Tellus (1988), *40A*, 81–95

Short-range evolution of small perturbations in a barotropic model

By JEAN-FRANÇOIS LACARRA and OLIVIER TALAGRAND, *Laboratoire de Météorologie Dynamique, 24, rue Lhomond, 75231 Paris Cedex 05, France*

(Manuscript received 14 August 1986; in final form 27 April 1987)

ABSTRACT

The short-range evolution of small initial errors is numerically investigated with an *f*-plane shallow-water model. It is shown that this evolution can be approximated by a linearized model for meteorologically realistic situations, and for ranges of up to about 48 hours. The results are consistent with a description of the slow manifold as an attracting set along which the dynamics of the flow is dominated by an instability process. As a consequence of the relatively large time scale for the meteorologically significant components of the flow, the linear model valid for short periods can be further simplified to a constant coefficient model describing only the evolution of the large-scale components of the error. The possible implications of this result for the improvement of assimilation procedures are briefly discussed.

1. Introduction

An element still missing in numerical weather prediction is an estimate of the amplitude of the prediction error. The growth of error over periods of a few days or more, which is closely associated with the spectral interactions occurring in quasi-geostrophic turbulent flow, has been the subject of numerous studies, and there is now a wide consensus that the details of atmospheric flow cannot be deterministically predicted beyond a range of about 10 days. Less attention has been given, at least so far, to the growth of forecast error at short ranges, extending from a few hours to one or two days. From a theoretical point of view, error growth is related to the instabilities which develop in the atmospheric flow, and this by itself would justify a detailed study of the growth observed in operational forecasting model. From a practical point of view, a better knowledge of the error growth would help in assigning confidence limits to short-range numerical forecasts. These limits would depend, not only on the intrinsic quality of the model, but also on the accuracy with which the initial state of the forecast is known, and on the particularities of this initial state, such as for instance its closeness to conditions in which instabilities can occur. One methodical approach which has been proposed so far for establishing confidence limits on individual forecasts is Monte-Carlo prediction, in which the forecasting model is integrated from different initial conditions chosen so as to provide a fair representation of the uncertainty on the actual initial state of the atmospheric flow. The dispersion of the corresponding forecasts provides a direct estimation of the amplitude of the forecast error. However the numerical cost of this approach is obviously high and has so far forbidden its implementation in operational prediction. Simplified algorithms have been proposed for achieving the same basic goal at a lower cost (Leith, 1974; Hoffman and Kalnay, 1983; Thompson, 1986). It seems probable that techniques for assigning confidence limits to numerical forecasts will develop in the coming years. In this context, prior knowledge on the statistical properties of the growth of forecast error, and on the dependence of that growth in the current situation of the flow will certainly be useful.

Another domain for which the same type of

knowledge would be very useful is assimilation of meteorological observations. In present operational practice, assimilation is implemented by applying successive updates, as new data become available, on one integration of the model carried out over the time interval within which observations have been performed. At each successive update, or *analysis*, the relative weights to be given to the current forecast on the one hand and to the available observations on the other, should depend on the relative quality of the forecast and on the observations. *Optimal analysis*, which is now the most commonly used assimilation method in major meteorological forecasting services, provides a systematic way for estimating these weights once the statistical variances and covariances of the forecast and observation errors are known. Kalman-Bucy filtering (e.g., Ghil et al., 1981), which generalizes optimal analysis to a dynamical system whose temporal evolution is governed by a linear system of differential equations, relates the statistics of forecast errors to the governing equations. In present practice however, the connection between the dynamics of the assimilating model and the analysis itself is rather loose. The growth of forecast errors is described by extremely simple empirical laws, which ignore most of the dynamics contained in the numerical model, in particular specific instability processes. It is probable that the quality of assimilation, like the quality of short-range forecasts, could be improved by using a more accurate description of the statistics of forecast error.

The primary incentive for the work described in this article was precisely the improvement of assimilation procedures. One specific question which has been studied is the following: can the growth of forecast error, in the conditions which occur in the practice of data assimilation, be satisfactorily described over short ranges (up to one or two days) by a linear system of differential equations? If this is the case, the theory of Kalman-Bucy filtering will be applicable, through use of the local appropriate linear system at every stage of the process, to the problem of assimilation of meteorological data with nonlinear models. The validity of that approach, called the *extended Kalman filtering* (Jazwinski, 1970; Ghil et al., 1982) has so far not been checked in the meteorological context. Another possible approach for data assimilation uses variational techniques (Le Dimet and Talagrand, 1986). In this context also, linearity of the evolution of the short range forecast error can be useful, in that it guarantees that the final analysis is to be legitimately considered as the conditional expectation of the current state of the atmospheric flow, knowing all the available observations (including previous forecasts).

It is known from the theory of differential equations that a small enough perturbation imposed on a solution of a dynamical system will evolve linearly over a time interval whose length will depend on the properties of the system. The corresponding linear system is called the *tangent linear system* in the vicinity of the solution under consideration, and is obtained by differentiating the equations describing the basic evolution of the system in the vicinity of that solution (e.g., Brauer and Nohel, 1969, p. 135–142). The question considered here can therefore be stated as follows: is the amplitude of forecast errors in data assimilation small enough so that its temporal evolution can be described, between successive updates, by the linear system tangent to the model? Another specific point considered in this article is the geostrophic or nongeostrophic character of the forecast error. In present operational analysis procedures, the forecast error is assumed to be geostrophic. This assumption has the advantage that the corrections added to the forecast are also geostrophic, and that only a small amount of gravity perturbations is in consequence introduced by the analysis, but it has apparently been checked directly only in rather special cases (see, e.g., Daley, 1980, appendix).

The general procedure followed in this study is outlined in Section 2. We describe the model used and its main properties in Section 3 and the numerical experiments in Section 4. Results are presented in Section 5 and discussed in Section 6. We show in Section 7 that, as a result of certain special features of the equations, we can simplify further the tangent linear model for the short-range evolution of the errors. Conclusions follow in Section 8.

2. Experimental procedure

The principle of the experimental procedure used in this article is extremely simple. A model

integration has first been performed, producing what will be called the *reference solution*. Various perturbations have then been added to the initial state of this reference solution, and the ensuing evolutions of the perturbations have been studied in detail, in particular in terms of their linear and geostrophic character. For the sake of simplicity, and also because errors in the initial conditions have in practice a much larger impact on short range forecasts than model errors, the experiments are of the "identical twin" type, i.e., a single model has been used and integrations performed with this unique model have been compared.

