

# The Inclusion of Non-conservative Forcing into a Conservative, Contour Advection Algorithm

David G. Dritschel  
*University of Warwick*

Maarten H. P. Ambaum  
*Royal Netherlands Meteorological Institute*

## Abstract

This paper describes a novel numerical approach for simulating fine-scale, dominantly advected fields, such as tracers, in layerwise two-dimensional flows, the most important examples of which are the Earth’s atmosphere and oceans. In particular, we describe how one can modify an approach previously developed for perfectly conserved tracers so that it can handle predominantly large-scale, non-conservative forcing. Such forcing is called “diabatic” in the atmosphere and oceans and is mainly the result of absorption or emission of radiation by chemical constituents (*e.g.* carbon dioxide or ozone) in the atmosphere or Ekman pumping by wind-stress in the oceans. This forcing plays a fundamental role in shaping the observed atmospheric and oceanic circulation on long time scales. It “spins up” a circulation which is dynamically unstable, and sets off a sequence of events involving, indeed depending upon, complex, turbulent interactions, whose mean effect is to significantly modify the circulation that would otherwise develop in the absence of instability. It has become recognised that fine-scale structure is both abundant and intimately connected with the present circulation. However, the numerical models widely employed in studying the atmosphere and the oceans inadequately cope with this fine-scale structure; solution convergence is poor with increasing resolution, and the numerical cost is great. Here, we introduce a hybrid method that overcomes the difficulties of resolving fine-scale structure. This method combines a contour-based, grid-free numerical model with a conventional, grid-based one, and introduces a means for accumulating the forcing into the motion of the contours, via the grid-based model. The new approach could ultimately lead to significant improvements in modelling and understanding the atmospheric and oceanic circulation.

## 1 Introduction

Recently, a new algorithm was developed by Dritschel & Ambaum [7], who fused two radically different approaches for simulating conservative fields (purely “advected” fields  $q$  obeying the evolution equation  $\partial q/\partial t + \mathbf{u} \cdot \nabla q = 0$ ) in layerwise-two-dimensional flows, the most important examples of which are the Earth’s atmosphere and oceans. In such flows, the influence of rotation and stable density stratification force motion to be predominantly parallel to the stratification surfaces [14]. In the atmosphere any cross-surface motion is due to *diabatic* (entropy-changing) processes, most significantly thermal heating or damping (*c.f.* [12]). These diabatic processes, though weak and operating on a long time scale, to a great extent shape the observed large-scale circulation of the atmosphere. They destroy, weakly, the conservative evolution of a particularly fundamental field, the potential vorticity  $q$ , whose distribution largely determines all other dynamic fields (*e.g.* velocity, pressure and temperature — see [14]). Nevertheless, the potential vorticity still behaves like a nearly conserved tracer on time scales of a week or so, and as such, it is characterised by abundant fine-scale

structure [18, 25]. This fine-scale structure contributes only weakly to the induced velocity field, etc., but plays a crucial role in determining the form and location of the tropopause dividing weakly stratified air below about 10km from strongly stratified air above [1] as well as ozone depletion within the polar stratospheric vortex, a 6000km wide region of super-rotating air situated over the wintertime pole extending from a height of roughly 12km upwards to 60km [13, 15].

Present research relies strongly on numerical simulation to understand the factors contributing to the observed circulation, to ozone depletion, to climate change, etc. However, the numerical methods employed poorly resolve fine-scale structure, and serious questions have been raised about the research findings, in particular the poor convergence of the results with increasing numerical resolution [3, 6, 9, 10, 26]. The object of the present article is to introduce an alternative numerical method capable of resolving fine-scale behaviour at a tiny fraction of the cost that would be necessary using present numerical methods. This new method builds on the conservative algorithm introduced by Dritschel & Ambaum [7]. The details are described in the following section. In Section 3, numerical tests are presented for the relaxation of parallel flows to a prescribed equilibrium. In Section 4, non-parallel flows are considered, and the dependence on numerical parameters is exhibited. This study justifies the basic approach and pins down the optimal numerical parameter relationships. In Section 5, we describe future applications of the method.

## 2 The Structure of the Algorithm

The new algorithm combines a grid-based model for handling the diabatic input of potential vorticity (hereafter PV) and a contour-based model for the advective part of the evolution. This combination is necessary because a contour-based model alone could never account for diabatic processes in a part of the domain where it is impossible to accurately define a PV gradient or where the PV gradient vanishes. The former is the case wherever the contour density is low — a situation that will inevitably occur during the evolution. The latter occurs at extrema of the PV distribution.

The PV that builds up in the grid-based model through diabatic processes will attain ever smaller scales through advection. It is here important to stress that these processes are quite unlike those associated with ordinary (molecular) diffusion. The latter are efficient at destroying fine-scale structure, but in the real atmosphere, such diffusive effects are utterly negligible over the range of scales that could ever be conceivably modelled. Diabatic processes, by contrast, act most strongly at large scales, due to the broadly distributed nature of thermal heating in the atmosphere. They can either destroy or enhance PV. Essentially, they act to drag the PV back to some quasi-equilibrium profile (one which only slowly changes on the advective time-scale), with large scales being dragged back faster than small scales [2]. In fact, the relaxation of the small scales is so slow that they never come into equilibrium in reality, which is why fine-scale structure is so abundant in the atmosphere.

This natural tendency to produce fine-scale structure by advection strongly limits the success of grid-based models; however, this is where one can exploit the advantages of a contour-based model. In the latter, the PV advection is solved explicitly by advecting PV contours, *i.e.* by computing the trajectories of the (in practice finite) collection of points comprising each contour. Pure advection can be carried out numerically without any constraint on stability, and representing PV as contours permits one to resolve, at low computational cost, fine-scale structure such as sharp gradients and filaments at least an order of magnitude beyond the limits of grid-based methods. Eventually, fine-scale filamentary PV is removed by a topological reconnection scheme called “surgery” but the result is a much reduced dis-

sipation of PV and indeed a complete preservation of high-gradient structures (see [7] & refs).

