter for both buoyancy terms and sound wave terms in the regional weather forecast model COSMO.

13.6.4 Damping by Semi-Lagrangian Interpolation

Semi-Lagrangian schemes require spatial interpolations at every time step to determine the transported variables at the departure points. These interpolations provide a source of uncontrolled damping in model simulations. Conventional wisdom says that semi-Lagrangian approximations damp more over a given length integration when run with short time steps than when run with long time steps. The argument is that more interpolations are performed with the shorter time step, thus the net damping will be larger. The damping from interpolation generally increases as the wavelength decreases. Thus spectra of, for example, the kinetic energy or vertical

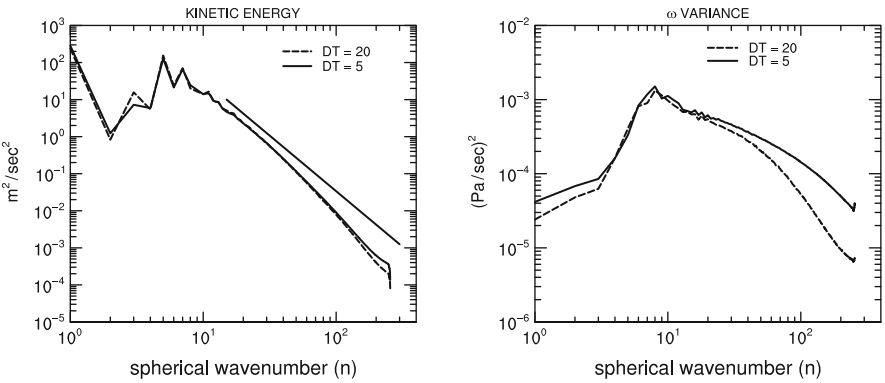

Fig. 13.23 Spectra as function of the spherical wavenumber (n) from TL255 CAM 3.1 semi-Lagrangian integrations with linear transform grid using 5 and 20 min time steps for (*left*) 250 hPa kinetic energy and (*right*) 500 hPa pressure vertical velocity (ω) variance

velocity variance, will be steeper approaching the truncation limit in an experiment using a short time step than in one with a long time step.

Figure 13.23, however, shows that this is not necessarily the case. This figure plots the 250 hPa kinetic energy and 500 hPa pressure vertical velocity variance spectra as a function of spherical wavenumber (n) from integrations with the semi-Lagrangian version of CAM 3.1 at TL255 truncation with 5 and 20 min time steps. This semi-Lagrangian model uses an optional linear (TL) unaliased grid of approximately 0.7° (or 78 km at the equator) which is the same as the quadratically unaliased grid used by the T170 Eulerian spectral model. The linear grid is defined to be the minimum grid required for transformations of a field from spectral space to grid point space and back again to spectral without loss of information. With such a grid, only linear terms are unaliased (Williamson 1997). The numerical algorithms are detailed in Collins et al. (2004). The simulations presented in this subsection are for an aqua-planet (Williamson 2008a) but in our experience, except possibly for the long waves, the shape of spectra in aqua-planet simulations is the same as in earth-like simulations. For both variables the spectra fall off faster for the long time step than for the short time step.

So where does conventional wisdom go wrong? It does not take into account that the short and long time step departure or interpolation points are not at the same relative location in a grid interval and thus the damping rate for a single interpolation is not the same. Figure 13.24a shows the response function of selected wavelengths for cubic Lagrange polynomial interpolation as used in CAM 3.1 SLD as a function of relative position in the grid interval (β) following Williamson and Laprise (2000). For all waves, the amplitude damping increases from the edge of the grid interval ($\beta = 0$ or 1), where it is zero, to the center of the interval ($\beta = 0.5$) where the damping of each wave is greatest, with the $2\Delta x$ wave annihilated there. As a specific example of relative damping assume the long time step is four times the short time step, and that the short time step yields a departure point location with $\beta = 0.125$.

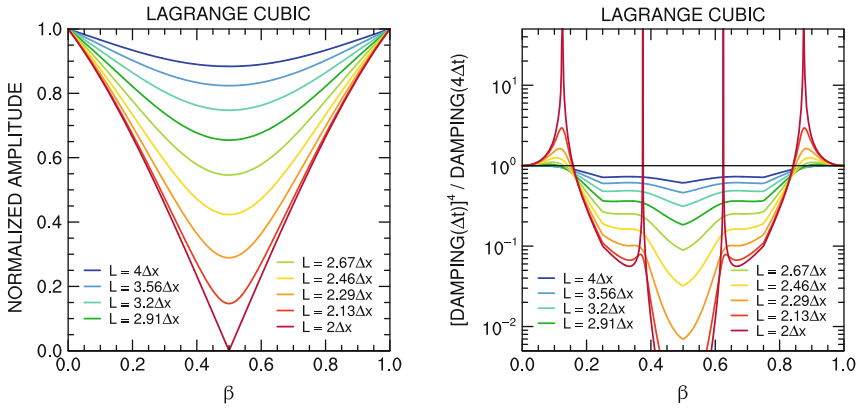

Fig. 13.24 (*Left*) Response function for Lagrange cubic interpolation as function of location within grid interval for selected wavelengths. (*Right*) Ratio of response function for short time step to the power 4 to the response function for long time step as a function of the location in the grid interval of the short time step. The long time step is four times the short time step

The long time step then has a departure point with $\beta = 0.5$ and the amplitude of the $2\Delta x$ wave is zero after interpolation. The amplitude after one short time step ($\beta = 0.125$) is 0.805. However four such interpolations are required to reach the same forecast time as one long time step, therefore the total damping with the short time step is $0.805^4 = 0.42$. Clearly, in this extreme example, the damping is less with the short time step than with the long one.

The general situation is shown in Fig. 13.24b. Here and in the following we consider the situation with the long time step being four times the short one. Later, specific results from the semi-Lagrangian CAM 3.1 at TL255 truncation will use 5 and 20-min time steps. Figure 13.24b plots the ratio of the short time step damping to the fourth power to the damping of the long time step as a function of β which is the location in the grid interval for the short time step. There is a region with $\beta <\sim 0.15$ and a mirror one $\beta >\sim 0.85$ in which the net damping from the short time step is less than that from the long. Elsewhere ($\sim 0.15 < \beta <\sim 0.85$) the net damping from the short time step is greater. The only exception being the $2\Delta x$ wave at $\beta = 0.375$ and 0.625 where the corresponding long time step is $\beta = 1.5$ and 2.5, both of which are equivalent to $\beta = 0.5$ with zero response function. With smaller time step ratios the zero crossings shift inward toward $\beta = 0.5$ with a structure similar to Fig. 13.24b except the secondary interior $2\Delta x$ ratio of ∞ does not occur (not shown). With a time step ratio of 3, the damping ratio for the $2\Delta x$ wave crosses 1 at β around 0.2 and 0.8, and with a time step ratio of 2, the ramping ratio for the $2\Delta x$ wave crosses 1 at β around 0.3 and 0.7.

Therefore, the overall damping in a semi-Lagrangian integration will depend on the population of departure points. Figure 13.25 shows the frequency distribution of the departure point locations from the 5 and 20 min TL255 integrations. The frequency distribution is calculated over all grid points at the 250 hPa model level

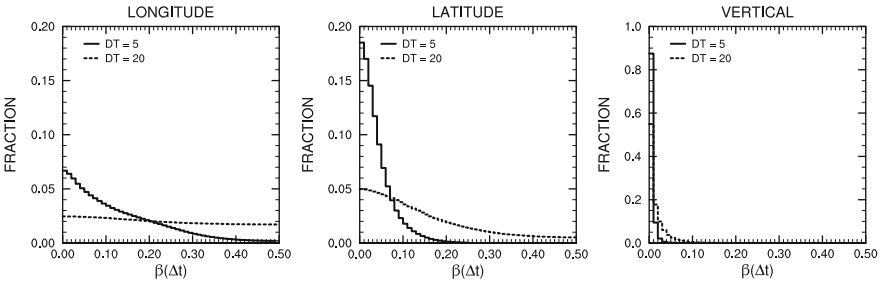

Fig. 13.25 Fraction of departure points falling in 0.01 bins ranging from 0.0 to 0.5 of the grid interval from semi-Lagrangian integrations with $\Delta t = 5$ and 20 min for (*left*) longitude, (*center*) latitude and (*right*) vertical directions

(corresponding to the spectra in Fig. 13.23) every 6 h for 125 days. Because of natural symmetries only β in the range 0–0.5 need be considered; other values fold back into this range. CAM 3.1 SLD uses a tensor product interpolation in which interpolations are done sequentially in longitude, in latitude and then in the vertical coordinate. The frequency distributions are shown for each coordinate. The plots show the fraction of the time the departure point falls in each 0.01 interval from 0.0 to 0.5. We consider the 5 min case first. In the vertical almost all departure points have $\beta < 0.05$. In latitude almost all departure points have $\beta < 0.25$, with only a small subset between 0.15 and 0.25. In longitude, on the other hand, departure points have values as large as $\beta = 0.5$, although the fraction is small approaching 0.5. Five minutes is not a particularly long time step for TL255 resolution. In fact it is the time step normally used for the Eulerian version of CAM 3.1 at T170 truncation. The longitudinal frequency distribution for integrations with 20 min is relatively flat decreasing from 0.022 for the first bin $\beta = [0.0, 0.01)$ to 0.019 for the last bin $\beta = (0.49, 0.05]$. (Recall there are 50 bins, $50 \times 0.01 = 0.5$) The 20 min latitudinal frequency distribution is much less steep than the 5 min one, starting at 0.05 for the first bin, decreasing to 0.02 at $\beta = 0.2$ and continuing to around 0.005 for the last bin. The 20 min vertical frequency distribution starts at 0.55 at the first bin and is close to zero by $\beta = 0.1$

The frequency distribution for the ratio of the damping of the short time step to that of the long time step from the CAM 3.1 SLD integrations is shown in Fig. 13.26 for the $2\Delta x$ wave. To calculate the frequency distribution, equivalent 20-min departure points were calculated based on the 5-min ones captured from the model integration. The ratio of the damping to the fourth power of each 5-min departure point to the damping from the matching 20-min departure point was calculated. The frequency distribution for the log of the ratio is plotted in Fig. 13.26 for the $2\Delta x$ wave. A value of 0 is neutral, positive values imply less damping for the short time step and negative values imply more damping for the short time step. To avoid including the neutral 0 values in either the positive or negative bin, they are given a special bin of their own. This bin contains the fraction of values within rounding of zero and is indicated by the dot plotted at zero abscissa. Before discussing Fig. 13.26

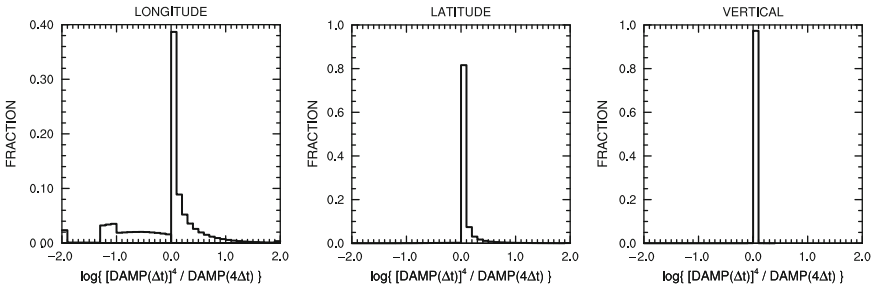

Fig. 13.26 Fraction of log of ratio of short time step damping to fourth power to long time step damping for $2\Delta x$ wave in 0.1 bins from -2.0 to 2.0 for (*left*) longitude, (*center*) latitude and (*right*) vertical interpolations

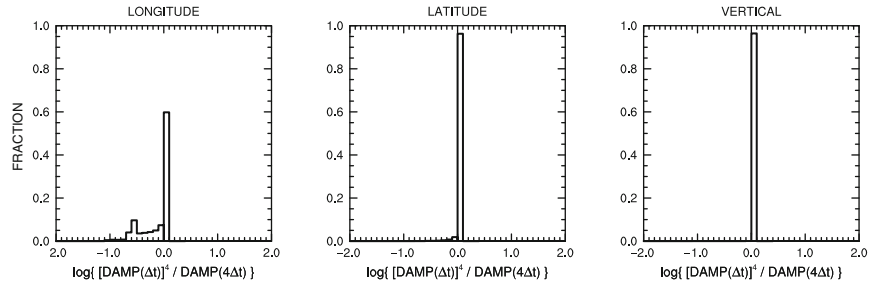

Fig. 13.27 Fraction of log of ratio of short time step damping to fourth power to long time step damping for $2.7\Delta x$ wave in 0.1 bins from -2.0 to 2.0 for (*left*) longitude, (*center*) latitude and (*right*) vertical interpolations

we note that when we base the calculation on the sampled departure points from the 20-min CAM 3.1 SLD experiment and calculate the equivalent 5-min departure points, the plot of the damping ratio is virtually indistinguishable from Fig. 13.26.

The frequency distribution for longitude in Fig. 13.26 is non-zero for both positive and negative log of ratios, indicating there are points where the short time step damps less and points where it damps more. The frequency distribution for latitude however indicates the short time step always damps less, there are no negative values. The frequency distribution for the vertical trajectory also indicates the short time step always damps less; however almost all values are in the first positive bin, and the remainder have the log of the ratio within rounding of zero (the dot in the figure.)

Figure 13.27 shows the same frequency distribution but now for the $2.7\Delta x$ wave. In longitude, at 60% of the points the short time step damps less, but the damping ratio is between 1 and 1.26 (the first positive bin with log ranging from 0 to 0.1). At the remainder of the points the long time step damps less, and the ratio of damping is primarily distributed over bins with ratios 0.2 to 1 (log ranging from -0.7 to 0) with some ratios from 0.08 to 0.2 (log ranging from -1.1 to -0.7). In latitude, at

over 95% of the points the small time step damps less. As in longitude, the positive log damping ratio is in the first positive bin, where the damping ratio is between 1 and 1.26. In the vertical, the small time step damps less at 97% of the points. The ratio at the remaining points is within rounding of 1 (i.e., log is 0).