Although this procedure is numerically very similar to Monte Carlo forecasting, the results are interpreted in a different way. In the Monte Carlo approach, which is intrinsically statistical, each individual trajectory must be associated with some weight representing the probability of the corresponding initial state. The averages obtained from a number of different trajectories can then be interpreted as mathematical expectations. An example is the study made by Lorenz (1982), based on the set obtained by taking the differences between a number of one-day forecasts performed at the European Centre for Medium Range Weather Forecasts and the corresponding analyses. On the contrary, the point of view taken here is deterministic. The statistical mean state is replaced by the reference state considered as being the true state while each of the other initial states is interpreted as the sum of that true state plus an error or perturbation. The quantities in which we will be interested will not be ensemble averages, but quantities associated to individual trajectories.

3. The model

The numerical model used for the experiments described below is governed by the shallow-water equations on a plane rotating with constant angular velocity:

$$\begin{cases} \dfrac{\partial \vec{V}}{\partial t} + (f+\zeta)\vec{z} \times \vec{V} + \vec{\nabla}\left(\varphi + \dfrac{V^2}{2}\right) + \nu\Lambda^4 \vec{V} = 0, \\ \dfrac{\partial \varphi}{\partial t} + \vec{\nabla}\cdot(\varphi \vec{V}) + \nu\Lambda^4 \varphi = 0, \end{cases} \tag{1}$$

where

$\vec{V}$: horizontal velocity
φ: free surface potential
f: vorticity of the basic rotation (Coriolis parameter)
$\vec{z}$: vertical unit vector
ζ: relative vorticity in the rotating frame
ν: diffusion coefficient

It is well-known that two different physical interpretations can be given to eqs. (1). In the first interpretation they describe the evolution of a three-dimensional incompressible fluid with horizontal velocity independent of the vertical coordinate, φ being the potential of the free surface. In the second interpretation, they describe the evolution of a two-dimensional compressible fluid, φ representing then the density of the fluid.

The model is pseudo-spectral, which means that the different variables are expanded in truncated Fourier series, the spatial derivatives being computed in Fourier space while quadratic advective terms are evaluated in physical space by use of fast Fourier transforms.

The dissipation terms in (1) are intended primarily to represent the effect of the unresolved scales of motion. The presence of a dissipative term in the height equation is justified by the fact that height is closely linked to vertical velocity in the three-dimensional interpretation of the shallow-water equations. Dissipation represented as above by a fourth-power Laplacian is highly selective in wave number, and has been successfully used in the case of incompressible two-dimensional turbulence for parameterization of subgrid scale effects (Basdevant et al., 1981).

While the total energy and the total potential enstrophy of the flow are two invariants of the undiscretized eqs. (1), spatial discretization resulting from truncation in Fourier space can conserve only energy and only at the expense of aliasing errors, which act as large scale sources balancing the exchanges of energy with subgrid scales. This is the case in the model used here. As for the effect of temporal discretization, numerical tests have shown that the loss of energy remain small (less than one percent in relative value for the loss of energy) in integrations of a few days which are of interest for us.

Eqs. (1), when linearized about a state of rest, possess for each Fourier component the three classical eigenmodes: one low-frequency Rossby mode (stationary in the case considered here of a constant basic rotation) and two high-frequency inertia-gravity modes. When the amplitude of the fields become large, the linear space spanned by the Rossby modes is no longer invariant for the evolution equations, and the concept of a *slow manifold* has been introduced by Leith (1980) in order to account for the fact that atmospheric flow mostly consists of low-frequency motions. This slow manifold is tangent at the state of rest to the linear space of Rossby modes and is an invariant of the evolution equations, i.e., the model trajectory defined by an initial state lying on the slow manifold will itself lie on that manifold and will not in addition exhibit fast oscillations. From a theoretical point of view, the existence of such a manifold is a rather controversial theoretical question (Errico, 1982; Vautard and Legras, 1986). In practice, the subset of model states obtained by applying the normal mode initialization algorithm of Machenhauer (1977) approximately satisfies the properties of the slow manifold for integrations of a few days. It is only an approximation since numerical integration shows that this set is not exactly invariant, but fast oscillations are removed as can be seen from Fig. 1 which represents the temporal evolution of the free surface potential deviation from the mean at a given grid point in two integrations performed respectively with and without a normal mode initialization.

For the experiments described in this article, the model domain D was taken to be a square with side $L = 6400$ km and with periodic boundary conditions, the Coriolis parameter f was taken equal to 10^{-4} s^{-1} and the mean free surface potential Φ_0 to 10^5 m^2 s^{-2}. The diffusion coefficient ν depends on the resolution in such a way as to damp the largest wavenumber components with an e-folding time of about 4 hours. Two resolutions have been used, corresponding to a 32×32 and a 64×64 point grid, respectively. The grid mesh is therefore 200 km in the first case and 100 km in the second. The main reason for using two different resolutions was to study the effect of truncation and dissipation.

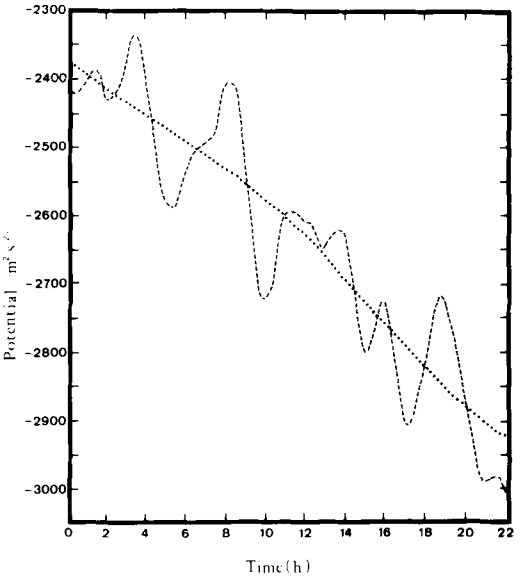


Fig. 1. Temporal evolution of the deviation from the mean value of the free surface potential at one grid point in: (a) an integration performed from an exactly geostrophic initial state (dashed line) and (b) an integration performed from a state obtained by performing a nonlinear normal mode initialization on the initial state of (a) (dotted line).

4. Experiments

4.1. Reference solution

The reference solution has been determined in such a way as to possess a very low level of inertia-gravity oscillations, to have a small Rossby number ($\mathrm{Ro} \leqslant 0.2$) and to consist mostly of statistically equilibrated large-scale motions. The model was first integrated for one month, starting from a large-scale Rossby wave to which a wide-band spectrum random noise had been added. A nonlinear normal mode initialization was then performed, followed by a one-week integration and a new initialization, and finally by the reference integration itself, whose length is equal to two days.