The crux of the algorithm is that it is possible to use a grid-based model to explicitly account for the diabatic input of PV, and to *transfer* this PV to a contour-based model, before it mixes down to scales that are too small to be accurately resolved by the grid-based model; the grid-based model is used as a temporary reservoir of “diabatic PV.”

Here is how to do this. The PV field  $q$  is written as the sum of a gridded field  $q_d$  and a contoured field  $q_a$ . The contoured field will be written as

$$q_a(x, y, t) = q_0(t) + q_1(t)h(x, y, t) .$$

Here  $q_0$  and  $q_1$  are functions of time alone, and  $h$  is the field represented by contours with unit PV jumps.

The basic equation to be solved is

$$\frac{Dq}{Dt} = \mathcal{S} , \tag{2.1}$$

with  $D/Dt = \partial/\partial t + \mathbf{u} \cdot \nabla$  and  $\mathcal{S}$  the diabatic source term. This source term is assumed to contain predominantly large scales, so that it can be accurately represented on a (moderately-coarse) grid. The velocity field  $\mathbf{u}$  is obtained by whatever standard method used by the grid-based model —  $\mathbf{u}$  itself is not a fine-scale field [7, 16]. Here, in the simple single-layer (two-dimensional) model that we are using for illustration,  $\mathbf{u}$  is obtained from the total PV by “inversion”, *i.e.* by inverting the “Helmholtz” operator

$$\nabla^2 \psi - \frac{1}{L_R^2} \psi = q$$

for the streamfunction  $\psi^1$  and calculating the incompressible velocity field from

$$u = -\frac{\partial \psi}{\partial y} \quad ; \quad v = \frac{\partial \psi}{\partial x}$$

(this is the simplest version of the widely-used “quasi-geostrophic” model [5, 11, 12, 14] and is sufficient to demonstrate the potential utility of the new algorithm in a far wider context).

Eq. 2.1 can be split into the following set of equations,

$$\frac{Dh}{Dt} + \mathbf{u}_d \cdot \nabla h = 0 , \tag{2.2a}$$

$$\frac{Dq_d}{Dt} = \mathcal{S} - \frac{Dq_a}{Dt} = \mathcal{S} - \dot{q}_0 - \dot{q}_1 h + q_1 \mathbf{u}_d \cdot \nabla h , \tag{2.2b}$$

where an extra “diabatic velocity”  $\mathbf{u}_d$  has been introduced. The first equation is solved by contour advection, *i.e.*

$$\frac{d\mathbf{x}}{dt} = \mathbf{u} + \mathbf{u}_d ,$$

which, apart from advection errors arising from using a finite time step, is accurate down to the surgical scale (the scale below which the surgery scheme starts to act, normally a tenth of the grid scale). The second equation is solved on the grid with a “semi-Lagrangian” algorithm [4, 20, 22, 23], the method most widely used now in weather forecasting. The semi-Lagrangian method, like contour advection, uses particle trajectories (now backwards in

<sup>1</sup> $L_R$  is the so-called Rossby radius of deformation, a fundamental length scale dependent on the external rotation and stratification of the fluid. In the atmosphere at mid-latitudes it is some 1500km.

time) to find  $q_d$  on grid points at the next time step. An important difference is that the semi-Lagrangian scheme requires interpolation of  $q_d$ , and this is strongly diffusive for fields that readily generate fine-scale structure, like the total PV field here (see [8] for a comparison between contour advection and the semi-Lagrangian method). This is the reason why we have used contour advection to represent as much of the fine-scale character of  $q$  as possible; the new algorithm attempts to keep  $q_d$  broad, thereby permitting it to be accurately represented on a moderate-resolution grid. The formation of small scales in the  $q_d$  field can be suppressed by having the right-hand side of Eq. 2.2b be equivalent to some damping  $\mathcal{D}$  on  $q_d$ . In fully grid-based models, this damping is often a hyperviscosity (*i.e.*  $\nabla^{2p}q$  with  $p > 1$ ), but we will take a linear damping for reasons given below.

With this damping term defined, we can rewrite Eq. 2.2b as

$$\frac{Dq_d}{Dt} = q_1 \mathbf{u}_d \cdot \nabla h - \mathcal{W} + \mathcal{D} , \quad (2.3a)$$

$$\mathcal{W} = \dot{q}_0 + \dot{q}_1 h + \mathcal{D} - \mathcal{S} . \quad (2.3b)$$

Now the introduction of  $\mathbf{u}_d$  becomes clear: we will choose  $\mathbf{u}_d$  such that  $q_1 \mathbf{u}_d \cdot \nabla h$  compensates for  $\mathcal{W}$  as much as possible. In this case the right-hand side of the  $q_d$  equation has become a damping term. At the same time  $\mathbf{u}_d$  will be chosen smooth so that advection errors in the  $h$  field will be minimized. The quantities  $\dot{q}_0$  and  $\dot{q}_1$  are also chosen to achieve maximal compensation between  $q_1 \mathbf{u}_d \cdot \nabla h$  and  $\mathcal{W}$ .

Theoretically the optimal choice for  $\mathbf{u}_d$  would be

$$\mathbf{u}_d = \frac{\mathcal{W}}{q_1} \frac{\nabla h}{|\nabla h|^2} ,$$

but this is ill-defined for vanishing  $\nabla h$ , or where  $\nabla h$  itself is ill-defined because of a low contour density. The first problem is solved by desingularizing the denominator in the expression for  $\mathbf{u}_d$ . The second problem is solved by smoothing the terms in the expression for  $\mathbf{u}_d$  with a Fourier filter (see below) such that  $\mathbf{u}_d$  is defined everywhere. The value of  $\nabla h$  is determined on the grid after converting the contoured  $h$  field to gridded values using the fast domain-filling technique described in [7].

