Adjoint of a non-hydrostatic model

P.Bernardet, S.Farges, K.Yassine, CNRM, Météo-France

April 27, 1993

Abstract

The form of the adjoint equations depends upon the scalar product chosen. We argue that, for adjoint model assimilation, the adequate scalar product, giving convergent gradients, is based upon the covariances of the guess errors. We give an estimate for the condition number of the minimization problem. Our non-hydrostatic model evolves under the anelastic constraint; this constraint should be used at observation times, instead of applying a filter at initial time as would be the case for a compressible model.

However, for the adjoint equations to be compared to the direct ones, the appropriate scalar product is based upon energy. Non-hydrostatic equations are usually written in a form closer to the original Euler equations of fluid dynamics than hydrostatic ones; it gives an opportunity to have a closer look at the adjoint equations and their relationship to the direct ones. We discuss the behaviour of advection terms and the exchange between buoyancy and vertical advection of reference potential temperature.

Conservation of energy establishes adjoint relationships between the discretized operators of the direct equation. Among other things, this leads to a symmetric formulation of the pressure equation. While it is not automatic that the adjoint of a discretized operator converges, so it does with a careful choice of discretization. The pressure equation with this symmetric discretization is easily solved by an iterative procedure.

The development of the adjoint is eased as many operators are already present in the direct model.

1 Introduction

During the Pyrex experiment over the French Pyrenees mountains some stationary non-hydrostatic flows have been observed. We wish to apply the interpolation principles of Rabier & Bernardet (1991) to these data. That is, we wish to minimize the model tendency while enforcing the state of the model to agree with the observations. An important part of the signal collected by the research planes during the experiment is due to lee waves. The resolution of this interpolation problem by gradient methods makes the need for the adjoint of a non-hydrostatic model with orography.

The construction of the functional to minimize does not make reference to a scalar product; it is the notion of gradient which points to an underlying scalar product we have to choose; choosing it is of no consequence upon the minimum; however the right choice will ease the minimization process. We should not forget that what we try to analyze are continuous fields of velocity, temperature represented through a discretization process. We should then ensure that the scalar product chosen is such that the gradients are sufficiently regular and that they are not misrepresented through discretization.

On the other hand, the equations of motion conserve some form of energy, and scalar products which bear a relation to it seem more natural in order to interpret the dynamics of the adjoint equations; at this stage the equations should make no reference to a particular reference system or discretization process.

We choose the anelastic type of model; this is more convenient as it simplifies the model state and as we disregard acoustic waves; we have to look how this anelastic constraint conveys to the adjoint and how it is taken into account at observation time.

Hamiltonian mechanics are a method to deduce energetically consistent model equations from an approximation to the energy; here, we adopt a more limited point of view to show how the notion of adjoint helps to design a consistant set of direct equations.

Spatial discretization uses the C-Arakawa grid; the interest here is that, in contrast with a Galerkine approximation, some choice is left for the discretization of the operators; even if the direct equations do not conserve energy, we can construct their adjoint; however, enforcing conservation shows that some of the discretized operators are linked together by adjoint relationships; as a consequence, the elliptic equation for pressure involves a symmetric operator; then its solution can be found by a standard preconditioned conjugate gradient algorithm.

So the plan of this paper is the following:

1. independently of the type of model, we discuss adjoint model assimilation as an interpolation problem and what scalar product one should use to get a well behaved minimization

2. we introduce the anelastic equations and use the notion of adjoint operator in the design

3. we try to interpret the adjoint equations and discuss what scalar product is the most appropriate

4. we use the notion of adjoint for the discretization of the direct model and show that a minimum effort is necessary to get the adjoint once we have the direct

2 Scalar product for adjoint model assimilation

The variational methods seek to blend in a single assimilation step observations gathered during a finite period; the general problem is cumbersome but admits several simplifications; the adjoint model method of assimilation neglects the model error; it is applicable whenever the spatial scales of the deviations between the analysis and the guess $\mathbf{x}_g$ have their evolution accurately predicted during the analysis period; of course, finer scales remain unpredictable whatever the accuracy of the model for a too long assimilation period, but scales described by the operational rawinsonde network are used accurately by the current models.

In this section, we examine the importance of the scalar product chosen to define the descent direction during adjoint model assimilation.

2.1 The variational problem

The state of a model $\mathbf{x}$ is composed of several scalar fields; for the vertical plane anelastic model we are to use, we have horizontal and vertical speed, and potential temperature: $\mathbf{x} = \begin{pmatrix} \vec{u} \\ \theta \end{pmatrix}$. Then we seek the state of the model $\mathbf{x}_0$ at the initial time t_0 of the assimilation period. We then construct a single scalar functional J where all sources of disappointment are gathered:

- a smoothing term[1]expresses that we want the initial fields $\mathbf{x}_0 = \mathbf{x}(t_0)$ to be as smooth as the spatial covariances C of errors of the guess $\mathbf{x}_{g0} = \mathbf{x}_g(t_0)$ entitles us to expect;

- an observation term states that we expect the true state of the atmosphere at location m_k and time t_k, represented in the model by $\mathbf{x}(m_k, t_k)$, and the K observations $\tilde{\mathbf{x}}_k$ to be close; we assume that observations are not correlated, so departures from observations are weighted by the variance of the observation error o_k^2.

Here, we can introduce the "observation operators" H_k. In the case of pointwise observations, they are defined by $H_k\mathbf{x} = \mathbf{x}(m_k, t_k)$. Thus we write:

$$J(\mathbf{x}) = 1/2 \int_\Omega (\mathbf{x}_0 - \mathbf{x}_{g0})\, C^{-1}\, (\mathbf{x}_0 - \mathbf{x}_{g0}) + 1/2 \sum_{k=1,K} o_k^{-2}\, (H_k\mathbf{x} - \tilde{\mathbf{x}}_k)^2 \qquad (2.1)$$

If the guess and the observations are all taken at the same time, minimizing J amounts to the variational formulation equivalent to optimum interpolation. Otherwise, to relate the

[1]The smoothing term is written in 2.1 in a concise way. If $\mathbf{x}$(m) is a scalar depending upon the space location m, then C or C^{-1} are scalars depending upon two space markers m and n; the smoothing term should be written: $\int(\mathbf{x}(m) - \mathbf{x}_g(m)).C^{-1}(m,n)(\mathbf{x}(n) - \mathbf{x}_g(n))dmdn$. C and C^{-1} satisfy: $\forall \mathbf{x} : \int C(m,n)C^{-1}(n,p)\mathbf{x}(p)dndp = \mathbf{x}(m)$;

When $\mathbf{x}$ is a vector field, C is a matrix depending upon two space markers and the dot in the smoothing term refers to a scalar product.

If $\mathbf{x}$ is a scalar and space markers are discretized, $\mathbf{x}(m_i) = \mathbf{x}_i$ is a vector and C is a covariance matrix. If $\mathbf{x}$ is a vector, C is a matrix of matrices.

different observations times we rely on the fact that the atmospheric state x(m,t) should satisfy the non-linear model equations N, assumed accurate; this strong constraint is embodied in the Lagrangian $\mathcal{L}$:

$$\mathcal{L}(\mathbf{x},\mathbf{y}) = J(\mathbf{x}) + \int_{\Omega\times[t_0,T]} \mathbf{y}.\left(\frac{\partial \mathbf{x}}{\partial t} - N\mathbf{x}\right)$$

y is a Lagrange multiplier and Ω the domain of the model. The Euler equations list the conditions for J to be a minimum under the constraint.

2.2 Adjoint equations

Stating that the Lagrangian is stationary under variations of x and y results in the Euler equations; the first one relates to the variation in y and gives again the constraint:

$$\frac{\partial \mathbf{x}}{\partial t} = N\mathbf{x} \tag{2.2}$$

the second one relates to the variations in x:

$$\delta\mathcal{L}(\mathbf{x},\mathbf{y}) = \delta J(\mathbf{x}) + \int_{\Omega\times[t_0,T]} \mathbf{y}.\left(\frac{\partial \delta\mathbf{x}}{\partial t} - L\delta\mathbf{x}\right) \tag{2.3}$$

where L, linearization of N, is defined by: $N(\mathbf{x}+\delta\mathbf{x}) = N(\mathbf{x}) + L(\mathbf{x})\delta\mathbf{x}$; L is usually a differential operator; various integrations by parts lead to isolate δx in 2.3; more generally, we define a scalar product by a matrix operator M; for our anelastic model it will be:

$$\langle \mathbf{x};\mathbf{x}'\rangle_M = \int \mathbf{x}M^{-1}\mathbf{x}' = \int_{\Omega\times[t_0,T]} \bar{\rho}\left(\vec{u}\vec{u}' + \frac{g}{\theta\theta_z}\theta\theta'\right) \tag{2.4}$$

and recall that the definition of the adjoint of a linear operator L is: $\langle \mathbf{y}; L\delta\mathbf{x}\rangle_M = \langle L^*\mathbf{y}; \delta\mathbf{x}\rangle_M$ then taking the adjoint of the linear differential operator L is just making the required integration by parts; the second Euler equation is thus what is called the adjoint equation of 2.2:

$$\frac{\partial \mathbf{y}}{\partial t} + L^*\mathbf{y} = -o_k^{-2}H_k^*\left(H_k\mathbf{x} - \tilde{\mathbf{x}}_k\right)\delta_{t_k} \tag{2.5}$$

with the appropriate time boundary conditions:

$$\begin{cases} \mathbf{y}_0 & = MC^{-1}\mathbf{x}_0 \\ \mathbf{y}_T & = 0 \end{cases} \tag{2.6}$$

The adjoint variable y satisfies the adjoint equation L^* between observation times; at each observation time the adjoint variable undergoes a jump at the observation location; so the adjoint model has the task to integrate backwards the pointwise perturbations represented by the Dirac weights δ_i at the observation points m_i. Solving the assimilation problem amounts to find x and y satisfying 2.2 and 2.5 with the initial and final time boundary conditions 2.6.

2.3 Underlying interpolation

Some insight in the interpolation method can be gained if we do not follow the above approach. Instead, as model states $\mathbf{x}(t)$ satisfy 2.2 we eliminate $\mathbf{x}(t_k)$ in 2.1 in terms of $\mathbf{x}(t_0)$ thanks to the resolvent $\mathcal{N}$ defined such that $\vec{\mathbf{x}}(m,t_k) = \mathcal{N}_k \vec{\mathbf{x}}(m,t_0)$; then J is written:

$$J(\mathbf{x}_0) = 1/2 \int_\Omega (\mathbf{x}_0 - \mathbf{x}_g) C^{-1} (\mathbf{x}_0 - \mathbf{x}_g) + 1/2 \sum_k o_k^{-2} (H_k \mathcal{N}_k \mathbf{x}_0 - \tilde{\mathbf{x}}_k)^2 \qquad (2.7)$$

J will be minimum if its first variation is null; here we write δJ using the scalar product 2.4:

$$\delta J = \left\langle MC^{-1} (\mathbf{x}_0 - \mathbf{x}_g) ; \delta \mathbf{x}_0 \right\rangle_M + \sum o_k^{-2} \left\langle \mathcal{L}_k^* H_k^* (H_k \mathcal{N}_k \mathbf{x}_0 - \tilde{\mathbf{x}}_k) ; \delta \mathbf{x}_0 \right\rangle \qquad (2.8)$$

where we have introduced $\mathcal{L}$, the linearized resolvent:

$$\mathcal{N}_i(\mathbf{x} + \epsilon \delta \mathbf{x}) = \mathcal{N}_i \mathbf{x} + \epsilon \mathcal{L}_i \delta \mathbf{x} + o\left(\epsilon^2\right)$$

and its adjoint $\mathcal{L}^*$. In the case of pointwise observations, H_k^* represents a pointwise perturbation of the model state. Let us define the scalars λ_k by:

$$\lambda_k = -o_k^{-2} (H_k \mathcal{N}_k \mathbf{x}_0 - \tilde{\mathbf{x}}_k)$$

These scalars are functions of the solution $\mathbf{x}_0$ and are not known before the problem is solved; however the following relation is necessary for δJ to be null:

$$(\mathbf{x}_0 - \mathbf{x}_{g0}) = \sum_{k=1,K} \lambda_k C M^{-1} \mathcal{L}_k^* H_k^* = \sum_{k=1,K} \lambda_k C \mathcal{L}_k' H_k' \qquad (2.9)$$

$\mathcal{L}'$ is the transpose of $\mathcal{L}$; we have used the relation $\mathcal{L}' = M^{-1} \mathcal{L}^* M$; as expected, the choice of the scalar product which defines the adjoint has no influence upon the solution.

2.9 states that the analyzed deviation from the guess at initial time is the sum of the K functions of space $C\mathcal{L}_k' H_k'$ (these are the Green functions of the problem); we stress that the number K of observations in an assimilation problem is determined by the observation network, while N, number of model degrees of freedom should be chosen large enough so that the solution $\mathbf{x}$ converges; otherwise, in the case of a too dense observation network (compared to model resolution) some aliasing should occur.