The energy spectrum changed very rapidly in the first days of the one-month preliminary integration, but then quickly stabilized. The stabilized Rossby energy spectrum has a slope comprised between -3.5 and -4. The first of the two initializations performed in the preliminary integration eliminated the fast oscillations which were still present in the model fields, but the

second initialization produced almost no change. This observation is in agreement with previous results about normal mode initialization (e.g., Daley, 1981).

The energy spectrum and the streamfunction of the reference integration, averaged over the two days, are shown in Fig. 2. The mean quadratic velocity U is 20 m s^{-1}, the Rossby number can be estimated to lie between 0.15 and 0.2, and the Froude number ($U/\sqrt{\Phi_0}$) is about 0.1.

4.2. *Perturbations*

Three types of parameters have been taken into account in the choice of the perturbations to be added to the reference state:

the scale of the perturbation;

its physical type, i.e., whether it consists of Rossby or inertia-gravity motions;

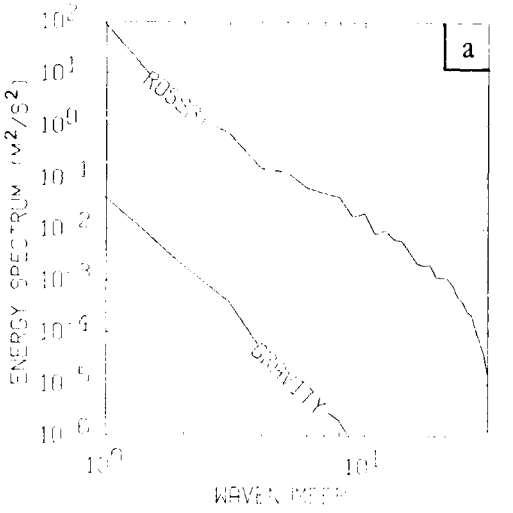
its amplitude.

As concerns the scale, we have considered three different classes of perturbations, sketched in Fig. 3. All three have energy spectra of the same shape as the basic flow, but extend respectively over the small scales (class 1), over the small and intermediate scales (class 2), and over the complete range of scales resolved by the model (class 3). Choices 1 and 2 are consistent with what happens in practical situations, where it is legitimate to consider that the real flow is exactly known in the large scales, but only statistically known in the small scales. Class 3 was used as an element of comparison and also because the possibility exists that the inhomogeneous spatial distribution of conventional and space borne observing systems might create systematic errors in large scales.

For each of these three classes, we have taken three different physical types, namely a pure Rossby perturbation, a pure inertia-gravity perturbation and a mixed perturbation, in which energy was equally distributed between Rossby and gravity modes. Nine directions have thus been defined in phase space, along which the amplitudes of the perturbations have been varied.

The amplitude λ of a perturbation $(\delta\vec{V}, \delta\varphi)$ has been defined as the square root of the total quadratic energy of the perturbation per unit mass of fluid, i.e., as:

$$\lambda = \left[\frac{1}{2L^2}\int_D \left(\delta V^2 + \frac{\delta\varphi^2}{\Phi_0}\right) \mathrm{d}x\,\mathrm{d}y\right]^{\frac{1}{2}}. \qquad (2a)$$



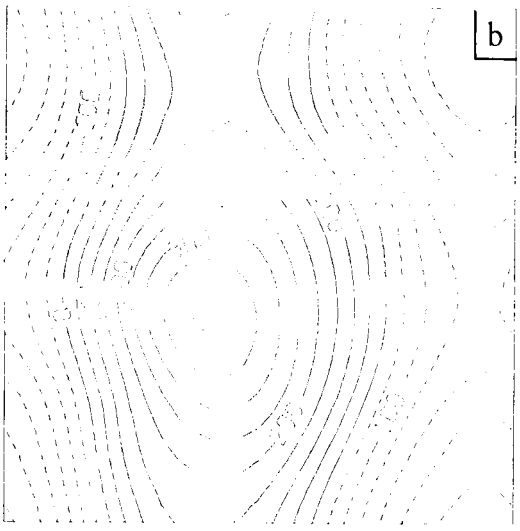


Fig. 2. (a) Energy spectrum (m^2/s^2) as a function of wavenumber (wavenumber unit: $2\pi/L$, L being the side of the domain) and (b) streamfunction (10^5 m^2/s) of the reference state (isoline interval $5\cdot10^6$ m^2/s).

Typical values for the uncertainty on the state of the atmospheric flow at 500 mb are in practice $\delta\varphi \sim 150$ m^2 s^{-2} and $|\delta\vec{V}| \sim 4$ m s^{-1}, which lead to $\lambda \sim 3$ m s^{-1}. In our experiments, five values of λ, in geometric progression with ratio $r = \frac{1}{4}$, have