The ensuing expression for  $\mathbf{u}_d$  becomes

$$\mathbf{u}_d = \frac{1}{q_1} \left( \frac{\mathcal{W} \nabla h}{|\nabla h|^2 + \epsilon^2} \right)_f , \quad (2.4)$$

where the subscript  $f$  denotes the Fourier filter, which makes  $\mathbf{u}_d$  a smooth field. The parameter  $\epsilon$  is related to a spatial scale  $\lambda_m = \pi/k_m$  by  $\epsilon = 1/\lambda_m$ , and it limits the maximum value of  $\mathbf{u}_d$  to

$$\max(\mathbf{u}_d) = \max \left( \frac{\mathcal{W}}{2q_1} \right)_f \lambda_m . \quad (2.5)$$

The filter  $f$  is of the form  $1/(1 + k^2/k_c^2)$  in Fourier space, with  $k$  the total wavenumber of the filtered mode and  $\lambda_c = \pi/k_c$  the filter scale. The filter is chosen this way because now the expected spectrum for  $\mathbf{u}_d$  is steeper than  $k^{-2}$ , so that  $\mathbf{u}_d$  may effectively be represented on a grid (just as the velocity  $\mathbf{u}$  itself, see [7, 16, 17, 24] for justification). Furthermore, this filter does not lead to fringes in the filtered field. From inspection of the equations it may be seen that the strain tensor  $\nabla \mathbf{u}_d$  is maximally proportional to  $k_m$  and inversely proportional to  $k_c$ . This is formalized in Appendix A, where it is shown that, for stability reasons that limit the norm of the strain tensor, we should choose  $k_m = k_c$ .

Now that we have found  $\mathbf{u}_d$  in terms of  $\mathcal{W}$ ,  $h$ ,  $q_1$ , and  $k_c$ , we still have to fix  $\dot{q}_0$  and  $\dot{q}_1$ . These will now be chosen such as to minimize the mean-square production of  $q_d$ , *i.e.* the functional

$$\int (q_1 \mathbf{u}_d \cdot \nabla h - \mathcal{W})^2 dx dy, \quad (2.6)$$

where the integral extends over the whole domain. In Appendix B it is shown that this minimization leads to a system of two linear equations with unknowns  $\dot{q}_0$  and  $\dot{q}_1$ .

In Eq. 2.3 it can be seen that the diabatic input of PV directly affects  $q_d$ , insofar as the diabatic velocity of the contours cannot compensate this diabatic input through the term  $q_1 \mathbf{u}_d \cdot \nabla h$ . The transfer of PV from the grid-based model to the contour-based model is effected through the action of  $\mathcal{D}$ . For suppose that at some moment the forcing is turned off. Then the  $q_d$  field will be damped with a strength  $\mathcal{D}$  and at the same time this PV is transferred to the  $q_a$  field through the advection of the contours by  $\mathbf{u}_d$  and through the contour level changes  $\dot{q}_0$  and  $\dot{q}_1$ . When  $q_d$  vanishes,  $\mathcal{W}$  vanishes, so that  $\mathbf{u}_d$  vanishes also and we end up with a fully adiabatic contour advection model with evolution equation  $Dh/Dt = 0$ .

As the transfer of PV from the grid-based to the contour-based model is determined by the strength of  $\mathcal{D}$ , we can find the optimal form of  $\mathcal{D}$  by the preferred behaviour of the transfer. For example, one should not choose a diffusion type of damping, because the transfer of this term to the  $q_a$  field would lead to a velocity associated with antidiffusion on  $q_d$ . This would make the scheme unstable whenever  $q_a$  and  $q_d$  would correlate in some regions. In our implementation, we take a linear damping of the form

$$\mathcal{D} = -q_d/\tau_{tr},$$

where  $\tau_{tr}$  is the transfer timescale. This timescale should be long enough to have the transfer mechanism work effectively, *i.e.* long enough to have significant contour gradients in  $q_a$  coinciding with PV structures in  $q_d$ . The timescale should be shorter, though, than the mix-down time of the velocity field  $\mathbf{u}$ . The optimal values for  $\tau_{tr}$  and the other numerical parameters are next deduced from direct simulation results in the next two sections.

### 3 Relaxation of Parallel Flows to Equilibrium

In testing the new model, we will employ a commonly-used simple form for the diabatic forcing [1, 2, 19], namely

$$\mathcal{S} = \frac{\psi - \bar{\psi}(x, y)}{\tau_r L_R^2}, \quad (3.1)$$

where  $\psi$  is the streamfunction defined above,  $\bar{\psi}(x, y)$  is a prescribed ‘‘radiative equilibrium’’ profile, and  $\tau_r$  is the relaxation time scale towards radiative equilibrium. For simplicity, we have developed the algorithm first for a doubly-periodic domain, of side length  $2\pi$ , though the extension to the more realistic spherical geometry is straightforward.

If one spectrally transforms this forcing, it becomes readily apparent that large scales relax back to equilibrium faster than do small scales. In fact, for a 2D wavenumber  $k$ , the relaxation rate of a spectral component of  $\psi$  back to equilibrium (ignoring nonlinear interactions) is simply  $1/(\tau_r(k^2 L_R^2 + 1))$ . In the one-dimensional, parallel flow problem considered in this section, nonlinear interactions cannot occur, so all flows do relax back to radiative equilibrium. This is not generally the case for 2D flows, since the radiative equilibrium profile is often unstable.