In summary, semi-Lagrangian integrations with short time steps do not necessarily damp shorter waves more than integrations with long time steps do. The different time steps yield different departure points and therefore different damping from the interpolations. The overall damping depends on the population of departure points which in turn depends on the atmospheric flow and model time step.

13.7 Fixers and Thoughts About Conservation Properties

It is generally desirable for a dynamical core to possess discrete analogues of the conservation properties of the adiabatic and frictionless continuous equations of motion as e.g., laid out in Chap. 11. However, the continuous equations possess an infinite number of invariants, such as mass, tracer mass, total energy, enstrophy and angular momentum just to name a few, whereas a numerical model can only conserve very few quantities. A straightforward way to ensure the conservation of an invariant is to choose it as a prognostic variable and utilize a flux-form finite volume discretization. Such a built-in conservation law is then a design feature of a dynamical core and its numerical scheme. This design decision needs to be carefully weighted against other beneficial properties like the computational efficiency or accurate wave dispersion characteristics.

Conservation can also be obtained through special mathematical properties of the numerical discretization. For example, spatial discretizations can be formulated so that they enforce the conservation of global integrals, such as mass, total energy and potential enstrophy (Arakawa 1966; Arakawa and Lamb 1981; Arakawa and Hsu 1990). This is also discussed in Chap. 12. The basic question is how accurately a dynamical core *needs* to capture various conservation properties and whether global conservation is sufficient or local conservation needs to be enforced. These issues are addressed in Thuburn (2008b) who gives guidance concerning the desirable conservation properties of GCMs.

In practice, there are many reasons why numerical models might lose even the most basic conservation properties of the continuous equations like the conservation of dry air mass, tracer mass or total energy. A prominent reason is that the equations are often not formulated in conservation form. But even if a conservation form is chosen the inevitable dissipation, either explicitly specified or inherent in the numerical schemes, and use of filters can violate the conservation in the discretized case (Takacs 1988). For example the conservation of mass principle is violated if the mass variable needs to be time-filtered or spatially filtered for numerical stability reasons. In addition, full GCMs contain physical parameterizations that represent the unresolved often dissipative processes at the subgrid-scale such as boundary layer turbulence. Kinetic energy is therefore generally lost due to dissipation which

might translate into a loss of total energy. Conservation must therefore be addressed not only in the dynamical core and its numerical discretization but also in full GCMs with physics packages.

If a conservation property is violated in a GCM, the global conservation can still be artificially recovered. This can be done through the use of so-called *fixers*. Fixers are modeling paradigms that allow an ad-hoc and a posteriori restoration of conserved quantities at each time step. There is no physical basis for such restorations other than that the conservation is a necessary or desirable characteristic of the GCM. A general expectation might be that the GCM simulations become more trustworthy if conservation properties are obeyed. This is especially true for the conservation of the dry air mass and the total energy which prevents the model climate from drifting into unrealistic states. However, the use of fixers does not imply that the physical processes and scale interactions are better represented. In addition, it is often imperative to fix unphysical negative tracer masses to prevent the model from "exploding" in the physical parameterizations.

This section discusses three types of a posteriori fixers that are broadly used in GCMs today. They are the mass fixers for dry air, filling algorithms for tracers and total energy fixers. Most often, the application of a fixer is an undocumented design feature of a GCM.

13.7.1 Dry Air Mass Fixer

In nature, dry air mass has no true physical sources and sinks, and is conserved regardless of diabatic or frictional processes. Conservation of dry air mass is probably the most fundamental conservation property that should be enforced in GCMs. In fact, the conservation of mass is paramount for long climate simulations where any drift in the total mass translates into a drift of the pressure distribution through the equation of state. This leads to spurious motions and artificial drifts of the simulated climate. For short weather predictions though, GCMs have put less emphasis on the conservation of mass. This design decision is probably justified since the drifts in the mass over short 10-day forecasts are generally negligible for practical purposes.

In the absence of sources and sinks, the mass of water vapor is conserved just as the mass of the dry atmosphere is. Since total air is a mixture of water vapor and dry air the conservation of both mass of water vapor and of dry air are often considered together. This is especially true if the moist surface pressure is the prognostic forecast variable. If a model prognoses the dry air pressure, a sole dry air mass fixer of course suffices.

The most popular dry air mass fixers are built upon an ad-hoc correction of the global surface pressure field regardless of the origin of the pressure drift which is often unknown. In models that predict $\ln p_s$ like CAM EUL or SLD this is done by adjusting the surface pressure at all grid points so that the gradient of the logarithm of the surface pressure field $\nabla \ln p_s$ is unaffected. The fixer thereby preserves the

gradients of the pressure gradient force. The latter is an important driver in the momentum equations that should not be arbitrarily changed. Such a mass fixer is for example documented in Williamson and Olson (1994). For brevity, we only present the design of the fixer for dry air masses. The extensions for moist air is shown in Williamson and Olson (1994), Collins et al. (2004) and Rasch et al. (1995).

Conceptually, the global dry air mass in hydrostatic models is represented by the global integral P of the dry surface pressure as given by

$$P(t) = \frac{1}{4\pi} \int_{-\pi/2}^{\pi/2} \int_{0}^{2\pi} p_{s\,dry}(\lambda, \phi, t) \cos\phi \, d\lambda d\phi \tag{13.126}$$

in spherical coordinates. The total mass, that needs to be conserved after each time step, is determined by the global mean surface pressure of the initial state $P(t = 0)$. Let t^+ denote the future time after completion of a time step but before the application of the mass fixer. The values of the surface pressure field at t^+ are therefore provisional. The surface pressure at the final future time step t^{j+1} is then fixed in the following way

$$p_s(\lambda, \phi, t^{j+1}) = M p_s(\lambda, \phi, t^+) \tag{13.127}$$

with the correction factor

$$M = \frac{P(t = 0)}{P(t^+)}. \tag{13.128}$$

Such a fixer is applied by default at each time step in NCAR's CAM Eulerian and semi-Lagrangian spectral transform dynamical cores. This formulation is only valid for hydrostatic dynamical cores.

To highlight the effects of the mass fixer on idealized dry dynamical core simulations we present results from both CAM EUL and CAM SLD (version 4) at the triangular truncation T85 ($\approx$ 156 km) with 26 levels (L26). In particular, the models were run with and without the mass fixer for 1,800 days and utilized the Held and Suarez (1994) forcing. All simulations start from identical initial conditions that contain a global mean dry surface pressure of 1,000 hPa. After 1,800 days, the Eulerian simulation without the mass fixer exhibits a global mean surface pressure of 999.9992 hPa which is a quite accurate. On the other hand, the SLD simulation without the mass fixer shows a steady, almost linear, increase in the amount of total mass. After 900 days the global mean surface pressure is 1005.29 hPa and increases to 1011.11 hPa by day 1,800. This is a substantial increase that would prevent credible climate simulations unless a mass fixer is employed. The mass is perfectly conserved in both dynamical core simulations with the mass fixer, as expected.

It is informative to evaluate the changes in the model climate due to the mass fixer. Such an assessment can reveal whether there are any systematic differences in the circulation when modeled with and without the mass fixer. Figures 13.28a-d and 13.29 show the zonal-mean 1200-day-mean temperature and zonal wind fields from the simulations with the mass fixer, and the differences between the runs with and without the mass fixer for both EUL and SLD. No total energy fixer is applied.

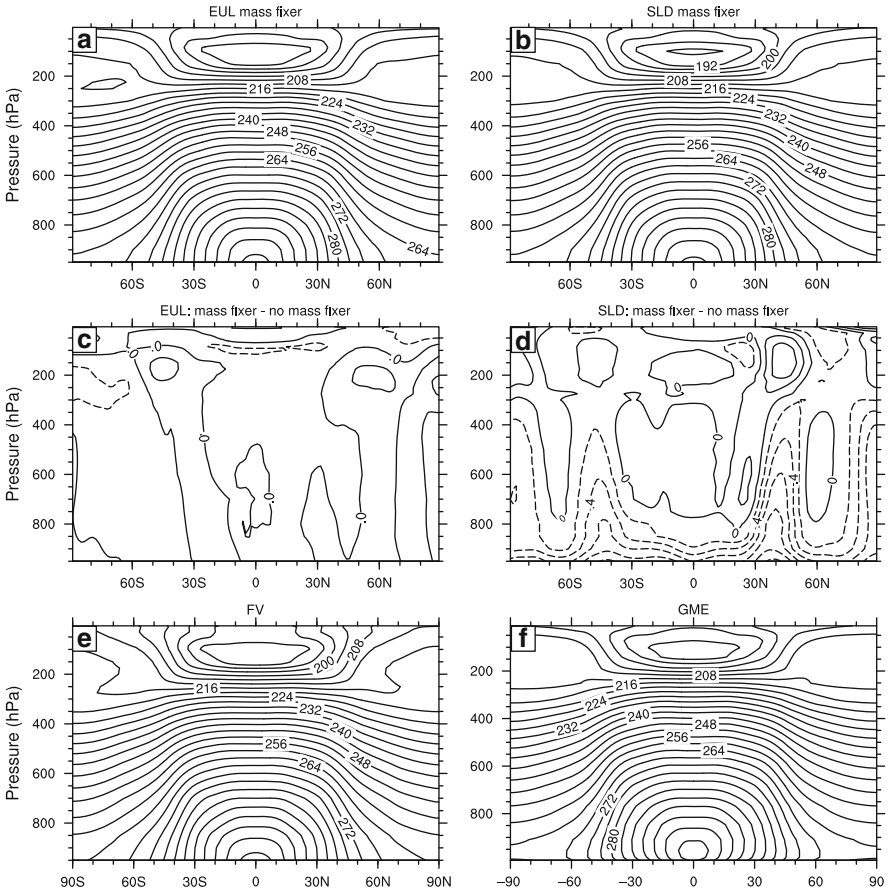

Fig. 13.28 Zonal-mean time-mean temperature field (K) forced with the Held-Suarez forcing: (**a**,**b**) CAM 4 EUL and CAM 4 SLD simulations at the resolution T85L26 with mass fixer, (**c**,**d**) difference between the EUL and SLD simulations with and without mass fixer, and comparisons to (**e**) CAM 4 FV at the resolution $1^\circ \times 1^\circ$ L26 and (**f**) GME at the resolution $ni = 64$ L19 ($\approx$ 120 km). (**a**–**d**) are 1200-day means, (**e**) is a 450-day mean, (**f**) 900-day mean. The contour intervals in (**a**,**b**,**e**,**f**) are 4 K and in (**c**,**d**) 0.2 K. Negative contours are *dashed*. The time step is $\Delta t = 600$ s (EUL) and $\Delta t = 1{,}800$ s (SLD)

The EUL simulation used the T85 default horizontal diffusion with the coefficients $K_4 = 1 \times 10^{15}$ m^4 s^{-1} and $K_2 = 2.5 \times 10^5$ m^2 s^{-1} whereas no explicit diffusion was utilized in the SLD run. Note again that the K_2 value is the base value at the third level from the top. It is doubled at the second level and doubled again at the top level. The 1200-day time averages incorporate day 600–1,800 of the model simulations. Furthermore, we show Held-Suarez temperature results from two additional models CAM 4 FV and GME without a mass fixer (Fig. 13.28e,f) to demonstrate that the EUL and SLD simulations are visually very similar to other models (Jablonowski

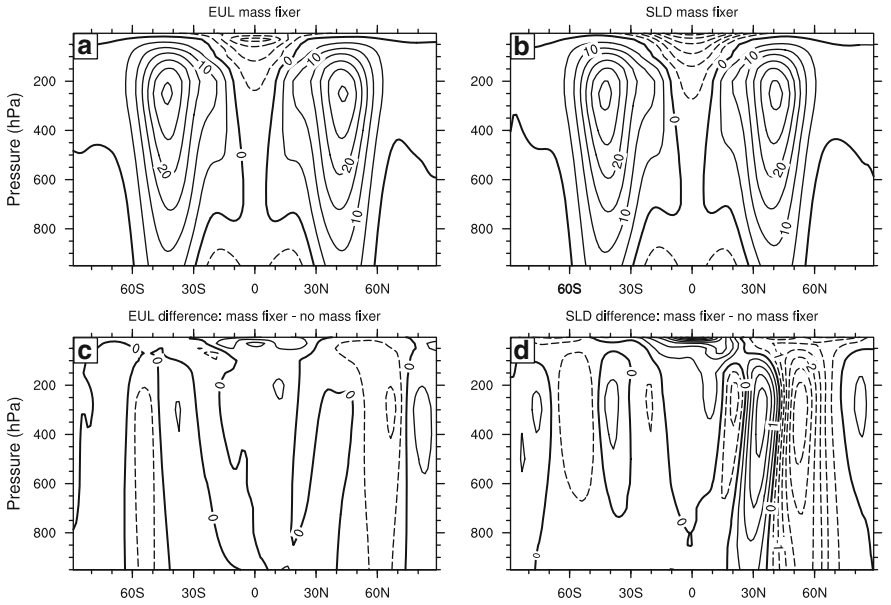

Fig. 13.29 Zonal-mean 1200-day-mean zonal wind (m/s) in CAM 4 EUL and CAM 4 SLD at the resolution T85L26 forced with the Held-Suarez forcing: (**a**,**b**) simulations with mass fixer, (**c**,**d**) difference between a simulation with and without mass fixer. The contour intervals are (**a**,**b**) 5 m/s and (**c**,**d**) 0.5 m/s. Negative contours are *dashed*. The time step is $\Delta t = 600$ s (EUL) and $\Delta t = 1{,}800$ s (SLD)

1998; Wan et al. 2008). FV is mass-conserving by design, and the small mass gain in GME over its 1440-day simulation period was on the order of 0.5 hPa.