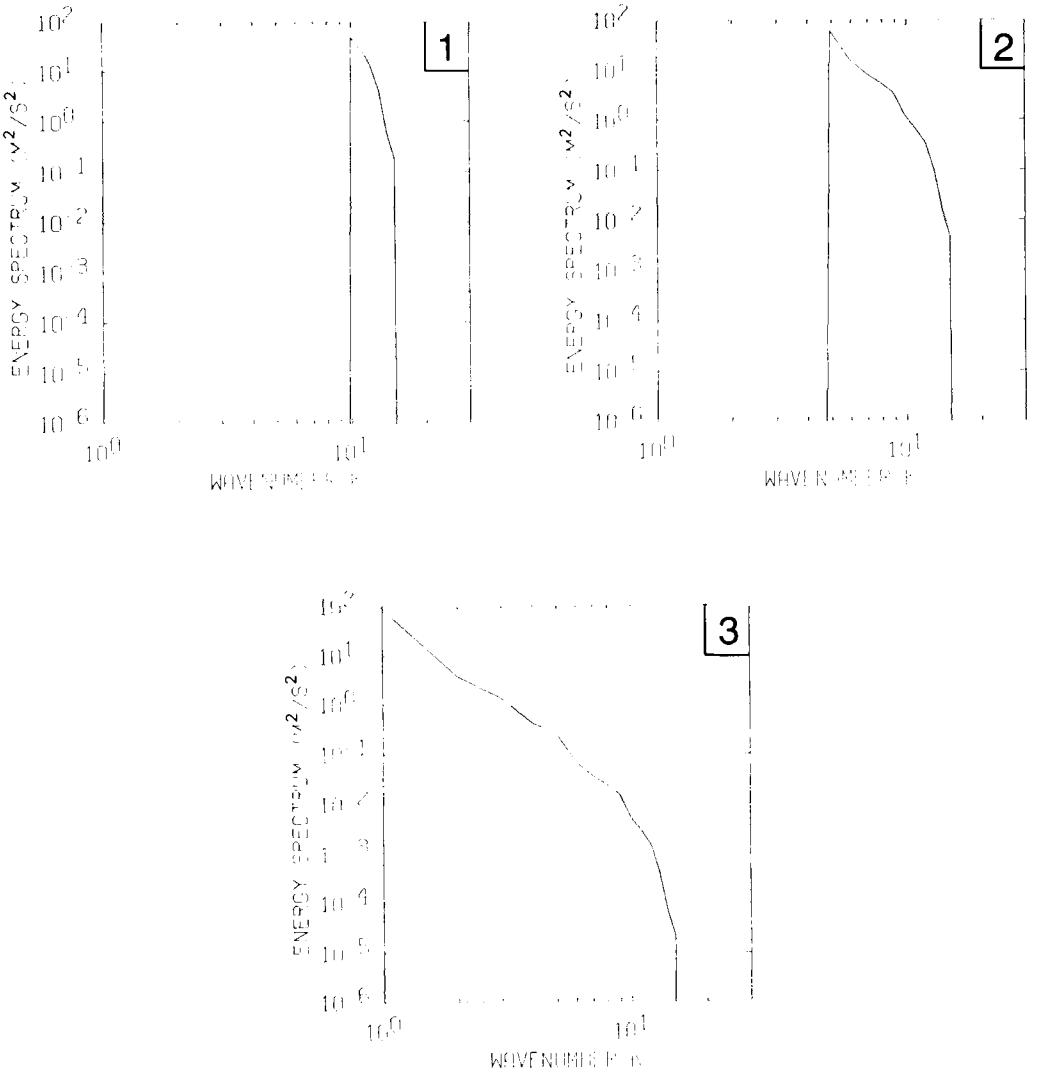


Fig. 3. Shapes of the initial spectra corresponding to the three classes of perturbations (1, 2, 3) described in the text.

been taken along each of the nine perturbation directions:

$$\lambda_i = r^i \lambda_0, \quad i = 0, 1, 2, 3, 4, \tag{2b}$$

with $\lambda_0 = 10\ \mathrm{m\,s^{-1}}$. The main reason for varying the perturbation energy was of course to check the linearity of the response for different amplitudes of the initial perturbations.

4.3. *Linearity*

Let us denote by F_T the mapping which to any initial state $X(0) = (\vec{V}(0), \varphi(0))$ associates the state $X(T) = (\vec{V}(T), \varphi(T))$ obtained at time T by integration of the basic eqs. (1). F_T is often called the *resolvent* of (1) between times 0 and T. What we will be interested in here is determining the

response $\delta X(T)$ resulting from an initial perturbation $\delta X(0)$ imposed on $X(0)$. If system (1) is written in the symbolic form $\mathrm{d}(X)/\mathrm{d}t = F(X)$, a perturbed solution $(X(t) + \delta X(t))$ satisfies the equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}(X(t) + \delta X(t)) = F(X(t) + \delta X(t))$$

$$= F(X(t)) + \frac{\partial F}{\partial X}(X(t))\,\delta X(t) + \mathrm{O}(\delta X(t)),$$

which, upon retaining only the first order terms in δX, becomes:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\delta X) = \frac{\partial F}{\partial X}(X(t))\,\delta X. \tag{3}$$

This linear system of equations is called the *tangent linear system* of (1) in the vicinity of the particular solution $X(t)$. It describes the temporal evolution of the perturbation $\delta X(t)$, to first order with respect to the initial perturbation $\delta X(0)$. The resolvent DF_T of (3) is the differential of the resolvent F_T of (1). Expliciting $\delta X(t) = (\delta\vec{V}(t), \delta\varphi(t))$ the tangent linear system reads:

$$\begin{cases} \dfrac{\partial \delta\vec{V}}{\partial t} + (f + \zeta(t))\vec{z} \times \delta\vec{V} + \delta\zeta\vec{z} \times \vec{V}(t) \\ \quad + \vec{\nabla}(\delta\varphi + \vec{V}(t)\cdot\delta\vec{V}) + \nu\Lambda^4\,\delta\vec{V} = 0 \\ \dfrac{\partial \delta\varphi}{\partial t} + \vec{\nabla}\cdot(\delta\varphi\vec{V}(t) + \varphi(t)\,\delta\vec{V}) + \nu\Delta^4\,\delta\varphi = 0. \end{cases} \tag{4}$$

There exists one tangent linear system for each solution $X(t)$ of (1). Except in the very special case when $(\vec{V}(t), \varphi(t))$ is a stationary solution of (1), the coefficients of (4) will vary in time. These coefficients will normally also vary in space with the consequence that, contrary to what happens with linear systems whose coefficients are constant in space, different spectral components of $(\delta\vec{V}(t), \delta\varphi(t))$ will mutually interact in the temporal evolution described by (4).

The linear approximation represented by (3) deserves some additional explanations. Let us consider in phase space a straight line Λ going through $X(0)$, and parallel to some unit energy vector e. The image $F_T(\Lambda)$ by the resolvent F_T will not in general be a straight line. But the tangent to $F_T(\Lambda)$ at point $X(T)$ is precisely the image of Λ by the resolvent DF_T of the tangent linear system (3) (see Fig. 4). The problem of the linearity of the response to a perturbation parallel to e is precisely to determine for which initial energy λ the image by (1) of the perturbation λe will be well approximated by $\lambda DF_T(e)$. The question of the global linearity of the response is reduced to a set of one dimensional problems, namely the comparisons for various directions e of the exact final perturbation:

$$\Lambda_{e,T}(\lambda) = F_T(X(0) + \lambda e) - F_T(X(0))$$

and of the corresponding linear approximation $\lambda DF_T(e)$.

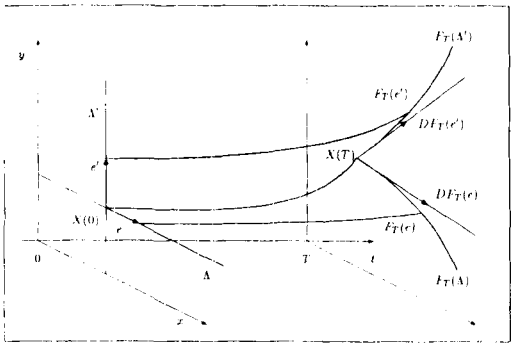


Fig. 4. Schematic representation of the temporal evolution of a dynamical system in a two-dimensional phase space. Two straight lines Λ and Λ' going through a given point $X(0)$, with unit vectors e and e' respectively, are transformed into curves $F_T(\Lambda)$ and $F_T(\Lambda')$ in the exact nonlinear evolution and into straight lines defined by $DF_T(e)$ and $DF_T(e')$ in the tangent linear evolution about the solution $X(t)$.

