Chapter 8

Atmospheric Data Assimilation: Conventional Methods

8.1 History

In this section we give a short review of the history of conventional atmospheric data assimilation methods. More detailed reviews can be found in Daley [39], in Ghil & Malanotte–Rizzoli [64], and in Harms et al. [73]. We call conventional methods the assimilation schemes currently used in operational centers, as for example at the National Centers for Environmental Prediction (NCEP[1]); Parrish & Derber [112]) in the United States and the European Centre for Medium–Range Weather Forecasts (ECMWF; Heckley et al. [74]), in England.

Weather forecasting started after the invention of synoptic charts (meteorological maps). Nowadays, data analysis by synoptic charts is known as *subjective analysis*, since these were designed by hand, and consequently were subjected to empiricism and skill. These charts are built by marking the magnitude of an observed quantity, in the locations where observations are made, on an ordinary geographic map, and by tracing contours between the marked points. Visual extrapolation and interpolation are made in order to allow for contours drawing. From these charts, many atmospheric conditions can be inferred, including conditions about variables not directly observed, by applying a series of rules based on geometric arguments (Bjerkenes [14]; see also Saucier [122] for a detailed explanation of these procedures). A meteorologist with great experience can then issue forecasts for one or two days based on these charts.

The advent of computers and the evolution of numerical analysis methods introduced a more rapid and consistent way of generating synoptic charts. The first *objective analysis*, as it was called, was produced by Panofsky [111]. He used a technique of fit by least squares in two dimensions. This technique consists basically in expanding the fields (variables), which are to be analyzed, in a series of polynomials about the observation point, minimizing the

[1] Old National Meteorological Center

square of their differences with the observed values. The expansion coefficients are then determined by inverting a matrix.

Furthermore, the availability of computers contributed to the arrival of a new area of research in the beginning of the 40's. Led by the great John von Neumann at the Institute of Advanced Studies in Princeton, a prominent group of meteorologists initiated what was later called numerical weather prediction (NWP). Initiated in the end of the 40's, in a rich collaborative environment that continued until the mid–50's, as described by Wiin–Neilsen [139], the first weather prediction was performed by Charney, Fjørtoft & Neumann [23]. This prediction was based on the numerical integration of the barotropic vorticity equation, as proposed by Charney.

The work of Panofsky [111] in objective analysis was motivated by the Princeton project and executed there. This procedure was perfected by ideas such as those of Bergthórsson & Döös [11], who used a numerical forecasts as the first guess to the analyzed field. This led to the successive corrections method that represents considerable computational improvement over the polynomial fit method. As described in Lorenc et al. [96], an improved version of this method is used today, and is called *analysis correction*, at the British forecast center (U.K. Meteorological Office).

The work of Eliassen [49], Gandin [57], Eddy [45, 46], and Phillips [116] introduced statistical interpolation ideas to atmospheric science problems. This procedure is analogous to the successive corrections procedure in the sense that the analyzed variable, w_j^a, at a point j, is still obtained by means of a linear combination between the forecast field w_j^f (here we use f for forecast), at the same point, and the increment (*innovation*) due to the observation at this point, according to the expression

$$w_j^a = w_j^f + \mathbf{k}_j^T(\mathbf{w}^o - \mathbf{H}\mathbf{w}^f) \tag{8.1}$$

where $\mathbf{w}^o$ is a m–vector representing m observations and the $m \times n$ matrix $\mathbf{H}$ corresponds to necessary interpolations in order to transfer information from the forecast n–vector $\mathbf{w}^f$, usually obtained on a regular grid, to the observation places. Analysis methods based on (8.1) are said to be univariate, since observations of a certain quantity corrects only equivalent quantities. That is, we can identify the vector w as being the temperature, so that the use of (8.1) does not affect the winds, but the temperature field alone.

The quantities $\mathbf{k}_j$ above are the weights given to the observation increments. These can be determined based on what we saw in the initial lectures using statistical concepts. For example, they can be determined by imposing the condition that the ensemble mean of the difference between the analysis and the true value of the analyzed quantity be minimum. In the method of least squares, the quantity to be minimized is a measure of Euclidean distance between the exact value and its estimate, whose solution produces an expression for the weights in (8.1),

$$[\mathbf{H}\mathbf{S}^f\mathbf{H}^T + \mathbf{R}]\mathbf{k}_j = \mathbf{s}_j^f \tag{8.2}$$

where $\mathbf{S}^f$ and $\mathbf{R}$ correspond to the forecast and observation error covariance matrices; $\mathbf{s}_j^f$ is a j^th column of the forecast error covariance matrix.

Physical constraints were incorporated into statistical analysis by Gandin [57] and applied to operational data assimilation systems by McPherson et al. [102] and Lorenc [95]. These

constraints are taken into consideration by means of an extension of the univariate form, briefly described above, for the multivariate case. In this last case, the analyzed state at a point is built as a combination of information about various variables in distinct points. For example, the forecast and analysis vectors above can be redefined to include winds and the mass fields, as $\mathbf{w}_j^{f,a} \equiv (u_j^{a,f}, v_j^{f,a}, h_j^{f,a})^T$, where u and v represent the two components of the wind in the zonal and meriodional directions, and, h represents the mass field. The calculation of the weights $\mathbf{k}_j$ is accomplished by following an expression similar to (8.2), except that now, the forecast error covariance matrix $\mathbf{S}^f$ contains terms of cross–covariances related to the cross–correlations among different variables. Therefore, the matrices in (8.2) are of larger dimension than in that of the univariate case. Therefore, the inclusion of physical constraints results in a computationally more intensive system of equations to be solved to produce an analysis.

To solve the system of equations (8.2), we assume the statistical properties of the forecast errors is known, that is, that the matrix $\mathbf{S}^f$ is known. In fact, this matrix is prescribed based on assumptions such as vertical separability and horizontal homogeneity and isotropy of error correlations. Observational studies such as those of Rutherford [119] and Schlatter [123], based on differences of observed and climatological forecast fields are designed to determine the statistics of forecast errors. Analytical expressions can be derived to parameterize error statistics, and the study of Balgovind et al. [6] (see Exercise 8.3) justifies some of these parameterizations by using a univariate model based on the quasi–geostrophic potential vorticity equation.

Both methods of univariate and multivariate statistical interpolation or, as more commonly known, *optimal interpolation*, are used nowadays in forecast centers such as NCEP and ECMWF. In some of its applications, forecast errors variances are taken as a function of a linear growth parameter, while the correlation fields are prescribed in an analytical form, by imposing geostrophic balance for the error fields (Schlatter et al. [124]), as the physical constraint.

As we mentioned previously, the major part of the computational load in optimal interpolation is primarily due to the matrix inversion in (8.2); for regions with high data density, this matrix can have a very large dimension. Defining a region of influence, to produce an analyzed variable at a certain point taking into consideration observations within a radius on influence only is one attempt to reduce the computational cost (Lorenc [95]). This technique is known as *data selection*, and in some cases it has unpleasant consequences (da Silva et al. [37]). Another technique, known as superobing, substitutes various observations occurring in nearby locations by a single observation, as for example, the mean of all the observations, with modified standard deviation. Superobing reduces the number of observations included in (8.2) (see Lorenc [95]).

Another factor responsible for high computational cost in operational analysis systems is the need for *quality control* of the data. Quality control systems are developed in order to eliminate and/or correct observations with gross errors, which are introduced by artificial means, distinct from the measurement process itself, such as data corruption due to transmission via the telecommunication network. The study of Hollingsworth et al. [80] shows that the final analysis result is very sensitive to quality control procedures. Other approaches to the quality control problem are found in the work of Lorenc & Hammon [97]

and Gandin [58].