In all of the simulations performed, the grid resolution was fixed at 128 in each direction (of course only one direction is required in the 1D simulations). This resolution is adequate to

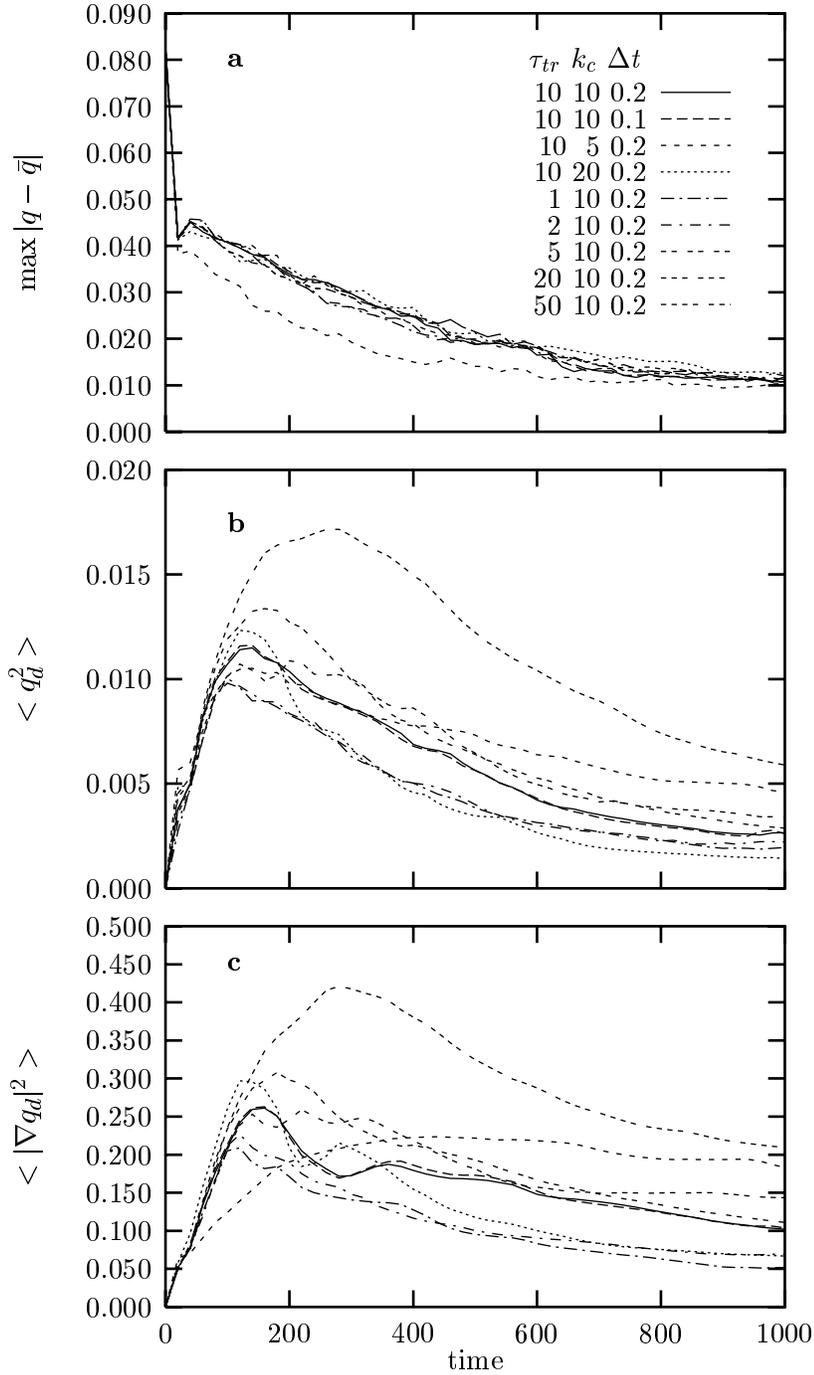


Figure 1: One-dimensional simulation results. (a)  $\max |q - \bar{q}|$  versus  $t$ , (b) r.m.s.  $q_d$  versus  $t$ , and (c) r.m.s.  $|\nabla q_d|^2$  versus  $t$ . The line types corresponding to the different parameter values are tabulated in panel (a) and are the same for panels (b) & (c).

accurately capture the fine-scale structure in the 2D simulations, as has been demonstrated by Dritschel *et al.* [8] for a nearly identical flow configuration.

The basic parallel flow is prescribed by the initial PV, a zig-zag profile,

$$q(x, y, 0) = \eta Q \operatorname{sgn}(y)(a - ||y| - a|) \quad (3.2)$$

for  $|y| < 2a$ , and  $q = 0$  otherwise. Note,  $q = \pm\eta Q$  at  $y = \pm a$ . As for the diabatic forcing, the radiative equilibrium profile  $\bar{\psi}$  is chosen to be the streamfunction one gets by inverting the above PV when  $\eta = 1$ . To force an adjustment,  $\eta \neq 1$  is used for  $q(x, y, 0)$ ; in fact, we have chosen  $\eta = 0.66845$ . This forces  $q$  to adjust to its equilibrium profile  $\bar{q} = Q \text{sgn}(y)(a - ||y| - a|)$ . We have taken  $a = 0.5$  and  $L_R = 0.5$  as in [8], as well as  $Q = \pi/2$  and  $\tau_r = 10$ .

The object in this section is to determine what choice of numerical parameters leads to the smoothest, most accurate adjustment to equilibrium. To that end, we varied the time step  $\Delta t$ , the filter wavenumber  $k_c$ , and the damping time scale  $\tau_{tr}$  in a series of simulations. No significant sensitivity was found for  $\Delta t$ , as shown below. The recommended time step, for accurate advection, is

$$\Delta t = \frac{\pi}{10 \max |q|} , \quad (3.3)$$

[7] where  $\max |q|$  is the maximum PV amplitude in the equilibrium profile, *i.e.*  $Q$  here. With  $Q = \pi/2$ , the above formula gives  $\Delta t = 0.2$ .