More generally one can write the Taylor expansion of $\Lambda_{e,T}(\lambda)$ in powers of λ:

$$\Lambda_{e,T}(\lambda) = \lambda e_1 + \frac{\lambda^2}{2} e_2 + \frac{\lambda^3}{6} e_3 + \cdots \tag{5}$$

where

$$e_i = D^{(i)} F_T(X(0))(\underbrace{e, e, \ldots}_{i}),$$

$D^{(i)}$ being a synthetic notation for the differential of order i with respect to the initial state $(\vec{V}(0), \varphi(0))$ at time 0. For a system discretized to k variables x_j, $D^{(i)}$ is made up of the $\binom{i+k-1}{k-1}$ partial derivatives of degree i

$$D^{i_1}_{x_1} \ldots D^{i_k}_{x_k} \quad \text{with } i_1 + \cdots + i_k = i.$$

The first order term corresponding to the linear approximation, our primary purpose in this article is to evaluate the relative magnitudes of the leading terms in (5). e_1 can be exactly computed by integration of system (4) between 0 and T (or more precisely, by integration of the linear equations obtained from the model discretized equations through the same process which led above from (1) to (4)). Exact computation of e_2 and of higher derivatives would also be possible in principle, although at a much higher numerical cost. In practice, finite difference estimates of e_1 and e_2 have been obtained as follows from the numerically determined differences $\Delta_{e,T}(\lambda_i)$

$$D1(i) = \frac{\Delta_{e,T}(\lambda_{i+1}) - r^2 \Delta_{e,T}(\lambda_i)}{(1-r)\lambda_{i+1}} = e_1 + O(\lambda_i^2),$$

$$D2(i) = 2\,\frac{\Delta_{e,T}(\lambda_{i+1}) - r\,\Delta_{e,T}(\lambda_i)}{r(r-1)\lambda_i^2} = e_2 + O(\lambda_i),$$

with $i = 0, 1, 2, 3$ and $r = \frac{1}{4}$.

These approximations are not the only possible ones. In particular, one could approximate e_2 by using λ_{i+1}, λ_i, λ_{i-1}. It can be shown that this would basically be equivalent to the above expression. The accuracy of the *a priori* best estimates $D1(3)$ and $D2(3)$ has been checked by comparing $\Delta_{e,T}(\lambda)$ with the quadratic expression:

$$\Delta Q_{e,T}(\lambda) = \lambda \cdot D1(3) + \frac{\lambda^2}{2} D2(3).$$

It can also be checked by computing

$$(D2(2) - D2(3)) = \tfrac{2}{3} r^2 (1 - r^2)\lambda_0 e_3 + O(\lambda),$$

which gives an estimate of the relative size of the third-order term in (5) and defines the domain of validity of the second-order approximation.

As for the domain of validity of the first order approximation, the ratio $R = 2|D1(3)|/|D2(3)|$ gives an upper bound of the values of λ leading to a linear evolution since for $\lambda = R$ the first and second order terms in (5) are of the same magnitude. For each individual wavenumber k we will use the corresponding spectral ratio $R(k) = 2|e_1(k)|/|e_2(k)|$.

5. Results

5.1. *Quadratic approximation*

For the perturbation amplitudes which have been considered ($\lambda \leqslant 10\ \mathrm{m\,s^{-1}}$), the third-order term in (5) is always negligible in comparison with the second-order term for forecast ranges of 24 hours or less (less than 10% for large and medium scales perturbations). It becomes comparable to the second-order term at a range of about 48 hours, and then only in the smallest scales of the motion.

5.2. *Linear approximation*

For a forecast range of 24 hours, and for pure Rossby perturbations, R varies between 10 $\mathrm{m\,s^{-1}}$ for large-scale perturbations and 0.2 $\mathrm{m\,s^{-1}}$ for small-scale perturbations. For gravity perturbations, the value of R, which does not vary significantly with the scale of the perturbation, is about 15 $\mathrm{m\,s^{-1}}$.

Comparison of the spectra of e_2 and e_1 shows that, for all types of perturbations, the former is much flatter than the latter. The shape of $R(k) = 2|e_1(k)|/|e_2(k)|$ will then be similar to the shape of $e_1(k)$. Since the spectrum of the error $\Delta_{e,T}$ is the same as the spectrum of e_1 in the linear regime, it results that $R(k)$ will be maximum in the scales where the error is itself maximum. This is clearly visible in Fig. 5, relative to an intermediate scale perturbation, where the maximum

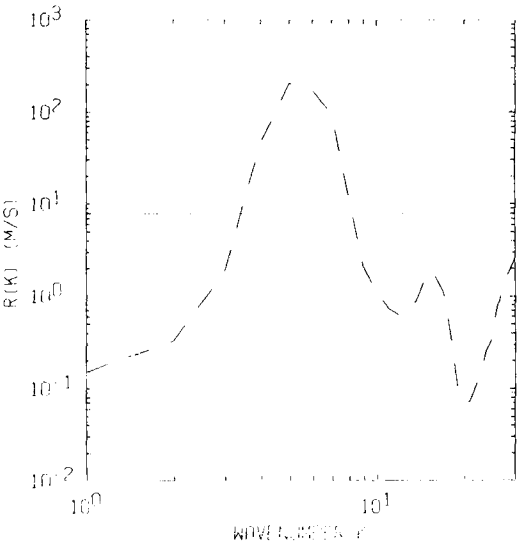


Fig. 5. Coefficient $R(k)$ (m/s) for a medium scale gravity perturbation (dashed line). The horizontal solid line represents the global coefficient R.

of $R(k)$ coincides with the maximum of the initial perturbation (see Fig. 3.2). This means that the error will behave more linearly in the scales where it is larger. This result, which may seem paradoxical, is due to the fact that the relative influence of the nonlinear terms, whose amplitude is almost independent of the wave number, is larger in the scales where the error is small.

5.3. Spectral evolution of the perturbations

Our numerical results show that a Rossby perturbation does not create gravity components and vice-versa. In particular, an initially geostrophic perturbation remains geostrophic. This fact, which is obvious for small perturbations added to a state of rest, was not *a priori* obvious for perturbations added to a fully developed meteorological flow.