The greatest disadvantage of the optimal interpolation method is that the forecast error covariance is *prescribed* in a relatively arbitrary manner. Although the errors possess dynamic balance, the dynamic balance is not exactly satisfied by the governing equations, moreover these errors do not propagate in any way. Explicit use of the governing equations to derive error statistics is the main context of advanced methods for data assimilation, to be discussed in the following lecture. Among other things, the advanced methods produce error statistics with appropriate balance.

8.2 Initialization

Intialization is the procedure by which gravity waves are filtered out from the initial conditions to allow for an evolution *practically* free of fast components. We say practically because in the general nonlinear case, the evolution of the initialized state (after initialization has been performed) still contains allows for fast components of the system to be excited, since the all modes interact due to nonlinearity. For the linear case, however, various initialization procedures exist and consist of the motivation for extensions to the nonlinear case.

Initialization is not a data assimilation method, however, it is a fundamental ingredient for atmospheric forecast obtained through computer models. The majority of assimilation methods used operationally use some type of initialization method for the fields to be used as initial condition in the forecast model. It is worth saying that the need to initialize the analyzed fields, that is, fields obtained after after the assimilation of observations, could be, in some cases, eliminated if the assimilation procedure was done "correctly". In other words, referring to the existing balance due to dynamical processes, it is possible, in some cases, to produce analyzed fields which are automatically balanced, without us having to use an explicit initialization procedure. The word "correct" used above, refers to assimilation schemes in which the initialization procedure is embedded in its structure. This will become more clear as we progress. It is good to underline that we are referring only to balances due to dynamic processes, such as geostrophic balance; balances due to physical processes, such as heat exchange in the atmospheric system, are much more complicated to incorporate automatically in data assimilation procedure. The major part of this topic goes beyond what we intend to cover in these lectures, and we will see in the following sections only superficially what we have just mentioned.

One way of describing the initialization problem is to imagine that in the solution space of the equations that govern atmospheric motion there exists a subspace of slow solutions called $\mathcal{S}$, which is free of high frequency waves (gravity waves — with potential to destroy possible weather forecasts). In fact, there are various definitions for what it is understood for the slow subspace (slow manifold; see Boyd [17]) but we will not go into these details in what follows. Initialization can be seen as the process of projecting, in some way, a general state represented by the n–vector $\mathbf{w}^a$, at a specific time, onto this slow manifold. Through concepts of linear algebra, we know that there is a matrix (operator) of projection $\mathbf{\Pi}$, of

dimension $n \times n$, which satisfies the following conditions

$$\begin{aligned} Range\,\mathbf{\Pi} &= \mathcal{S}\,, && (8.3a)\\ \mathbf{\Pi}^2 &= \mathbf{\Pi}\,, && (8.3b)\\ (\mathbf{E\Pi})^T &= \mathbf{E\Pi}\,, && (8.3c) \end{aligned}$$

(e.g., Halmos [69], Section 75). This matrix is known as the **E**–orthogonal projection matrix onto $\mathcal{S}$, or simply the orthogonal projection in when **E** is the identity matrix. Here, **E** is assumed to be positive definite and symmetric.

Let $\mathbf{V}_S$ be an $n \times n_S$ matrix, with columns built from the slow eigenvectors of a dynamical system, that is, the slow normal modes which total n_S. Hence, the **E**–orthogonal projector is given by:

$$\mathbf{\Pi} = \mathbf{V}_S \left(\mathbf{V}_S^T \mathbf{E} \mathbf{V}_S\right)^{-1} \mathbf{V}_S^T \mathbf{E}\,, \tag{8.4}$$

which can be verified by substitution in the conditions (8.3).

Making use of this projector we can describe the initialization problem as a least squares problem: find the n–vector $\mathbf{w}^i$ in $\mathcal{S}$ that is as close as possible to the general vector $\mathbf{w}^a$. In other words, we want to minimize the following functional

$$\eta = (\mathbf{w}^i - \mathbf{w}^a)^T \mathbf{E} (\mathbf{w}^i - \mathbf{w}^a) \tag{8.5}$$

where **E** is a symmetric and positive definite weight, or re–scaling, matrix. As we will show, the solution of this problem is, as we can expect, given uniquely by

$$\mathbf{w}^i = \mathbf{\Pi} \mathbf{w}^a\,. \tag{8.6}$$

Posed in this way, the initialization problem is know as *linear variational normal mode initialization* (e.g., Daley [40], Temperton [127]).

To prove the result above, notice that certainly $\mathbf{w}^i \equiv \mathbf{\Pi} \mathbf{w}^a \in \mathcal{S}$. Moreover, a general element of $\mathcal{S}$ can be written as

$$\mathbf{w} = \mathbf{w}^i + \epsilon = \mathbf{\Pi} \mathbf{w}^a + \epsilon \tag{8.7}$$

where

$$\epsilon = \mathbf{\Pi} \epsilon\,, \tag{8.8}$$

since $\epsilon = \mathbf{w} - \mathbf{w}^i$ should be in $\mathcal{S}$. Now,

$$\begin{aligned} \eta = \eta(\mathbf{w}) &= (\mathbf{w} - \mathbf{w}^a)^T \mathbf{E} (\mathbf{w} - \mathbf{w}^a)\\ &= \left[(\mathbf{\Pi} - \mathbf{I})\mathbf{w}^a + \epsilon\right]^T \mathbf{E} \left[(\mathbf{\Pi} - \mathbf{I})\mathbf{w}^a + \epsilon\right]\\ &= \left[(\mathbf{\Pi} - \mathbf{I})\mathbf{w}^a\right]^T \mathbf{E} \left[(\mathbf{\Pi} - \mathbf{I})\mathbf{w}^a\right] + \epsilon^T \mathbf{E} \epsilon + 2\delta\,, \end{aligned} \tag{8.9}$$

where

$$\begin{aligned} \delta &= \epsilon^T \mathbf{E} (\mathbf{\Pi} - \mathbf{I}) \mathbf{w}^a\\ &= (\mathbf{\Pi}\epsilon)^T \mathbf{E} (\mathbf{\Pi} - \mathbf{I}) \mathbf{w}^a\,, \end{aligned} \tag{8.10}$$

according to (8.8). Based on (8.3c) and on the fact that $\mathbf{E}^T = \mathbf{E}$ it follows that

$$\begin{aligned}
\delta &= \epsilon^T \mathbf{\Pi}^T \mathbf{E}(\mathbf{\Pi} - \mathbf{I})\mathbf{w}^a \\
&= \epsilon^T (\mathbf{E}\mathbf{\Pi})^T (\mathbf{\Pi} - \mathbf{I})\mathbf{w}^a \\
&= \epsilon^T \mathbf{E}\mathbf{\Pi}(\mathbf{\Pi} - \mathbf{I})\mathbf{w}^a \\
&= \epsilon^T \mathbf{E}(\mathbf{\Pi}^2 - \mathbf{\Pi})\mathbf{w}^a \\
&= 0\,, \qquad (8.11)
\end{aligned}$$

where (8.3b) was used to obtain the last equality. By (8.9) we have

$$\begin{aligned}
\eta(\mathbf{w}) &= (\mathbf{w}^i - \mathbf{w}^a)^T \mathbf{E}(\mathbf{w}^i - \mathbf{w}^a) + \epsilon^T \mathbf{E}\epsilon \\
&= \eta(\mathbf{w}^i) + \epsilon^T \mathbf{E}\epsilon \\
&\geq \eta(\mathbf{w}^i)\,, \qquad (8.12)
\end{aligned}$$

where the equality prevails if and only if $\epsilon = \mathbf{0}$, since $\mathbf{E}$ is positive defined. Hence, $\mathbf{w}^i = \mathbf{\Pi}\mathbf{w}^a$ is the unique minimizer.