Fig. 1 shows these Green functions for the anelastic model linearized around rest state and C isotropic when t_k varies. Fig. 2 is obtained with an aspect ratio of 10 between the horizontal and vertical scales.

The analysis at final time is:

$$\mathbf{x}_T = \mathbf{x}_{gT} + \sum \lambda_k \mathcal{N}_T C \mathcal{L}_k^* H_k^* \qquad (2.10)$$

If the resolvent $\mathcal{N}$ and its linearization $\mathcal{L}$ are identity (this is what occurs if $T = t_0$), we have optimum interpolation with covariance C:

$$\mathbf{x}_T = \mathbf{x}_{gT} + \sum \lambda_k C H_k^* \qquad (2.11)$$

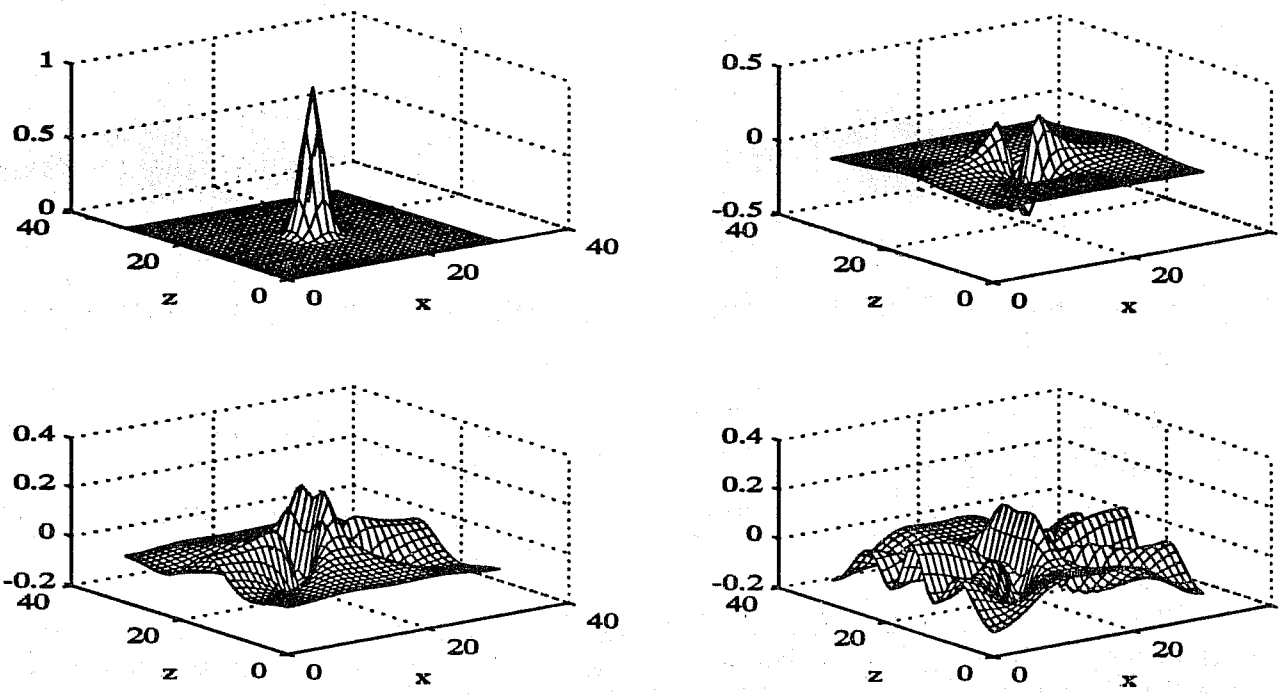

Figure 1: Green functions 2.9 for the biperiodic (x,z) anelastic model linearized around rest state and reference potential temperature; an observation of temperature is placed at the center of the domain for a time-lag $t_k - t_0 = 0,\ \pi/N,\ 2\pi/N,\ 4\pi/N$ with N the Brunt-Vaissala frequency

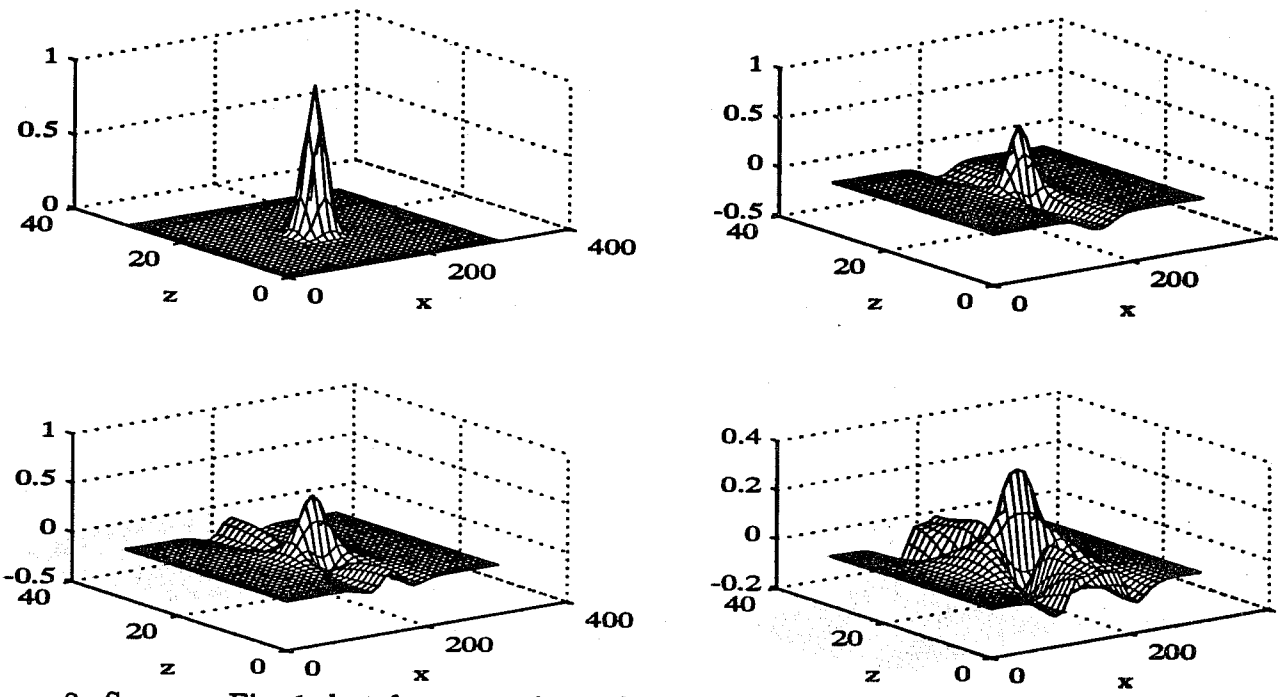

Figure 2: Same as Fig. 1, but for a covariance function with the horizontal scale stretched by a factor of 10

Comparing 2.10 and 2.11 we see that the covariance used to determine the Green functions with nontrivial $\mathcal{N}$ is $\mathcal{N}C\mathcal{L}^*$.

In the simple linear, homogeneous case described in fig. 1, there is no difference between the Green functions 2.10 and 2.11[2]. Differences would come for example from amplification processes absent here.

The adjoint model, instead of transporting the covariance matrix forwards as does the Kalman filter, transports by $\mathcal{L}^*$ the deviations from the observations, analyzes them at time t_0 with covariance C, and transports the result forwards with the non-linear model. The advantage, compared with Kalman Bucy filtering, is computational economy; the difficulty, present also in Kalman-Bucy filtering, is we expect the discretized adjoint to be able to cope with pointwise perturbations H_k^* well enough so that the Green functions $C\mathcal{L}_k' H_k'$ are well approximated (for distributed observations, such as satellite soundings, the problem is alleviated); the direct model has to accurately predict the non-linear evolution of the smoothed fields 2.9, which might be a far more easy job.

The simple example of fig. 1 shows that, in a first approximation, dispersion is used to relate obervations made at different times but has no net effect upon the evolution of the covariances. Covariances are advected in the absence of amplification processes.

[2]This will be shown in the Interpretation of adjoint section

2.4 Consistant scalar product

To solve the assimilation problem is to determine the minimum for 2.7; it is found by following the gradients of J, or combinations of the successive gradients; by inspection of δJ we find:

$$\nabla J = C^{-1}(\mathbf{x}_0 - \mathbf{x}_g) - \sum \lambda_k \mathcal{L}'_k H'_k \tag{2.12}$$

so, comparing with the solution form 2.9 we find that the gradient is not in the K-dimensional subspace where the solution lies.

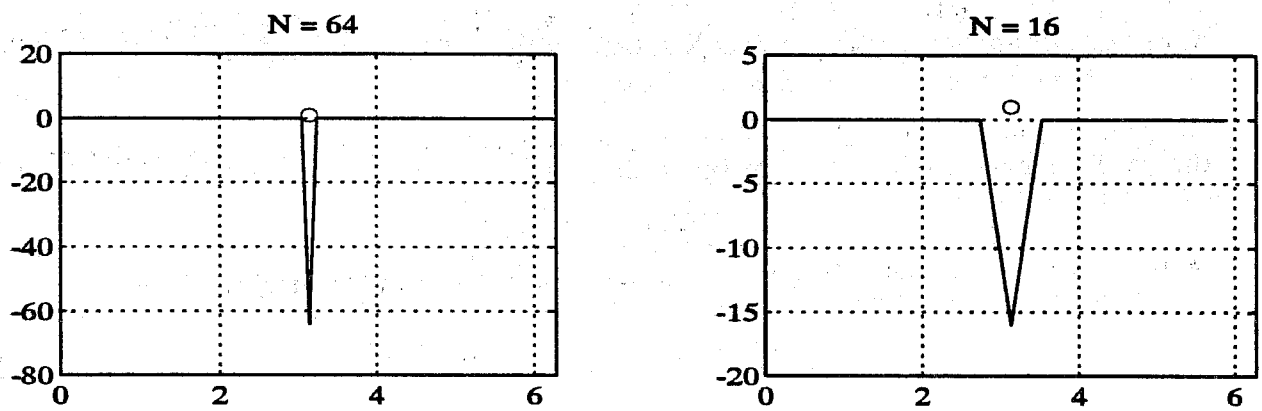

Figure 3: Gradient ∇J with one observation marked by a circle and a Gaussian covariance of guess error. Left: with N=64 modes, Right: with N=16 modes

Let us take the example of the identity model ($\mathcal{L}$=I) and K=1 pointwise observation at location m_1; Figure 3 shows that the first gradient $\nabla J = H'_1$, obtained with $x_0 = x_g$, is the discretization of the Dirac distribution $\delta(m - m_1)$ corresponding to a pointwise modification of the guess; its norm increases with model truncation[3]: the problem is meaningless in its continuous representation.

We conclude that there is no convergence of the gradient 2.12 for pointwise observations when model truncation increases.

For a discretized model, however, the minimization process converges, beginning by the smallest scales; but when the model truncation increases, the minimization becomes harder and harder to perform (Fig 4); with N=64 modes, 25 iterations are what is needed to get half the way towards the solution.

To get around that problem, we have to change the descent directions. The gradient depends upon the scalar product chosen; any descent direction can be made aligned with

[3]Let $\mathbf{x}$ be represented in a monodimensional scalar grid-point model over a periodic domain $m \in 0, 2\pi$ with N points $m_i = 2i\pi/N$ by a column vector $\mathbf{x}_i$; the scalar product is defined by $\langle \mathbf{x}; \mathbf{x}' \rangle = \frac{2\pi}{N} \sum_{n=1,N} x_i x'_i$; if $\mathbf{x}$=1, its norm is 2π independently of truncation N; then an observation in m_i is made through the line vector $H^j = \delta^j_i$; the gradient is the column vector $H'_j = \frac{N}{2\pi}\delta^i_j$; the norm of the gradient is $\frac{N}{2\pi}$ and increases with N

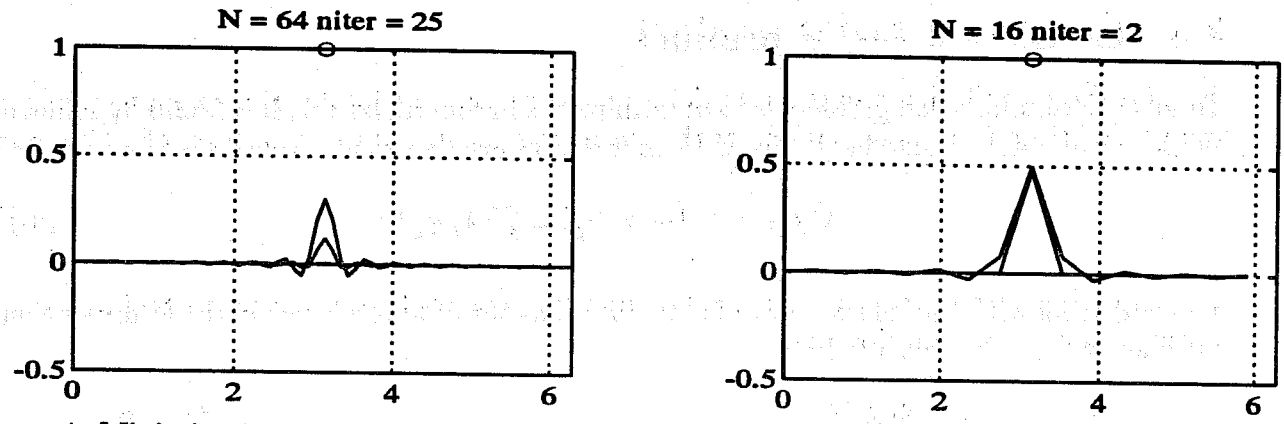

Figure 4: Minimization with gradient ∇J. Left: with N=64 modes, Right: with N=16 modes; + mark the exact solution, circles the observations

the gradient by a proper choice of the scalar product[4]; let us define a new scalar product, related to M, by:

$$\langle \vec{x}; \vec{y} \rangle_M = \left\langle M^{-1}\vec{x}; \vec{y} \right\rangle \tag{2.13}$$

and choose M to be equal to guess error correlations: M=C; the gradient becomes:

$$\nabla_C J = (\mathbf{x}_0 - \mathbf{x}_g) - \sum \lambda_k C \mathcal{L}'_k H'_k \tag{2.14}$$

As is shown in figure 5, the gradient converges[5]as soon as model truncation gives a good spatial representation of the guess correlation error;

If the deviation from the guess $\mathbf{x}_0^i - \mathbf{x}_g$ at iteration i and the previous gradients $(\nabla_C J)^j$, $j < i$ are already in the K-dimensional subspace, then the present gradient $(\nabla_C J)^i$, the next descent direction $\mathbf{d}^i$ and the next deviation from the guess $\mathbf{x}_0^{i+1} - \mathbf{x}_g$ lie also in that space. This shows that, if $\mathbf{x}_0^0 = \mathbf{x}_g$, then all the minimization is made in the K-dimensional subspace.