As regards the energy growth, there is a crucial difference between the Rossby and gravity perturbations. Large-scale Rossby perturbations grow exponentially, with an amplitude doubling time of about two days. For intermediate-scale Rossby perturbations, there is initially a transfer of energy to larger scales which then grow at an exponential rate. Thus intermediate scale Rossby perturbations have the particular property of exciting larger scales. This is associated with a change in the shape of the perturbation spectrum, the relative importance of the larger scales tending to increase in the first day of the forecast. Inertia-gravity perturbations on the contrary tend to decrease at an exponential rate which is relatively small in large scales (one week halving time) and more rapid in small scales. The e-folding time of dissipation varies as k^{-8} with wavenumber k, and it is only in the smaller scales that dissipation can contribute to this decrease. In addition, inertia-gravity perturbations do not excite scales which were not originally excited, and there is no significant change in the shape of their spectrum over the time periods considered here. Keeping in mind what has been said about the spectral variations of $R(k)$, it is seen that it is in the large scales that the evolution of Rossby perturbations will be most linear, while for inertia-gravity perturbations the evolution will be most linear in the initially most excited scales.

A consequence of the fact that Rossby perturbations are amplified while gravity perturbations are not is that the former will tend to progressively dominate the latter. Thus, even though, as said above, there is no interaction between the two kinds of perturbations, any initial error will in fact tend to become more and more geostrophic. This is consistent with the assumption of geostrophy usually made on the forecast error in operational analysis. It is also consistent with results obtained by Daley (1980) on perturbations with only kinetic initial energy where it was seen that the growth of the error was concentrated in the Rossby component.

Finally, except for a more rapid decrease of the intermediate scale gravity perturbations with the low-resolution grid, no difference has been observed between the results obtained with the two grids. This more rapid decrease is due to the dissipation, which is more active at low resolution.

6. Interpretation

If, as our results strongly suggest, the evolution of the Rossby perturbations is independent of the evolution of the gravity perturbations, it should be possible to describe the former as the evolution of a two-dimensional incompressible flow. This is so because the interaction between Rossby modes conserves the linearized potential vorticity, which means that the equation for the streamfunction Ψ of a Rossby perturbation reads:

$$\frac{\partial((\nabla^2 - \lambda_D^2)\Psi)}{\partial t} + J(\Psi, \nabla^2\Psi) = 0, \tag{6}$$

where J is the Jacobian and λ_D the inverse of the Rossby deformation radius.

This is supported by the evolution of the streamfunction of a Rossby perturbation. Since the basic flow is almost meridional, the regions where the perturbation extracts kinetic energy from the basic flow are characterized by the condition $(\partial V/\partial x \cdot uv \leqslant 0)$, where V is the meridional component of the basic flow velocity and (u, v) the components of the perturbation velocity (e.g., Pedlosky, 1979, Subsection 7.3). It is seen in Fig. 6 that these regions (framed) coincide with the regions where the basic flow presents a strong velocity shear, i.e., with the saddle points of the reference streamfunction (see Fig. 2b). This is exactly what occurs in barotropic instability of a

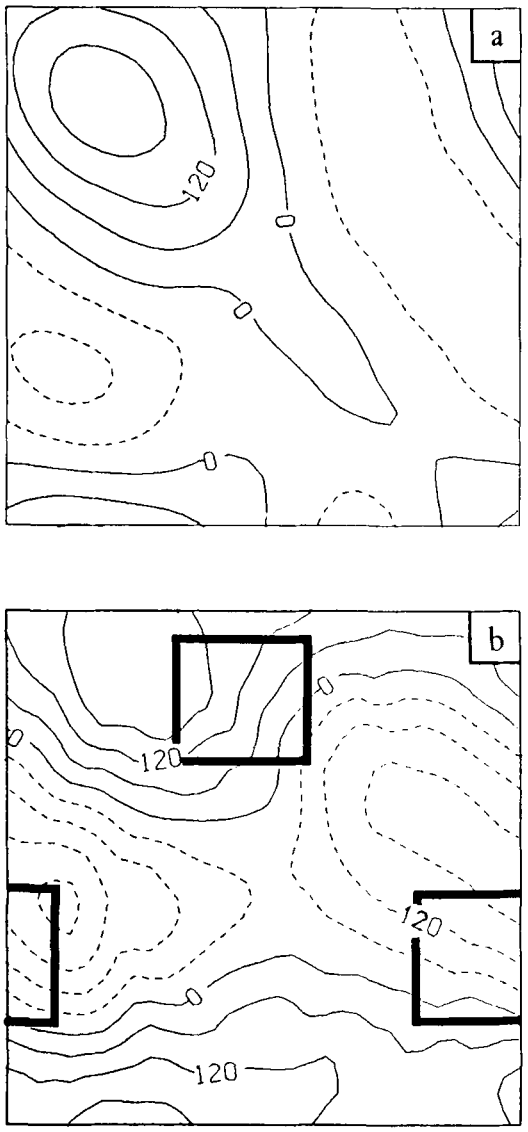


Fig. 6. Streamfunction of a large scale perturbation, initially (a), and after a one day integration (b) (arbitrary units). The isoline interval is the same in the two figures. Framed regions represent the part where the perturbation extracts energy from the basic flow.

two-dimensional incompressible flow. It is also consistent with the experiments of Balgovind et al. (1983) where the statistical evolution equation of the forecast error for a 24–36 hour range is dominated by the potential enstrophy conservation equation.

As for the behaviour of the inertia-gravity perturbations, it cannot of course be interpreted in the context of incompressible turbulence. When evaluating the importance of the various terms that contribute to the evolution of the gravity modes, we found that the leading term is the *wave term*, i.e., the term obtained by linearization around the state of rest. This is not surprising since the linear frequency of gravity modes is very high. In addition, the Froude number being small, the additional linear term resulting from the interaction with the basic flow will itself be small. It then results that, as when the basic state is the state of rest, gravity perturbations will produce fast oscillations. This is consistent with the logic of nonlinear normal mode initialization, for which any perturbation of the initialized state in the gravity mode direction gives rise to fast oscillation. Nonlinear normal mode initialization can be successful only if the velocity is not too large (e.g., Thaning, 1984), which precisely ensures that the linear wave term dominates the evolution of a gravity perturbation. This would probably not be the case if the basic state had a large Rossby number. All this explains why gravity perturbations behave more linearly than Rossby perturbations. It leads us to think that the evolution of the latter would be more linear in the presence of a β-effect, which produces a nonzero linear wave term in the corresponding evolution equation. It also explains why the interaction between the basic flow and gravity perturbations does not produce any transfer between different Fourier components: because of the fast oscillations of the gravity perturbations, the interaction term is on the average almost zero.