Results for  $k_c = 5, 10$  &  $20$ ,  $\tau_{tr} = 1, 2, 5, 10$  &  $20$  and  $\Delta t = 0.2$  &  $0.1$  are shown next in figures 1(a)–(c). Fig. 1(a) shows  $\max |q - \bar{q}|$  versus time  $t$ , Fig. 1(b) shows the root-mean-square (r.m.s.) value of  $q_d$  versus  $t$ , and Fig. 1(c) shows the r.m.s. value of  $\nabla q_d$  versus  $t$  (see caption for a description of the line styles used to differentiate the results).

The decay of  $\max |q - \bar{q}|$  varies little across the range of numerical parameters investigated. Within numerical errors, this quantity should behave the same in all simulations. Even so, it is evident that large values of  $\tau_{tr}$  and small values of  $k_c$  lead to significant numerical errors. This is borne out in Figs. 1(b) & (c), which display much more sensitive measures of the numerical solution accuracy. In fact, the algorithm is designed to minimise the r.m.s. value of  $\nabla q_d$  (Fig. 1(c)) for *given*  $k_c$ ,  $\tau_{tr}$  and  $\Delta t$ ; the above results show how one can best choose these global parameters to minimise the r.m.s. value of  $\nabla q_d$  overall. In fact, Figs. 1(b) & (c) collectively point to  $\tau_{tr} = 1$  or  $2$  and  $k_c = 10$  or  $20$  being optimal, with no significant dependence on the time step  $\Delta t$ . Hence, the optimal timescale over which to transfer  $q_d$  to  $q_a$  is approximately the characteristic time-scale  $T \equiv \pi / \max |q|$ . The optimal filter wavenumber is approximately one tenth of the maximum wavenumber. We next consider the much richer two-dimensional flow to see if these parameter relationships continue to hold.

## 4 Non-Parallel Flows and Numerical Parameter Choices

In this section, the initial PV distribution considered in the previous section is modified by replacing  $y$  in Eq. 3.2 by

$$\hat{y} = y + c_m \sin mx + c_n \sin nx . \quad (4.1)$$

Furthermore,  $\eta = 1$  is used now and the “radiative-equilibrium” flow  $\bar{\psi}$  is determined by inverting this initial, non-parallel profile of PV. In Eq. 4.1, we have taken  $m = 2$ ,  $c_2 = -0.05$  and  $n = 3$ ,  $c_3 = 0.05$ . All other physical parameters are the same as previously.

The initial flow is unstable despite the effect of the relaxation term  $\mathcal{S}$ , as Fig. 2 illustrates. However, one sees that the instability collapses and a new nearly parallel flow emerges, which is itself unstable. The flow destabilizes and recovers repeatedly, though non-periodically, and this can be ascribed to the effect of the relaxation term. With  $\mathcal{S} = 0$ , the initial flow breaks down and spreads throughout the domain (not shown).

The dependence on the numerical parameters is discussed next. The time evolution of r.m.s.  $q_d$  and  $\nabla q_d$  for  $k_c = 10, 20$  &  $40$  and  $\tau_{tr} = 0.5, 1, 2, 3, 4, 5$  &  $10$  (all for  $\Delta t = 0.2$ ) are shown in figures 3(a) & (b). These results confirm that  $\tau_{tr} = 1$  or  $2$  and  $k_c = 10$  or  $20$  (the latter being marginally better) give the best results in terms of keeping  $q_d$  a smooth,