As the conjugate gradient method is exact in a number of steps equal to the dimension space, the mimimum will be attained in at most K steps when the evolution model is linear; in the presence of 1 observation, 1 descent only along 2.14 will be necessary, as shown in Fig. 6.

---

[4]It is not difficult to show that, if d is a descent direction, then we can choose M to define the scalar product such that the gradient is -d.

Let us construct an orthonormal basis $\{e_n\}$ with e_1 aligned with ∇J: $\nabla J = g_1 e_1$ and e_2 in $\{d, \nabla J\}$: $\mathbf{d} = d_1 e_1 + d_2 e_2$

d is a descent direction if $\langle d; \nabla J \rangle = d_1 g_1 < 0$; the gradient with the scalar product defined by M as in 2.13 is $\nabla_M J = M \nabla J$; any matrix $M = \begin{pmatrix} -\frac{d_1}{g_1} & -\frac{d_2}{g_1} & & .. \\ -\frac{d_2}{g_1} & \alpha & & .. \\ & & 1 & .. \\ .. & .. & .. & .. & .. \\ & & & .. & 1 \end{pmatrix}$ with $\alpha > -\frac{d_2^2}{d_1 g_1}$ is positive definite and satisfies $M\nabla J = -\mathbf{d}$

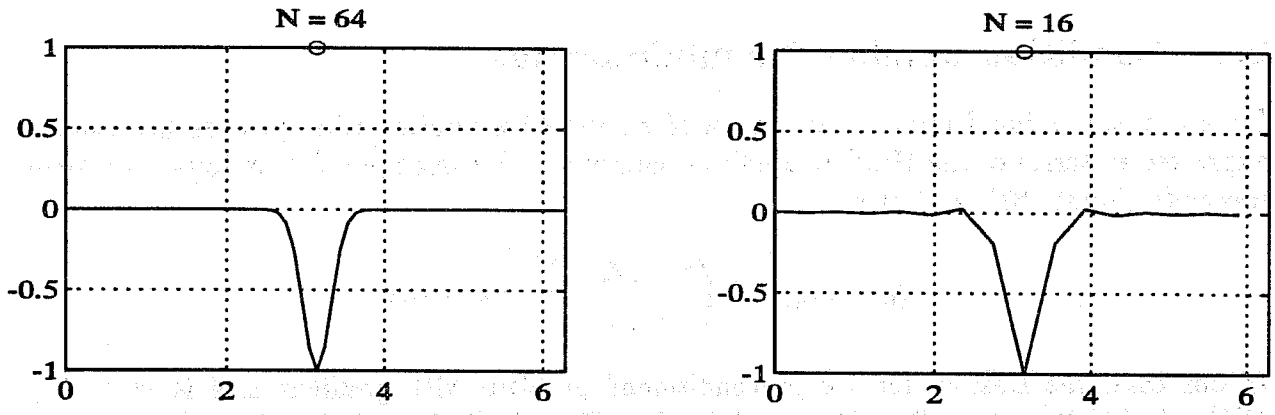

Figure 5: Gradient C∇J with one observation and a Gaussian covariance of guess error. Left: with N=64 modes, Right: with N=16 modes

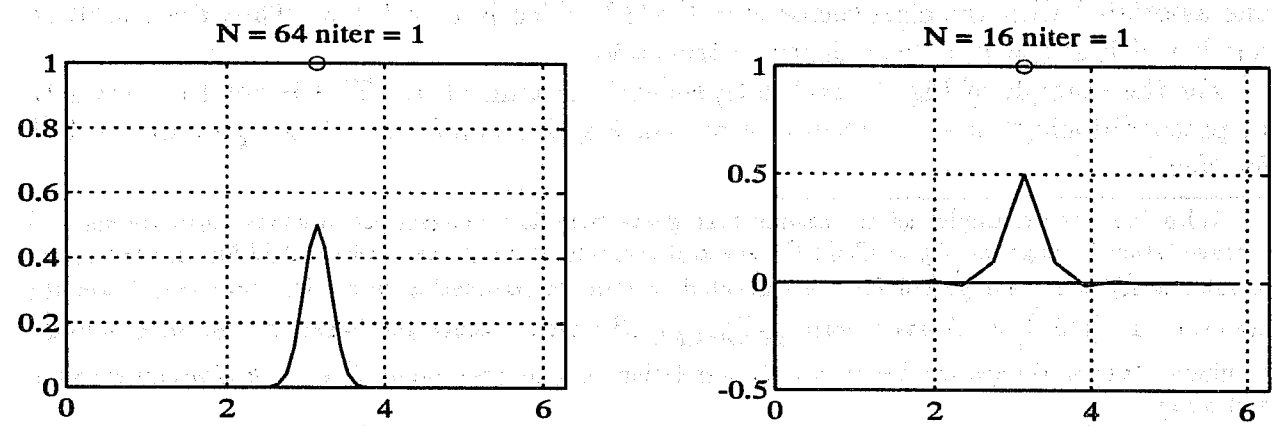

Figure 6: Minimization with gradient C∇J. Left: with N=64 modes, Right: with N=16 modes

We thus find that the scalar product appropriate for the adjoint model assimilation problem should be based upon the correlations of the guess errors at initial time.

Nevertheless, we shall see that for the adjoint equations to reveal their meaning the scalar product used to define $\mathcal{L}^*$ should be based upon the energy of the system.

2.5 Condition number for minimization

Bounds upon residual error $\mathbf{e}_n$ at step n of an iterative minimization process are usually expressed in terms of the Hessian condition number κ. For example, for conjugate gradient methods (Golub, 83), we have:

$$\langle \mathbf{e}_n; \mathcal{H}\mathbf{e}_n \rangle < 4 \left(\frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}} \right)^{2n+2} \langle \mathbf{e}_0; \mathcal{H}\mathbf{e}_0 \rangle$$

In our case, the Hessian for the preconditioned problem with gradient 2.14 is $\mathcal{H} = I + C^{1/2}H'O^{-1}HC^{1/2}$ where H and O are defined as $H = \sum_k H_k \mathcal{L}_k$ and $O = diag(o_k^2)$.

Practically, when C is homogeneous, convolution by C to get the gradient 2.14, or deconvolution to compute the scalar product 2.13 is easy: it amounts to a multiplication (division) in spectral space; for a spectral model, the matrix C is diagonal when it applies to spectral components.

Let us show that κ is simply related to the ratio of the guess error to the observation error and that κ converges as well as model truncation increases.

First, in the case of K=1 observation, $\mathcal{H}$ is identity with a rank one modification; let $\lambda = O^{-1}HCH'$ be the ratio of the discretized variance of guess error at the observation point and time HCH' to the observation error variance O; the eigenvalues of $\mathcal{H}$ are 1, except the one associated with the eigenvector $w = C^{1/2}H'$ which is $\mu = 1 + \lambda$. Thus the condition number of $\mathcal{H}$ is $\kappa = \mu$, independently of truncation.

For the example of Fig. 4, κ=2 independently of truncation. This is not the case with no preconditioning[6]: $\kappa \rightarrow \infty$ when N$\rightarrow \infty$. For Fig. 4, κ=53 for N=16 and goes to $\kappa \approx 10^{18}$ for N=64.

---

[5]With the same example, let us assume that guess error is homogeneous in space; then its spectral representation is diagonal: $\hat{C}_{l'}^{l} = C_l \delta_{l'}^{l}$; the spectral representation of the gradient 2.12 for an observation in m=0 is $\hat{H}_l^{l'} = 1$, as expected for the truncated spectral representation of a Dirac function. Then the gradient 2.14 $\left(\hat{C}\hat{H}'\right)_l = C_l$ has a norm $\frac{1}{2\pi}\sum_{l=0,N/2} C_l^2$ which is convergent when N$\rightarrow \infty$ with ordinary requirements upon the spectral behaviour of a correlation function (as soon as $C_l^2 < \frac{A}{l^\alpha}$, A arbitrary constant and α>1).

[6]A sketch of the proof is as follows: let us take the case of no observations; then $\mathcal{H} = C^{-1}$. As C is a symmetric bounded operator its eigenvalues μ_i can be ordered in a sequence converging to 0: $\frac{\mu_1}{\mu_i} \rightarrow \infty$ *when* $i \rightarrow \infty$; its highest eigenvalues μ_i, $i = 1, N$ are approximated by those μ_i' of the discretized problem, so $\frac{\mu_1'}{\mu_N'} \rightarrow \infty$ *when* $N \rightarrow \infty$

This result can be generalized for the case of K>1 observations; let (v,λ) be one of the K solutions to the generalized eigenvalue problem:

$$HCH'v = \lambda Ov$$

Then $C^{1/2}H'v$ is an eigenvalue of $\mathcal{H}$ for $\mu = 1 + \lambda$.

For the case of Fig. 7, κ=11 and 5.44.

With preconditioning by C we stay in the K-dimensional subspace; in fact, it is the condition number of $\mathcal{H}$ restricted to it which is relevant, as is shown by the case of 1 observation. $\mathcal{H}$ is identity modified by a rank K perturbation, so its eigenvalues are all 1 except for the K ones given above.

Thus the condition number of the restricted Hessian is:

$$\kappa = \frac{1 + \lambda_{max}}{1 + \lambda_{min}}$$

When observations are of equal variance and decorrelated, the optimum is still found in one step.

Practically, an estimation for κ can be easily computed; we have the bound:

$$\lambda < max_i \left(\sum_k \frac{|HCH'|_{ik}}{O_{kk}} \right) \tag{2.15}$$

A special case deserves attention: the monovariate case when L is identity, observations pointwise and C with positive coefficients; it is then easy to compute this bound: we construct the weighted sum of Green functions centered at the observations:

$$\sigma(m) = \sum_k \frac{C(m, m_k)}{O_{kk}} \tag{2.16}$$

λ is bounded by the maximum of this field: $\lambda < \sigma(m_k)\ \forall k$; no easy minoration is found for λ_{min}; so the majoration for κ is then:

$$\kappa < 1 + max_k \sum_{k'} \frac{C(m_k, m_{k'})}{O_{k'k'}}$$

σ(m) is in fact just the gradient with observations equal to -1.

Convergence in one minimization step does not ordinary occur with uneven spacing or observation errors; the gradient is strong where observations are precise or clustered (see Fig. 7); Convergence is known to occur first upon the leading eigenvectors of the Hessian. (see Fig 8 for our test-case), and convergence has to occur first upon quite fine scales in data rich areas before any modification to the guess is made in data poor areas. We would like instead convergence to occur for a given wave-length more uniformly over the model domain.

This shows that the nature of the remaining difficulty is to be found in uneven data density.