As an example of the initialization procedure, we follow Daley [39] and consider the system of shallow water equations linearized about the state of rest and dealt with here earlier in Section 6.5. For that, we introduce first an operator $\mathbf{F}$ corresponding to the Fourier transform in two spatial dimensions. Thus, formally a vector $\mathbf{w}(x, y, t)$ can be transformed according to

$$\hat{\mathbf{w}}(k, \ell, t) = \mathbf{F}[\mathbf{w}(x, y, t)] \qquad (8.13)$$

The Fourier operator $\mathbf{F}^{-1}$ "=" $\mathbf{F}^*$ which correspond to the inverse Fourier transform is such that

$$\mathbf{w}(x, y, t) = \mathbf{F}^*[\mathbf{w}(k, \ell, t)] \qquad (8.14)$$

where $*$ indicates the conjugate transpose of an operator. Consequently, we can write the matrix $\hat{\mathbf{V}}_S$ from the previous lecture as,

$$\hat{\mathbf{V}}_S \equiv \mathbf{F}[\mathbf{V}_S] \qquad (8.15)$$

Furthermore, we define the Fourier component $\hat{\mathbf{E}}$ of the matrix $\mathbf{E}$ above as:

$$\hat{\mathbf{E}} \equiv \mathbf{F}(\mathbf{F}^*\mathbf{E})^T \qquad (8.16)$$

Then, formally again, by inserting the appropriate identity in expression (8.4) for the projector $\mathbf{\Pi}$ we have

$$\begin{aligned}
\mathbf{\Pi} &= (\mathbf{F}^*\mathbf{F})\mathbf{V}_S \left[\mathbf{V}_S^T (\mathbf{F}\mathbf{F}^*)\mathbf{E}(\mathbf{F}^*\mathbf{F})\mathbf{V}_S\right]^{-1} \mathbf{V}_S^T (\mathbf{F}^*\mathbf{F})\mathbf{E}(\mathbf{F}^*\mathbf{F})\,, \\
&= \mathbf{F}^*\hat{\mathbf{V}}_S \left(\hat{\mathbf{V}}_S^T \hat{\mathbf{E}}\hat{\mathbf{V}}_S\right)^{-1} \hat{\mathbf{V}}_S^T \hat{\mathbf{E}}\mathbf{F}\,, \\
&= \mathbf{F}^*\hat{\mathbf{\Pi}}\mathbf{F} \qquad (8.17)
\end{aligned}$$

where

$$\hat{\mathbf{\Pi}} \equiv \hat{\mathbf{V}}_S \left(\hat{\mathbf{V}}_S^T \hat{\mathbf{E}}\hat{\mathbf{V}}_S\right)^{-1} \hat{\mathbf{V}}_S^T \hat{\mathbf{E}}\,, \qquad (8.18)$$

Explicitly, we recall that, in the Fourier space, the solution of the system of shallow water equations in time t is given by

$$\hat{\mathbf{w}}(t) = \hat{\mathbf{V}}\mathbf{\Lambda}(t) \;=\; \hat{\mathbf{V}} \begin{pmatrix} c_- e^{i\sigma^- t} 0 & 0 & \\ 0 & c_0 & 0 \\ 0 & 0 & c_+ e^{i\sigma^+ t} \end{pmatrix} \tag{8.19}$$

where $\hat{\mathbf{V}}$ is the eigenvector matrix in (7.43), which for convenience we write as in Daley [39]:

$$\hat{\mathbf{V}} \;=\; \begin{pmatrix} G_\psi^- & R_\psi & G_\psi^+ \\ G_\chi^- & R_\chi & G_\chi^+ \\ G_\phi^- & R_\phi & G_\phi^+ \end{pmatrix}, \tag{8.20}$$

and where the coefficients $\hat{\mathbf{c}} = (\hat{c}_-, \hat{c}_0, \hat{c}_+)^T$ can be determined from the initial condition, as in (7.47), and which we repeat for ease of reference once again:

$$\hat{\mathbf{c}} = \hat{\mathbf{V}}^T \begin{pmatrix} \hat{\psi}^a(0) \\ \hat{\chi}^a(0) \\ \hat{\phi}^a(0) \end{pmatrix} \tag{8.21}$$

for $\hat{\mathbf{w}}^a(0) = [\hat{\psi}^a(0), \hat{\chi}^a(0), \hat{\phi}^a(0)]^T$, a general initial state, meaning an analysis.

Therefore, according to (8.21) and (8.19), if a general initial condition has fast components, the solution for all times t_k will also have fast components unless the coefficients $\hat{c}_-$ and $\hat{c}_+$ are zero. The goal of the initialization procedure is to reconstruct the initial state $\mathbf{w}^a(0)$ as a (initialized) state $\hat{\mathbf{w}}^i(0) \equiv [\hat{\psi}^i(0), \hat{\chi}^i(0), \hat{\phi}^i(0)]^T$, free of fast components, so that its evolution (8.19) is also free of fast (gravity) waves.

Moreover, in the case of simple linear systems, the slow subspace coincides with the geostrophic space (rotational), and the "matrix" $\mathbf{V}_S$ of slow eigenvectors is given by

$$\hat{\mathbf{V}}_S = \begin{pmatrix} R_\psi \\ R_\chi \\ R_\phi \end{pmatrix} \tag{8.22}$$

for a number $n_S = 1$ of slow vectors.

Choosing the weighting matrix $\hat{\mathbf{E}}$ as a diagonal matrix (in Fourier space) and representing it by $\hat{\mathbf{E}} = diag(w_\psi, w_\chi, w_\phi)$ the kernel $\left(\hat{\mathbf{V}}_S^T \hat{\mathbf{E}} \hat{\mathbf{V}}_S\right)^{-1}$ of the $\mathbf{E}$–orthogonal projector can then be calculated according to

$$\begin{aligned} \left(\hat{\mathbf{V}}_S^T \hat{\mathbf{E}} \hat{\mathbf{V}}_S\right)^{-1} &= \left[\begin{pmatrix} R_\psi & R_\chi & R_\phi \end{pmatrix} \begin{pmatrix} w_\psi & 0 & 0 \\ 0 & w_\chi & 0 \\ 0 & 0 & w_\phi \end{pmatrix} \begin{pmatrix} R_\psi \\ R_\chi \\ R_\phi \end{pmatrix} \right]^{-1} \\ &= \frac{1}{w_\psi R_\psi^2 \,+\, w_\chi R_\chi^2 \,+\, w_\phi R_\phi^2} \\ &= \frac{k+1}{k w_\psi \,+\, w_\phi} \end{aligned} \tag{8.23}$$

where we make explicit use of the elements of the matrix $\hat{\mathbf{V}}_S$. Therefore, the projector $\hat{\mathbf{\Pi}}$ takes the form

$$\begin{aligned}\hat{\mathbf{\Pi}} &= \frac{k+1}{kw_\psi + w_\phi}\begin{pmatrix} R_\psi \\ R_\chi \\ R_\phi \end{pmatrix}\begin{pmatrix} R_\psi & R_\chi & R_\phi \end{pmatrix}\begin{pmatrix} w_\psi & 0 & 0 \\ 0 & w_\chi & 0 \\ 0 & 0 & w_\phi \end{pmatrix} \\ &= \frac{1}{kw_\psi + w_\phi}\begin{pmatrix} kw_\psi & 0 & \sqrt{k}w_\phi \\ 0 & 0 & 0 \\ \sqrt{k}w_\psi & 0 & w_\phi \end{pmatrix} \end{aligned} \tag{8.24}$$

Applying this projector to the general, non–initialized, vector $\mathbf{w}^a$ as in (8.6), we have

$$\hat{\mathbf{w}}^i \equiv \begin{pmatrix} \hat{\psi}^i \\ \hat{\chi}^i \\ \hat{\phi}^i \end{pmatrix} = \frac{1}{kw_\psi + w_\phi}\begin{pmatrix} kw_\psi\hat{\psi}^a + \sqrt{k}w_\phi\hat{\phi}^a \\ 0 \\ \sqrt{k}w_\psi\hat{\psi}^a + w_\phi\hat{\phi}^a \end{pmatrix} \tag{8.25}$$

The projector for which the diagonal elements of the matrix $\hat{\mathbf{E}}$ are unity, that is, $w_\phi = w_\chi = w_\phi = 1$ is called *slow orthogonal projector*. A slow state generated by means of this projector is one corresponding to zero (divergence) velocity potential $\hat{\chi} = 0$, and stream and geopotential functions given by:

$$\hat{\psi}^i = \frac{k\hat{\psi}^a + \sqrt{k}\hat{\phi}^a}{k+1} \tag{8.26a}$$

$$\hat{\phi}^i = \frac{\sqrt{k}\hat{\psi}^a + \hat{\phi}^a}{k+1} \tag{8.26b}$$

respectively.