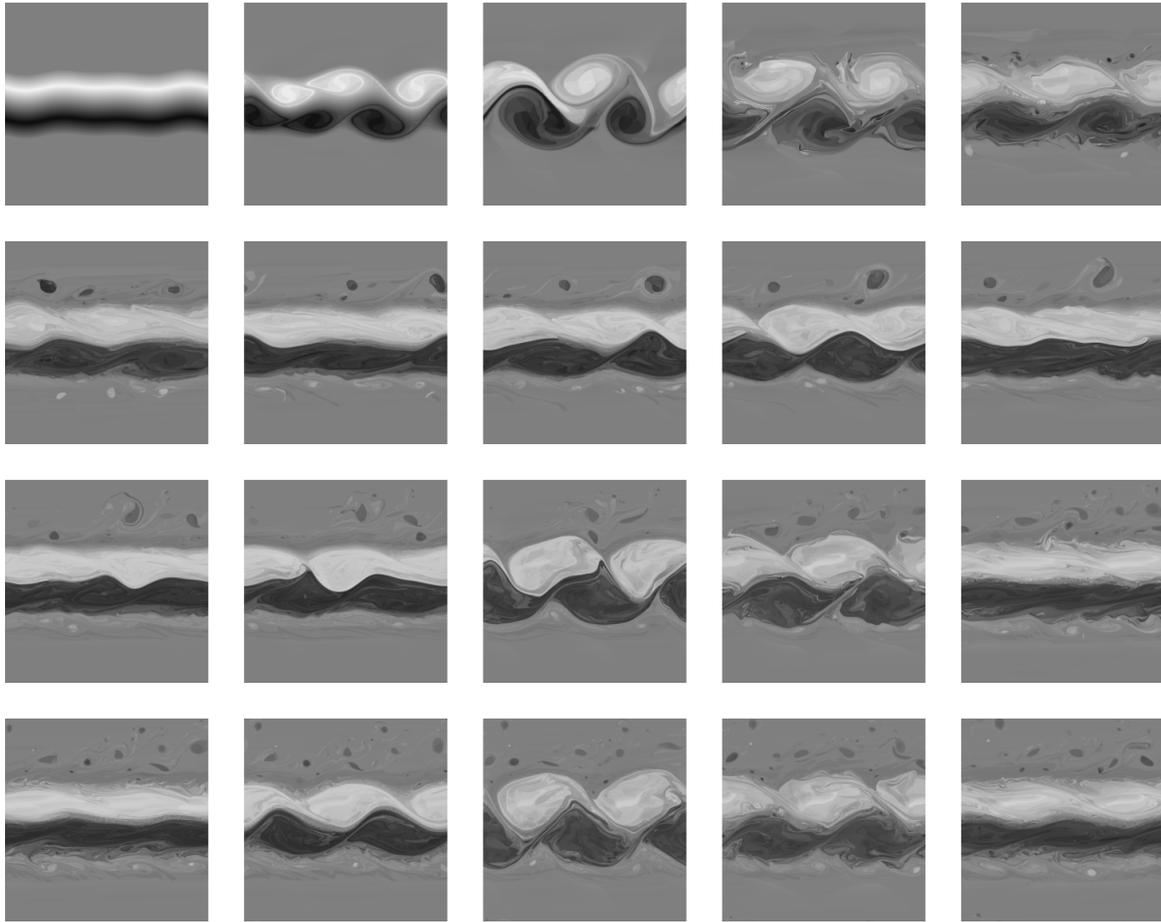


Figure 2: Evolution of the PV field in the two-dimensional simulation. Here, a 256 by 256 grid was used,  $k_c = 20$ ,  $\tau_{tr} = 1$  and  $\Delta t = 0.2$ . Time (in “days”) advances from left to right and then downwards. The times shown are 0, 10, 20, 40 and 60 (top panel), 100, 140, 160, 180 and 200 (second panel), 240, 260, 280, 300 and 320 (third panel), and 340, 360, 380, 400 and 420 (bottom panel).

low-amplitude field. In particular, either too small or too large  $\tau_{tr}$  leads to poorer results. These results justify the choice of the filter wavenumber  $k_c$  and transfer timescale  $\tau_{tr}$

$$k_c = 0.15k_{max} \quad , \quad \tau_{tr} = \frac{\pi}{Q} \quad (4.2)$$

or up to a factor of two smaller.

## 5 Future Work

The feasibility of including diabatic forcing in a contour-based numerical algorithm opens up a timely opportunity for greatly improved modelling of atmospheric and oceanic phenomena. The new algorithm results in much less dissipation of fine-scale, principally-advected flow structures, such as fronts in the atmosphere and their analogues in the oceans. The proliferation of these structures in the real atmosphere and oceans, and generally in planetary

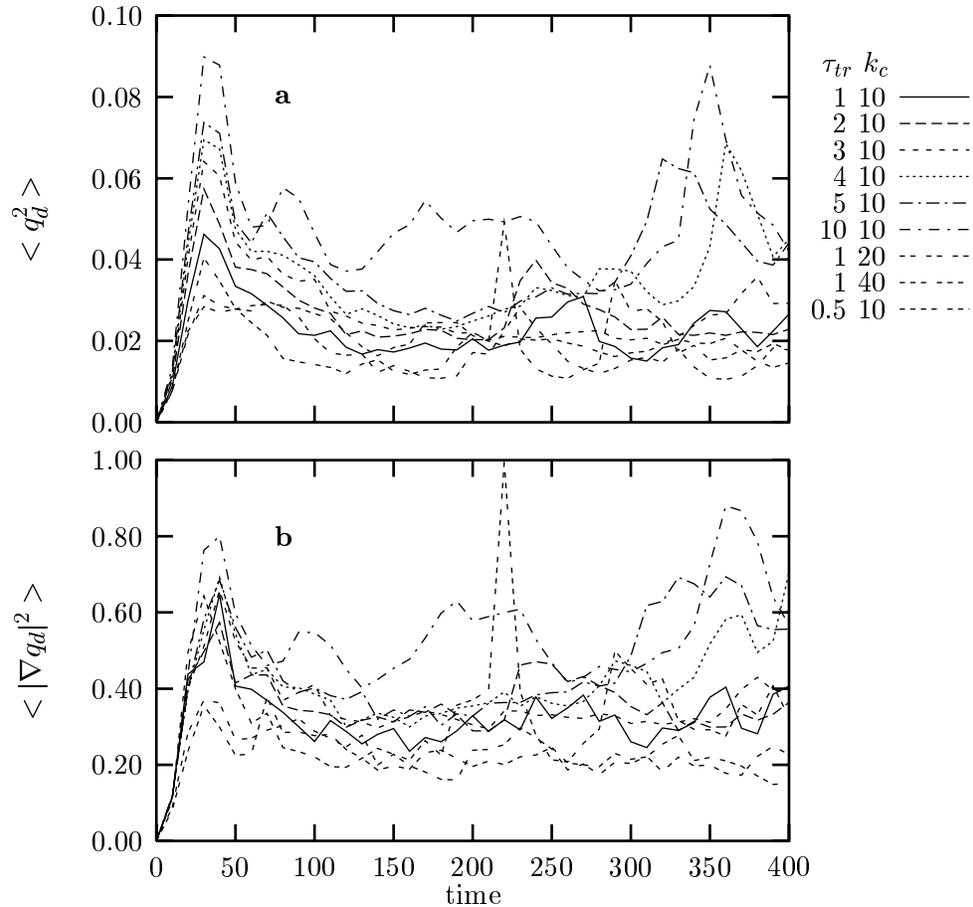


Figure 3: Two-dimensional simulation results. (a) r.m.s.  $q_d$  versus  $t$ , and (b) r.m.s.  $\nabla q_d$  versus  $t$ . The line types corresponding to the different parameter values are tabulated in panel (a) and are the same for panel (b).

atmospheres, is a consequence of the practically negligible diffusion in these flows. Rather, dissipation (and forcing) occurs principally at large scales, *e.g.* by solar radiation. These flows are consequently difficult to model using conventional, grid-based numerical models, which cannot cope with the natural tendency for the formation of fine-scale structure. The new algorithm represents a breakthrough in this respect.