The temperature and zonal wind difference plots for the EUL simulation shown in Figs. 13.28c and 13.29c suggest that the mass fixer has negligible impact on the mean EUL circulation. The temperature distributions and positions of the midlatitudinal jets in both hemispheres are almost identical in both runs. This is somewhat expected since the Eulerian simulation hardly lost mass over the 1800-day simulation so the impact of the mass fixer should be minimal. A clearer modulation of the mean circulation can be seen in the corresponding SLD runs (Figs. 13.28d and 13.29d). The simulation with the mass fixer appears to be slightly warmer in the upper atmosphere whereas the lower troposphere and the midlatitudes up to 400 hPa show systematically colder temperatures. In particular, the SLD mass fixer run is about 0.4–0.8 K colder in the levels below 800 hPa. The jets in the zonal wind field are slightly shifted equatorwards in the SLD mass fixer run (Fig. 13.29d). However, it is unclear whether these shifts are statistically significant without further investigations. As an aside, the differences in the zonal wind fields between the EUL and SLD runs (Figs. 13.29a,b) close to the model top are caused by the sponge-layer diffusion in the Eulerian model which is discussed in detail in Sect. 13.4.5.

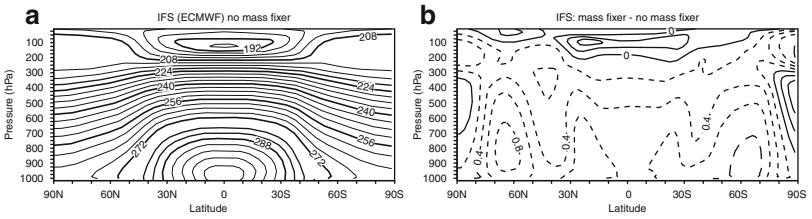

Fig. 13.30 Zonal-mean 1000-day-mean temperature field (K) in ECMWF's IFS model cyle 18R3 at the resolution T63L31 forced with the Held-Suarez forcing: (**a**) simulations without the mass fixer, (**b**) difference between a simulation with and without mass fixer. The contour intervals are (**a**) 4 K and (**b**) 0.2 K. Negative contours are *dashed*

To put the CAM 4 SLD changes due to the mass fixer into perspective we also show results from an older version of ECMWF's model IFS (cycle 18R3, November 1997). The dynamical core of IFS is a two-time-level semi-implicit semi-Lagrangian spectral transform model and therefore very similar to CAM SLD. The Held-Suarez test was run for 1,200 days with and without a mass fixer at the triangular truncation T63 ($\approx$ 210 km) on a reduced Gaussian grid with 31 vertical levels (L31). The specific design of the reduced grid and semi-Lagrangian model are explained in Hortal and Simmons (1991) and Hortal (2002), and are not of interest for the following discussion. Here, we solely concentrate on the effect of the mass fixer on the model climate. The IFS mass fixer follows the identical design principle as CAM's EUL/SLD mass fixer. Similar to CAM SLD, the *unfixed* IFS shows a systematic, almost linear, increase in mass over the 1200-day forecast period. The rate is $+0.012\%$ per 10 days. Assuming an initial global surface pressure of 1000 hPa this amounts to about 1014.4 hPa after 1,200 days. This increase in mass is resolution dependent. At the higher triangular truncation T106 ($\approx$ 125 km) the rate is reduced to $+0.0079\%$ per 10 days which yields a global surface pressure of 1009.5 hPa by day 1200. The changes in IFS's mass are thereby slightly higher but comparable to the changes of the mass in CAM 4 SLD.

Figure 13.30 shows the zonal-mean 1000-day-mean temperature distribution of the IFS run without the mass fixer and the temperature differences between the runs with and without the mass fixer. The overall temperature distribution in Fig. 13.30a resembles the CAM EUL and SLD runs with similar peak temperatures. The temperature difference in Fig. 13.30b also exhibits some structural resemblance to the SLD difference plot (Fig. 13.28d). The IFS mass fixer run is systematically colder throughout the lower and middle troposphere, and warmer near the poles and near the tropopause. The cold temperature difference peaks in midlatitudes with a magnitude of 0.8 K. Note, that the use of the mass fixer in SLD and IFS would be paramount for long climate simulations. Most likely, the warming and cooling signatures in the runs without the mass fixer are entirely spurious and related to the unphysical gain in mass. As an aside, the design of IFS's dynamical core has only slightly changed in comparison to more current hydrostatic versions of IFS (ECMWF 2010).

13.7.2 Filling Algorithms for Tracers

Mass conservation is one of the most important design aspects of tracer transport algorithms. If a scheme is nonconservative, it can significantly underestimate or overestimate the concentration of trace gases in long time integrations. This is particularly true if the transported quantity has a large residence time in the atmosphere like methane or nitrous oxide.

An additional desirable characteristic of transport schemes is monotonicity and thereby the prevention of non-physical under- and overshoots in the solution. They can lead to negative trace constituents or even the supersaturation in the humidity field. In particular, negative mixing ratios are undesirable since physical parameterizations cannot deal with e.g., negative moisture quantities. Ideally, both monotonic (also called shape-preserving or non-oscillatory) and mass-conserving advection schemes should therefore be used to assure the physical consistency of the advection process, and prevent negative tracer constituents from occurring in the first place. Examples are the flux-corrected advection scheme of Zalesak (1979) or the mass-conservative and monotonic semi-Lagrangian advection scheme by Lauritzen et al. (2010b). Alternatively, positive definite advection algorithms can be employed that, at least, prevent negative undershoots. A comprehensive overview of possible advection algorithms and their characteristics is given by Rood (1987) and Chap. 8.

In case negative tracer constituents occur during a model integration an a posteriori borrowing and filling algorithm is most often employed to *fix* the negative tracer mass. The basic idea is that a grid box with negative tracer values is filled to a minimum small positive value, and an equivalent amount is subtracted (borrowed) from other grid cells. This ensures that the total constituent mass remains the same. However, it does not eliminate overshoots or undershoots that are associated with non-negative parts of the field. Such a fixer is characterized as a *conservative* fixer. Fixers might also be used as *positivity* fixers that only obey a positive-definite constraint but neglect global conservation.

The mixing induced by both types of fixers can trigger nonlinear interactions. For example, Rasch and Williamson (1990a) showed that positivity fixers can greatly influence the water vapor budget in a spectral transport scheme due to the interactions of the fixed specific humidity field with physical parameterizations. This was especially true in the polar regions which are rather dry. The fixer operated as a moisture transport algorithm, yielding a local moisture source, and although it only brought the moisture up to a positive minimum value at negative points, it increased the overall moisture in the polar regions with strong impact on the clouds and precipitation.

Examples of tracer filling algorithms for the use in GCMs are presented in Mahlmann and Sinclair (1977), Royer (1986), Rasch and Williamson (1990a) and Rasch and Williamson (1991). They can be either local or global, and resemble a nonlinear diffusive process. Both, monotonicity constraints and filling algorithms are designed to control numerical dispersion errors, and could also be viewed as an implicit or explicit numerical filter, respectively. Note that even tracer advection algorithms that are strictly monotonic in one dimension might lead to violations of

the monotonicity when applied in multiple dimensions. Therefore, even schemes with monotonicity constraints such as the finite volume advection algorithm in the CAM FV dynamical core can trigger small under- and overshoots in multiple dimensions as reported in Lin and Rood (1996). Therefore, CAM FV applies a local filling algorithms to eliminate negative tracer masses as documented in Neale et al. (2010).

Another highly desirable property of a tracer advection scheme is that it should be consistent with the mass continuity equation. This is for example outlined in Lin and Rood (1996), Jöckel et al. (2001), Gross et al. (2002), Satoh et al. (2008) and Chap. 8, and is not discussed in detail here. As an aside, the expression "borrow and fill" is somewhat misleading since the amount taken from a neighboring grid cell is never given back to the cell it is taken from. A more appropriate description is "take and fill" or even "steal and fill" (Fedor Mesinger, personal communication). However, we stick with the expression "borrow and fill" in this chapter since it is widely used throughout the literature.

13.7.2.1 Local Filling Algorithms

Local filling algorithms are "borrow and fill" fixers that try to borrow mass primarily from the four nearest cells in the east, west, north, and south directions at the same model layer. However, if there is insufficient mass they might also borrow from a level below or above. Most often, the filling algorithm starts downward from the model top as described in Rasch and Williamson (1990a). It is not straightforward to write down a concise set of equations to describe a local borrower scheme. Therefore, we only describe the underlying ideas and present two variants documented in the literature. Note that there are many additional variants in practice, but borrowing schemes are not necessarily documented and should ideally be avoided with the help of improved tracer advection algorithms.

Variant 1 After identifying a grid cell with negative tracer mass the surrounding neighbors with positive tracer masses are determined and an equal percentage is borrowed from each. The negative mass is set to small minimum value. If there is insufficient mass in the neighboring cells, no action is taken and the next point is evaluated. The borrowing might be limited to one-third of the total mass available in the neighboring cells as suggested by Reames and Zapotocny (1999).

Variant 2 Reames and Zapotocny (1999) also tested a borrowing scheme that weighted the borrowed mass by the amount available in the neighboring boxes and by the velocity components toward these grid cells. This extends a suggestion by Mahlmann and Sinclair (1977) who argued that borrowing should first come from a neighboring grid cell that is downstream. Only if there is not enough mass in the downstream direction, mass is borrowed from an upstream point or even a more distant point.

The local filling algorithm can still leave residual negative values in case immediate neighbours do not have enough mass to fill a point. An additional global borrowing scheme can then be used to remove this residual as explained next. Of course, a

global borrowing scheme can also be used by itself without a local filling algorithm. As pointed out by Rood (1987) a characteristic of local filling algorithms is that many decisions have to be made. Therefore, a local filling scheme is quite expensive from a computational viewpoint. It can also violate the monotonicity of the tracer field.

13.7.2.2 Global Filling Algorithms

Global "borrow and fill" algorithms are less time-consuming but require global communication on parallel computing architectures. There are two classes of global borrowing schemes which can be characterized as a subtractive and multiplicative method as outlined by Rood (1987). First, the total area- or volume-weighted negative (N) and positive (P) tracer masses are computed and negative tracer values are set to a small minimum value. In the subtractive method, a fraction of the extra mass N is then subtracted from all grid points from which mass can be subtracted without creating new negative values. This subtraction needs to obey the constraint that the total mass is constant after the correction which could require some further searching. In the multiplicative method, each positive tracer value q^+ is replaced by $(1 - N/P)q^+$ which automatically conserves the global integral of the constituent mass and guarantees that the new tracer distribution is positive definite. The correction can either be applied level-by-level or globally in case the volume-weights are computed (Rasch and Williamson 1990b). Such a multiplicative fixer was for example used by Zubov et al. (1999) and Reames and Zapotocny (1999) .

Rasch and Williamson (1991) found that this variational adjustment of the tracer mass does neither improve nor degrade the accuracy of an unadjusted tracer transport scheme, but that it merely enforces the conservation and positivity constraints. However, Jöckel et al. (2001) argued that simple mass fixer algorithms for tracers have severe disadvantages since they either violate the monotonicity requirement or introduce non-physical transport. It is therefore best to select a conservative tracer transport scheme that is also consistent with the mass continuity equation as outlined in Chap. 8. It means that the discretized tracer transport scheme should reduce to the discretized continuity equation for air if the tracer mixing ratio is unity. The discretization then mimics the characteristics of the continuous equations.

13.7.3 Total Energy Fixers

Global invariants like the total energy provide useful constraints on the design of numerical schemes which makes a built-in conservation of total energy a desirable choice for GCMs. As suggested by Thuburn (2008b) the conservation of total energy and enstrophy in full GCMs might even play a major role in accurately capturing nonlinear transfers to small scales. However in practice, many aspects of today's GCMs and in particular their dynamical cores are not energy-conserving. This

includes the horizontal diffusion in case of neglected dissipative heating, the spatial and time filters, the time differencing, inherent numerical dissipation, or the discretization technique for the energy conversion term. The latter aspect is emphasized in Simmons and Burridge (1981) who developed an energy and angular-momentum conserving vertical finite-difference scheme.

Total energy conservation is paramount for long climate runs to prevent drifts in the mean circulation (Boville 2000). In early GCMs though, energy conservation was not a significant design consideration. For example the energy imbalance in an early version of NCAR's Community Climate Model CCM0 was $10\,\mathrm{W\,m^{-2}}$ as reported in Williamson (1988). The loss in energy originated primarily from inconsistencies in the vertical numerical approximations. In later versions, energy conservation received more attention, as e.g., version 1 of the Community System Model (CSM1). It conserved energy to about $0.4\,\mathrm{W\,m^{-2}}$ as analyzed by Boville and Gent (1998) but this level of conservation was partly due to a cancellation of errors. Today, total energy conservation has become a serious concern since even small imbalances of order $0.4\,\mathrm{W\,m^{-2}}$ can cause spurious long-term trends in multicentury coupled ocean-atmosphere simulations.

In general, the variation of energy during a simulation can manifest itself as either a gain or loss of total energy. Most often though, energy is lost which is mainly attributable to the kinetic energy dissipation. Kinetic energy dissipation is due to explicit horizontal diffusion, inherent dissipation in the numerical approximations and filters. In practice, it averages to an energy loss of about $2\,\mathrm{W\,m^{-2}}$ in the three CAM dynamical cores EUL, SLD and FV when applied at typical climate resolutions (Williamson 2007). This amount is clearly not negligible and only the contribution from the explicit horizontal diffusion can be analytically quantified. Some models therefore include a frictional heating term associated with the explicit horizontal momentum diffusion (Collins et al. 2004) as also outlined in Sect.13.3.7. A thorough review of the kinetic energy dissipation in NCAR's CAM model and the required compensating heating is provided in Boville and Bretherton (2003). As an aside, Bowler et al. (2009) estimated that the energy dissipation due to the interpolation error alone in the semi-Lagrangian advection scheme in the UK Met Office model is about $0.75\,\mathrm{W\,m^{-2}}$. They suggested using a stochastic kinetic energy backscatter scheme to reintroduce the missing energy from the explicit horizontal diffusion and semi-Lagrangian interpolations, partly into the resolved scales of the flow. An alternative energy backscattering method was also presented by Shutts (2005).

In order to maintain the energy balance, the globally averaged inherent dissipative heating can be determined via a residual calculation. The heating is then added to the temperature field in the thermodynamic equation. This can be done as either a globally uniform heating or cooling, or another ad hoc function. The choice of such ad hoc functions is arbitrary but there are adequate and inadequate choices. As revealed in Williamson et al. (2009) with the help of an idealized dynamical core test, an inadequate "bad" energy fixer has detrimental effects on the circulation. This was not obvious by a pure inspection of the ad hoc correction algorithm and not immediately obvious in full GCM runs with physical parameterizations.