Notice that this operation leaves the geostrophic modes unaltered. That is, for the case in which the general initial state is geostrophically balanced, we have that $\hat{\psi}^a = \sqrt{k}\hat{\phi}^a$, and from the expressions above it follows that the initialized state is given by,

$$\hat{\psi}^i = \sqrt{k}\hat{\phi}^a = \hat{\psi}^a \tag{8.27a}$$

$$\hat{\phi}^i = \hat{\phi}^a \tag{8.27b}$$

Therefore, the initialized state $\hat{\psi}^i = \sqrt{k}\hat{\phi}^i$, is also geostrophically balanced.

In fact, from (8.26) it follows that

$$k\hat{\psi}^i + \sqrt{k}\hat{\phi}^i = k\hat{\psi}^a + \sqrt{k}\hat{\phi}^a \tag{8.28}$$

Thus, observing that $\hat{q} = k\hat{\psi} + \sqrt{k}\hat{\phi}$ is the expression for the quasi–geostrophical potential vorticity (see Exercise 7.1), we see that $\hat{q}^i = \hat{q}^a$ means that the slow orthogonal projector keeps this quantity conserved.

The slow orthogonal projector used above can be written as

$$\hat{\mathbf{\Pi}}_{\|} = \frac{1}{k+1}\begin{pmatrix} k & 0 & \sqrt{k} \\ 0 & 0 & 0 \\ \sqrt{k} & 0 & 1 \end{pmatrix} \tag{8.29}$$

Other possible choices for the weights w_ψ and w_ϕ produce the following projectors:

$$\hat{\mathbf{\Pi}}_g = \begin{pmatrix} 0 & 0 & \sqrt{k} \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{8.30}$$

for $w_\psi = 0$ e $w_\phi = 1$, and

$$\hat{\mathbf{\Pi}}_r = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1/\sqrt{k} & 0 & 0 \end{pmatrix} \tag{8.31}$$

for $w_\psi = 1$ and $w_\phi = 0$. An initialized state generated by means of $\hat{\mathbf{\Pi}}_g$ is referred to as a state with *geopotential constraint*, since this projector maintains the geopotential component $\hat{\phi}^a$ of a general state $\mathbf{w}^a$ unaltered; analogously, a state generated by the projector $\hat{\mathbf{\Pi}}_r$ is said to be a state of *rotational wind constraint*, since in this case, it is stream function $\hat{\psi}^a$ is unaltered.

The projector $\hat{\mathbf{\Pi}}_g$ is adequate when the geopotential component of the general state $\mathbf{w}^a$ is the one in which we have greater dependability (less error) than the other state components; the projector $\hat{\mathbf{\Pi}}_r$, however, is convenient when we have great reliability in the component of the stream function. In the general case, in which the reliability of the components of an arbitrary state $\mathbf{w}^a$ is not well defined, the orthogonal projector $\hat{\mathbf{\Pi}}_{||}$ is the most adequate. Obviously, when the component of divergent winds $\hat{\chi}$ is of great reliability, we have to choose one of the projectors above. In fact, if we do not know anything about the geopotential and stream function components, we can do nothing in terms of initialization for the case of the shallow water equations linearized about the state of rest.

It is important to notice that, in the example we are considering in this section, the normal modes of the system are orthogonal. In the more general case in which these modes are not orthogonal, basically due to the fact that the matrix $\hat{\mathbf{L}}$ of Section 6.5 is not symmetric due to general basic flows, more general projectors can be obtained. In these cases we should make use of the a bi–orthonormal set of modes that can be obtained through the use of eigenvectors of the adjoint matrix of $\hat{\mathbf{L}}$ (e.g., Ghil [61], for more details in the context of shallow water model in one dimension, linearized about a constant jet, see also Exercise 7.2).

8.3 Dynamic Relaxation

Dynamic relaxation, also called nudging or Newtonian relaxation, is a data assimilation procedure continuous in time. The observations are introduced in the governing equations by means of forcing terms added to the equations, in order to "pull" (relax) the fields in the direction of the observations. Dynamic relaxation is employed during a period of time known as the *pre–forecast*, so that at the end of this period the solution is as close as we want to the observations; from that time on, it is possible to produce regular forecasts. One of the advantages of this procedure is that the initial fields, from which the forecasts are issued, are automatically in dynamic balance at the end of the assimilation (relaxation) period.

More clearly, following the treatment of Haltiner & Williams [70], the data assimilation procedure by dynamic relaxation consists of the following steps:

1. Specify the initial condition, at time $t_0 - T$, when the evolution of the pre–forecast begins, where T represents the period of forecast time and t_0 the instant of time when the forecasts are to be issued from.

2. Solve the governing equations during the interval of time $[t_0 - T, t_0]$, including the forcing terms to relax the solution in the direction of the observations (or analysis).

3. Arriving at time t_0, evolve the governing equations, without the forcing terms, up to the time t of the desired forecast.

In general, the evolution of any prognostic quantity, at a mesh point, where an observation supposedly exists, can be represented by the equation:

$$\frac{\partial w}{\partial t} = f(\mathbf{w}) + \gamma(w^o - w) \tag{8.32}$$

where w is the scalar quantity of interest, f is a function of the vector state $\mathbf{w}$ of the system, which includes the terms of the governing dynamics, the last term is a component of the forcing term, added to the governing equations during the pre–forecast period, and includes the observation w^o, with relaxation parameter γ. Written in this form, the equation above presumes the availability of the observation at the mesh point of interest, and eventually at all grid points. Since observations are rarely available at grid points, it is best to replace the observation w^o by the analyzed value w^a. In this way the relaxation expression can be written as

$$\frac{\partial w}{\partial t} = f(\mathbf{w}) + \gamma(w^a - w) \tag{8.33}$$

The intention of the method can be understood from a simple example, by considering $f = 0$ in the equation above. Assuming that the observation w^o is independent of time, and integrating (8.32) from $t_0 - T$ to t_0 we have:

$$\begin{aligned} w &= w_0 e^{-\gamma T} + \gamma w^a e^{-\gamma t_0} \int_{t_0 - T}^{t_0} e^{\gamma s}\, ds \\ &= w_0 e^{-\gamma T} + (1 - e^{-\gamma T}) w^a \end{aligned} \tag{8.34}$$

where w_0 is solution at time $t_0 - T$. Therefore, as the relaxation interval T increases, the solution approaches the value of the analysis w^a (observation w^o). In practice, the interval T is fixed and the relaxation parameter γ is chosen in order to relax the solution more rapidly, or more slowly, in the direction of the analyses (observations).