Presently an effort is made to set up comparative experiments with semi-Lagrangian models, in order to assess the relative gain of the new method. At the same time the algorithm is being extended to represent more realistic descriptions of the dynamics of the atmosphere. First steps are the inclusion of spherical geometry, multilayered versions of the algorithm, non-linear dependency of the velocity field on the PV, or a primitive equation version of the algorithm. These steps could ultimately lead to a General Circulation Model, with the diabatic CASL algorithm as its dynamical kernel.

## References

- [1] M. H. P. Ambaum. On the formation of the tropopause. *J. Atmos. Sci.*, 54:555–568, 1997.

- [2] J. L. Anderson. A simulation of blocking with a forced barotropic model. *J. Atmos. Sci.*, 52:2593–2608, 1995.
- [3] C. Appenzeller, H. C. Davies, and W. A. Norton. Fragmentation of stratospheric intrusions. *J. Geophys. Res.*, 101:1435–1456, 1996.
- [4] J. R. Bates, Y. Li, A. Brandt, S. F. McCormick, and J. Ruge. A global shallow-water numerical model based on the semi-lagrangian advection of potential vorticity. *Q. J. Roy. Meteorol. Soc.*, 122:1981–2005, 1995.
- [5] J. G. Charney. On the scale of atmospheric motion. *Geophys. Publ.*, 17 No. 2, 1948. 17pp.
- [6] W. K. Dewar and P. D. Killworth. Do fast gravity-waves interact with geostrophic motions? *Deep-sea Res., Part I*, 42:1063–1081, 1995.
- [7] D. G. Dritschel and M. H. P. Ambaum. A contour-advective semi-lagrangian algorithm for the simulation of fine-scale conservative fields. *Q. J. Roy. Meteorol. Soc.*, 123:1097–1130, 1997.
- [8] D. G. Dritschel, L. M. Polvani, and A. R. Mohebalhojeh. The contour-advective semi-lagrangian algorithm for the shallow water equations. *Mon. Wea. Rev.*, 1998. (submitted).
- [9] S. Edouard, B. Legras, F. Lefèvre, and R. Eymard. The effect of small-scale inhomogeneities on ozone depletion. *Nature*, 384:444–447, 1996.
- [10] S. Edouard, B. Legras, and V. Zeitlin. The effect of dynamical mixing in a simple model of the ozone hole. *J. Geophys. Res.*, 101(D11):16771–16779, 1996.
- [11] A. E. Gill. *Atmosphere-Ocean Dynamics*. Academic Press, 1982. 662pp.
- [12] J. R. Holton. *An introduction to dynamic meteorology (3rd edition)*. San Diego, Academic Press, inc., 1982. 511pp.
- [13] J. R. Holton, P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister. Stratosphere–troposphere exchange. *Revs. Geophys.*, 33(4):403–439, 1995.
- [14] B. J. Hoskins, M. E. McIntyre, and A. W. Robertson. On the use and significance of isentropic potential vorticity maps. *Q. J. Roy. Meteorol. Soc.*, 111:877–946, 1985.
- [15] M. E. McIntyre. The stratospheric polar vortex and sub-vortex: fluid dynamics and midlatitude ozone loss. *Phil. Trans. Roy. Soc. London*, 352:227–240, 1995.
- [16] J. Methven and B. J. Hoskins. On the advection of high resolution tracers by low resolution winds. *J. Atmos. Sci.*, 1998. (submitted).
- [17] W. A. Norton. Breaking rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a technique for advecting material contours. *J. Atmos. Sci.*, 51:654–673, 1994.
- [18] R. A. Plumb, D. W. Waugh, R. J. Atkinson, M. R. Schoeberl, L. R. Lait, P. A. Newman, E. V. Browell, A. Simmons, M. Loewenstein, and D. W. Toohey. Intrusions into the lower stratospheric arctic vortex during the winter of 1991/1992. *J. Geophys. Res.*, 99:1089–1106, 1994.

- [19] L. M. Polvani, D. W. Waugh, and R. A. Plumb. On the subtropical edge of the stratospheric surf zone. *J. Atmos. Sci.*, 52:1288–1309, 1994.
- [20] H. Ritchie, C. Temperton, A. Simmons, M. Hortal, T. Davies, D. Dent, and M. Harmud. Implementation of the semi-lagrangian method in a high resolution version of the ecmwf forecast model. *Mon. Wea. Rev.*, 123:489–514, 1995.
- [21] P. K. Smolarkiewicz and J. A. Pudykiewicz. A class of semi-lagrangian approximations for fluids. *J. Atmos. Sci.*, 49:2082–2096, 1992.
- [22] A. Staniforth and J. Côté. Semi-lagrangian integration schemes for atmospheric models — a review. *Mon. Wea. Rev.*, 11(9):2206–2223, 1991.
- [23] C. Temperton and A. Staniforth. An efficient two-time-level semi-lagrangian semi-implicit integrating scheme. *Q. J. Roy. Meteorol. Soc.*, 113:1025–1040, 1987.
- [24] D. W. Waugh and R. A. Plumb. Contour advection with surgery: a technique for investigating finescale structure in tracer transport. *J. Atmos. Sci.*, 51:530–540, 1994.
- [25] D. W. Waugh, R. A. Plumb, R. J. Atkinson, M. R. Schoeberl, L. R. Lait, P. A. Newman, M. Loewenstein, D. W. Toohey, L. M. Avallone, C. R. Webster, and R. D. May. Transport of material out of the stratospheric arctic vortex by rossby wave breaking. *J. Geophys. Res.*, 99:1071–1078, 1994.
- [26] C. Wunsch and D. Stammer. The global frequency-wavenumber spectrum of oceanic variability estimated from TOPEX/POSEIDON altimetric measurements. *J. Geophys. Res.*, 100(C12):24895–24910, 1996.