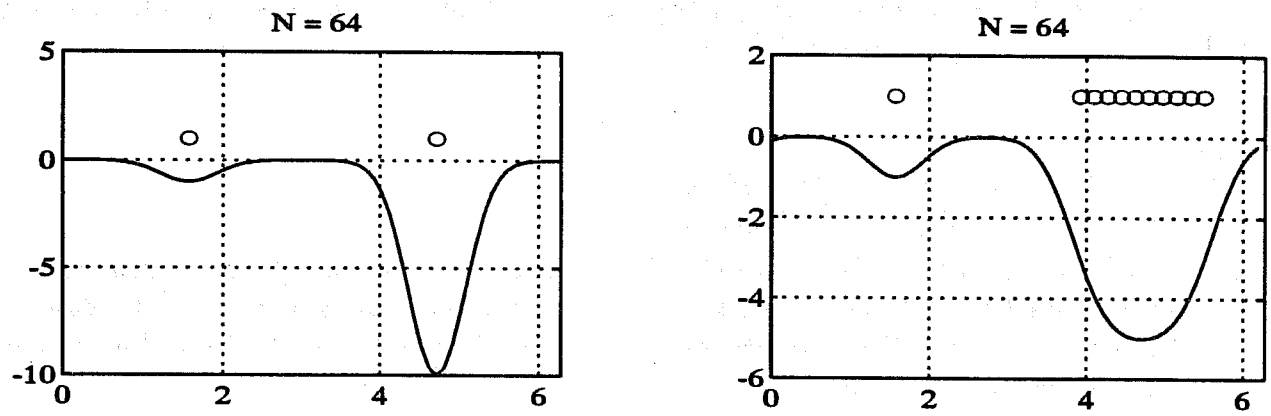

Figure 7: gradient C∇J. Left: with $o_1^{-2} = 1$ *and* $o_2^{-2} = 10$, Right: with uniform observation error

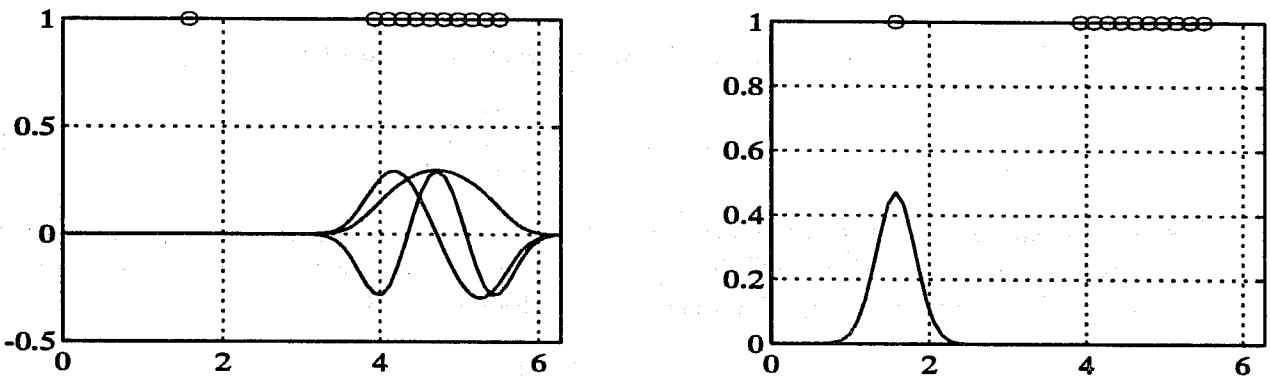

Figure 8: Leading Eigenvectors of the Hessian. Left: 3 leading ones, Right: the fourth one

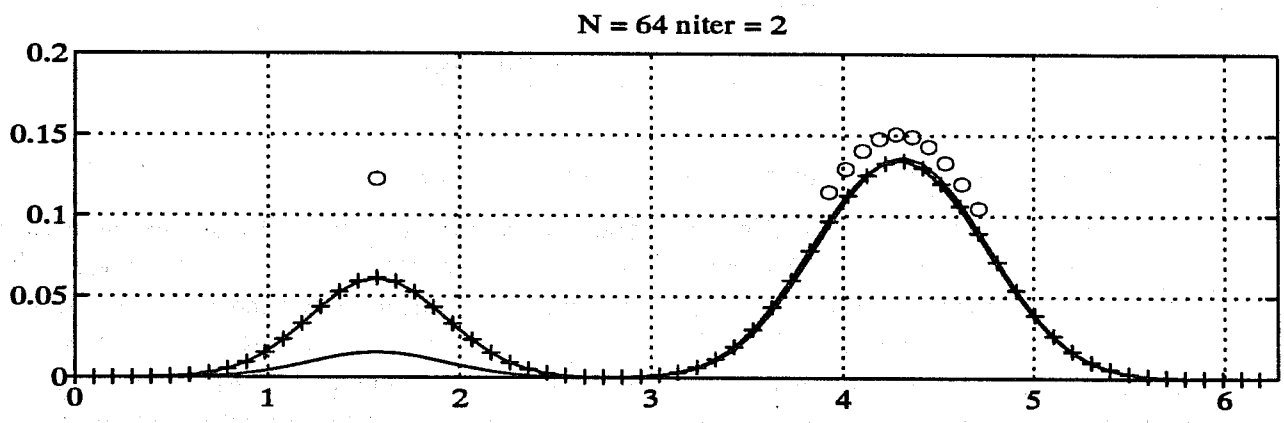

Figure 9: Minimization with gradient C∇J

To get around, we have to abandon the clean framework just developed; practically, to avoid costly operations, preconditioning should be restricted to multiplications by diagonal or other sparse matrices, either in spectral space (this gives convolution in real space), or real space (this is modulation); the preconditioning to get C∇J in 2.14 enters the first category; here we want modulation according to data density; it is known pointwise from 2.16; the gradients for isolated observations have some characteristic half-width α (coming mainly from the decorrelation length of guess errors); in order for the gradients to be not too distorted, the modulation should have scales superior to α. We propose to consider modulation by the inverse of:

$$\sigma(m) = 1 + \sum_i C_b(m, m_i) \times \sum_k \frac{|HCH'|_{ik}}{O_{kk}} \tag{2.17}$$

where C_b is a convolution operator of half-width β. The above formula is not as terrible as it looks; it boils down to 2.16 in the special case where the formula is valid and β=α.

The gradient is thus:

$$\nabla_{\sigma J} J = \sigma^{-1} \times \left(\mathbf{x}_0 - \mathbf{x}_g - \sum \lambda_k C \mathcal{L}'_k H'_k \right) \tag{2.18}$$

We then obtain a better convergence in the first descent steps; convergence is, of course, comparatively deteriorated in further iterations: with the gradient 2.18 the solution cannot any longer be reached in K steps due to the distortion of the Green functions; The optimum for our test problem has been found for the convolution length β=1.5α; for β=∞ we recover the gradient 2.14.

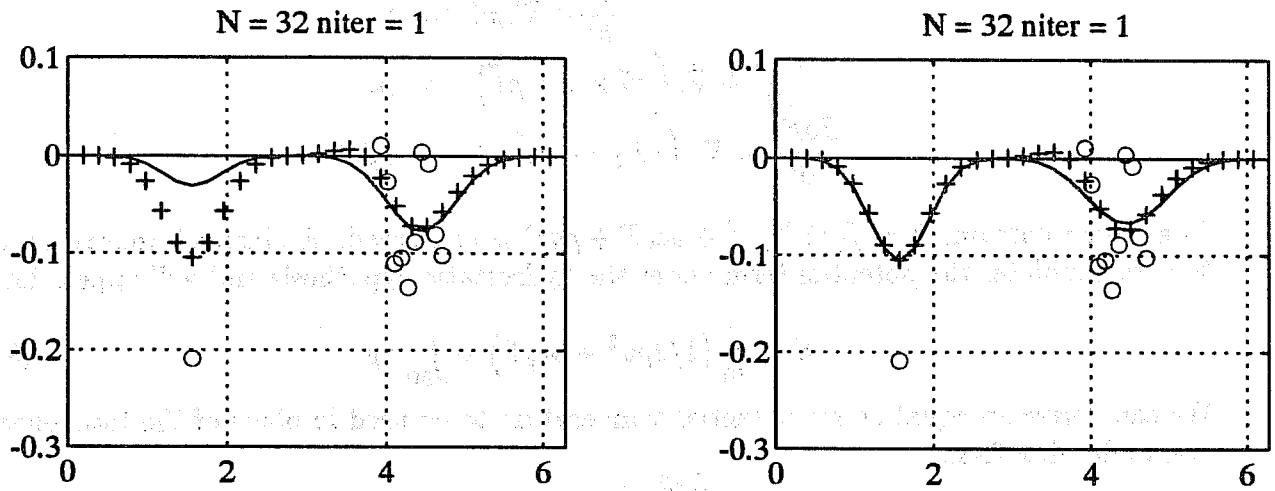

Figure 10: Minimization with gradient C∇J (b=∞) and $\sigma^{-1} \times C\nabla J$ (β=1.5)

$itern^{\circ}$	0	1	2	3
$C\nabla J(\beta=\infty)$	1.71	0.58	0.007	0.00036
$\sigma^{-1}\times C\nabla J(\beta=1.5)$	1.71	0.035	0.0031	0.00084

Table 1: Distance to solution during minimization according to type of preconditioning

3 The non-hydrostatic set of equations

As the most widely used system for variational assimilation is based upon the hydrostatic equations, it might be helpful to recall in this paper the derivation of the anelastic equations with some emphasis upon the energetics.

3.1 The Euler equations

We adopt the tensorial notations by Bonnet, Luneau (89); once acquainted with them it is possible to write advection terms in flux form in a concise way. The advection of the vector $\vec{u}$ by the advecting field $\vec{v}$ is expressed with the help of the divergence of the tensor $\vec{u}\otimes\vec{v}$; its components in the orthogonal cartesian system $\vec{e}_i$ are given by:

$$\vec{\nabla}.\vec{u}\otimes\vec{v}=\frac{\partial u^i v^j}{\partial x^j}\vec{e}_i$$

The Euler equations express the conservation of mass, momentum, and "total energy" $E_T = 1/2\vec{u}.\vec{u}+c_vT$:

$$\begin{aligned}\frac{\partial\rho}{\partial t}+\nabla.\rho\vec{u} &= 0\\ \frac{\partial\rho\vec{u}}{\partial t}+\vec{\nabla}.\left(\rho\vec{u}\otimes\vec{u}+p\vec{\vec{I}}\right) &= \rho\vec{g}\\ \frac{\partial\rho E_T}{\partial t}+\nabla.\left(\rho E_T\vec{u}+p\vec{\vec{I}}.\vec{u}\right) &= \rho\vec{g}.\vec{u}\end{aligned}$$

In a closed domain, $E=\int_\Omega(1/2\rho u^2+\rho c_vT+\rho gz)$ is conserved. A classical approximation to it manipulates the potential term under the hydrostatic hypothesis and will appear later:

$$E=\int_\Omega\left(1/2\rho u^2+\rho c_pT\right)+\int_{\partial\Omega}zp \tag{3.1}$$

We can derive an equation for potential temperature to be used in place of the total energy E_T; under flux form:

$$\frac{\partial\rho\theta}{\partial t}+\nabla.\rho\vec{u}\theta=0$$

3.2 The anelastic system

The compressible system necessitates some tricks in the temporal discretization to cope with the rapidly moving sound waves; we focus here on the essential phenomena and simplify the model state by considering a filtered system.

We define for ρ, θ and p reference values $\bar{\rho}(z)$, $\bar{\theta}(z)$, $\bar{p}(z)$ in hydrostatic balance: $d\bar{p} = \bar{\rho}gdz$; deviations from them are primed quantities: $\rho(x,z) = \bar{\rho}(z) + \rho'(x,z)$, etc, and assume $\rho' << \bar{\rho}$ **(anelastic hypothesis)**.

Perturbation in density ρ' is neglected everywhere, except in the buoyancy term where it is multiplied by g; g is assumed greater than the accelerations. The filtering of sound waves result from the omission of the local derivative in the continuity equation.

From the linearized state equation we have:

$$\frac{p'}{\bar{p}}\left(1+\frac{R}{c_p}\right) = \frac{\rho'}{\bar{\rho}} + \frac{\theta'}{\bar{\theta}}$$

The anelastic system is thus:

$$\begin{aligned} \nabla.\bar{\rho}\vec{u} &= 0 \\ \frac{\partial\bar{\rho}\vec{u}}{\partial t} + \vec{\nabla}.(\bar{\rho}\vec{u}\otimes\vec{u}) &= -\vec{g}\frac{\bar{\rho}\theta'}{\bar{\theta}} - \vec{P} \\ \frac{\partial\bar{\rho}\theta}{\partial t} + \nabla.(\bar{\rho}\theta\vec{u}) &= 0 \end{aligned}$$

where we have gathered in $\vec{P}$ all the terms containing pressure.

3.3 Pressure term and conservation of energy

In this section we want to show by a simple algebraic argument that, as pressure does not appear in the θ equation, then the form of the pressure term is determined by the expression of kinetic energy.