Another example can be presented by returning to the system of equations in Section 6.5. When we introduce the dynamic relaxation terms referring to the analyzed fields u^a, v^a and ϕ^a, of winds and geopotential, respectively, we have:

$$\frac{\partial u}{\partial t} - f_0 v + \frac{\partial \phi}{\partial x} - \gamma_u(u^a - u) = 0 \tag{8.35a}$$

$$\frac{\partial v}{\partial t} + f_0 u + \frac{\partial \phi}{\partial y} - \gamma_v (v^a - v) = 0 \tag{8.35b}$$

$$\frac{\partial \phi}{\partial t} + \Phi(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}) - \gamma_\phi (\phi^a - \phi) = 0 \tag{8.35c}$$

where, in this case, the relaxation parameters γ_u, γ_v, and γ_ϕ are in principle distinct. The analytic study of this problem can be found in Hoke & Anthes [81], for the unidimensional case.

To simplify the problem mathematically, let us follow Daley [39] where only the geopotential variable is relaxed (and observed), that is, $\gamma_u = \gamma_v = 0$. In this way, the equations for u and v are identical to those in Section 6.5, without the forcing term, thus we can use in their places the corresponding equations for vorticity and divergence. So, the system of equations to be solved becomes:

$$\frac{\partial \nabla^2 \psi}{\partial t} + f_0 \nabla^2 \chi = 0 \tag{8.36a}$$

$$\frac{\partial \nabla^2 \chi}{\partial t} - f_0 \nabla^2 \psi + \nabla^2 \phi = 0\,, \tag{8.36b}$$

$$\frac{\partial \phi}{\partial t} + \Phi \nabla^2 \chi + \gamma_\phi \phi = \gamma_\phi \phi^a \tag{8.36c}$$

Let us assume that the exact solution of the evolution of the fields $\psi = \psi^t, \chi = \chi^t$ and $\phi = \phi^t$ follows the system of original equations (7.29), without the forcing term, and that the state to be analyzed evolves according to the forced system of equations so that $\psi = \psi^i, \chi = \chi^i$ and $\phi = \phi^i$, are now the quantities which satisfy (8.36). Notice that both systems of equations (7.29) and (8.36) are linear. Defining the initialization errors in an usual way:

$$\tilde{\psi} = \psi^i - \psi^t \tag{8.37a}$$

$$\tilde{\chi} = \chi^i - \chi^t \tag{8.37b}$$

$$\tilde{\phi} = \phi^i - \phi^t \tag{8.37c}$$

we can identify the system of equations (8.36) as describing the evolution of errors. That is, the initialization errors evolve according to

$$\frac{\partial \nabla^2 \tilde{\psi}}{\partial t} + f_0 \nabla^2 \tilde{\chi} = 0 \tag{8.38a}$$

$$\frac{\partial \nabla^2 \tilde{\chi}}{\partial t} - f_0 \nabla^2 \tilde{\psi} + \nabla^2 \tilde{\phi} = 0\,, \tag{8.38b}$$

$$\frac{\partial \tilde{\phi}}{\partial t} + \Phi \nabla^2 \tilde{\chi} + \gamma_\phi \tilde{\phi} = \gamma_\phi \tilde{\phi}^a \tag{8.38c}$$

where $\tilde{\phi}^a$ is the analysis error in the geopotential field..

The system above has constant coefficients, as the system in Section 6.5, and we can once again solve it by normal modes. We introduce a transformation equivalent to that in (7.34), but now for the error fields:

$$\begin{pmatrix} \tilde{\psi}(x,y,t) \\ \tilde{\chi}(x,y,t) \\ \tilde{\phi}(x,y,t) \end{pmatrix} = \begin{pmatrix} \hat{\psi}(t) \\ i\hat{\chi}(t) \\ f_0\sqrt{k}\hat{\phi}(t) \end{pmatrix} \exp\left\{ i \left[\frac{(mx + ny)}{a} \right] \right\} \tag{8.39}$$

where m, n and a have the same meaning as in previously, and where the constant k is defined as in (7.35). Notice that the amplitudes $\hat{\psi}(t), \hat{\chi}(t)$ and $\hat{\phi}(t)$ in this case are not necessarily the same as those in Section 6.5, although they are represented by the same symbols. Furthermore, we decompose the error in the analysis ϕ^a in a similar manner:

$$\tilde{\phi}^a(x, y, t) \;=\; f_0\sqrt{k}\hat{\phi}^a(t)\exp\left\{i\left[\frac{(mx \;+\; ny)}{a}\right]\right\} \tag{8.40}$$

Therefore, the system of equations (8.38) is reduced to an ordinary non-homogeneous differential equation:

$$\frac{d\hat{\mathbf{w}}(t)}{dt} \;+\; if_0\hat{\mathbf{L}}'\hat{\mathbf{w}}(t) \;=\; \gamma\hat{\mathbf{w}}^a\,, \tag{8.41}$$

where the vector $\hat{\mathbf{w}} \equiv (\hat{\psi}, \hat{\chi}, \hat{\phi})^T$, the matrix $\hat{\mathbf{L}}'$ is given by

$$\hat{\mathbf{L}}' = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -\sqrt{k} \\ 0 & -\sqrt{k} & -i\gamma_\phi/f_0 \end{pmatrix}. \tag{8.42}$$

and the non–homogeneous part of the system of equations is given by the the analysis error vector $\hat{\mathbf{w}}^a \equiv (0, 0, \hat{\phi}^a)^T$. The solution of (8.41) consists of the solution of the homogeneous part of the system of equations plus the particular solution, that is,

$$\hat{\mathbf{w}}(t_0) = e^{-if_0\hat{\mathbf{L}}'T}\hat{\mathbf{w}}(t_0 - T) \;+\; \hat{\mathbf{w}}_p(t_0) \tag{8.43}$$

where $\hat{\mathbf{w}}(t_0 - T)$ is the initial condition vector and $\hat{\mathbf{w}}_p(t_0)$ represents a particular solution.

The interest here is to study of the behavior of the wave frequencies of this modified system. These frequencies are given by the eigenvalues of the matrix $\hat{\mathbf{L}}'$. However, now this matrix is imaginary, so its eigenvalues are also imaginary. It is simple to verify that the secular equation is:

$$\sigma_\ell^3 \;+\; i\frac{\gamma_\phi}{f_0}\sigma^2 \;-\; (1 \;+\; k)\sigma_\ell \;-\; i\frac{\gamma_\phi}{f_0} \;=\; 0 \tag{8.44}$$

whose solutions can be written as,

$$\sigma \;=\; Re(\sigma) \;+\; iIm(\sigma) \tag{8.45}$$

where $Re(\sigma)$ and $Im(\sigma)$ are the real and imaginary parts of the eigenvalues. The real part represents, as before, the frequencies of oscillation, while the imaginary part represents decaying, or growing, modes.

For $k = 0$, the squares of (8.44) are

$$\sigma_R \;=\; -i\frac{\gamma_p hi}{f_0} \tag{8.46a}$$

$$\sigma_G^{\pm} \;=\; \pm 1 \tag{8.46b}$$

So that the errors in the rotational mode for $k = 0$ decrease with time, while the errors in the gravitational modes are oscillatory. For $k > 0$, it is possible to show that, when $\gamma_\phi/f_0 > 0$, we have $Im(\sigma) < 0$. This means that the procedure of dynamic relaxation introduces decaying modes, except in the case of the two inertial–gravity modes for $k = 0$. Also, $Re(\sigma_R) = 0$, for $k \geq 0$, meaning that the frequencies of the rotational modes are not modified by relaxation procedure. (see Daley [39], pp. 359–360, for an approximate calculation of the frequencies for the case $k > 0$.)