The progressive damping of high-frequency gravity motions is in agreement with the classical description of geostrophic adjustment and with the existence of an attracting slow manifold towards which the orbits of the system tend to converge in phase space. In addition, the fact that the Rossby components of intermediate and large scales are amplified without interacting with the gravity components strongly suggests that the structure of the phase flow in the vicinity of the reference solution is hyperbolic, i.e., that the slow is split asymptotically between contracting and expanding directions. If this is true of every trajectory in phase space, the theory of

dynamical systems (Eckmann and Ruelle, 1985) predicts that the local stable and unstable manifolds will be approximately tangent at the reference trajectory to the gravity modes and to the Rossby modes of intermediate and large scale respectively. The attractor being the union of the unstable manifolds, it will be tangent at each point to the linear space spanned by these Rossby modes.

In summary, the dynamics of Rossby perturbations is dominated by a linear instability process which transfers energy to large scales. This process is perfectly explained by the classical phenomenology of two-dimensional incompressible turbulence, and must be related in some way to the local unstable directions of the slow manifold. The residue obtained by subtracting the linear process from the whole map $\Delta_{e,T}$ results from highly nonlinear phenomena which saturate at a relatively small level.

7. Variability of the tangent linear system

Using the tangent linear system for describing the temporal evolution of the error (e.g., in an assimilation process) can be useful in practice only if this does not increase too much the required amount of computation. In this respect, it is important to investigate two particular points. (1) The sensitivity of the tangent linear system to the particular reference solution in the vicinity of which it is defined. (2) The influence of the temporal variations of that solution on the resolvent of the tangent linear system.

A very powerful tool for sensitivity studies which involve the integration of a numerical model is the *adjoint* of that model (Le Dimet and Talagrand, 1986). For a given reference solution, the adjoint system is a linear system of differential equations whose resolvent is the adjoint (in the sense of the theory of linear operators) of the resolvent of the tangent linear system (4). The adjoint equations provide an extremely efficient way for computing the gradient of a given output parameter of the model with respect to the input parameters such as, for instance, the initial conditions.

The adjoint system of (4) reads:

$$\begin{cases} -\dfrac{\partial \delta' \vec{V}}{\partial t} - (f+\zeta(t))\vec{z}\times\delta'\vec{V} \\ \quad -\vec{z}\times\vec{\nabla}(\vec{V}(t)\times\delta'\vec{V}) - (\vec{\nabla}\cdot\delta'\vec{V})\vec{V}(t) \\ \quad -\dfrac{\varphi(t)}{\Phi_0}\vec{\nabla}(\delta'\varphi) + \nu\Delta^4\delta'\vec{V} = 0, \\ -\dfrac{\partial \delta'\varphi}{\partial t} - (\vec{\nabla}\cdot\delta'\varphi)\cdot\vec{V}(t) - \Phi_0\vec{\nabla}\cdot(\delta'\vec{V}) \\ \quad +\nu\Delta^4\delta'\varphi = 0, \end{cases} \tag{7}$$

where, in order to avoid any confusion, primes (′) have been used for denoting the variables of this new system. The form of (7) depends on the definition of a scalar product on the space of perturbations $(\delta\vec{V}, \delta\varphi)$. The scalar product used here is the scalar product associated with the quadratic form (2a), which measures the total energy of the perturbation $(\delta\vec{V}, \delta\varphi)$. The interested reader can find in Talagrand and Courtier (1987) indications on how the adjoint system is obtained from the tangent linear system.

Because of its very nature, the adjoint system proceeds from the output of the model to its input, which means that it must be integrated *backwards* in time. This does not create any theoretical or numerical problem. In particular it is seen that dissipation terms in (7) remain dissipative for backward integration (this can be easily understood if one notes that a dissipated variable must have a small derivative with respect to the model's initial conditions).

For the adjoint model as for the linearized model, we are in effect dealing with the discretized form of eq. (7) whose resolvent is thus a finite dimensional operator which we can represent by a matrix M_T^*. This matrix is the adjoint of the resolvent matrix M_T of the discretized analog of the tangent linear system (4) for the scalar product (2a).

It is easy to see that the directions in which the resolvent of the tangent linear system is most sensitive to a change in the reference solution are the directions in which an initial error is most rapidly amplified in the linear approximation. In order to identify these directions, we have determined the initial perturbations which grow

most rapidly for a given forecasting range T. These perturbations maximize the amplification factor

$$A(e) = \frac{\langle M_T \cdot e, M_T \cdot e \rangle}{\langle e, e \rangle} = \frac{\langle S_T \cdot e, e \rangle}{\langle e, e \rangle},$$

where $S_T = M_T^* M_T$. Geometrically the equation $\langle S_T \cdot e, e \rangle = 1$ defines an hyperellipsoid the smallest axis of which maximizes A. It is well-known that this axis is associated with the largest eigenvalue of the symmetric operator S_T. The fastest growing perturbations are the eigenvectors of S_T associated with the largest eigenvalues.

The very form of the operator S_T shows that, for given e, the determination of $S_T \cdot e$ requires one integration of the direct model between 0 and T, followed by one backward integration of the adjoint model between T and 0. Following Urban (1985), we have used to so-called Lanczos algorithm based on the Householder reduction method for symmetric matrices (e.g., Parlett, 1980) in order to determine the dominant eigenvectors. This algorithm does not require the computation of the whole matrix S_T but only the vectors $S_T \cdot e$ for a few appropriately selected perturbations e's.

The results obtained are at first sight surprising. For very short ranges, the growth is larger than the growth corresponding to the exponential 48-hour doubling time found above (for a 12-hour forecast range, the dominant eigenvalue is equal to 5.2, which would correspond to a doubling time of about 10 hours for the error amplitude), but is also less than exponential (for a 24-hour forecast range, the dominant eigenvalue is only 14 rather than 27.04). This can be explained as follows. Suppose that the resolvent can be represented as the integration of a linear system with constant coefficients. The eigenvectors have then an exponential behaviour. If we consider the limit as the time goes to infinity of the vector that grows fastest, we obtain the eigenvector associated with the eigenvalue having the largest real part, that is the first Lyapunov vector. But if instead, we are interested in finite time growth, the fact that the eigenvectors are not necessarily orthogonal allows vectors other than eigenvectors to grow fastest. In more mathematical terms, the eigenvectors of the symmetric operator S_T are different from the eigenvectors of the direct resolvent DF_T because the transposed resolvent DF_T^* does not commute with DF_T. We show in the appendix an example of a linear differential system of dimension 2 with constant coefficients having such a behaviour.