Let us gather in $\vec{S}$ buoyancy and advection terms, so that the momentum equation is:

$$\frac{\partial\bar{\rho}\vec{u}}{\partial t} = \vec{S} - \bar{\rho}\nabla p' \tag{3.2}$$

and define kinetic energy as:

$$E_c = \int_\Omega (1/2\bar{\rho}\vec{u}.\vec{u})$$

Its tendency is:

$$\frac{dE_c}{dt} = -\left\langle\frac{\bar{\rho}\theta'\vec{g}.\vec{u}}{\bar{\theta}}\right\rangle + <\vec{u}.\vec{P}>$$

The first term is conversion from kinetic energy to potential energy; normally, pressure does not appear in the potential energy, so the contribution $\langle\vec{u}.\vec{P}\rangle$ has no other to cancel with; for it to be null we require that:

$$\forall\vec{u}; \nabla.\bar{\rho}\vec{u} = 0, \vec{u}.\vec{n} = 0 \Rightarrow \left\langle\bar{\rho}\vec{u}; \frac{\vec{P}}{\bar{\rho}}\right\rangle = 0$$

It is a standard theorem[7]that with the above condition there exist a scalar p such that:

$$\frac{\vec{P}}{\bar{\rho}} = \nabla(p)$$

The momentum equation then writes:

$$\frac{\partial \bar{\rho}\vec{u}}{\partial t} + \vec{\nabla}.(\bar{\rho}\vec{u} \otimes \vec{u}) = -\vec{g}\frac{\bar{\rho}\theta'}{\bar{\theta}} - \bar{\rho}\nabla\frac{p'}{\bar{\rho}}$$

and, with the definition of Exner pressure $\pi = T/\theta$ and the relation $\partial\bar{\pi}/\partial z = -g/\bar{\theta}$ we verify that the system then conserves a total energy of the form 3.1 without the boundary term:

$$E = \int_\Omega (1/2\vec{u}.\vec{u} + c_p\bar{\pi}\theta)\,\bar{\rho}dxdz$$

This form of the pressure term is identical to the one used by Lipps, Hemler (82) or Lipps (90) for models expressed in Exner pressure; the pressure term used by Durran[8] can

[7]The matrix form of this theorem is more familiar and will be employed later for the discretized model; if the discretized divergence is represented by a matrix A and the gradient by the adjoint matrix $-A^*$, then the above theorem amounts to the fact that Ker(A) and $Im(A^*)$ are supplementary: if $\vec{u} \in Ker(A)$ and $\vec{P} \perp \vec{u}$, then $\vec{P} \in Im(A^*)$

[8]In Durran (89) one uses a pseudo-density:

$$\rho^* = \frac{\bar{\rho}\bar{\theta}}{\theta}$$

the anelastic constraint has the form:

$$\nabla.\bar{\rho}\bar{\theta}\vec{u} = 0$$

In contrast with the energy 3.1, no combination is made of potential and internal energy involving the hydrostatic approximation, but the pseudo-density is used:

$$E = \int \rho^* dxdz \left(1/2u^2 + 1/2w^2 + gz + c_v\frac{\bar{T}}{\bar{\theta}}\theta\right)$$

The pressure term can be found from the expression of kinetic energy and constraint only; let us write the momentum equation as:

$$\frac{\partial \vec{u}}{\partial t} = adv + buoy - \vec{P}$$

the pressure term produces no energy if:

$$\forall \vec{u}: \ \nabla.\bar{\rho}\bar{\theta}\vec{u} = 0 \Rightarrow \int \rho^*\vec{u}.\vec{P} = \int (\bar{\rho}\bar{\theta}\vec{u}).\frac{\vec{P}}{\theta} = 0$$

Then, in the same way, there exists a scalar π such that:

$$\vec{P} = c_p\theta\nabla\pi$$

which is precisely the form appearing in this paper.

be guessed in the same way from his continuity equation and approximation to the kinetic energy.

Besides closing the energetics, this form of the pressure term will simplify the variational problems to be found later.

For further reference, let us introduce the modified anelastic equations of Ogura & Phillips; they do not ensure conservation of energy due to a different pressure term $\vec{P} = \nabla p$

3.4 Elliptic problem for pressure

We have two prognostic equations; pressure here enters as a Lagrange multiplier so that the anelastic constraint is satisfied; taking the divergence of the momentum equation 3.2 we eliminate the temporal derivative thanks to the continuity equation and obtain an elliptic equation for p:

$$\nabla.\bar{\rho}\nabla p' = \nabla.\vec{S}$$

The pressure term keeps the divergence null, but also maintains the boundary conditions for $\vec{u}$. In a closed domain the wind is tangent to the boundary; if $\vec{n}$ is the unit vector orthogonal to the boundary, the condition $\frac{\partial \vec{u}}{\partial t}.\vec{n} = 0$ translates into Neumann boundary conditions for p: $\vec{n}.\nabla p' = \vec{n}.\vec{S}$; for lateral open boundaries the Orlanski radiative condition $\frac{\partial \vec{u}.\vec{n}}{\partial t} + c\frac{\partial \vec{u}.\vec{n}}{\partial x} = 0$ still leads to the specification of $\frac{\partial \vec{u}}{\partial t}.\vec{n}$ and to a Neumann problem for p.

4 The adjoint equations

While it is sometimes advocated that the adjoint refers only to the discretized form of the equations, we think it worth to have a look to the continuous problem.

4.1 The anelastic constraint

Let the linearized evolution operator be $L = \begin{pmatrix} L_u \\ L_\theta \end{pmatrix}$. We will write ∇p instead of $\begin{pmatrix} \nabla p \\ 0 \end{pmatrix}$. The model equations in the Lipps-Hemler form can thus be written as:

$$\begin{aligned} \frac{\partial \bar{\rho}\mathbf{x}}{\partial t} &= L\bar{\rho}\mathbf{x} - \bar{\rho}\nabla p \\ p &= (\nabla.\bar{\rho}\nabla)^{-1}\nabla.L_u\bar{\rho}\mathbf{x}; \bar{\rho}\partial p/\partial\vec{n} = \vec{n}.L_u\bar{\rho}\mathbf{x} \end{aligned}$$

taking the scalar product by $\mathbf{x}^*$ we integrate by parts to get $\mathbf{x}$ isolated on the left hand side; we are left with

$$\left\langle \nabla.\bar{\rho}\vec{u}^*; (\nabla.\bar{\rho}\nabla)^{-1}\nabla.L_u\bar{\rho}\mathbf{x} \right\rangle$$

We thus have to know the adjoint of the inverse operator with inhomogeneous boundary conditions. To avoid this difficulty, two approaches are possible: the one in which we define the model state space, the other in which we come back to the Lagrangian formalism.

4.1.1 Lagrangian method:

We will incorporate in our Lagrangian explicitly the anelastic constraint; here we have provided with observations at only final time.

$$L(\mathbf{x},\mathbf{x}^*,p,\pi)=\int_\Omega \mathbf{x}_0 C^{-1}\mathbf{x}_0+\int_Q\left(\frac{\partial\bar\rho\mathbf{x}}{\partial t}-N\bar\rho\mathbf{x}+\bar\rho\nabla p\right)\vec v-\int_Q\bar\rho\vec u.\nabla\pi+\sigma^{-2}\left(H\mathbf{x}-\tilde{\mathbf{x}}\right)^2$$

which should be stationary for all variations in $\mathbf{x},\mathbf{y}=\begin{pmatrix}\vec v\\ \theta^*\end{pmatrix},p,\pi$, resulting in the Euler equations:

$$\begin{aligned}
\frac{\partial\bar\rho\mathbf{x}}{\partial t}-N\bar\rho\mathbf{x}+\bar\rho\nabla p &= 0 & &\\
\nabla.(\bar\rho\vec u) &= 0 & &\vec u.\vec n=0\\
\frac{\partial\bar\rho\mathbf{y}}{\partial t}+\bar\rho L^*\mathbf{y}+\bar\rho\nabla\pi &= 0 & &\\
\nabla.\bar\rho\vec v &= 0 & &\vec v.\vec n=0
\end{aligned}$$

plus adequate boundary conditions in time.

The adjoint variable evolves with the same constraint as the direct one and the role of the pressure π as a Lagrange multiplier is here evident[9].

4.1.2 Projection method

We have to recognize that not all velocities $\vec u$ of $\mathcal{E}$, space of vector functions of position qualify to be part of a model state. They have to belong to a subspace $\mathcal{F}$ defined by:

$$\begin{aligned}
\nabla.\bar\rho\vec u &= 0\\
\vec u.\vec n|_\Gamma &= 0\ (\textit{rigid boundaries})
\end{aligned}$$

The pressure term corresponds to a projection $\mathcal{P}:\mathcal{E}->\mathcal{F}$ of the model state; $\mathcal{P}$ is a projector (a projector satisfies to $\mathcal{P}^2=\mathcal{P}$): $\vec u$ is left unchanged if it already is in $\mathcal{F}$. The temperature component of the model state is left unchanged by $\mathcal{P}$.

We have to consider the immersion $\mathcal{I}$ from $\mathcal{F}$ into $\mathcal{E}$ and distinguish formally the two spaces even if $\mathcal{F}\subset\mathcal{E}$; If L is the linearized evolution operator for a time-step the complete

[9]Would we have taken the modified anelastic form of the pressure term for the direct ($\vec P=\nabla p$), the resulting adjoint equations would have been:

$$\begin{aligned}
\frac{\partial\bar\rho\mathbf{y}}{\partial t}+\bar\rho L^*\mathbf{y}+\nabla\pi &= 0 & &\\
\nabla.\vec v &= 0 & &\vec v.\vec n=0
\end{aligned}$$

Then the adjoint variable would evolves with a different constraint from the direct one: the speed would be non-divergent, not the momentum density.

evolution operator writes: $\mathcal{P}L\mathcal{I}$; the adjoint will be: $\mathcal{I}^*L^*\mathcal{P}^*$ according to the scheme:

$$\begin{array}{ccccccccc} & \mathcal{F} & & \mathcal{F} & & & \mathcal{F} & & \mathcal{F} & \\ \mathcal{I} & \downarrow & & \uparrow & \mathcal{P} & \mathcal{I}^* & \uparrow & & \downarrow & \mathcal{P}^* \\ & \mathcal{E} & \rightarrow & \mathcal{E} & & & \mathcal{E} & \leftarrow & \mathcal{E} & \\ & & L & & & & & L^* & & \end{array}$$

The adjoint of $\mathcal{I}$ is the orthogonal projection upon $\mathcal{F}$ by:

$$\mathcal{I}(\vec{u}) - \vec{u} = 0 \Rightarrow \forall \vec{u} \in \mathcal{F} : \langle \vec{u}; \vec{v} - \mathcal{I}^*\vec{v} \rangle = 0$$

The Lipps-Hemler[10]form of pressure gives the following projection:

$$\vec{u} \in \mathcal{E} \rightarrow \mathcal{P}(\vec{u}) = \vec{u} - \vec{\nabla}\left(\nabla.\bar{\rho}\vec{\nabla}\right)^{-1}\nabla.\bar{\rho}\vec{u}$$

We verify $\mathcal{P}$ is orthogonal, so $\mathcal{I}^* = \mathcal{P}$, provided enters the scalar product under the form of kinetic energy:

$$\langle \vec{u}; \vec{v} \rangle = \int_\Omega \bar{\rho}\vec{u}.\vec{v}dxdz$$

The "lagrangian" and projection methods, are slightly different for the modified anelastic equations

In the projection method, the adjoint variable $\vec{v}$ evolves in $\mathcal{F}$ according to:

$$\frac{\partial \mathbf{y}}{\partial t} = -\mathcal{I}^*L^*\mathcal{P}^*\mathbf{y}; \nabla.\bar{\rho}\vec{v} = 0$$

while in the Lagrangian method the adjoint variable $\mathbf{z} = \begin{pmatrix} \vec{w} \\ \tau \end{pmatrix}$ evolves in $\mathcal{D}$:

$$\frac{\partial \mathbf{z}}{\partial t} = -\mathcal{P}^*L^*\mathbf{z}; \nabla.\vec{w} = 0$$

The two projections $\mathcal{I}^*$ and $\mathcal{P}^*$ are both perpendicular to $\mathcal{F}$; we have in fact $\mathcal{P}^*\mathcal{I}^* = \mathcal{P}^*$

If the relation $\mathbf{z} = \mathcal{P}^*\mathbf{y}$ is valid at starting time, it will be conserved:

$$\frac{\partial \mathcal{P}^*\mathbf{y}}{\partial t} = -\mathcal{P}^*L^*\mathcal{P}^*\mathbf{y} = -\mathcal{P}^*L^*\mathbf{z}$$

This shows the two methods are equivalent; the second one has one projection less.

[10]For the Shallow-convection form, $\mathcal{P}$ is defined by:

$$\vec{u} \in \mathcal{E} \rightarrow \mathcal{P}(\vec{u}) = \vec{u} - \bar{\rho}^{-1}\vec{\nabla}\Delta^{-1}\nabla.\bar{\rho}\vec{u}$$

and is not orthogonal; different projectors are necessary for the adjoint and for the direct.

4.1.3 Observation term and boundary conditions

The adjoint equations with the observation terms and boundary conditions in the presence of the constraint are:

$$\frac{\partial \mathbf{y}}{\partial t} + \mathcal{I}^* L^* \mathcal{P}^* \mathbf{y} = -o_k^{-2} \mathcal{I}^* H_k^* \left(H_k \mathcal{I} \mathbf{x} - \tilde{\mathbf{x}}_k \right) \delta_{t_k} \tag{4.1}$$

with the appropriate time boundary conditions:

$$\begin{cases} \mathbf{y}_0 & = \mathcal{I}^* C^{-1} \mathcal{I} \mathbf{x}_0 \\ \mathbf{y}_T & = 0 \end{cases} \tag{4.2}$$

to be compared to 2.5 and 2.6. This shows two things: the constraint is used at observation time and is part of the observation operator; the observation of one component of speed will be used to modify both components, without any recourse to the initial time covariances of the departures from the guess.