8.4 Optimal Interpolation

As discussed in the introduction of this lecture, data assimilation by the method of optimal interpolation (OI) uses an expression as in (8.1) to update instantaneously the values of the variables of the system at a mesh point (analysis point). Contrary to the relaxation method seen in the previous section, OI is an intermittent assimilation method, since it is used only at synoptic times, that is, instants of time considered standard in meteorology, such as 00, 06, 12, and 18 GMT (Greenwich Mean Time). During the 6 hours period, between the synoptic times, the state of the atmosphere evolves by means of a system of equations discretized in space and time, representing a model of general circulation of the atmosphere. These general circulation models produce a base state (do not confuse it with basic state), or forecast state $\mathbf{w}_k^f$, at time t_k. The correction due to the availability of observations can be obtained on the basis of the methods discussed in Lecture 4, through the formula

$$\mathbf{w}_k^a \;=\; \mathbf{w}_k^f \;+\; \tilde{\mathbf{K}}_k(\mathbf{w}_k^o - \mathbf{H}_k\mathbf{w}_k^f) \tag{8.47}$$

where the notation used here is the same as in previous lectures. As seen in the introduction of this lecture, the weights $\tilde{\mathbf{K}}_k$ in the OI method are obtained by means of the expression

$$\tilde{\mathbf{K}}_k \;=\; \mathbf{S}_k^f\mathbf{H}_k^T(\mathbf{H}_k\mathbf{S}_k^f\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \tag{8.48}$$

where we use a "tilde" over the weighting matrix to highlight the fact that these is not the Kalman gains. The reason for this being that the forecast error covariance matrix $\mathbf{S}_k^f$ is *specified*, rather than predicted according to the Kalman filer equations. As we have seen in Lecture 5, for the case of linear systems, the calculation of the forecast error covariance matrix involves an enormous computational cost. As a matter of fact, the reasons to avoid explicit calculation of $\mathbf{S}_k^f$ go beyond the computational issue. They are also attributed to the nonlinearity of governing equations, as well as to lack of knowledge of quantities such as the modeling and observation error covariance matrices. In OI the elements of the matrix $\mathbf{S}^f$ are specified based on statistical evaluations and dynamic constraints, as described below.

In a relatively general way, the state vector at a point $\mathbf{r} = (\lambda, \varphi, p)$, at a certain instant of time, encompasses the wind vector and the geopotential function $\mathbf{w}(\mathbf{r}) = (u, v, \phi)^T(\mathbf{r})$, where for the moment we omit the time index. Thus, the error covariance matrix $\mathbf{S}^f$ between two points $\mathbf{r}_i$ e $\mathbf{r}_j$ is given by

$$\mathbf{S}^f(\mathbf{r}_i, \mathbf{r}_j) = \mathcal{E}\{\tilde{\mathbf{w}}(\mathbf{r}_i)\tilde{\mathbf{w}}^T(\mathbf{r}_j)\} \tag{8.49}$$

where $\tilde{\mathbf{w}}(\mathbf{r}) = \mathbf{w}^f(\mathbf{r}) - \mathbf{w}^t(\mathbf{r})$ is the forecast error, for $\mathbf{w}^t(\mathbf{r})$ representing the real value of the state of the atmosphere. Therefore, we can decompose $\mathbf{S}^f$ as

$$\mathbf{S}^f(\mathbf{r}_i, \mathbf{r}_j) \;\equiv\; \begin{pmatrix} S^{f|uu}(\mathbf{r}_i, \mathbf{r}_j) & S^{f|uv}(\mathbf{r}_i, \mathbf{r}_j) & S^{f|u\phi}(\mathbf{r}_i, \mathbf{r}_j) \\ S^{f|vu}(\mathbf{r}_i, \mathbf{r}_j) & S^{f|vv}(\mathbf{r}_i, \mathbf{r}_j) & S^{f|v\phi}(\mathbf{r}_i, \mathbf{r}_j) \\ S^{f|\phi u}(\mathbf{r}_i, \mathbf{r}_j) & S^{f|\phi v}(\mathbf{r}_i, \mathbf{r}_j) & S^{f|\phi\phi}(\mathbf{r}_i, \mathbf{r}_j) \end{pmatrix} \tag{8.50}$$

where $S^{f|\cdot\cdot}(\mathbf{r}_i, \mathbf{r}_j)$ are the cross–covariance functions defined in analogy to (8.49), that is,

$$S^{f|uu}(\mathbf{r}_i, \mathbf{r}_j) = \mathcal{E}\{\tilde{u}(\mathbf{r}_i)\tilde{u}(\mathbf{r}_j)\} \tag{8.51}$$

where $\tilde{u}(\mathbf{r}) = u^f(\mathbf{r}) - u^t(\mathbf{r})$ represents the forecast error in the variable u, at point $\mathbf{r}$, and similarly for the rest of the error cross–covariances in (8.50).

The geopotential–geopotential forecast error covariance function $S^{f|\phi\phi}$, can be written in terms of the correlation and variance fields as

$$S_{ij}^{f|\phi\phi} \equiv S^{f|\phi\phi}(\mathbf{r}_i, \mathbf{r}_j) = \sigma_i^{\phi}\sigma_j^{\phi}C_{ij}^{\phi\phi}, \tag{8.52}$$

using a compact notation for the error standard deviation $\sigma_i^{\phi} = \sigma^{\phi}(\mathbf{r}_i)$ and the correlation, $C_{ij}^{\phi\phi} = C^{\phi\phi}(\mathbf{r}_i, \mathbf{r}_j)$. In conventional OI, the variance and correlation fields for the geopotential error are specified empirically as seen below, and the remainder of the cross correlations is specified by imposing the geostrophic constraint; that is, by assuming that the prediction error fields are geostrophically balanced.

In conventional applications of OI, we assume the field of standard deviation of geopotential errors to be independent of the coordinates in the horizontal $\mathbf{s} \equiv (\lambda, \varphi)$. Hence, the standard deviations of geopotential errors

$$\sigma_i^{\phi} = \sigma^{\phi}(\mathbf{r}_i) = \sigma^{\phi}(p_i) \tag{8.53}$$

are taken as a function of the pressure (height) levels alone. Moreover, the geopotential–geopotential correlation field is considered to be horizontally homogeneous, and separable from the vertical components, that is,

$$C_{ij}^{f|\phi\phi} = C^{\phi\phi}(\mathbf{s}_i - \mathbf{s}_j)V^{\phi\phi}(p_i, p_j) \tag{8.54}$$

Here we recall the notion of homogeneous random fields introduced in Lecture 2. Finally, we impose the hypothesis of horizontal isotropy, so that we can write

$$C_{ij}^{\phi\phi} = C^{\phi\phi}(s_{ij} = |\mathbf{s}_i - \mathbf{s}_j|)\,. \tag{8.55}$$

The effects of the homogeneity hypothesis were carefully studied in Cohn & Morone [32], for the case of spherical geometry. In particular, these authors observed that, in certain cases, the hypothesis that the standard deviations are independent of the horizontal coordinate is responsible for up to 30% errors in the real value of the standard deviations. The separability hypothesis of the correlations field in the vertical is currently seen as one of the main barriers to accurately forecast dramatic atmospheric events such as strongly baroclinic systems. Recent research has concentrated in eliminating, or at least, relaxing some of these hypothesis. Examples of these efforts are the work of Bartello & Mitchell [7] in non–separable covariance fields, and those of Gaspari & Cohn [59] in specification of non–homogeneous correlation fields.