To remedy the effect, Williamson et al. (2009) recommended using only very simple corrections like uniform adjustments at all grid points in the global domain. We discuss the "wrong" and "right" choices below, but first start with a brief review of the total energy equation.

13.7.3.1 Different Forms of the Total Energy Equation

The total energy equation for dry air can be obtained by adding the kinetic and potential energy equations to the first law of thermodynamics. The derivations of these equations may be found in Gill (1982). Here, we present the form of the total energy equation for adiabatic and hydrostatic dynamical cores that utilize the primitive equations.

As pointed out by Staniforth et al. (2003) total energy is formally only conserved if the model employs a rigid lid as the upper boundary condition. Such a rigid lid needs to be fixed in time and space but is allowed to vary with latitude. Models with elastic isobaric lids, like the popular choice of the upper boundary at constant pressure, are formally non-energy-conserving. But an energy-like invariant exists that gives these approaches merit (Kasahara 1974; Laprise and Girard 1990). The specific form of the total energy equation is tightly linked to the choice of the vertical coordinate. This is outlined in Arakawa and Suarez (1983) and briefly shown for pressure- and height-based vertical coordinates below.

In the continuous primitive equations with the hybrid pressure-based vertical coordinate η (Simmons and Burridge 1981) total energy is conserved if the following relationship holds (Laprise and Girard 1990)

$$\frac{\partial}{\partial t}\int_A\int_{\eta_{top}}^{\eta_s}\left(\frac{\mathbf{v}^2}{2}+c_pT\right)\frac{\partial p}{\partial \eta}\,d\eta\,dA=-\int_A\left(\Phi_s\frac{\partial p_s}{\partial t}-\Phi_{top}\frac{\partial p_{top}}{\partial t}\right)dA. \quad (13.129)$$

This equation is valid in the absence of diabatic and frictional effects. Φ_s, p_s and Φ_{top}, p_{top} are the geopotential and pressure at the surface and the model top, c_p is the specific heat of dry air at constant pressure and $\mathbf{v}=(u,v)$ stands for the horizontal velocity vector with the zonal and meridional wind components u and v. Furthermore, T symbolizes the temperature, p is the pressure, and t denotes the time. The integrals span the 3D and 2D domains where A symbolizes the horizontal area of the sphere. The vertical integral is bounded by the value η_s at the surface and η_{top} at the model top. Here, η_s is identical to unity and η_{top} is equivalent to p_{top}/p_0 with reference pressure $p_0=1000\,\mathrm{hPa}$. Note that η_{top} is non-zero for constant $p_{top}>0\,\mathrm{hPa}$. A constant pressure at the model top ensures the global conservation of total energy in the continuous equations and simplifies the 2D integral. Equation (13.129) then becomes

$$\frac{\partial}{\partial t}\left\{\int_A\frac{1}{g}\left[\left(\int_{\eta_{top}}^{\eta_s}\left(\frac{\mathbf{v}^2}{2}+c_pT\right)\frac{\partial p}{\partial \eta}\,d\eta\right)+\Phi_s p_s\right]dA\right\}=0. \quad (13.130)$$

Here we divided (13.129) by the gravity g to recover energy units (Kasahara 1974). This expression is equivalent to $\partial(TE)/\partial t = 0$ where TE symbolizes the global integral of the total energy as shown by the term in the curly bracket in (13.130). In the semi-discrete system with $\partial p/\partial \eta \approx \Delta p/\Delta \eta$ and $d\eta \approx \Delta \eta$, the domain-integrated total energy TE is given by

$$TE = \int_A \frac{1}{g}\left[\left(\sum_{k=1}^{K_{max}} \left(\frac{u_k^2 + v_k^2}{2} + c_p T_k\right)\Delta p_k\right) + \Phi_s p_s\right] dA. \tag{13.131}$$

The summation index k indicates the vertical index of a full model level with the maximum level number K_{max} near the surface. The pressure difference Δp_k is defined as

$$\Delta p_k = p_{k+1/2} - p_{k-1/2} = p_0\, \Delta A_k + p_s\, \Delta B_k \tag{13.132}$$

with $\Delta A_k = A_{k+1/2} - A_{k-1/2}$ and $\Delta B_k = B_{k+1/2} - B_{k-1/2}$. As an example, the discrete positions of the hybrid coefficients $A_{k+1/2}$ and $B_{k+1/2}$ at the model interface levels for the CAM EUL, SLD and FV dynamical cores (versions 3.1 and 4) are listed in Jablonowski and Williamson (2006b). $\Delta \eta_k$ is given by $\Delta \eta_k = \eta_{k+1/2} - \eta_{k-1/2} = \Delta A_k + \Delta B_k$. Note that the form of the domain-integrated total energy equation TE in the optional CAM 5 dynamical core HOMME (see (12.8) in Chap. 12) differs from (13.130). The main difference is that HOMME utilizes the parameter $[c_p^* = c_p + (c_{pv} - c_p)q]$ instead of c_p where c_p^* symbolizes the specific heat of moist air at constant pressure, c_{pv} denotes the specific heat of water vapor at constant pressure and q stands for the specific humidity.

If hydrostatic models with pressure-based σ coordinates like

$$\sigma = \frac{p - p_{top}}{p_s - p_{top}} \tag{13.133}$$

are considered (Phillips 1957; Kasahara 1974) the global integral of the dry total energy becomes

$$TE = \frac{1}{g} \int_A \int_{\sigma_{top}}^{\sigma_s} \left(\frac{\mathbf{v}^2}{2} + c_p T + \Phi_s\right) \frac{\partial p}{\partial \sigma}\, d\sigma\, dA \tag{13.134}$$

$$\approx \frac{1}{g} \int_A \left[\sum_{k=1}^{K_{max}} \left(\frac{u_k^2 + v_k^2}{2} + c_p T_k + \Phi_s\right)\Delta p_k\right] dA. \tag{13.135}$$

The pressure differences Δp_k are determined by

$$\Delta p_k = p_{k+1/2} - p_{k-1/2} = (\sigma_{k+1/2} - \sigma_{k-1/2})\,(p_s - p_{top}). \tag{13.136}$$

Note that the conservation of energy requires $\sigma_{top} = \sigma(p = p_{top}) = 0$ which is guaranteed in (13.133). The lower boundary at the surface is $\sigma_s = \sigma(p = p_s) = 1$.

As noted in Kasahara (1974) the global integral of the total energy for hydrostatic models with a pure height coordinate z in the vertical direction is represented by

$$TE = \int_A \int_{z_s}^{z_{top}} \left(\frac{\mathbf{v}^2}{2} + c_v T + gz \right) \rho \, dz \, dA \tag{13.137}$$

$$\approx \int_A \left[\sum_{k=1}^{K_{max}} \left(\frac{u_k^2 + v_k^2}{2} + c_v T_k + g z_k \right) \rho_k \, \Delta z_k \right] dA \tag{13.138}$$

where Δz_k symbolizes the height thickness of a layer with $\Delta z_k = z_{k-1/2} - z_{k+1/2}$ and z_{top} stands for the height of the model top. The quantities $\rho \, \mathbf{v}^2/2$, $\rho \, c_v T$ and $\rho g z$ are the kinetic, internal and potential energy per unit volume, respectively. c_v is the specific heat of dry air at constant volume and defined by $c_v = c_p - R_d$. R_d is the gas constant for dry air and ρ denotes the density which is defined by the ideal gas law $\rho = p/(R_d \, T)$. z_s is the height of the orography. If height-based orography-following coordinates like

$$\xi = \frac{z_{top} - z}{z_{top} - z_s} \tag{13.139}$$

are used the domain integral of the dry total energy transforms to

$$TE = \int_A \int_{\xi_{top}}^{\xi_s} \left(\frac{\mathbf{v}^2}{2} + c_v T + gz \right) \rho \, d\xi \, dA \tag{13.140}$$

$$\approx \int_A \left[\sum_{k=1}^{K_{max}} \left(\frac{u_k^2 + v_k^2}{2} + c_v T_k + g z_k \right) \rho_k \, \Delta \xi_k \right] dA \tag{13.141}$$

with the lower and upper integration limits $\xi_s = \xi(z = z_s)$ and $\xi_{top} = \xi(z = z_{top})$. $\Delta \xi_k = \xi_{k+1/2} - \xi_{k-1/2}$ is the thickness of a layer in the transformed ξ-coordinate.

As an aside, the full 3D velocity vector $\mathbf{v}_{3D} = (u, v, w)$ needs to be used for the computation of the kinetic energy in nonhydrostatic dynamical cores. A discussion of the total energy equation for dry shallow- and deep-atmosphere nonhydrostatic equation sets is provided in Staniforth et al. (2003). For moist dynamical cores the assessment of the total energy needs to be adjusted. Then p represents the pressure of the moist air and $\rho = p/(R_d \, T_v)$ is the moist density which utilizes the virtual temperature T_v. Other modifications might include the use of moist physical constants such as c_p^* as indicated above for the model HOMME. Satoh et al. (2008) discussed a particular form of the moist total energy equation for the nonhydrostatic dynamical core NICAM (their appendix B). Unfortunately, GCMs use many different and often undocumented approaches to calculating the total energy of moist air. For example, some models include the latent heat contributions Lq in TE where L symbolizes the latent heat of vaporization. Some models also add the energy contributions from cloud liquid water or even cloud ice. On the other hand, some models

do not include either of these. In this chapter we avoid this confusion by focusing on the dry dynamical core. This allows us to assess the impact of the total energy fixer for dry air in isolation.

13.7.3.2 Ad Hoc Corrections of Total Energy

We now present three ad hoc functions for the total energy correction that we label FIXER 1–3, and demonstrate that FIXER 1 leads to bad results. This is shown via dry idealized baroclinic wave simulations (Jablonowski and Williamson 2006a) with NCAR's CAM semi-Lagrangian spectral transform dynamical core at the resolution T170 L26. First, we assess the spirit of the three total energy fixers and review how they are applied. This is also discussed in Williamson et al. (2009).

Let $\left(\hat{T}^+, \hat{\mathbf{v}}^+, \hat{p}_s^+\right)$ symbolize the temperature, horizontal wind vector and surface pressure at the end of a time step and $\left(T^-, \mathbf{v}^-, p_s^-\right)$ denote the values at the beginning of the time step. In total energy-conserving model formulations the residual

$$RES = \widehat{TE}^+ - TE^- \tag{13.142}$$

would be zero if there are no diabatic sources and sinks. Note that Collins et al. (2004) describe how to include sources and sinks which is not discussed for the dynamical cores considered here. However, if GCMs with physical parameterizations or idealized forcings like the Held and Suarez (1994) forcing are utilized these sources and sinks need to be added. According to the total energy equation for hybrid coordinates (13.131) $\widehat{TE}^+$ is given by

$$\widehat{TE}^+ = \int_A \frac{1}{g} \left\{ \left[\sum_{k=1}^{K_{max}} \left(\frac{(\hat{\mathbf{v}}_k^+)^2}{2} + c_p \hat{T}_k^+ \right) (p_0 \, \Delta A_k + \hat{p}_s^+ \, \Delta B_k) \right] + \Phi_s \hat{p}_s^+ \right\} dA. \tag{13.143}$$

The equation for TE^- is identical to (13.143) except the superscript $+$ is replaced by superscript $-$, and the hat ($\hat{}$) is removed from TE, $\mathbf{v}$, T and p_s.

In general, RES is not zero due to the inherent and explicitly imposed diffusion processes in the dynamical cores. Therefore, modifications can be made to the provisional forecast values $\left(\hat{T}^+, \hat{\mathbf{v}}^+, \hat{p}_s^+\right)$. This adjustment yields updated values $\left(T^+, \mathbf{v}^+, p_s^+\right)$ which, if substituted for the provisional values in (13.142), yield a zero residual. This is the underlying concept of the energy fixer.

The form of the energy fixer used with the CAM semi-Lagrangian model only modifies the temperature. This could be interpreted as diffusive heating in case of a positive temperature adjustment. However, in case of cooling no such physical analogy can be drawn. The future wind and surface pressure fields are set to $\mathbf{v}^+ = \hat{\mathbf{v}}^+$ and $p_s^+ = M \hat{p}_s^+$ where M symbolizes the mass fixer if applied (see also Sect. 13.7.1).

We now define FIXER 1. Its temperature modifications are proportional to the magnitude of the local change in T at that time step and are given by

$$T^{+}(\lambda,\phi,\eta) = \hat{T}^{+}(\lambda,\phi,\eta) + \beta_1 |\hat{T}^{+}(\lambda,\phi,\eta) - T^{-}(\lambda,\phi,\eta)|. \tag{13.144}$$

This temperature adjustment follows the spirit of the water vapor fixer developed by Rasch and Williamson (1991) and Williamson and Rasch (1994) for predecessors of CAM 3. Its physical motivation follows the argument that if the change in water vapor is small, then it is most likely not responsible for a lack of conservation and therefore the effect of the fixer should be small. The constant β_1 is determined by replacing $\hat{T}^{+}$ with T^{+} in (13.143) and setting $RES = 0$ in (13.142). Equation (13.144) is then substituted for T^{+} in (13.143). Lastly, (13.143) and the equation for TE^{-} are plugged into (13.142) which is solved for β_1.

Alternatively, two other total energy fixers are suggested. FIXER 2 is given by

$$T^{+}(\lambda,\phi,\eta) = \hat{T}^{+}(\lambda,\phi,\eta) + \beta_2 \tag{13.145}$$

and FIXER 3 is formulated as

$$T^{+}(\lambda,\phi,\eta) = (1+\beta_3)\,\hat{T}^{+}(\lambda,\phi,\eta)\,. \tag{13.146}$$

The energy fixer FIXER 2 changes the provisional temperature by a constant, whereas FIXER 3 changes it proportionally. Both constants β_2 and β_3 are determined as described above for FIXER 1 and β_2 is shown in Williamson et al. (2009). FIXER 2 is adopted in the EUL and SLD dynamical cores of NCAR's CAM 3.1 GCM and was used operationally for climate applications. The FV dynamical core in CAM 4 and CAM 5 applies a slightly different total energy fixer that is described in Neale et al. (2010).