The equivalent for the hydrostatic models would be to assume the model trajectory stays on the "slow manifold" upon which the initialization process projects and make one initialisation step at each observation time.

4.2 Interpretation of the adjoint equations

Scalar product:

To define an adjoint, we need a quadratic norm; the exact energy is not quadratic; instead, we will use the quadratic aproximation 2.4 valid only for small departures of θ from $\bar{\theta}$.

Let u,w,θ be the basic flow for linearization (different from the reference state $\bar{\rho}, \bar{p}, \bar{\theta}$ necessary in the anelastic approximation), u',w',θ' the perturbations from that basic flow and u^*, w^*, θ^* the adjoint perturbations.

The linearized equations, in the Lipps Hemler form are:

$$\begin{aligned} \frac{\partial \bar{\rho} \vec{u}'}{\partial t} &= -\bar{\rho} \vec{u} . \nabla \vec{u}' - \bar{\rho} \vec{u}' . \nabla \vec{u} - \frac{\vec{g} \bar{\rho} \theta'}{\bar{\theta}} - \bar{\rho} \nabla p \\ \frac{\partial \bar{\rho} \theta'}{\partial t} &= -\bar{\rho} \vec{u} . \nabla \theta' - \nabla . \bar{\rho} \vec{u}' \theta \end{aligned}$$

Upon scalar product with the dual variables, $\vec{u}^*, \theta^*$, reordering the adjoint equations are:

$$\begin{aligned} \frac{\partial \bar{\rho} \vec{u}^*}{\partial t} &= -\nabla . \bar{\rho} \vec{u}^* \otimes \vec{u} + \left(\nabla \otimes \vec{u}^* \right) . \bar{\rho} \vec{u} + \frac{g \bar{\rho} \theta^*}{\bar{\theta}} \frac{\nabla \theta}{\bar{\theta}_z} - \bar{\rho} \nabla p \\ \frac{\partial \bar{\rho} \theta^*}{\partial t} &= -\nabla . \bar{\rho} \vec{u} \theta^* - \bar{\rho} \vec{u}^* . \nabla \bar{\theta} + \bar{\rho} \theta^* \vec{u} . \nabla ln \left(\bar{\theta}^2 \right)_z \end{aligned}$$

The terms in the **adjoint momentum equation** are:

- $\nabla . \bar{\rho} \vec{u}^* \otimes \vec{u}$, advection of the velocity perturbation, identical in the direct
- $\frac{g \bar{\rho} \theta^*}{\bar{\theta}} \frac{\nabla \theta}{\bar{\theta}_z}$, buoyancy term modified from the direct by the factor $\frac{\nabla \theta}{\bar{\theta}_z}$

- $(\nabla \otimes \vec{u}^*).\bar{\rho}\vec{u}$, adjoint of the perturbation of advection in the direct and hardly readable; part of it is absorbed by the pressure term; it helps going to the vorticity formulation; with:

$$\bar{\rho}\vec{u} = k \times \nabla\psi; \quad \omega = \nabla \times \vec{u}$$

retaining only the advection terms in the vorticity equation:

$$\frac{\partial \omega'}{\partial t} = -J\left(\psi, \frac{\omega'}{\bar{\rho}}\right) - J\left(\psi', \frac{\omega}{\bar{\rho}}\right) ....$$

Upon scalar product by ψ^* , neglecting the gradients of $\bar{\rho} : \bar{\rho} = 1$, the adjoint of the above terms is:

$$\frac{\partial \omega^*}{\partial t} = -J(\psi, \omega^*) - ....$$

only the advection of ω^* remains in the adjoint; the clumsy term has vanished.

The terms in the **adjoint thermodynamic equation** are:

- $\nabla.\bar{\rho}\vec{u}\theta^*$, advection of the θ perturbation, identical to the direct
- $\bar{\rho}\theta^*\vec{u}.\nabla ln\left(\bar{\theta}^2\right)_z$, residual of the previous term if $\bar{\theta}^2$ is not linear
- $\bar{\rho}\vec{u}^*.\nabla\bar{\theta}$, perturbation of the advection of $\bar{\theta}$ (of θ in the direct).

We notice that the buoyancy term from the direct is approximately exchanged with the advection of mean temperature from the adjoint.

The adjoint tries to explain the perturbations at final time in a minimal normfashion; for a scalar τ passively advected the linearized equation would be:

$$\frac{\partial \tau'}{\partial t} = -\bar{\rho}u.\nabla\tau' - \bar{\rho}u'.\nabla\tau$$

and a smoothing term for τ would have to appear in the assimilation problem. Two possibilities are given to explain the oberved perturbation: either assume the perturbation τ' was advected, thus the adjoint equation for τ:

$$\frac{\partial \tau^*}{\partial t} = -\bar{\rho}\vec{u}.\nabla\tau^*$$

or modify the basic flow advecting the perturbation, thus the following contribution to the evolution of adjoint velocity:

$$\frac{\partial \vec{u}^*}{\partial t} = -\tau^*\nabla\tau$$

This term, when the passive scalar is potential temperature, is the buoyancy term in the adjoint momentum equation.

So there is not complete identity between the adjoint and the direct; fisrt the advection of momentum has not the same form; second, the exchange between buoyancy and advection of mean temperature is not exact; this has to be linked with the fact that the linearized equations conserve the energy we have defined only when u is uniform and $\bar{\theta}\bar{\theta}_z$ constant.

This was the case in the example of Fig. 1. Let us determine the conditions upon the covariances C so that they are left unchanged by the linear model evolution: $\mathcal{L}C\mathcal{L}' = C$.

As there is no amplification of perturbations for the norm M corresponding to energy: $\forall \mathbf{x} : \langle \mathbf{x}; \mathbf{x} \rangle_M = \langle \mathcal{L}\mathbf{x}; \mathcal{L}\mathbf{x} \rangle_M$, so $\mathcal{L}^*\mathcal{L} = I$;

From the properties of the adjoint under a change of scalar product, the condition we seek is:

$$\mathcal{L}C\mathcal{L}' = \mathcal{L}CM^{-1}\mathcal{L}^*M = C$$

which shows that CM^{-1} has to commute with $\mathcal{L}$; so C must share with $\mathcal{L}$ its eigenspaces.

Due to the symmetries in $\mathcal{L}$ and M, the form for C can be given. $\mathcal{L}$ and M are homogeneous in space so they are block-diagonal in spectral space; let $\vec{k} = (k, m)$ denote the wave-number; each block $\mathcal{L}_{\vec{k}}$ of $\mathcal{L}$ is a matrix whose order is the number of independent variables of the model, here 2 (we can choose vorticity and potential temperature); $\mathcal{L}_{\vec{k}}$ is antisymmetric with pure imaginary eigenvalues $-iku \pm \lambda_k$ and complex eigenvectors, distinct for each wave-number. CM^{-1} has to share the block decomposition, so CM^{-1} is homogeneous in space. Each block of CM^{-1} is symmetric, so its eigenvectors are real and cannot be those of $\mathcal{L}_{\vec{k}}$ unless $C_{\vec{k}}M_{\vec{k}}^{-1} = \alpha_{\vec{k}}I$. So:

$$C_{\vec{k}} = E\left(\begin{pmatrix}\omega_{\vec{k}}\\ \theta_{\vec{k}}\end{pmatrix}\begin{pmatrix}\omega_{\vec{k}}\\ \theta_{\vec{k}}\end{pmatrix}'\right) = \alpha_{\vec{k}}\begin{pmatrix}\vec{k}.\vec{k} & 0\\ 0 & \dfrac{\bar{\theta}\bar{\theta}_z}{g}\end{pmatrix}$$

where $\alpha_{\vec{k}}$ is a different (positive) constant for each wave-number.

In other words, C is homogeneous, there is no covariance between model variables θ and ω, the covariance function for θ can have any admissible shape, but the covariance function for ω is deduced from the one for θ to make equal covariances in the kinetic energy part and potential part. This is not surprising, since the only terms left in our linearized equations reflect the exchange between kinetic and potential energy of the perturbation.

One can then wonder whether there exist a scalar product for which the adjoint equations are identical to the direct ones so that $\mathcal{L}^*\mathcal{L} = I$. The answer is we need the basic state to be stationary and stable; if the basic state is stationary, then there exist a quantity quadratic at first order in u', θ' called pseudo-energy $\mathcal{A}$ (Scinocca & Shepherd, 91). Its quadratic part A is conserved by (and defines) the linearized equations. When the stationary state is stable, A can be used to define the scalar product. It can be shown that the adjoint equations are then identical to the direct ones.

However, conditions for a covariance matrix to be left unchanged under transport by the model linearized around a stable stationary state are still to be found.

5 Discretized direct model

5.1 Generalized coordinates

We need to modify the cartesian system of coordinates in order to ease the uneven boundary conditions produced by orography. Instead of applying a vertical change of coordinates only, a general one can be applied. As is the case for most hydrostatic models, the model state is expressed in terms of the horizontal components of speed. However, the change of coordinates is explicitly apparent in the use of the contravariant speeds for advection.

The drawback is we need some algebra to get to eqn. 5.1

An adequate treatment of the subject is to be found in Dutton (86); we summarize the only part necessary here.

The terrain following coordinate defined in relation with an underlying cartesian coordinate system $\left(\vec{i}_i, i=1,2\right) = \left(\vec{i},\vec{j}\right)$; the new system is denoted by overbars, and we define an indice such that: $(\bar{x}^i, i=1,2) = (\bar{x},\bar{z})$. The vectors tangent to the isolines of $\bar{x}^i$ given by:

$$\left\{\begin{array}{c}\vec{e}_1\\ \vec{e}_2\end{array}\right\} = \left(\begin{array}{cc}\frac{\partial x}{\partial \bar{x}} & \frac{\partial z}{\partial \bar{x}}\\ \frac{\partial x}{\partial \bar{z}} & \frac{\partial z}{\partial \bar{z}}\end{array}\right)\left\{\begin{array}{c}\vec{i}\\ \vec{j}\end{array}\right\} = J_{\bar{x}}^{x}\left\{\begin{array}{c}\vec{i}\\ \vec{j}\end{array}\right\}$$

are called the covariant basis; the vectors $\vec{e}^i = \nabla \bar{x}^i$ are orthogonal to the isolines and are called the contravariant basis. The components of a velocity vector upon the tangent basis transform as the contravariant basis under a change of coordinates and are thus called contravariant components.

A surface integral can be calculated in the new coordinate system:

$$\int f(x,y)dxdy = \int g^{1/2} f(\bar{x},\bar{z})d\bar{x}d\bar{z}$$

$g^{1/2}$ is expressed in terms of the metric tensor:

$$g^{1/2} = det\left(J_{\bar{x}}^{x}\right)$$

Inverting $J_{\bar{x}}^{x}$ we get reverse relations:

$$\left\{\begin{array}{c}\vec{i}\\ \vec{j}\end{array}\right\} = C\left\{\begin{array}{c}\vec{e}_1\\ \vec{e}_2\end{array}\right\}; \quad C = g^{-1/2}\left(\begin{array}{cc}\frac{\partial z}{\partial \bar{z}} & -\frac{\partial x}{\partial \bar{z}}\\ -\frac{\partial z}{\partial \bar{x}} & \frac{\partial x}{\partial \bar{x}}\end{array}\right)$$

and C is used to get the contravariant components from the cartesian ones:

$$\vec{u} = \bar{u}^i\vec{e}_i = u^i\vec{i}_i => \bar{u}^i = Cu^i$$

An infinitesimal vector known by its contravariant components will have the length:

$$\vec{u}.\vec{u} = \bar{u}^i g_{ij}\bar{u}^j$$

with the metric tensor g defined by:

$$g_{ij} = (J_{\bar{x}}^{x})^t J_{\bar{x}}^{x} = \frac{\partial x^k}{\partial \bar{x}^i}\frac{\partial x^k}{\partial \bar{x}^j}$$

The gradient of a scalar has simple covariant components:

$$\vec{e}_i.\vec{\nabla}p = p_{,i} = \frac{\partial p}{\partial \bar{x}^i}$$

On the contrary, the derivatives of vector components are not usually transforming like tensors; we have to resort to "covariant derivatives" where "Christoffel symbols" intervene; they are defined as:

$$\bar{u}^i_{,j} = \frac{\partial \bar{u}^i}{\partial \bar{x}^j} + \Gamma^i_{jl}\bar{u}^l$$

Some of the Christoffel symbols are null, so that the divergence, in terms of the contravariant or the cartesian components is simply given by:

$$\nabla.\vec{u} = g^{-1/2}\frac{\partial g^{1/2}\bar{u}^j}{\partial \bar{x}^j} = g^{-1/2}\frac{\partial C^{ij}g^{1/2}u_j}{\partial \bar{x}^i}$$

The cartesian components of the gradient can be found upon multiplying the above by p and integration by parts; from:

$$\int g^{1/2}p\nabla.\vec{u}d\bar{x}d\bar{z} = -\int g^{1/2}\vec{u}.\nabla p d\bar{x}d\bar{z} = -\int g^{1/2}u_j C^{ij}\frac{\partial p}{\partial \bar{x}^i}d\bar{x}d\bar{z}$$

we deduce, taking $\vec{u} = \vec{i}_j$:

$$\vec{i}_j.\nabla p = -C^{ij}\frac{\partial p}{\partial \bar{x}^i}$$

They can also be found by applying the divergence formula:

$$\vec{i}_j.\nabla p = \nabla.\vec{i}_j p = g^{-1/2}\frac{\partial C^{ij}g^{1/2}p}{\partial \bar{x}^i}$$

The fact that the two expressions are sought to be equivalent in their discretized form is what is coined "metric identities".