Focusing our attention on the conventional procedure of OI, consider the geostrophic balance relations among the variables u_i, v_i and ϕ_i:

$$u_i = \alpha_i \frac{\partial \phi_i}{\partial \varphi_i} \tag{8.56a}$$

$$v_i = \beta_i \frac{\partial \phi_i}{\partial \lambda_i}\,, \tag{8.56b}$$

for,

$$\alpha_i = -\frac{1}{f_i a} \tag{8.57a}$$

$$\beta_i = \frac{1}{f_i a \cos\varphi_i}\,, \tag{8.57b}$$

where a is the earth radius and $f_i = 2\Omega \sin\varphi_i$ is the Coriolis parameter. Notice that the expressions above apply only to mid–latitudes.

Assuming that the ensemble mean represents an approximation of the real variables $\mathbf{w}^t$, after applying the ensemble mean operator to (8.56), we can write

$$u_i^t = \mathcal{E}\{u_i\} = \alpha_i \frac{\partial \mathcal{E}\{\phi_i\}}{\partial \varphi_i} \tag{8.58a}$$

$$v_i^t = \mathcal{E}\{v_i\} = \beta_i \frac{\partial \mathcal{E}\{\phi_i\}}{\partial \lambda_i} \tag{8.58b}$$

and subtracting (8.58) from (8.56) we obtain the geostrophic relation among the errors in the variables u, v and ϕ:

$$\tilde{u}_i = \alpha_i \frac{\partial \tilde{\phi}_i}{\partial \varphi_i} \tag{8.59a}$$

$$\tilde{v}_i = \beta_i \frac{\partial \tilde{\phi}_i}{\partial \lambda_i}\,. \tag{8.59b}$$

From these relations it follows that

$$S_{ij}^{f|u\phi} = \mathcal{E}\{\tilde{u}_i\tilde{\phi}_j\} = \alpha_i \frac{\partial}{\partial \varphi_i}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60a}$$

$$S_{ij}^{f|\phi u} = \mathcal{E}\{\tilde{\phi}_i\tilde{u}_j\} = \alpha_j \frac{\partial}{\partial \varphi_j}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60b}$$

$$S_{ij}^{f|v\phi} = \mathcal{E}\{\tilde{v}_i\tilde{\phi}_j\} = \beta_i \frac{\partial}{\partial \lambda_i}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60c}$$

$$S_{ij}^{f|\phi v} = \mathcal{E}\{\tilde{\phi}_i\tilde{v}_j\} = \beta_j \frac{\partial}{\partial \lambda_j}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60d}$$

$$S_{ij}^{f|uv} = \mathcal{E}\{\tilde{u}_i\tilde{v}_j\} = \alpha_i\beta_j \frac{\partial}{\partial \varphi_i \partial \lambda_j}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60e}$$

$$S_{ij}^{f|vu} = \mathcal{E}\{\tilde{v}_i\tilde{u}_j\} = \alpha_j\beta_i \frac{\partial}{\partial \lambda_i \partial \varphi_j}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60f}$$

$$S_{ij}^{f|uu} = \mathcal{E}\{\tilde{u}_i\tilde{u}_j\} = \alpha_i\alpha_j \frac{\partial}{\partial \varphi_i \partial \varphi_j}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60g}$$

$$S_{ij}^{f|vv} = \mathcal{E}\{\tilde{v}_i\tilde{v}_j\} = \beta_i\beta_j \frac{\partial}{\partial \lambda_i \partial \lambda_j}\mathcal{E}\{\tilde{\phi}_i\tilde{\phi}_j\}\,, \tag{8.60h}$$

where all the covariances are written as a function of the error covariance function ϕ–ϕ given in (8.52). The complete forecast error covariance matrix can be written symbolically as

$$\mathbf{S}_{ij}^{f} = \begin{pmatrix} G_i^u \\ G_i^v \\ 1 \end{pmatrix} \left[\begin{pmatrix} G_j^u \\ G_j^v \\ 1 \end{pmatrix} S_{ij}^{f|\phi\phi} \right]^T \tag{8.61}$$

where G_m^u and G_m^v are differential operators defined as

$$G_m^u \equiv \alpha_m \frac{\partial}{\partial \varphi_m} \tag{8.62a}$$

$$G_m^v \equiv \beta_m \frac{\partial}{\partial \lambda_m} \tag{8.62b}$$

for $m = i, j$.

By means of a limiting procedure, it is possible to show (see Cohn & Morone [32]) that the forecast error standard deviations for the winds are given by

$$\sigma_i^u = \sigma_i^\phi \, |\alpha_i| \left[\lim_{p_j \to p_i} \frac{\partial \log C^{\phi\phi}}{\partial \varphi_i \partial \varphi_j} \right]^{1/2} , \tag{8.63a}$$

$$\sigma_i^v = \sigma_i^\phi \, |\beta_i| \left[\lim_{p_j \to p_i} \frac{\partial \log C^{\phi\phi}}{\partial \lambda_i \partial \lambda_j} \right]^{1/2} , \tag{8.63b}$$

as a function of the geopotential height (co)variance at a point.

Consider a simple example the case of specifying the matrix $\mathbf{S}_k^f$ for on a plane atmosphere. This eliminates the need to specify the vertical correlations $V^{\phi\phi}$ in (8.54). Therefore, the state vector $\mathbf{w}^{f,a}$ in consideration is similar to that considered in the previous section, and in the end of the previous lecture, for the shallow water system of equations. To simplify the problem further, we treat here the case of a β–plane, with latitude and longitude represented by the dependent variables x and y, respectively. In this case, we notice that the constants α_i and β_i defined above are substituted by

$$\alpha_i = -\frac{1}{f_i} , \tag{8.64a}$$

$$\beta_i = \frac{1}{f_i} , \tag{8.64b}$$

where now $f_i = f_0 + \beta y_i$ is the Coriolis parameter. In practice, the state vector is treated on a grid, and therefore the derivatives seen above should be interpreted as finite differences.

A common model for the geopotential–geopotential correlation function is the Gaussian model, that is,

$$C_{ij}^{\phi\phi} = \exp\left(-\frac{b}{2} s_{ij}^2 \right) , \tag{8.65}$$

where, s_{ij} now is the distance between two points, (x_i, y_i) and (x_j, y_j) on the plane,

$$s_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 , \tag{8.66}$$

and b is an empirical constant proportional to the inverse of the decorrelation distance. Therefore, the derivative of $C^{\phi\phi}$ with respect to the variable ξ_k can be written as:

$$\frac{\partial C_{ij}^{\phi\phi}}{\partial \xi_k} = -\frac{b}{2} C_{ij}^{\phi\phi} \frac{\partial s_{ij}^2}{\partial \xi_k} \tag{8.67}$$

where ξ represents either x or y, and k represents either i or j. It is clear that, according to the definition (2.52), the covariance $\mathbf{S}^{f|\phi\phi}$ represents an isotropic field (consequently, homogeneous).