Despite the pseudo-physical justification for FIXER 1, it has detrimental effects on the flow fields in the baroclinic wave test case which which had not been obvious in long Earth-like simulations. Figure 13.31 shows the surface pressure field at day 10 from the SLD baroclinic wave simulation with FIXER 1 and FIXER 2. The two are clearly different, whereas the FIXER 2 simulation is visually indistinguishable

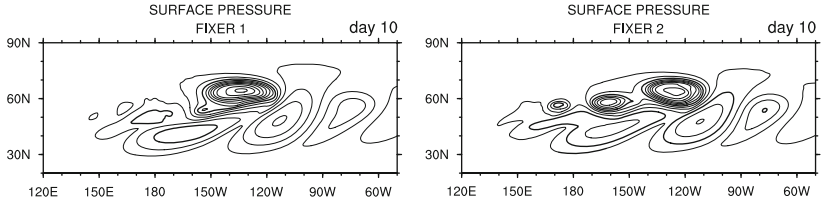

Fig. 13.31 Surface pressure at day 10 from the CAM semi-Lagrangian spectral dynamical core at the resolution T170L26 with energy fixer FIXER 1 and FIXER 2. Contour interval is 7.5 hPa with the 980 hPa and 1,010 hPa contours *thicker*. The time step is $\Delta t = 900$ s

from FIXER 3, from simulations without an energy fixer or from CAM EUL and FV reference simulations as shown in Williamson et al. (2009) and Jablonowski and Williamson (2006a,b). Therefore, FIXER 1 should not be used in practice. Here, we list it for demonstration purposes to raise awareness for potential problems with energy fixers.

The negative effect of FIXER 1 can be understood when recognizing that most of the loss of total energy is associated with damping of the wind field and the reduction of kinetic energy. However, with semi-Lagrangian approximations the damping is due in large part to the interpolations fundamental to the numerical method. There is no physical argument to justify making the fixer proportional to the change in temperature as tried in FIXER 1. Although the changes of the energy by the fixers are very small, the energy change due to FIXER 1 is systematic and obviously accumulates to a significant error. The short dynamical core run was able to isolate this problem. In full GCMs with physical parameterizations such detrimental effects of the total energy fixer are masked by many other processes and harder to identify.

Note again that the application of fixers might be unavoidable in long climate simulations to prevent unphysical signals from systematic mass or energy drifts. This is a strong argument for mass-conserving designs of dynamical cores to at least avoid mass-fixing the GCM. The latter is conceptually straightforward and some examples of mass-conserving dynamical cores include FV, FVcubed, HOMME, ICON, NICAM and WRF. NICAM also employs a total energy conservation form (Satoh 2002, 2003). Such a choice of the prognostic equation for total energy written in flux form automatically guarantees the conservation of total energy provided no explicitly added dissipation or filters are utilized. However, built-in total energy conservation in full GCMs with explicit diffusion or filters, and physical parameterizations is harder to accomplish and might not easily be achievable. Today, most climate models use a total energy fixer which is often undocumented and sometimes called with a different name. For example, the Unified Model (UM) has a total energy "correction" which is occasionally, e.g. once a day, applied (Terry Davies, personal communication). It is viewed as a correction that substitutes all missing physics processes.

13.8 Final Thoughts

The intention of this chapter was to remind the GCM modeling community of the many, sometimes hidden, diffusion processes in the dynamical cores of atmospheric general circulation models. There is no universal theory that guides the design of subgrid-scale diffusion, dissipation, mixing, damping, smoothing, filters or fixers, or however we name the many ad hoc mechanisms in GCMs. They are needed to keep the simulation stable or to satisfy important physical properties, and the hope is that they capture and mimic in some unknown way the true processes at the unresolved subgrid scale. There is no physical basis though, that dissipation can accomplish this. Therefore, a lesser goal is that the dissipative processes keep

a GCM simulation stable and promote its physical realism, while doing no harm to the resolved fluid flow. The latter is a practical approach, and this chapter highlighted the many considerations that contribute to the practical designs of diffusion, filters and fixers in GCMs. In practice, diffusion processes and filtering in atmospheric models are based on a subjective determination of when the noise and short waves have been sufficiently damped while minimizing the damping of longer wave modes. This might be described as the "Art of Filtering". The selected filters are tuned for the grid resolutions and flow regimes, and this tuning is based on modeling experience.

There is no numerical scheme or diffusive process that incorporates an optimal solution to the filtering problem at all scales. An optimal formulation at a particular resolution or for a specific flow regime might fail if conditions and scales change. Adequate diffusion and filtering mechanisms are scale-dependent. For example, 2D divergence damping is considered adequate for large-scale hydrostatic motions where rotational motions are of main interest. However, small-scale mesoscale regimes are highly influenced by divergent motions, and 2D divergence damping might become detrimental at those scales in nonhydrostatic models. This remains to be seen as we enter the nonhydrostatic modeling era with future global grid resolution near the kilometer scale.

In general, there is no "right" or "wrong" solution to the subgrid-scale filtering problem. All approaches shown in this chapter have been tried in GCMs, but some have more merit or physical motivations than others. There might also be approaches that are clearly adequate or inadequate for a particular model design. In practice, GCMs apply a potpourri of damping mechanisms, either explicitly or implicitly in the numerical scheme. They act in concert, interact in nonlinear ways, and the causes and effects are usually hard to isolate individually. This chapter provided a comprehensive and systematic evaluation of the many dissipative processes in GCMs and evaluated their impact on the flow with the help of idealized test cases. It is difficult to judge how generally applicable the results of our study are with respect to full GCMs with physical parameterizations. We would expect though that many of the conclusions transfer and that new dissipative processes like boundary layer diffusion make the picture even more complex. This will require another round of evaluations, but to say the least, the dynamical core assessments might have given us new guidance and clues for the evaluations of full weather and climate models.

Acknowledgments We would like to thank Jerry Olson (NCAR) for developing the required semi-Lagrangian CAM 3.1 codes and for running the experiments required in Sects. 13.3.8 and 13.6.4. We would also like to thank Mike Blackburn (University of Reading) for discussions on the University of Reading spectral model and for pointing out relevant papers. We thank Fedor Mesinger (NCEP) and the second reviewer for their very insightful suggestions. DLW was partially supported by the Office of Science (BER), U.S. Department of Energy, Cooperative Agreement No. DE-FC02-97ER62402. CJ was supported by the Office of Science (BER), U.S. Department of Energy, Award No. DE-FG02-07ER64446. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Appendix: Overview of Selected Dynamical Cores

This chapter has featured examples from many GCMs. Here, we briefly review the basic characteristics of their dynamical cores and give pointers to their primary references. Note that the grid staggering options mentioned below are shown in Randall (1994) or Chap. 3.

CAM EUL The Eulerian (EUL) dynamical core is the default in the Community Atmosphere Model CAM version 3.1 at the National Center for Atmospheric Research (NCAR). It is optional in the later versions CAM 4 or 5 (Neale et al. 2010). This hydrostatic shallow-atmosphere dynamical core is formulated in vorticity-divergence form and based on the traditional leapfrog three-time-level, semi-implicit spectral transform approximations (Machenhauer 1979). A quadratically-unaliased Gaussian transform grid with horizontal triangular truncation is utilized (Collins et al. 2004). In the vertical direction, centered finite differences are used. All prognostic variables are co-located.

CAM SLD The semi-Lagrangian (SLD) dynamical core is an optional dynamical core in NCAR's CAM model (versions 3.1, 4 or 5). It utilizes the primitive equations and is based on two-time-level, semi-implicit semi-Lagrangian spectral transform approximations with quasi-cubic Lagrangian polynomial interpolants. A triangular truncation is adopted which can work both on a quadratically-unaliased transform grid or linear Gaussian grid (Williamson and Olson 1994; Collins et al. 2004). In our experiments here, a Gaussian quadratic transform grid was chosen unless noted otherwise. All prognostic variables are co-located.

COSMO The acronym COSMO stands for "Consortium for Small-scale Modeling" and denotes the nonhydrostatic regional weather prediction model at the German Weather Service (former name was LM which is the abbreviation for "Lokal Modell"). COSMO is a deep-atmosphere, finite-difference model on a staggered C-grid with a split-explicit temporal discretization (Doms and Schättler 2002; Gassmann and Herzog 2007; Baldauf 2010). Vertically traveling sound waves are handled implicitly.

ECHAM5 This model has been developed at the Max-Planck Institute for Meteorology (MPI) in Hamburg, Germany (Roeckner et al. 2003). It utilizes a semi-implicit Eulerian spectral transform dynamical core with triangular truncation and a Gaussian quadratic transform grid. The shallow-atmosphere dynamical core is hydrostatic and formulated in vorticity-divergence form. A three-time-level leapfrog time-stepping scheme is employed. All prognostic variables are co-located.

FV (also CAM FV) The mass-conservative Finite-Volume (FV) dynamical core employs the vector-invariant form of the primitive equations. It is written in flux form that is built upon a 2D shallow-water approach in the horizontal plane (Lin and Rood 1997) on a latitude–longitude grid with D-grid staggering. The vertical discretization utilizes a ÔLagrangian control-volumeÕ principle with conservative vertical remapping steps (Lin 2004). An explicit two-time-level time-stepping scheme is employed. The FV dynamical core is the default in

NCAR's model CAM version 4 or 5 (Neale et al. 2010), and is optional in CAM 3.1 (Collins et al. 2004). It is shared with NASA's GEOS5 model (Rienecker et al. 2008) and the NOAA Geophysical Fluid Dynamics Laboratory's (GFDL) atmospheric model AM2.1. A shallow water version and 3D variant of the FV model with an adaptive grid was also implemented by Jablonowski (2004) and Jablonowski et al. (2006, 2009).

FVcubed This variant of the FV model on an almost uniform-resolution cubed sphere grid with D-grid staggering is employed at the NASA Goddard Space Flight Center (GSFC) and GFDL (Putman and Lin 2007, 2009). It is part of the most recent internal version of NASA's GEOS model (version 6) and GFDL's internal release called "Riga". The cubed-sphere version has slightly different inherent dissipation characteristics e.g., due to the use of alternative limiters in the finite-volume algorithm. The FVcubed model also features an optional nonhydrostatic extension.

GEOS The Goddard Earth Observing System (GEOS) model has been under development at the NASA Goddard Space Flight Center. A commonly used version is GEOS5 which is documented in Rienecker et al. (2008). This version utilizes the FV dynamical core on a latitude–longitude grid as described above. Older GEOS versions like GEOS2 were based on the momentum equations in momentum form and utilized a finite-difference method on a latitude–longitude grid with a staggered C-grid arrangement. An explicit time three-time-level time-stepping scheme was used in GEOS2. A comprehensive model description can be found in Suarez and Takacs (1995). GEOS2 has also been used in a stretched-grid variant as documented in Fox-Rabinovitz et al. (1997). The forthcoming GEOS6 version will be based on the FVcubed dynamical core.

GME This primitive equation based dynamical core in vector-invariant form has been developed at the German Weather Service (DWD). It applies a finite-difference approximation with local spherical basis functions at each grid point. The horizontal grid is based on an icosahedral grid. An Arakawa-A grid staggering is chosen that places the prognostic variables at the vertices of the triangles. The semi-implicit numerical scheme is second-order accurate and applies a classical leapfrog three-time-level approach (Majewski et al. 2002, 2008)

HOMME The High Order Method Modeling Environment (HOMME) model is an optional hydrostatic dynamical core in NCAR's model CAM version 5 (Neale et al. 2010). It is a spectral element cubed-sphere dynamical core in momentum form with fourth-order polynomials inside each element (Thomas and Loft 2005; Taylor et al. 2007, 2009). The spectral element method is compatible, making the method elementwise mass-conservative (see Chap. 12 and Taylor and Fournier (2010)). The default time-stepping scheme is explicit and utilizes the three-time-level leapfrog method. Other time-stepping options are also available. All prognostic variables are co-located.

ICON The ICOsahedral Nonhydrostatic general circulation model ICON is a finite-difference model in momentum form and currently under development at MPI and DWD in Germany. It utilizes a dual icosahedral and hexagonal grid with C-grid staggering, is mass-conservative and employs a semi-implicit

three-time-level leapfrog time-stepping scheme. The shallow water and hydrostatic version are documented in Rípodas et al. (2009) and Wan (2009).

IFS This Integrated Forecasting System (IFS) is used for weather predictions at the European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, U.K.. It is a primitive equation based two-time-level semi-implicit semi-Lagrangian spectral transform dynamical core with a linear Gaussian transform grid and triangular truncation (Hortal 2002; ECMWF 2010). A finite-element discretization is employed in the vertical direction (Untch and Hortal 2004). All prognostic variables are co-located.

NICAM The Nonhydrostatic Icosahedral Atmospheric Model (NICAM) is a deep-atmosphere model which has been developed at the Frontier Research Center for Global Change (FRCGC), the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) and the Center for Climate System Research (CCSR) at the University of Tokyo, Japan. It is a finite-difference model written in mass-, momentum- and total energy-conserving form that utilizes a time-splitting scheme (Satoh 2002, 2003; Tomita and Satoh 2004; Satoh et al. 2008). All prognostic variables are co-located (Arakawa A-grid). NICAM's icosahedral grid configuration is optimized for uniformity by using a so-called "spring dynamics" grid.

UM The Unified Model (UM) is a mass-conserving nonhydrostatic GCM for weather and climate assessments that has been developed at the UK Met Office in Exeter, U.K.. It is built upon a finite-difference, deep-atmosphere, two-time-level, semi-implicit, semi-Lagrangian dynamical core in momentum form on a latitude–longitude grid (Davies et al. 2005; Staniforth and Wood 2008). This dynamical core is also called "New Dynamics". The prognostic variables are placed on a staggered C-grid. A comprehensive model description can be found in Staniforth et al. (2006).

WRF The Weather Research and Forecasting (WRF) model developed at NCAR is mostly used as a limited-area model. It is formulated in mass- and momentum-conserving form. WRF is a nonhydrostatic, deep-atmosphere, finite-difference model on a staggered C-grid with a split-explicit temporal discretization (Skamarock and Klemp 2008; Skamarock et al. 2008). Vertically traveling sound waves are handled implicitly. WRF can utilize a variety of map transformations and grids, e.g., the latitude–longitude grid for global applications.