The cartesian components of moment are scalars; they result of the scalar product $\bar{\rho}u_j = \vec{i}_j.\bar{\rho}\vec{u}$; considering that $\vec{i}$ is constant:

$$\vec{i}.\left(\vec{\nabla}.(\bar{\rho}\vec{u}\otimes\vec{v})\right) = \nabla.\left(\left(\bar{\rho}\vec{u}.\vec{i}\right)\vec{v}\right)$$

This equality is used to express the advection terms in the model; no Christoffel symbol will thus be necessary.

5.2 Component equations

Let U be $g^{1/2}$ times the contravariant components of the moment; it is obtained from the cartesian components of speed through:

$$U^i = g^{1/2}\bar{\rho}C^{ij}u_j$$

The continuity equation writes:

$$g^{1/2}\nabla.\bar{\rho}\vec{u} = \frac{\partial U^i}{\partial \bar{x}^i} = 0$$

$g^{1/2}$ times the advection of horizontal momentum writes:

$$g^{1/2}\nabla.(\vec{u}.\vec{i})\bar{\rho}\vec{u} = \frac{\partial uU^i}{\partial \bar{x}^i}$$

We are now able to write the component form of the equations:

$$\begin{aligned}
\frac{\partial g^{1/2}\bar{\rho}u}{\partial t} &= -\frac{\partial uU^i}{\partial \bar{x}^i} - \vec{i}.\vec{P} \\
\frac{\partial g^{1/2}\bar{\rho}w}{\partial t} &= -\frac{\partial wU^i}{\partial \bar{x}^i} + g\frac{g^{1/2}\bar{\rho}\theta'}{\bar{\theta}} - \vec{k}.\vec{P} \\
\frac{\partial g^{1/2}\bar{\rho}\theta}{\partial t} &= -\frac{\partial \theta U^i}{\partial \bar{x}^i} \\
0 &= \frac{\partial U^i}{\partial \bar{x}^i}
\end{aligned}$$

These equations appear as the advection of 3 scalars coupled through the buoyancy and pressure terms. Compared with hydrostatic models in general vertical coordinate they are more simple.

5.3 Discretization of the equations

The precise discretization might be important for energy conservation and stability. For Galerkine methods, once the finite elements chosen, the discretization follows and energy conservation is automatic. Here, with a grid-point method, conservations are easier to obtain when one is aware of the notion of adjoint operators; particularly, this notion is not evident for the extrapolation operators at the boundaries.

The Arakawa C-grid is used; cartesian and contravariant components of speed are defined at the same points:

$$\begin{array}{ccccccc}
 & w & & w & & z = & \Delta z \\
 & p & u & p & & z = & 1/2\Delta z \\
\cdots & w & \cdots & w & \cdots & z = & 0 \\
 & p & u & p & & z = & -1/2\Delta z
\end{array}$$

that is w is defined at the domain boundaries, and pressure is supposed to be defined one rank outside boundaries so that the normal gradient is defined at the w points on the boundary. w is defined at K levels; u,θ at K-1 levels; pressure is used at K+1 levels.

I found no easy way to infer the discretization of the advection terms from the divergence equation. The advection terms are discretized as follows, with less averaging operators than would be deduced from a strict application of the discretized divergence:

$$\begin{aligned}
\frac{\partial g^{1/2}\bar{\rho}u}{\partial t} &= -\delta_x\left(m_x U.m_x u\right) - \delta_z\left(m_x W.m_x u\right) - \vec{i}.\left(\vec{B}+\vec{P}\right) \\
\frac{\partial g^{1/2}\bar{\rho}w}{\partial t} &= -\delta_x\left(m_z e_U U.m_x w\right) - \delta_z\left(m_z e_W W.m_z e_w w\right) - \vec{k}.\left(\vec{B}+\vec{P}\right) \\
\frac{\partial g^{1/2}\bar{\rho}\theta}{\partial t} &= -\delta_x\left(U.m_x\theta\right) - \delta_z\left(W.m_z\theta\right)
\end{aligned}$$

with the classical Schumann operators:

$$\delta_x U = \frac{U(\bar{x}+1/2\Delta\bar{x},\bar{z}) - U(\bar{x}-1/2\Delta\bar{x},\bar{z})}{\Delta\bar{x}}, \quad m_x U = \frac{U(\bar{x}+1/2\Delta\bar{x},\bar{z}) + U(\bar{x}-1/2\Delta\bar{x},\bar{z})}{2}$$

$\vec{B}$ is the buoyancy term, $\vec{P}$ is the pressure term.

In order to calculate the vertical advection at the boundary, extrapolations e_U, e_W, e_w, e_θ of w, θ, U, W are necessary.

We have to specify the discretized relation between cartesian and generalized components; we will choose:

$$\begin{pmatrix} U \\ W \end{pmatrix} = C \begin{pmatrix} g^{1/2}\bar{\rho}u \\ g^{1/2}\bar{\rho}w \end{pmatrix}, \quad C = \begin{pmatrix} g^{-1/2}\dfrac{\partial z}{\partial \bar{z}} & -g^{-1/2}\dfrac{\partial x}{\partial \bar{z}}m_{xz} \\ -g^{-1/2}\dfrac{\partial z}{\partial \bar{x}}m_{xz}e_u & g^{-1/2}\dfrac{\partial x}{\partial \bar{x}} \end{pmatrix}$$

Another extrapolation, for u, is present at the boundaries.

So we have to verify conservation of kinetic energy by the advection terms; we have also to discover the precise form of the buoyancy and pressure term.

5.4 Conservation of energy

5.4.1 Discretized energy

We wish our equations to conserve energy; at least, advection terms should conserve kinetic energy; several obvious choices are possible, according to the weight a we give to the w component on the boundary; defining the two scalar products as:

$$\langle u;v\rangle_u = \sum_{k=1/2,K-1/2}^{i} u_{ik}v_{ik} \quad \langle w;\omega\rangle_w = \sum_{k=0,K}^{i} a_k w_{ik}\omega_{ik}$$

so that:

$$E_c = 1/2\left\langle g^{1/2}\bar{\rho}u;u\right\rangle_u + 1/2\left\langle g^{1/2}\bar{\rho}w;w\right\rangle_w$$

and similarly:

$$E_p = 1/2 \left\langle g^{1/2} \bar{\rho} c_p \bar{\pi}; \theta \right\rangle_\theta$$

The choice $a_k = 1$ everywhere gives more symmetric formulations. The choice $a_0 = a_K = 1/2, a_k = 1$ otherwise mimicks the definition one would obtain from finite elements.

5.4.2 Conservation by advection terms

If the advection by $\vec{v}$ was antisymmetric in the advected quantity $\vec{u}$ (advection under the form $\bar{\rho}\omega \times \vec{u}$, with $\omega = \nabla \times \vec{v}$), then conservation of energy would not resort to the non-divergence of $\vec{v}$. Here, the advection term is under flux form, so conservation occurs only when the advecting field is non-divergent. The discretization of divergence and advection thus are interdependant; we use in the continuous equations the identity:

$$\vec{u}.\left(\vec{\nabla}.\vec{u} \otimes \bar{\rho}\vec{v}\right) = 1/2\vec{\nabla}.\vec{u}^2\bar{\rho}\vec{v} + 1/2\vec{u}^2\nabla.\bar{\rho}\vec{v}$$

This identity is not true in the discretized case due to the presence of averaging operators; one resorts to integrations by parts and to identities such as:

$$\langle \delta_x u; v \rangle = -\langle u; \delta_x v \rangle ; \quad \langle m_x u; v \rangle = \langle u; m_x v \rangle \quad \langle m_x u; \delta_x u \rangle = \left\langle 1/2 \delta_x u^2 \right\rangle$$

to show that:

$$\frac{dE_c}{dt} = 1/2 \left\langle m_x u^2; \delta_x U + \delta_z W \right\rangle_u + 1/2 \left\langle m_z w^2; \delta_x U + \delta_z W \right\rangle_w - \left\langle \vec{u}; \frac{\vec{g}\bar{\rho}\theta'}{\bar{\theta}} + \vec{P} \right\rangle$$

thanks to a divergence-free extrapolation of U and W, and a copy of w from k=0 to k=-1 at the bottom, and a similar extrapolation at the top.

This confirms that the discretization of the continuity equation necessary for conservation is:

$$\delta_x U + \delta_z W = 0$$

5.4.3 Conservation by the pressure term

We wish also the pressure term not to create kinetic energy, as in the continuous case:

$$\forall u, w : (\delta_x; \delta_z)\, C \begin{pmatrix} g^{1/2}\bar{\rho}u \\ g^{1/2}\bar{\rho}w \end{pmatrix} = 0 \; then \; \left\langle \vec{u}; \vec{P} \right\rangle = 0$$

$\vec{P}$ is orthogonal to the kernel of the above discretized divergence operator; $\vec{P}$ should be in the image of the adjoint operator of the discretized divergence: there exist a scalar field $p^* = p/\bar{\rho}$ such that:

$$\vec{P} = g^{1/2}\bar{\rho}C^* \begin{pmatrix} \delta_x p^* \\ \delta_z^* p^* \end{pmatrix} \tag{5.1}$$

where δ_z^* represents the opposite of the adjoint of δ_z. No problem occurs in the x-direction due to periodicity. These adjoints will be discussed later.

The discretized Poisson equation giving pressure is:

$$(\delta_x;\delta_z)\,C g^{1/2}\bar{\rho}C^*\begin{pmatrix}\delta_x p^*\\ \delta_z^* p^*\end{pmatrix} = (\delta_x;\delta_z)\,C\vec{S}$$

where $\vec{S}$, again, represents the tendancy terms other than pressure in the moment equation.

So the discretization of advection determines the discretization of the continuity equation, which in turn determines the discretization of the gradient for the pressure term. This results in a symmetric discretized Laplacian, especially once one has recognized that the gradient uses the adjoint of the extrapolation e_u present in C.

So far, nothing requires the type of extrapolation for u; we might believe that the choice should be guided by sole accuracy.

In fact, we know that the adjoint of a convergent direct operator is not necessarily convergent; especially extrapolations at the boundary create specific problems.

The adjoint of the divergence leads to a converging pressure term only with a specific choice of the scalar product and extrapolation for u. If the extrapolation for u is derived from $\frac{\partial u}{\partial \vec{e}_2} = 0$ and the weights $a_0 = a_K = 1/2, a_k = 1$ for the w scalar product are retained, the pressure term is convergent; if a higher order accuracy is sought by linearly extrapolating u, then the pressure term we derive is not convergent; the only choice is then to give up with energy conservation and symmetry.

So far, this model has proved to compare well with analytical solutions or Long's model in extreme situations and is not really sensitive to the type of extrapolations.

5.4.4 Buoyancy term

We can recognize that the form of the buoyancy term has to be derived from the discretization of the thermodynamic equation. We should have:

$$\left\langle c_p\pi;(\delta_x\ \delta_z)\begin{pmatrix}m_x\theta\\ m_z\theta\end{pmatrix} C g^{1/2}\bar{\rho}\vec{u}\right\rangle + \left\langle \vec{u};\vec{B}\right\rangle = 0$$

By adjonction of all the operators:

$$\vec{B} = g^{1/2}\bar{\rho}C^*\begin{pmatrix}m_x\theta\delta_x c_p\pi\\ m_z^*\theta\delta_z^* c_p\pi\end{pmatrix} \tag{5.2}$$

We see that there is no reason for the buoyancy term to be exactly vertical, except for a judicious choice of the reference state. An extrapolation for π and θ is required for the tendancy at the boundaries, but is determined by the adjoint operators m_z^* and δ_z^*.

We thus have seen that in an anelastic model, despite the choice of a staggered grid and general coordinates, energy conservation by the advection terms can still be enforced; no degrees of freedom are allowed for the discretization of the pressure or buoyancy terms.

5.5 The pressure solver:

In accord with the above consideration, we adopt adjoint discretizations for the gradient and the divergence; the benefit is our discretized Laplacian is then symmetric and can be solved, after proper preconditioning, by simple and efficient iterative methods.