## Appendix A Relation of parameters to Lifshitz number

The “stability” of an advection scheme is related to the so-called Lifschitz-number [21]  $L$  which is defined as the norm of the strain matrix multiplied by the timestep

$$L = \Delta t \|\nabla \mathbf{v}\| ,$$

with  $\mathbf{v}$  the advecting velocity. Here “stability” does not refer to the absence of numerical instabilities, but rather to the absence of trajectory crossings during one timestep  $\Delta t$ . The criterion for a stable advection scheme is then

$$L \lesssim 1 .$$

The Lifschitz number should pertain to the total advective velocity, so  $\mathbf{v} = \mathbf{u} + \mathbf{u}_d$ . As we can only give a rather crude a priori estimate of  $\mathbf{u}_d$  we will demand that the Lifschitz number  $L_d$  pertaining to  $\mathbf{u}_d$  is smaller than the Lifschitz number  $L_u$  pertaining to  $\mathbf{u}$  alone, so  $L_d < L_u$ . The norm of the strain matrix  $\nabla \mathbf{u}$  can be estimated as the inverse of the eddy turnover time  $\tau_e$ , so  $L_u \approx \Delta t / \tau_e$ . The upper bound of the Lifschitz number  $L_d$  can be estimated by using Eq. 2.5 to find an estimate of the maximum norm of the strain matrix of  $\mathbf{u}_d$ : the velocity field  $\mathbf{u}_d$  is proportional to the magnitude of  $\mathcal{W}_f$  and its gradient is limited by the Fourier filter with filter wavenumber  $k_c$ . Therefore the norm of the strain matrix of  $\mathbf{u}_d$  is approximately limited by  $(k_c/k_m) \max(\mathcal{W}_f)/q_1$ . The magnitude of  $\mathcal{W}_f$  will be minimised during the simulation because of the minimisation of the functional in Eq. 2.6, which is more-or-less proportional

to  $\mathcal{W}$ . But  $\mathcal{W}$  can never fully vanish because in Eq. 2.3b the term  $\dot{q}_1 h$  makes discrete jumps over contours with magnitude  $\dot{q}_1$  whereas all other terms are smooth fields. This means that  $\mathcal{W}$  can only be reduced to values of about  $\dot{q}_1$ . As  $\dot{q}_1$  is inversely proportional to the number of contours  $m$ , we then have that  $\mathcal{W}$  is limited by a value of the order of  $\max(\mathcal{S})/m$ . The magnitude of  $\max(\mathcal{S})$  is for a radiative forcing inversely proportional to the radiative timescale  $\tau_r$  and proportional to the PV contrast in the fluid which is equal to  $m\dot{q}_1$ . So  $\max(\mathcal{W})/\dot{q}_1 = \max(\mathcal{S})/m\dot{q}_1 = 1/\tau_r$ . This then leads to an estimated upper bound on  $L_d$  of

$$L_d \lesssim \frac{k_c}{k_m} \frac{\tau_e}{\tau_r} \frac{\Delta t}{\tau_e} ,$$

where we have introduced the eddy turnover time  $\tau_e$  for convenience. Using the previous estimate for  $L_u$ , it then immediately follows that

$$L_d \lesssim \frac{k_c}{k_m} \frac{\tau_e}{\tau_r} L_u .$$

As normally  $\tau_e < \tau_r$ , it is save to choose  $k_c \sim k_m$ . An interesting consequence is that now the estimated upper bound on  $L_d$  is independent of  $k_c$ .

## Appendix B Determination of $\dot{q}_0$ and $\dot{q}_1$

We want to determine  $\dot{q}_0$  and  $\dot{q}_1$  from minimizing the functional in Eq. 2.6. Now introducing

$$\mathbf{g} = \frac{\nabla h}{|\nabla h|^2 + \epsilon^2} \text{ and } \mathcal{W}_0 = \mathcal{D} - \mathcal{S} ,$$

we can rewrite the functional in Eq. 2.6 as

$$\int [\dot{q}_0(\mathbf{g}_f \cdot \nabla h - 1) + \dot{q}_1((h\mathbf{g})_f \cdot \nabla h - h) + (\mathcal{W}_0\mathbf{g})_f \cdot \nabla h - \mathcal{W}_0]^2 \text{ dxdy} ,$$

which is of the form

$$\int [\dot{q}_0 A + \dot{q}_1 B - C]^2 \text{ dxdy} ,$$

with  $A$ ,  $B$ , and  $C$  known fields. This functional can be minimized with respect to variations in  $\dot{q}_0$  and  $\dot{q}_1$ , by having the partial derivatives of the functional with respect to these two quantities equal to zero. This leads to two linear equations in the unknowns  $\dot{q}_0$  and  $\dot{q}_1$ . Writing the domain integral of any function  $\phi$  as  $\langle \phi \rangle$ , we can write these two linear equations as

$$\begin{aligned} \langle A^2 \rangle \dot{q}_0 + \langle AB \rangle \dot{q}_1 &= \langle AC \rangle , \\ \langle AB \rangle \dot{q}_0 + \langle B^2 \rangle \dot{q}_1 &= \langle BC \rangle . \end{aligned}$$

This set of equations is always solvable because  $A \neq B$  whenever contours are present.