References

Andrews DG, Mahlman JD, Sinclair RW (1983) Eliassen-Palm diagnostics of wave-mean flow interaction in the GFDL "SKYHI" general circulation model. J Atmos Sci 40:2768–2784

Arakawa A (1966) Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. J Comput Phys 1:119–143

Arakawa A, Hsu YJG (1990) Energy conserving and potential-enstrophy dissipating schemes for the shallow water equations. Mon Wea Rev 118:1960–1969

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

Arakawa A, Suarez MJ (1983) Vertical differencing of the primitive equations in sigma coordinates. Mon Wea Rev 111(1):34–45
Asselin R (1972) Frequency filter for time integrations. Mon Wea Rev 100(5):487–490
Baldauf M (2010) Linear stability analysis of Runge-Kutta based partial time-splitting schemes for the Euler equations. Mon Wea Rev 138:4475–4496
Bates JR, Semazzi FHM, Higgins RW, Barros SRM (1990) Integration of the shallow water equations on the sphere using a vector semi-Lagrangian scheme with a multigrid solver. Mon Wea Rev 118:1615–1627
Bates JR, Moorthi S, Higgins RW (1993) A global multilevel atmospheric model using a vector semi-Lagrangian finite difference scheme. Part I: Adiabatic formulation. Mon Wea Rev 121:244–263
Becker E (2001) Symmetric stress tensor formulation of horizontal momentum diffusion in global models of atmospheric circulation. J Atmos Sci 58:269–282
Becker E (2003) Frictional heating in global climate models. Mon Wea Rev 131:508–520
Becker E, Burkhardt U (2007) Nonlinear horizontal diffusion for GCMs. Mon Wea Rev 135: 1439–1454
Bermejo R, Staniforth A (1992) The Conversion of Semi-Lagrangian Advection Schemes to Quasi-Monotone Schemes. Mon Wea Rev 120:2622–2632
Black TL (1994) The new NMC mesoscale Eta model: Description and forecast examples. Weather and Forecasting 9:265–284
Boer GJ, Denis B (1997) Numerical covergence of the dynamics of a GCM. Climate Dynamics 13:359–374
Boer GJ, Shepherd TG (1983) Large-scale two-dimensional turbulence in the atmosphere. J Atmos Sci 40:164–184
Bonaventura L, Ringler T (2005) Analysis of discrete shallow-water models on geodesic Delaunay grids with C-type staggering. Mon Wea Rev 133:2351–2373
Bourke W, McAvaney B, Puri K, Thurling R (1977) Global modeling of atmospheric flow by spectral methods. In: Chang J (ed) Methods in Computational Physics, vol 17, Academic Press, pp 267–324
Boville BA (1986) Wave-mean flow interactions in a general circulation model of the troposphere and stratosphere. J Atmos Sci 43:1711–1725
Boville BA (1991) Sensitivity of simulated climate to model resolution. J Climate 4:469–485
Boville BA (2000) Toward a complete model of the climate system. In: Mote P, O'Neill A (eds) Numerical modeling of the global atmosphere in the climate system, NATO Science Series C: Mathematical and Physical Sciences, vol 550, Kluwer Academic Publishers, pp 419–442
Boville BA, Baumhefner DP (1990) Simulated forecast error and climate drift resulting from the omission of the upper stratosphere in numerical models. Mon Wea Rev 118:1517–1530
Boville BA, Bretherton CS (2003) Heating and kinetic energy dissipation in the NCAR Community Atmosphere Model. J Climate 16:3877–3887
Boville BA, Gent PR (1998) The NCAR Climate System Model, version one. J Climate 11:1115–1130
Boville BA, Randel WJ (1986) Observations and simulation of the variability of the stratosphere and troposphere in January. J Atmos Sci 43:3015–3034
Bowler NE, Arribas A, Beare SE, Mylne KR, Shutts GJ (2009) The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. Quart J Roy Meteor Soc 135:767–776
Boyd J (1996) The Erfc-Log filter and the asymptotics of the Euler and Vandeven sum accelerations. In: Proc. of the Third International Conference on Spectral and High Order Methods, pp 267–276, ed. by A. V. Ilin and L. R. Scott, Houston J. Mathematics, Houston, Texas
Boyd J (1998) Two comments on filtering (artificial viscosity) for Chebyshev and Legendre spectral and spectral element methods: Preserving boundary conditions and interpretation of the filter as a diffusion. J Comput Phys 143:283–288
Burkhardt U, Becker E (2006) A consistent diffusion dissipation parameterization in the ECHAM climate model. Mon Wea Rev 134:1194–1204

Canuto C, Hussaini MY, Quarteroni A, Zang TA (1987) Spectral Methods in Fluid Dynamics. Springer, 600 pp.

Carpenter RL, Droegemeier KK, Woodward PR, Hane CE (1990) Application of the Piecewise Parabolic Method (PPM) to meteorological modeling. Mon Wea Rev 118:586–612

Chen TC, Wiin-Nielsen A (1978) Nonlinear cascades of atmospheric energy and enstrophy in a two-dimensional spectral index. Tellus 30:313–322

Colella P, Sekora MD (2008) A limiter for PPM that preserves accuracy at smooth extrema. J Comput Phys 227:7069–7076

Colella P, Woodward PR (1984) The Piecewise Parabolic Method (PPM) for gas-dynamical simulations. J Comput Phys 54:174–201

Collins WD, Rasch PJ, Boville BA, Hack JJ, McCaa JR, Williamson DL, Kiehl JT, Briegleb BP, Bitz CM, Lin SJ, Zhang M, Dai Y (2004) Description of the NCAR Community Atmosphere Model (CAM3.0). NCAR Technical Note NCAR/TN-464+STR, National Center for Atmospheric Research, Boulder, Colorado, 214 pp., available from http://www.ucar.edu/library/collections/technotes/technotes.jsp

Cordero E, Staniforth A (2004) A problem with the Robert-Asselin time filter for three-time-level semi-implicit semi-Lagrangian discretizations. Mon Wea Rev 132:600–610

Côté J, Gravel S, Staniforth A (1995) A generalized family of schemes that eliminate the spurious resonant response of semi-Lagrangian schemes to orographic forcing. Mon Wea Rev 123:3605–3613

Côté J, Gravel S, Méthot A, Patoine A, Roch M, Staniforth A (1998a) The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. Mon Wea Rev 126(6):1373–1395

Côté J, Gravel S, Méthot A, Patoine A, Roch M, Staniforth A (1998b) The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. Mon Wea Rev 126(6):1373–1395

Davies T, Cullen MJP, Malcolm AJ, Mawson MH, Staniforth A, White AA, Wood N (2005) A new dynamical core for the Met Office's global and regional modelling of the atmosphere. Quart J Roy Meteor Soc 131(608):1759–1782

Deardorff JW (1970) A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. J Fluid Mechanics 41:453–480

Déqué M, Cariolle D (1986) Some destabilizing properties of the Asselin time filter. Mon Wea Rev 114:880–884

Dey CH (1978) Noise suppression in a primitive equation prediction model. Mon Wea Rev 106:159–173

Doms G, Schättler U (2002) A description of the nonhydrostatic regional model LM. Part I: Dynamics and numerics. Consortium for Small-Scale Modelling (COSMO) LM F90 2.18, German Weather Service, Offenbach, Germany, 134 pp., available from http://www.cosmo-model.org

Dubal M, Wood N, Staniforth A (2004) Analysis of parallel versus sequential splittings for time-stepping physical parameterizations. Mon Wea Rev 132:121–132

Dubal M, Wood N, Staniforth A (2006) Some numerical properties of approaches to physics dynamics coupling for NWP. Quart J Roy Meteor Soc 132:27–42

Dudhia J (1993) A nonhydrostatic version of the Penn State-NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. Mon Wea Rev 121:1493–1513

Durran DR (1999) Numerical methods for wave equations in geophysical fluid dynamics, First edn. Springer, 465 pp.

Durran DR (2010) Numerical methods for fluid dynamics: with applications in geophysics, Second edn. Springer, 516 pp.

Durran DR, Klemp JB (1983) A compressible model for the simulation of moist mountain waves. Mon Wea Rev 111:2341–2361

ECMWF (2010) Part III: Dynamics and numerical procedures. IFS Documentation - Cy36r1, Operational implementation 26 January 2010, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, England, 29 pp., available from http://www.ecmwf.int/research/ifsdocs/CY36r1/index.html

Enomoto T, Kuwano-Yoshida A, Komori N, Ohfuchi W (2008) Description of AFES 2: Improvements for high-resolution and coupled simulations. In: Hamilton K, Ohfuchi W (eds) High resolution numerical modelling of the atmosphere and ocean, Springer, pp 77–97

Farge M, Sadourny R (1989) Wave-vortex dynamics in rotating shallow water. J Fluid Mech 206:433–462

Fiedler BH (2000) Dissipative heating in climate models. Quart J Roy Meteor Soc 126:925–939

Fox-Rabinovitz MS, Stenchikov GL, Suarez MJ, Takacs LL (1997) A finite-difference GCM dynamical core with a variable-resolution stretched grid. Mon Wea Rev 125:2943–2968

Fudeyasu H, Wang Y, Satoh M, Nasuno T, Miura H, Yanase W (2008) Global cloud-system-resolving model NICAM successfully simulated the lifecycles of two real tropical cyclones. Geophys Res Lett 35:L22,808

Gassmann A, Herzog HJ (2007) A consistent time-split numerical scheme applied to the non-hydrostatic compressible equations. Mon Wea Rev 135:20–36

Gelb A, Gleeson JP (2001) Spectral viscosity for shallow water equations in spherical geometry. Mon Wea Rev 129:2346–2360

GFS (2003) The GFS atmospheric model. NCEP Office Note 442, National Centers for Environmental Prediction (NCEP), Environmental Modeling Center, Camp Springs, Maryland, 14 pp., available from http://www.emc.ncep.noaa.gov/officenotes/index.shtml see also http://wwwt.emc.ncep.noaa.gov/?branch=GFS

Gill AE (1982) Atmosphere-ocean dynamics. Academic Press, 768 pp.

Giraldo FX, Rosmond TE (2004) A scalable Spectral Element Eulerian Atmospheric Model (SEE-AM) for NWP: Dynamical core tests. Mon Wea Rev 132:133–153

Giraldo FX, Hesthaven JS, Wartburton T (2002) Nodal high-order discontinous Galerkin methods for the shallow water equations. J Comput Phys 181:499–525

Gordon CT, Stern WF (1982) A description of the GFDL global spectral model. Mon Wea Rev 110:625–644

Gravel S, Staniforth A, Côté J (1993) A stability analysis of a family of baroclinic semi-Lagrangian forecast models. Mon Wea Rev 121:815–824

Griffies SM, Hallberg RW (2000) Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale Eddy-permitting ocean models. Mon Wea Rev 128:2935–2946

Gross ES, Bonaventura L, Rosatti G (2002) Consistency with continuity in conservative advection schemes for free-surface models. Int J Numer Meth Fluids 38:307–327

Haltiner GJ, Williams RT (1980) Numerical prediction and dynamic meteorology. John Wiley & Sons, 477 pp.

Hamilton K, Takahashi YO, Ohfuchi W (2008) Mesoscale spectrum of atmospheric motions investigated in a very fine resolution global general circulation model. J Geophys Res 113:D18,110

Hamming RW (1977) Digital filters. Prentice Hall, 226 pp.

Held IM, Suarez MJ (1994) A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. Bull Amer Meteor Soc 75(10):1825–1830

Hortal M (2002) The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model. Quart J Roy Meteor Soc 128(583):1671–1687

Hortal M, Simmons AJ (1991) Use of reduced grids in spectral models. Mon Wea Rev 119:1057–1074

Huang H, Stevens B, Margulis SA (2008) Application of dynamic subgrid-scale models for large-eddy simulation of the daytime convective boundary layer over heterogeneous surfaces. Boundary-Layer Meteorology 126:327–348

Huynh HY (1996) Schemes and constraints for advection. In: 15th International Conference on Numerical Methods in Fluid Dynamics, Monterey June 24–28, 1996, CA, USA

Jablonowski C (1998) Test der Dynamik zweier globaler Wettervorhersagemodelle des Deutschen Wetterdienstes: Der Held-Suarez Test. Master's thesis, University of Bonn, Germany, Department of Meteorology, 140 pp.

Jablonowski C (2004) Adaptive grids in weather and climate modeling. PhD thesis, University of Michigan, Ann Arbor, MI, Department of Atmospheric, Oceanic and Space Sciences, 292 pp.

Jablonowski C, Williamson DL (2006a) A baroclinic instabilitiy test case for atmospheric model dynamical cores. Quart J Roy Meteor Soc 132(621C):2943–2975

Jablonowski C, Williamson DL (2006b) A baroclinic wave test case for dynamical cores of General Circulation Models: Model intercomparisons. NCAR Tech. Note NCAR/TN-469+STR, National Center for Atmospheric Research, Boulder, Colorado, 89 pp., available from http://www.ucar.edu/library/collections/technotes/technotes.jsp

Jablonowski C, Herzog M, Penner JE, Oehmke RC, Stout QF, van Leer B, Powell KG (2006) Block-structured adaptive grids on the sphere: Advection experiments. Mon Wea Rev 134:3691–3713

Jablonowski C, Oehmke RC, Stout QF (2009) Block-structured adaptive meshes and reduced grids for atmospheric general circulation models. Phil Trans R Soc A 367:4497–4522

Jakimow G, Yakimiw E, Robert A (1992) An implicit formulation for horizontal diffusion in gridpoint models. Mon Wea Rev 120:124–130

Jakob R, Hack JJ, Williamson DL (1993) Solutions to the shallow-water test set using the spectral transform method. NCAR Tech. Note NCAR/TN-388+STR, National Center for Atmospheric Research, Boulder, Colorado, 82 pp., available from http://www.ucar.edu/library/collections/technotes/technotes.jsp

Janjić ZI (1990) The step-mountain coordinate: Physical package. Mon Wea Rev 118:1429–1443

Jöckel P, von Kuhlmann R, Lawrence MG, Steil B, Brenninkmelter CAM, Crutzen PJ, Rasch PJ, 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. Quart J Roy Meteor Soc 127:1035–1052

Kalnay E (2003) Atmospheric modeling, data assimilation and predictability. Cambridge University Press, 341 pp.