5.5.1 Boundary conditions

We can modify the divergence operator so that boundary conditions are incorporated. For simplicity we show how in the one-dimension case: all fields will depend on z only. Then the matrix of the discretized gradient is rectangular K×(K+1):

$$\left.\frac{\partial p}{\partial z}\right|_{k=0;K} = \frac{p_{k+1/2} - p_{k-1/2}}{\Delta z}$$

The adjoint of the above is a rectangular matrix (K+1)×K; two boundary terms appear:

$$\begin{aligned} \nabla.\vec{u}|_{K+1/2} &= \frac{-w_K}{\Delta z} \\ \nabla.\vec{u}|_{k=1/2,K-1/2} &= \frac{\delta_z w}{\Delta z} \\ \nabla.\vec{u}|_{-1/2} &= \frac{w_0}{\Delta z} \end{aligned}$$

The modified divergence is null if the divergence is null in the inner domain, and if the boundary condition w=0 is satisfied on the boundary. In two or more dimensions we recognize that we discretize in fact the operator $\vec{v} -> \nabla.\vec{v} - \vec{v}.\vec{n}\delta_\Sigma$ where δ_Σ is Dirac masses distributed on the boundary Σ.

5.5.2 Iterative solution

To solve:

$$\nabla.(\bar{\rho}\vec{u} - \bar{\rho}\nabla p) = 0, \bar{\rho}\frac{\partial p}{\partial \vec{n}}(k = 1, K) = \bar{\rho}\vec{u}.\vec{n}$$

we minimize the quadratic functional:

$$J(p) = \int g^{1/2}\nabla p\,(1/2\bar{\rho}\nabla p - \bar{\rho}\vec{u})\,d\bar{x}d\bar{z}$$

Its gradient for the Euclidean scalar product is:

$$\nabla J(p) = -\nabla.\bar{\rho}\nabla p + \nabla.\bar{\rho}\vec{u} + \delta_\Sigma\,(\bar{\rho}\nabla p - \bar{\rho}\vec{u})\,.\vec{n}$$

The gradient is null when both the divergence is null and the boundary conditions are satisfied.

Such an iterative method has an efficiency decreasing quadratically with the number of points in each direction on the grid, or with the aspect ratio $\Delta x/\Delta z$; we precondition with the symmetric operator A:

$$Ap = \frac{\partial^2 p}{\partial \bar{x}^2} + \frac{\partial^2 p}{\partial \bar{z}^2}$$

This operator is discretized in the same manner as the Laplacian, with the change C=I; vertical and horizontal directions are decoupled; after horizontal Fourier transform we are left with a tridiagonal matrix on the vertical to invert for each horizontal mode. The operator A is negative definite, fast to invert and so qualifies as a preconditioning for our problem.

The preconditioned gradient is discretized as follows:

$$\nabla J = \left(\delta_x^2 + \delta_z^2\right)^{-1} (\delta_x; \delta_z)\, C \left(g^{1/2} \bar{\rho} C^* \begin{pmatrix} \delta_x p^* \\ \delta_z p^* \end{pmatrix} - g^{1/2} \bar{\rho} \begin{pmatrix} u \\ w \end{pmatrix} \right)$$

The preconditioned operator $A^{-1}\Delta$ has a condition number independent of the truncation and aspect ratio; it seems for one of the scalar products to depend linearly upon the dynamics of $g^{1/2}$, that is, with a Gal-Chen grid, of $(1 - h/H)^{-1}$ where h is the maximum mountain height and H the maximum domain height. It is found that the condition number depends strongly upon the type of extrapolation chosen for speed u; it is always low for realistic mountain heights and shapes. In Fig. 11 we show a projected velocity field for an extreme geometry.

5.6 Discretization of the linear and adjoint models

The complete expression of a discretized adjoint is normally cumbersome; in our case, in spite of the grid-point formulation, things are comparatively simple; the linearized direct model is composed of five separate steps:

1. conversion from cartesian to contravariant components; there the metric terms are entirely hidden
2. advection of three scalar perturbations
3. perturbation advection

   step 2 and 3 for the linear model are obtained from the non-linear model: we cast the advection term as a bilinear operator

$$\frac{\partial \mathbf{x}}{\partial t} = -\mathcal{B}\left(e_U U; e_{\mathbf{x}} \mathbf{x}\right)$$

   acting upon the extrapolated fields; in linearized form:

$$\frac{\partial \mathbf{x}'}{\partial t} = -\mathcal{B}\left(e_U U; e_{\mathbf{x}} \mathbf{x}'\right) - \mathcal{B}\left(e_U C \vec{u}'; e_{\mathbf{x}} \mathbf{x}\right)$$

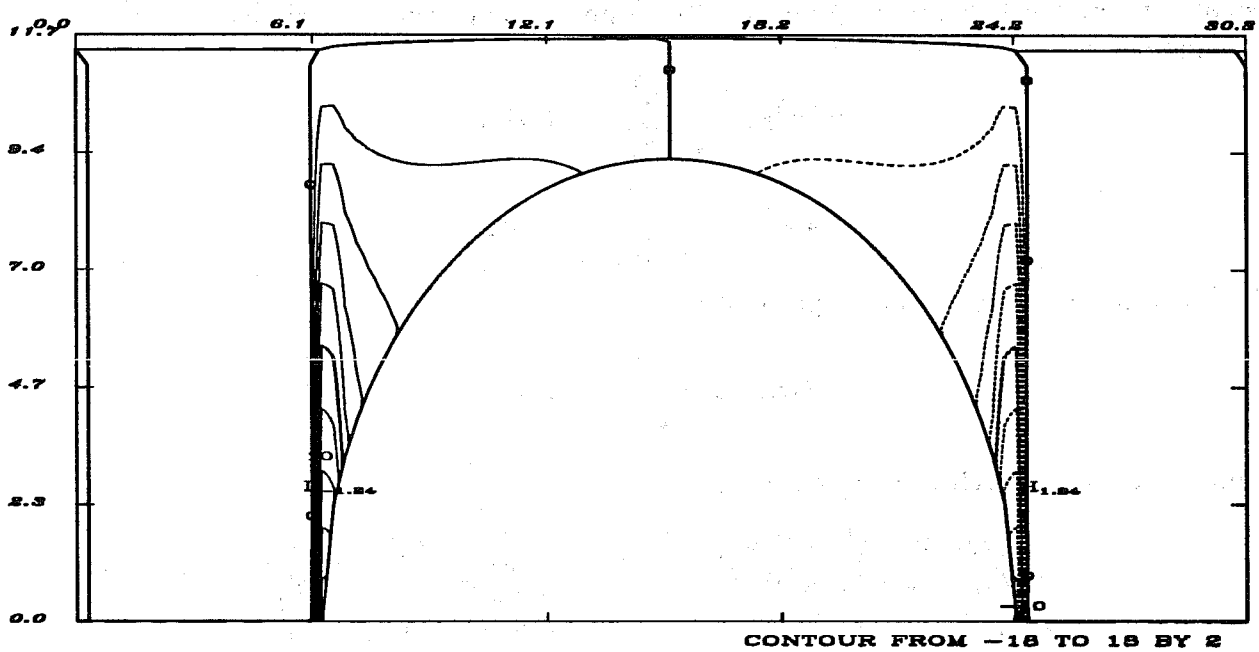

Figure 11: Projection of a constant u=18m/s velocity field; shown is the resulting field of w. The domain is periodic with $\Delta x = \Delta z$; the orography is half a circle occupying 3/4 of the height.

$\mathcal{B}$ has right and left adjoints, so the adjoint advection has two terms:

$$\frac{\partial \mathbf{x}^*}{\partial t} = -e_{\mathbf{x}}^* \mathcal{B}_d^* (e_U U; \mathbf{x}^*) - C^* e_U^* \mathcal{B}_g^* (\vec{u}^*; e_{\mathbf{x}} \mathbf{x})$$

4. buoyancy calculation

5. projection by the pressure term to satisfy the continuity equation

so the linearized model require no extra development; the adjoint model is made by combination of the adjoint of these five steps; however:

- the adjoint of step 1 was necessary for the construction of the direct model
- perturbation advection of potential temperature has for adjoint the buoyancy operator
- step 2 could have been made self-adjoint by an adequate choice of discretization
- step 5 is self-adjoint

The adjoint necessitates only the definition of the bilinear energy scalar product; however, due to the various extrapolations in the direct model, we are led to define some more scalar products to disentangle the adjoints of the Schuman operators defined in the direct model.

Due to the removal of lateral boundary conditions by horizontal periodicity, only extrapolations in the vertical are present. One defines scalar products labeled p and we on extended spaces:

$$\langle p;\pi\rangle_p = \sum_{k=-1/2,K+1/2}^{i} p_{ik}\pi_{ik}$$

$$\langle w;\omega\rangle_{we} = \sum_{k=-1,K+1}^{i} w_{ik}\omega_{ik}$$

We choose for example the scalar product for w as:

$$\langle w;\omega\rangle_w = \sum_{k=0,K}^{i} w_{ik}\omega_{ik}$$

Adjunction rules for Schuman operators are:

$$\begin{array}{ll} \langle u;\delta_z w\rangle_u = -\langle \delta_z z(u);w\rangle_w & \langle u;,m_z w\rangle_u = \langle m_z z(u);w\rangle_w \\ \langle w;\delta_z p\rangle_w = -\langle \delta_z z(w);p\rangle_p & \langle w;m_z p\rangle_w = \langle m_z z(w);p\rangle_p \\ \langle p,\delta_z w\rangle_p = -\langle \delta_z z(p);w\rangle_{we} & \langle p,m_z w\rangle_p = \langle m_z z(p);w\rangle_{we} \end{array}$$

z(.) is extension by 0, r(.) extension by copy; for example:

$$z(w)_k = w_k, k = 0, K$$

$$z(w)_{-1} = z(w)_{K+1} = 0$$

We thus found:

$$m_z^*. = m_z z(.), \qquad \delta_z^*. = \delta_z z(.)$$

Direct operators shrink the vertical dimension space, adjoints expand it:

$$m_z : \mathcal{R}^{K+e+1} \to \mathcal{R}^{K+e}, \qquad m_z^* : \mathcal{R}^{K+e} \to \mathcal{R}^{K+e+1}$$

with e=-1,0,1,2 for u,w,p,we

In a similar manner, all extrapolation operators have an adjoint which shrinks the dimension space.

It has been found that extension operators encountered are amenable to some factorizations like:

$$w + e_w^*\omega = e_w^*(z(w) + \omega)$$

With the above notations, extrapolations are decoupled and can easily be adapted to a change of scalar product for w or a change in boundary conditions.

Writing the discretized adjoint is straightforward, albeit still painful. Let us detail the two advection terms:

Perturbation advection:

$$\frac{\partial u^*}{\partial t} = -e_u^*\left(z.m_x\left(m_x U.\delta_x u^*\right) + m_z^*\left(m_x W.\delta_z^* u^*\right)\right)$$

$$\frac{\partial w^*}{\partial t} = -e_w^* \left(z.m_x \left(m_z e_U U.\delta_x w^* \right) + m_z^* \left(m_z e_W W.\delta_z^* w^* \right) \right)$$

Advection by perturbation:

$$\frac{\partial \vec{u}^*}{\partial t} = -C^* \begin{pmatrix} m_x \left(m_x u.\delta_x u^* \right) + e_U^* m_z^* \left(m_x w.\delta_x w^* \right) \\ m_x \left(m_z u.\delta_z^* u^* \right) + e_W^* m_z^* \left(m_z e_w w.\delta_z^* w^* \right) \end{pmatrix}$$

The adjoint of C is:

$$\vec{u}^* = C^* \vec{U}^* = \begin{pmatrix} \dfrac{\partial z}{\partial \bar{z}} & -e_u^* m_{xz}^* \dfrac{\partial x}{\partial \bar{z}} \\ -m_{xz}^{\pm} \dfrac{\partial z}{\partial \bar{x}} & \dfrac{\partial x}{\partial \bar{x}} \end{pmatrix} \vec{U}^*$$

The advection of temperature and buoyancy terms receive a similar treatment.

Bibliographie:

Bonnet, Luneau(83):Les théories de la Dynamique des Fluides, Cepadues

Bernard-Boussières (91): Représentation de la topographie dans un modèle bidimensionnel non-hydrostatique de convection, Rapport de stage, CNRM

Estrade, Kadri (91): Préconditionnement du Laplcien pour un modèle anélastique plan vertical CNRM/ENM, Juillet 92

S.Farges (91): Adjoint d'un modèle anélastique plan vertical, CNRM/ENM, 92.

G.Gollub, G.Meurant (83): Résolution numérique des grands systèmes linéaires, Eyrolles

Lipps, Hemler, 82: A scale analysis of deep moist convection and some related numerical calculations JAS39, 2192-2210

Scinocca,J.F., Shepherd,T.G., Non-linear wave-activity conservation laws and Hamiltonian structure for the two-dimensional anelastic equations, **JAS49**, 5-27

F.Rabier, P.Bernardet (91): Variational analysis of orographic waves, Beitr.Phys.Atmosph, 64, n°3, pp207-217

Wilhemson, Ogura (1972) The pressure perturbation and the numerical study of a cloud, **JAS29**, 1295-1307