Substituting the expression for the distance s_{ij}^2 in (8.67), and expressions (8.60) we have

$$S_{ij}^{f|u\phi} \, / \, S_{ij}^{f|\phi\phi} = -\alpha_i \, (y_i - y_j) \, b , \tag{8.68a}$$

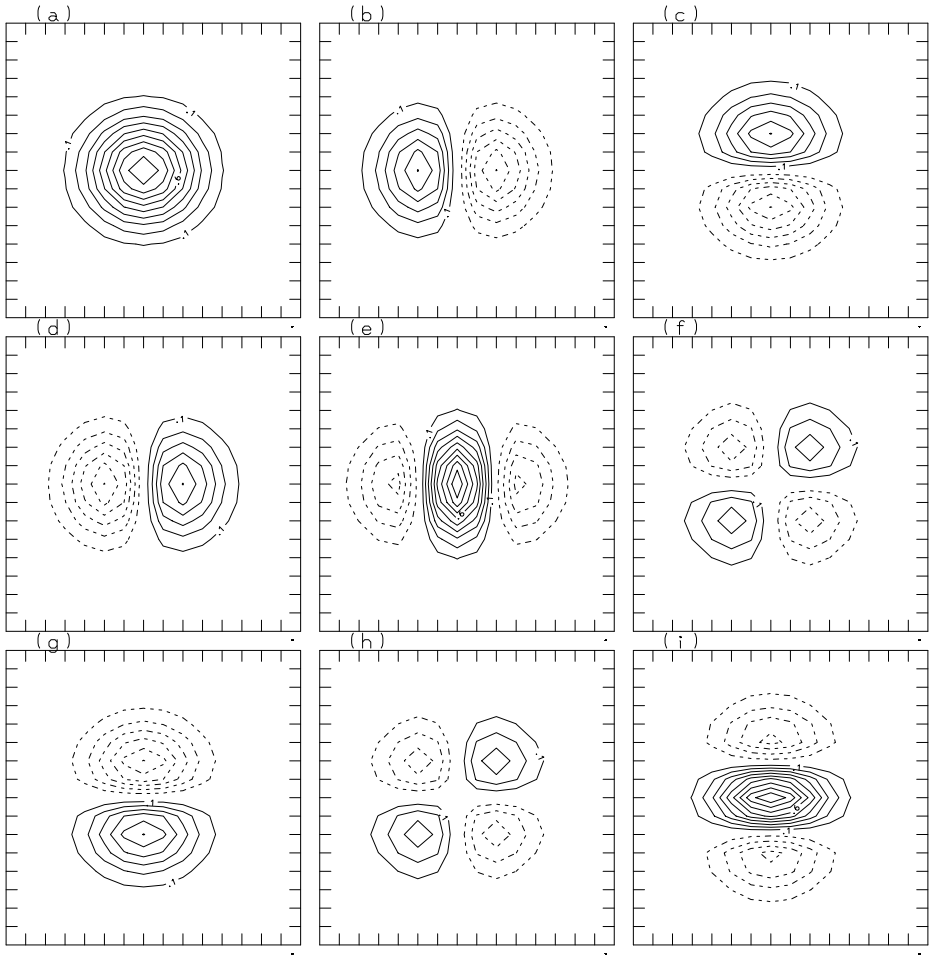

Figure 8.1: Point–correlations at the center of a square domain using the geostrophic balance relation; the equations obtained in this section: (a) ϕ–ϕ; (b) ϕ–v; (c) ϕ–u; (d) v–ϕ; (e) v–v; (f) v–u; (g) u–ϕ; (h) u–v; and (i) u–u.

$$S_{ij}^{f|\phi u} / S_{ij}^{f|\phi\phi} = \alpha_j (y_i - y_j) b , \tag{8.68b}$$

$$S_{ij}^{f|v\phi} / S_{ij}^{f|\phi\phi} = -\beta_i (x_i - x_j) b , \tag{8.68c}$$

$$S_{ij}^{f|\phi v} / S_{ij}^{f|\phi\phi} = \beta_j (x_i - x_j) b , \tag{8.68d}$$

$$S_{ij}^{f|uv} / S_{ij}^{f|\phi\phi} = \alpha_i \beta_j (x_i - x_j) (y_i - y_j) b^2 , \tag{8.68e}$$

$$S_{ij}^{f|vu} / S_{ij}^{f|\phi\phi} = \alpha_j \beta_i (x_i - x_j) (y_i - y_j) b^2 , \tag{8.68f}$$

$$S_{ij}^{f|uu} / S_{ij}^{f|\phi\phi} = \alpha_i \alpha_j \left[1 - b(x_i - x_j)^2\right] b , \tag{8.68g}$$

$$S_{ij}^{f|vv} / S_{ij}^{f|\phi\phi} = \beta_i \beta_j \left[1 - b(y_i - y_j)^2\right] b . \tag{8.68h}$$

Also, substituting (8.65) in the expressions (8.63), the standard deviation for the forecast errors in the winds becomes

$$\sigma_i^u = \sigma_i^\phi \sqrt{b} |\alpha_i| , \tag{8.69a}$$

$$\sigma_i^v = \sigma_i^\phi \sqrt{b} |\alpha_i| . \tag{8.69b}$$

All the quantities are now written as functions of the standard deviation of the forecast errors in geopotential heights. The correlation formulas obtained above are shown on Fig. 8.1 for a square cartesian plane, at its mid–point.

Exercises

1. Based on the shallow–water system considered in this chapter, while we studied the initialization problem, show that the quantity corresponding to the quasi–geostrophic potential vorticity:

   $$q = \nabla^2 \psi + f_0 - \frac{f_0}{\Phi} \phi$$

   is conserved. Furthermore, using the Fourier transform introduced in (7.34), show that we can write

   $$\hat{q} = k \hat{\psi} + \sqrt{k} \hat{\phi}$$

   (Note: The quantity q above refers to the complete field, that is, basic state plus perturbations.)

2. Consider the shallow–water equations in Exercise 7.2. Assume these equations are applied to a doubling periodic channel; using Fourier transform for u', v' e ϕ', obtain the matrix $\hat{\mathbf{L}}$ for the corresponding system.

3. Balgovind et al. [6] proposed a simple model to describe the spatial structure of forecast geopotential error fields for time scales of one to two days. Based on barotropic potential vorticity conservation equation, these researchers arrived to the following stochastic equation for the errors $\tilde{\phi} = \tilde{\phi}(x, y)$ in the geopotential field, in a tangent plane:

   $$L\tilde{\phi} = F(x, y)$$

   where the operator L is defined as

   $$L(x, y) \equiv (\nabla^2 - \alpha^2)$$

and ∇ is the Laplacian in two Cartesian dimensions x–y, x and y representing longitude and latitude, respectively, and $\alpha > 0$ being a constant related with the atmosphere depth. The term F is the stochastic forcing, with known statistics:

$$\begin{aligned} \mathcal{E}\{F(\mathbf{s})\} &= 0 \\ \mathcal{E}\{F(\mathbf{s}_1, \mathbf{s}_2)\} &= \sigma^2 \delta(\mathbf{s}_1 - \mathbf{s}_2) \end{aligned}$$

and represents the uncertainties in the model. Here, $\mathbf{s} = (x, y)$, $\mathbf{s}_1$, $\mathbf{s}_2$ are two points in the x-y plane, and σ is the noise F variance, assumed known. Moreover, define the error covariance function $P(\mathbf{s}_1, \mathbf{s}_2)$ as

$$P(\mathbf{s}_1, \mathbf{s}_2) \equiv \mathcal{E}\{\tilde{\phi}(\mathbf{s}_1)\tilde{\phi}(\mathbf{s}_2)\}\,.$$

Hence, show that:

(a) The equation for the spatial correlation structure ρ is

$$L_1 L_2 \rho(\mathbf{s}_1, \mathbf{s}_2) = \sigma \delta(\mathbf{s}_1 - \mathbf{s}_2)$$

For that, assume that the geopotential error variance field is identical to the forcing variance field F.

(b) Considering the unidimensional case, that is, when $\mathbf{s} \to x$, and assuming the correlation field is homogeneous, the correlation function ρ satisfies the following equation:

$$\left(\frac{d^2}{ds^2} - \alpha^2\right)^2 \rho(s) = \delta(s)$$

where $s \equiv |x_1 - x_2|$.

(c) The solution to the equation in the previous item can be written as:

$$\rho(x_1, x_2) = \rho(x_1 - x_2) = \frac{1}{4\alpha^3}\left(1 + \alpha|x_1 - x_2|\right) e^{-\alpha|x_1 - x_2|}$$

(Hint: Use symmetric Fourier transform.) This model for error correlations is sometimes called second order autoregressive, a more complicated version of which is utilized in some operational OI systems.