Kalnay-Rivas E, Bayliss A, Storch J (1977) The 4th order GISS model of the global atmosphere. Beiträge zur Physik der Atmosphäre 50:299–311

Kasahara A (1974) Various vertical coordinate systems used for numerical weather prediction. Mon Wea Rev 102:509–522

Klemp JB, Durran DR (1983) An upper boundary condition permitting internal gravity wave radiation in numerical mesoscale models. Mon Wea Rev 111:430–444

Klemp JB, Lilly DK (1978) Numerical simulation of hydrostatic mountain waves. J Atmos Sci 35:78–107

Klemp JB, Skamarock WC, Dudhia J (2007) Conservative split-explicit time integration methods for the compressible nonhydrostatic equations. Mon Wea Rev 135:2897–2913

Klemp JB, Dudhia J, Hassiotis AD (2008) An upper gravity-wave absorbing layer for NWP applications. Mon Wea Rev 136:3987–4004

Knievel JC, Bryan GH, Hacker JP (2007) Explicit numerical diffusion in the WRF model. Mon Wea Rev 135:3808–3824

Koshyk JN, Boer GJ (1995) Parameterization of dynamical subgrid-scale processes in a spectral GCM. J Atmos Sci 52(7):965–976

Koshyk JN, Hamilton K (2001) The horizontal kinetic energy spectrum and spectral budget simulated by a high-resolution troposphere-stratosphere-mesosphere GCM. J Atmos Sci 58:329–348

Lander J, Hoskins BJ (1997) Believable scales and parameterizations in a spectral transform model. Mon Wea Rev 125:292–303

Laprise R, Girard C (1990) A spectral General Circulation Model using the piecewise-constant finite-element representation on a hybrid vertical coordinate system. J Climate 3:32–52

Lauritzen PH, Jablonowski C, Taylor MA, Nair RD (2010a) Rotated versions of the Jablonowski steady-state and baroclinic wave test cases: A dynamical core intercomparison. J Adv Model Earth Syst 2:Art. #15, 34 pp.

Lauritzen PH, Nair RD, Ullrich PA (2010b) A conservative semi-Lagrangian multi-tracer transport scheme (CSLAM) on the cubed-sphere grid. J Comput Phys 229:1401–1424

Laursen L, Eliasen E (1989) On the effects of the damping mechanisms in an atmospheric general circulation model. Tellus Series A 41:385–400

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. IV. A new approach to numerical convection. J Comput Phys 23:276–299
Leith CE (1971) Atmospheric predictability and two-dimensional turbulence. J Atmos Sci 28: 145–161
LeVeque RJ (2002) Finite Volume Methods for Hyperbolic Problems. Cambridge University Press, ISBN 0-521-00924-3, 558 pp.
Li Y, Moorthi S, Bates JR (1994) Direct solution of the implicit formulation of fourth order horizontal diffusion for gridpoint models on the sphere. NASA Technical Memorandum 104606, Vol. 2, NASA Goddard Space Flight Center, Greenbelt, Maryland, 42 pp.
Lin SJ (2004) A "vertically Lagrangian" finite-volume dynamical core for global models. Mon Wea Rev 132:2293–2307
Lin SJ, Rood RB (1996) Multidimensional flux-form semi-Lagrangian transport scheme. Mon Wea Rev 124:2046–2070
Lin SJ, Rood RB (1997) An explicit flux-form semi-Lagrangian shallow water model on the sphere. Quart J Roy Meteor Soc 123:2477–2498
Lin YL (2007) Mesoscale dynamics. Cambridge University Press, 630 pp.
Lindberg K, Alexeev VA (2000) A study of the spurious orographic resonance in semi-implicit semi-Lagrangian models. Mon Wea Rev 128:1982–1989
Lynch P, Huang X (1992) Initialization of the HIRLAM model using a digital filter. Mon Wea Rev 120:1019–1034
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 ICSU, Geneva, pp 121–275
MacVean MK (1983) The effects of horizontal diffusion on baroclinic development in a spectral model. Quart J Roy Meteor Soc 109:771–783
Mahlmann JD, Sinclair WW (1977) Tests of various numerical algorithms applied to a simple trace constituent air transport problem. In: Suffett IH (ed) Advances in Environmental Science and Technology, Part 1: The Fate of Pollutants in the Air and Water Environments, vol 8, Wiley, pp 223–252
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 130:319–338
Majewski D, Frank H, Liermann D (2008) GME User's Guide. Tech. Rep. corresponding to model version gmtri 2.17 and higher, German Weather Service DWD, Frankfurt, Germany, 70 pp.
McCorquodale P, Colella P (2010) A high-order finite-volume method for hyperbolic conservation laws on locally-refined grids. J. Comput. Phys., in review
McDonald A, Haugen J (1992) A two-time-level, three-dimensional semi-Lagrangian, semi-implicit, limited-area gridpoint model of the primitive equations. Mon Wea Rev 120(11):2603–2621
McDonald A, Haugen J (1993) A two time-level, three-dimensional, semi-Lagrangian, semi-implicit, limited-area gridpoint model of the primitive equations. Part II: Extension to hybrid vertical coordinates. Mon Wea Rev 121(7):2077–2087
McPherson RD, Stackpole JD (1973) Noise suppression in the eight-layer global model. Office Note 83, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, National Meteorological Center, 36 pp., available from http://www.emc.ncep.noaa.gov/officenotes/index.shtml
Mellor GL (1985) Ensemble average, turbulence closure. In: Manabe S (ed) Advances in Geophysics, Issues in Atmos. and Ocean Modeling. Part B: Weather Dynamics, Academic Press, pp 345–357
Mesinger F, Arakawa A (1976) Numerical methods used in atmospheric models. In: GARP Publication Series No. 17, vol 1, World Meteorological Organization, p 64, Geneva, Switzerland

Mesinger F, Jovic D (2002) The Eta slope adjustment: Contender for an optimal steepening in a piecewise-linear advection scheme? Comparison tests. NCEP Office Note 439, National Centers for Environmental Prediction (NCEP), Environmental Modeling Center, Camp Springs, Maryland, 29 pp., available from http://www.emc.ncep.noaa.gov/officenotes/index.shtml

Nair RD (2009) Diffusion experiments with a global discontinuous Galerkin shallow-water model. Mon Wea Rev 137:3339–3350

Nastrom GD, Gage KS (1985) A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. J Atmos Sci 42:950–960

Neale RB, Hoskins BJ (2000) A standard test for AGCMs including their physical parameterizations: I: The proposal. Atmos Sci Lett 1:101–107

Neale RB, Chen CC, Gettelman A, Lauritzen PH, Park S, Williamson DL, Conley AJ, Garcia R, Kinnison D, Lamarque JF, Marsh D, Mills M, Smith AK, Tilmes S, Vitt F, Cameron-Smith P, Collins WD, Iacono MJ, Rasch PJ, Taylor M (2010) Description of the NCAR Community Atmosphere Model (CAM 5.0). NCAR Technical Note NCAR/TN-XXX+STR, National Center for Atmospheric Research, Boulder, Colorado, draft, available from http://www.cesm.ucar.edu/models/cesm1.0/cam/

Nelson SP, Weible ML (1980) Three-dimensional Shuman filter. J Appl Meteor 19:464–469

Orr A, Wedi N (2009) The representation of non-orographic gravity waves in the IFS Part I: Assessment of the middle atmosphere climate with Rayleigh friction. ECMWF Technical Memorandum 592, European Centre for Medium-Range Weather Forecasts, Reading, U.K., 15 pp., available from http://www.ecmwf.int/publications/library/do/references/list/14

Orr A, Bechthold P, Scinocca J, Ern M, Janiskova M (2010) Improved middle atmosphere climate and forecasts in the ECMWF model through a non-orographic gravity wave drag parameterization. J Climate 23:5905–5926

Orszag SA (1971) On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. J Atmos Sci 28:1074–1074

O'Sullivan D, Dunkerton TJ (1995) Generation of inertia-gravity waves in a simulated life cycle of baroclinic instability. J Atmos Sci 52:3695–3716

Phillips N (1959) An example of non-linear computational instability. In: Bolin B (ed) The Atmosphere and Sea in Motion, Oxford University Press, pp 501–504

Phillips NA (1957) A coordinate system having some special advantages for numerical forecasting. J Meteor 14:184–185

Polvani LM, Kushner PJ (2002) Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. Geophys Res Lett 29(7):070,000–1

Polvani LM, Scott RK, Thomas SJ (2004) Numerically converged solutions of the global primitive equations for testing the dynamical core of atmospheric GCMs. Mon Wea Rev 132:2539–2552

Purser RJ (1987) The filtering of meteorological fields. J Climate and Appl Meteor 26:1764–1769

Purser RJ (1988) Degradation of numerical differencing caused by Fourier filtering at high latitudes. Mon Wea Rev 116:1057–1066

Purser RJ, Leslie LM (1994) An efficient semi-Lagrangian scheme using third-order semi-implicit time integration and forward trajectories. Mon Wea Rev 122(4):745–756

Putman WM, Lin SJ (2007) Finite-volume transport on various cubed-sphere grids. J Comput Phys 227:55–78

Putman WM, Lin SJ (2009) A finite-volume dynamical core on the cubed-sphere grid. In: Numerical Modeling of Space Plasma Flows: Astronum-2008, Astronomical Society of the Pacific Conference Series, vol 406, pp 268–276

Randall DA (1994) Geostrophic adjustment and the finite-difference shallow-water equations. Mon Wea Rev 122:1371–1377

Rasch PJ (1986) Toward atmospheres without tops: Absorbing upper boundary conditions for numerical models. Quart J Roy Meteor Soc 112:1195–1218

Rasch PJ, Williamson DL (1990a) Computational aspects of moisture transport in global models of the atmosphere. Quart J Roy Meteor Soc 116:1071–1090

Rasch PJ, Williamson DL (1990b) On shape-preserving interpolation and semi-Lagrangian transport. SIAM J Sci Stat Comput 11(4):656–687

Rasch PJ, Williamson DL (1991) The sensitivity of a general circulation model climate to the moisture transport formulation. J Geophy Res 96:13,123–13,137

Rasch PJ, Boville BA, Brasseur GP (1995) A three-dimensional general circulation model with coupled chemistry for the middle atmosphere. J Geophys Res 100:9041–9072

Raymond WH (1988) High-order low-pass implicit tangent filters for use in finite area calculations. Mon Wea Rev 116:2132–2141

Raymond WH, Garder A (1988) A spatial filter for use in finite area calculations. Mon Wea Rev 116:209–222

Raymond WH, Garder A (1991) A review of recursive and implicit filters. Mon Wea Rev 119: 477–495

Reames FM, Zapotocny TH (1999) Inert trace constituent transport in sigma and hybrid isentropic-sigma models. Part I: Nine advection algorithms. Mon Wea Rev 127:173–187

Rienecker MM, Suarez MJ, Todling R, Bacmeister J, Takacs L, Liu HC, Gu W, Sienkiewicz M, Koster RD, Gelaro R, Stajner I, Nielsen E (2008) The GEOS-5 data assimilation system – Documentation of versions 5.0.1 and 5.1.0. Technical Report Series on Global Modeling and Data Assimilation NASA/TM-2007-104606, Vol. 27, NASA Goddard Space Flight Center, Greenbelt, Maryland, 92 pp.

Rípodas P, Gassmann A, Förstner J, Majewski D, Giorgetta M, Korn P, Kornblueh L, Wan H, Zängl G, Bonaventura L, Heinze T (2009) Icosahedral Shallow Water Model (ICOSWM): results of shallow water test cases and sensitivity to model parameters. Geosci Model Dev 2:231–251

Ritchie H, Temperton C, Simmons A, Hortal M, Davies T, Dent D, Hamrud M (1995) Implementation of the Semi-Lagrangian Method in a High-Resolution Version of the ECMWF Forecast Model. Mon Wea Rev 123:489–514

Rivest C, Staniforth A, Robert A (1994) Spurious resonant response of semi-Lagrangian discretizations to orographic forcing: Diagnosis and solution. Mon Wea Rev 122:366–376

Robert A, Lépine M (1997) Anomaly in the behaviour of the time filter used with the leapfrog scheme in atmospheric models. In: Lin CA, Laprise R, Ritchie H (eds) Numerical Methods in Atmospheric and Oceanic Modeling: The André J. Robert Memorial Volume, Canadian Meteorological and Oceanographic Society / NRC Research Press, pp S3–S15

Robert AJ (1966) The Integration of a low order spectral form of the primitive meteorological equations. J Meteor Soc Japan 44(5):237–245

Roeckner E, Bäuml G, Bonaventura L, Brokopf R, Esch M, Hagemann MGS, Kirchner I, Kornblueh L, Manzini E, Rhodin A, Schlese U, Schulzweida U, Tompkins A (2003) The atmospheric general circulation model ECHAM5. Part I model description. Tech. Rep. 349, Max Planck Institute for Meteorology, 127 pp., available from http://www.mpimet.mpg.de/en/wissenschaft/modelle/echam/echam5.html

Rood RB (1987) Numerical advection algorithms and their role in atmospheric transport and chemistry models. Rev Geophys 25:71–100

Royer J (1986) Correction of negative mixing ratios in spectral models by global horizontal borrowing. Mon Wea Rev 114:1406–1410

Ruge JW, McCormick SF, Yee SYK (1995) Multilevel adaptive methods for semi-implicit solution of shallow-water equations on the sphere. Mon Wea Rev 123:2197–2205

Sadourny R (1975) The dynamics of finite-difference models of the shallow-water equations. J Atmos Sci 32:680–689

Sadourny R, Maynard K (1997) Formulations of lateral diffusion in geophysical fluid dynamics models. In: Lin CA, Laprise R, Ritchie H (eds) Numerical Methods in Atmospheric and Oceanic Modeling: The André J. Robert Memorial Volume, Canadian Meteorological and Oceanographic Society / NRC Research Press, pp 547–556

Sardeshmukh PD, Hoskins BI (1984) Spatial smoothing on the sphere. Mon Wea Rev 112:2524–2529

Satoh M (2002) Conservative scheme for the compressible nonhydrostatic models with the horizontally explicit and vertically implicit time integration scheme. Mon Wea Rev 130(5):1227–1245

Satoh M (2003) Conservative scheme for a compressible nonhydrostatic model with moist processes. Mon Wea Rev 131:1033–1050

Satoh M (2004) Atmospheric circulation dynamics and general circulation models. Springer (Praxis), 643 pp.

Satoh M, Matsuno T, Tomita H, Miura H, Nasuno T, Iga S (2008) Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. J Comput Phys 227:3486–3514

Schlesinger RE, Uccellini LW, Johnson DR (1983) The effects of the Asselin time filter on numerical solutions to the linearized shallow-water wave equations. Mon, Wea, Rev, 111:455–467

Schmidt GA, Ruedy R, Hansen JE, Aleinov I, Bell N, Bauer M, Bauer S, Cairns B, Canuto V, Cheng Y, Del Genio A, Faluvegi G, Friend AD, Hall TM, Hu Y, Kelley M, Kiang NY, Koch D, Lacis AA, Lerner J, Lo KK, Miller RL, Nazarenko L, Oinas V, Perlwitz J, Perlwitz J, Rind D, Romanou A, Russell GL, Sato M, Shindell DT, Stone PH, Sun S, Tausnev N, Thresher D, Yao MS (2006) Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. J Climate 19:153–192

Shapiro R (1970) Smoothing, filtering, and boundary effects. Rev Geophys and Space Phys 8(2):359–387

Shapiro R (1971) The use of linear filtering as a parameterization of atmospheric diffusion. J Atmos Sci 28:523–531

Shapiro R (1975) Linear filtering. Mathematics of Computation 29(132):1094–1097

Shepherd TG, Shaw TA (2004) The angular momentum constraint on climate sensitivity and downward Influence in the middle atmosphere. J Atmos Sci 61:2899–2908

Shepherd TG, Semeniuk K, Koshyk JN (1996) Sponge layer feedbacks in middle-atmosphere models. J Geophys Res 101:23,447–23,464

Shuman FG (1957) Numerical methods in weather prediction: II. Smoothing and filtering. Mon Wea Rev 85:357–361

Shuman FG (1969) On a special form for viscous terms. Office Note 32, U.S. Department of Commerce, Environmental Science Service Administration, Weather Bureau, 4 pp., available from http://www.emc.ncep.noaa.gov/officenotes/index.shtml

Shuman FG, Stackpole JD (1969) The currently operational NMC model, and results of a recent numerical experiment. In: Proc. WMO/IUGG Symp. Numerical Weather Prediction, Japan Meteorological Agency, Tokyo, Japan, pp II–85–98

Shutts G (2005) A kinetic energy backscatter algorithm for use in ensemble prediction systems. Quart J Roy Meteor Soc 131:3079–3102

Simmons AJ, Burridge DM (1981) An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. Mon Wea Rev 109:758–766

Skamarock WC (2004) Evaluating mesoscale NWP models using kinetic energy spectra. Mon Wea Rev 132:3019–3032

Skamarock WC, Klemp JB (1992) The stability of time-split numerical methods for the hydrostatic and the nonhydrostatic elastic equations. Mon Wea Rev 120:2109–2127

Skamarock WC, Klemp JB (2008) A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. J Comput Phys 227:3465–3485

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang XY, Wang W, Powers JG (2008) A description of the Advanced Research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR, National Center for Atmospheric Research, Boulder, Colorado, 113 pp., available from http://www.ucar.edu/library/collections/technotes/technotes.jsp

Smagorinsky J (1963) General circulation experiments with the primitive equations. I. The basic experiment. Mon Wea Rev 91:99–164

Smagorinsky J (1993) Some historical remarks on the use of nonlinear viscosities. In: Galperin B, Orszag SA (eds) Large Eddy Simulation of Complex Engineering and Geophysical Flows, Cambridge University Press, pp 3–36

St-Cyr A, Jablonowski C, Dennis J, Thomas S, Tufo H (2008) A comparison of two shallow water models with non-conforming adaptive grids. Mon Wea Rev 136:1898–1922

Staniforth A, Côté J (1991) Semi-Lagrangian integration schemes for atmospheric modeling - A review. Mon Wea Rev 119:2206–2223

Staniforth A, Wood N (2008) Aspects of the dynamical core of a nonhydrostatic, deep-atmosphere, unified weather and climate-prediction model. J Comput Phys 227(7):3445–3464

Staniforth A, White A, Wood N (2003) Analysis of semi-Lagrangian trajectory computations. Quart J Roy Meteor Soc 129(591):2065–2085

Staniforth A, White A, Wood N, Thuburn J, Zerroukat M, Cordero E, Davies T, Diamantakis M (2006) Joy of U.M. 6.3 - model formulation. Unified Model Documentation Paper No 15, UK Meteorological Service, Exeter, England, available from http://research.metoffice.gov.uk/research/nwp/publications/papers/unified_model/

Stephenson DB (1994) The impact of changing the horizontal diffusion scheme on the northern winter climatology of a general circulation model. Quart J Roy Meteor Soc 120:211–226

Suarez MJ, Takacs LL (1995) Documentation of the ARIES/GEOS dynamical core: Version 2. NASA Technical Memorandum 104606, Vol. 5, NASA Goddard Space Flight Center, Greenbelt, Maryland, 45 pp.

Takacs L, Sawyer W, Suarez MJ, Fox-Rabnovitz MS (1999) Filtering techniques on a stretched grid general circulation model. Tech. Rep. NASA/TM-1999-104606, Vol. 16, NASA Goddard Space Flight Center, Greenbelt, Maryland, 50 pp.

Takacs LL (1988) Effect of using a posteriori methods for the conservation of integral invariants. Mon Wea Rev 116:525–545

Takacs LL, Balgovind RC (1983) High-latitude filtering in global grid-point models. Mon Wea Rev 111:2005–2015

Takahashi YO, Hamilton K, Ohfuchi W (2006) Explicit global simulation of the mesoscale spectrum of atmospheric motions. Geophys Res Lett 33:L12,812

Tandon MK (1987) Robert's recursive frequency filter: A reexamination. Meteor Atmos Phys 37:48–59

Tanguay M, Yakimiw E, Ritchie H, Robert A (1992) Advantages of spatial averaging in semi-implicit semi-Lagrangian schemes. Mon Wea Rev 120:113–123

Tatsumi Y (1983) An economical explicit time integration scheme for a primitive model. J Meteorol Soc Japan 61:269–287

Taylor M, Tribbia J, Iskandarani M (1997) The spectral element method for the shallow water equations on the sphere. J Comput Phys 130:92–108

Taylor MA (2008) CAM/HOMME: Parallel scalability and aqua planet results for CAM on the cubed-sphere grid. Presentation at the Community Climate System Model (CCSM) Workshop in Breckenridge, CO, available online at http://www.ccsm.ucar.edu/events/ws.2008/Presentations/Tarn/SEWG/Taylor_ccsm08.pdf

Taylor MA, Fournier A (2010) A compatible and conservative spectral element method on unstructured grids. J Comput Phys 229:5879–5895

Taylor MA, Edwards J, Thomas S, Nair RD (2007) A mass and energy conserving spectral element atmospheric dynamical core on the cubed-sphere grid. Journal of Physics: Conference Series 78:012,074, available online at http://www.iop.org/EJ/toc/1742-6596/78/1

Taylor MA, St-Cyr A, Fournier A (2009) A non-oscillatory advection operator for the compatible spectral element method. In: Allen G (ed) Computational Science Ð ICCS 2009, Part II, Lecture Notes in Computer Science, vol 5545, Springer Berlin/Heidelberg, pp 273–282

Temperton C, Staniforth A (1987) An efficient two-time-level semi-Lagrangian semi-implicit integration scheme. Quart J Roy Meteor Soc 113:1025–1039

Terasaki K, Tanaka HL, Satoh M (2009) Characteristics of the kinetic energy spectrum of NICAM model atmosphere. SOLA 5:180–183

Thomas SJ, Loft RD (2005) The NCAR spectral element climate dynamical core: Semi-implicit Eulerian formulation. J Sci Comput 25:307–322

Thuburn J (1993) Use of a flux-limited scheme for vertical advection in a GCM. Quart J Roy Meteor Soc 119:469–487

Thuburn J (1997) TVD Schemes, Positive Schemes, and the Universal Limiter. Mon Wea Rev 125:1990–1993

Thuburn J (2008a) Numerical wave propagation on the hexagonal C-grid. J Comput Phys 227:5836–5858

Thuburn J (2008b) Some conservation issues for dynamical cores of NWP and climate models. J Comput Phys 227(7):3715–3730

Tomita H, Satoh M (2004) A new dynamical framework of nonhydrostatic global model using the icosahedral grid. Fluid Dyn Res 34:357–400

Tomita H, Miura H, Iga S, Nasuno T, Satoh M (2005) A global cloud-resolving simulation: Preliminary results from an aqua planet experiment. Geophys Res Lett 32:L08,805

Untch A, Hortal M (2004) A finite-element scheme for the vertical discretization of the semi-Lagrangian version of the ECMWF forecast model. Quart J Roy Meteor Soc 130:1505–1530

Vandeven H (1991) Family of spectral filters for discontinuous problems. J Sci Comput 6:159–192

Váňa F, Bénard P, Geleyn J, Simon A, Seity Y (2008) Semi-Lagrangian advection scheme with controlled damping: An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. Quart J Roy Meteor Soc 134:523–537

von Storch J (2004) On statistical dissipation in GCM-climate. Climate Dynamics 23:1–15

Wan H (2009) Developing and testing a hydrostatic atmospheric dynamical core on triangular grids. Reports on Earth System Science 65, Max Planck Institute for Meteorology, Hamburg, Germany, 153 pp., available from http://www.mpimet.mpg.de/en/wissenschaft/publikationen/berichte-erdsystemforschung.html

Wan H, Giorgetta MA, Bonaventura L (2008) Ensemble Held Suarez test with a spectral transform model: Variability, sensitivity, and convergence. Mon Wea Rev 136:1075–1092

Washington WM, Baumhefner DP (1975) A method of removing Lamb waves from initial data for primitive equation models. J Appl Meteor 14:114–119

White AA, B J Hoskins IR, Staniforth A (2005) Consistent approximate models of the global atmosphere: shallow, deep, hydrostatic, quasi-hydrostatic and non-hydrostatic. Quart J Roy Meteor Soc 131:2081–2107

Whitehead J, Jablonowski C, Rood RB, Lauritzen PH (2011) A stability analysis of divergence damping on a latitude-longitude grid. Mon. Wea. Rev., accepted (pending revisions)

Williams PD (2009) A proposed modification to the Robert-Asselin time filter. Mon Wea Rev 137:2538–2546

Williamson DL (1976) Linear stability of finite-difference approximations on a uniform latitude-longitude grid with Fourier filtering. Mon Wea Rev 104:31–41

Williamson DL (1978) The Relative Importance of Resolution, Accuracy and Diffusion in Short-Range Forecasts with the NCAR Global Circulation Model. Mon Wea Rev 106:69–88

Williamson DL (1988) The effect of vertical finite difference approximations on simulations with the NCAR Community Climate Model. J Climate 1(1):40–58

Williamson DL (1990) Semi-Lagrangian moisture transport in the NMC spectral model. Tellus 42A:413–428

Williamson DL (1997) Climate simulations with a spectral, semi-Lagrangian model with linear grids. In: Lin CA, Laprise R, Ritchie H (eds) Numerical Methods in Atmospheric and Oceanic Modeling: The André J. Robert Memorial Volume, Canadian Meteorological and Oceanographic Society / NRC Research Press, pp 280–292

Williamson DL (2007) The evolution of dynamical cores for global atmospheric models. J Meteorol Soc Japan 85B:241–269

Williamson DL (2008a) Convergence of aqua-planet simulations with increasing resolution in the Community Atmospheric Model, Version3. Tellus 60A:848–862

Williamson DL (2008b) Equivalent finite volume and Eulerian spectral transform horizontal resolutions established from aqua-planet simulations. Tellus 60A:839–847

Williamson DL, Browning GL (1973) Comparison of grids and difference approximations for numerical weather prediction over the sphere. J Appl Meteor 12:264–274

Williamson DL, Laprise R (2000) Numerical approximations for global atmospheric general circulation models. In: Mote P, O'Neill A (eds) Numerical modeling of the global atmosphere in the climate system, NATO Science Series C: Mathematical and Physical Sciences, vol 550, Kluwer Academic Publishers, pp 127–219

Williamson DL, Olson JG (1994) Climate simulations with a semi-Lagrangian version of the NCAR Community Climate Model. Mon Wea Rev 122(7):1594–1610

Williamson DL, Rasch PJ (1989) Two-dimensional semi-Lagrangian transport with shape-preserving interpolation. Mon Wea Rev 117:102–129

Williamson DL, Rasch PJ (1994) Water vapor transport in the NCAR CCM2. Tellus A 46:34–51

Williamson DL, Drake JB, Hack JJ, Jakob R, Swarztrauber PN (1992) A standard test set for numerical approximations to the shallow water equations in spherical geometry. J Comput Phys 102:211–224

Williamson DL, Olson JG, Boville BA (1998) A comparison of semi-Lagrangian and Eulerian tropical climate simulations. Mon Wea Rev 126:1001–1012

Williamson DL, Olson J, Jablonowski C (2009) Two dynamical core formulation flaws exposed by a baroclinic instability test case. Mon Wea Rev 137:790–796

Wood N, Diamantakis M, Staniforth A (2007) A monotonically-damping second-order-accurate unconditionally-stable numerical scheme for diffusion. Quart J Roy Meteor Soc 133:1559–1573

Xue M (2000) High-order monotonic numerical diffusion and smoothing. Mon Wea Rev 128:2853–2864

Xue M, Droegemeier KK, Wong V (2000) The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. Meteor Atmos Phys 75:161–193

Zalesak ST (1979) Fully multidimensional flux-corrected transport algorithms for fluids. J Comput Phys 31:335–362

Zubov VA, Rozanov EV, Schlesinger ME (1999) Hybrid scheme for three-dimensional advective transport. Mon Wea Rev 127:1335–1346