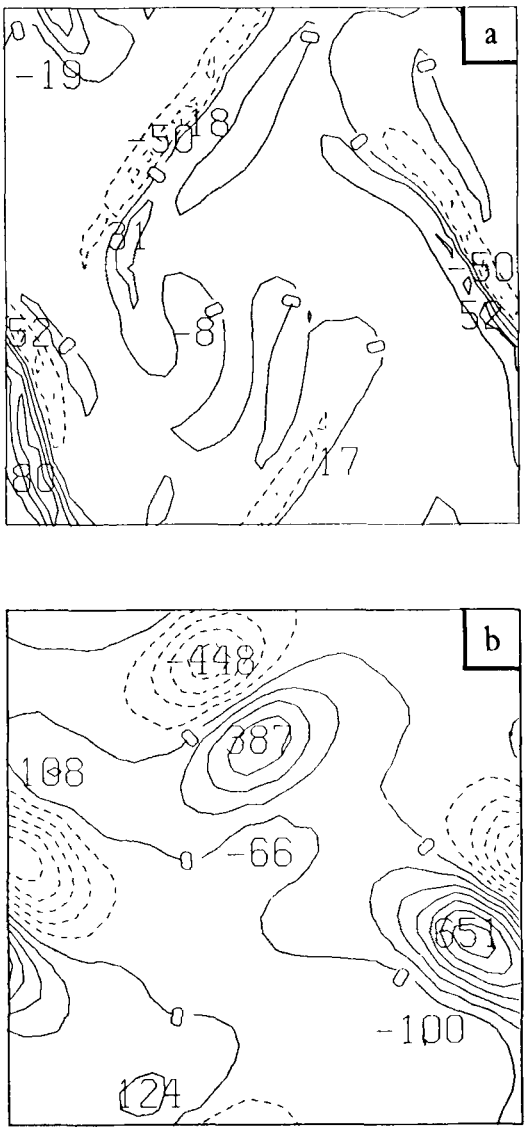


Fig. 7. Streamfunction of the fastest growing perturbation over a 24-hour period, at the beginning (a) and at the end (b) of the period. Extrema values are printed. The unit (arbitrary) is the same in both charts but the contour interval in (b) is four times the contour interval in (a).

The fastest growing perturbations are geostrophic and concentrated in small scales. Fig. 7

shows the streamfunction of the dominant perturbation for a 24-hour forecast range, at the initial and final times. The perturbation streamfunction shows dipoles in the vicinity of the saddle-nodes of the reference streamfunction, around which barotropic instability has already been seen to develop in the direct perturbation experiments. Similar situations have been observed in the evolution of Lyapunov vectors (Legras, pers. comm.) in the case of incompressible two-dimensional turbulence on the sphere. It is also seen in Fig. 7 that the scale of the dominant perturbation increases over the time range under consideration. This remains true for longer ranges, and the dominant eigenperturbation at 48 hours has a structure very similar to that obtained at 24 hours by direct explicit perturbation (Fig. 6).

The resolvent is therefore very sensitive to the location of the saddle-nodes of the streamfunction. But the modes which are most sensitive to the saddle-nodes are restricted to small scales and we have seen in the preceding section that if something can be useful for the analysis it is the large-scale part of the resolvent.

The fact that the growth of large-scale Rossby modes is exponential suggests that it is governed by a constant coefficient linear differential system. Indeed the evolution of the basic flow is slow and the coefficients of the tangent linear system (4) vary slowly in time. We have therefore performed a linear integration, the coefficients of the system (4) being now kept at their initial values. For large-scale Rossby waves, the results are very similar to those obtained in the direct perturbation experiments. In particular, the dominant eigenvectors have a two-day doubling time and are associated with a large-scale geostrophic flow whose pattern is such that kinetic energy is extracted from the basic flow.

These results suggest that it may be sufficient, in order to describe the temporal evolution of the error in an assimilation scheme, to use a simplified constant-coefficient version of the tangent linear system, restricted to large-scale Rossby modes. With respect to present assimilation procedures, the main gain achieved by proceeding so would be that the specific features of the current atmospheric flow would be explicitly taken into account, in particular in terms of the instabilities which are likely to develop. Additional work must of course be performed in order to define explicitly the most appropriate form of the simplified linear system describing the evolution of the forecast error.

8. Conclusion

The results presented and discussed in this article show that the short-time evolution of the dominant components of the forecast error can be approximated to a high degree of accuracy by a linear differential system. In addition, the temporal variation of the coefficients of that system can be ignored without significant degradation of the accuracy of the results.

However these results have been obtained with a rather simple numerical model which is far from containing all the features of atmospheric physics relevant for short-range forecasting. One question has to do with the use in this article of plane rather than spherical geometry. It has already been mentioned that the β-effect, by increasing the linearity of Rossby wave evolution, should make the linear approximation even more valid. The use of plane geometry makes somewhat arbitrary the choice of the ratio between the two intrinsic scales of the flow, namely the size of the periodic domain and the Rossby radius of deformation. Preliminary experiments have shown that reducing the relative size of the Rossby radius reduces in approximately the same proportion the scale of the most unstable modes. This is consistent with results recently obtained by Farge and Sadourny (1986) and with the interpretation these authors have given of the role of the inverse Rossby radius λ_D in eq. (6).

Atmospheric dynamics is dominated to a larger extent by baroclinic instability than by barotropic instability and an important question is of course whether the conclusions obtained here would remain valid for a more realistic baroclinic model. Although results obtained by Balgovind et al. (1983) with a primitive equation model indicate that the linearity hypothesis is coherent from a statistical point of view, further numerical experiments will obviously be necessary in order to answer precisely to that question. But the fact that baroclinic instability develops at a roughly exponential rate over time periods of about 48 hours can be taken as an indication that the results presented here would not be drastically altered with a more realistic primitive equation model.

9. Acknowledgements

This work was supported by a grant from the Centre National d'Etudes Spatiales. The computing facilities were allocated by the scientific council of the Centre de Calcul Vectoriel pour la Recherchie.

10. Appendix

We will present an example of a constant coefficient linear differential system for which the fastest growing mode over a given time interval is not the dominant eigenvector, the corresponding largest amplification factor being larger than the amplification factor associated to the largest eigenvalue.

We consider the 2×2 matrix A given by:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

and the associated linear differential system of equations $dX/dt = A \cdot X$. We look for the time behaviour of the Euclidean norm $\|X(t)\|^2 = x_1(t)^2 + x_2(t)^2$ where x_1 and x_2 are the two components of the vector X.

The eigenvalues λ_1, λ_2 and the corresponding unit eigenvectors v_1, v_2 are:

$$\lambda_1 = 1, \quad v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \lambda_2 = 0, \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

The resolvent of the system at time T is then easily obtained:

$$M_T = \begin{pmatrix} 1 & \exp(T) - 1 \\ 0 & \exp(T) \end{pmatrix}.$$

Computing the amplification factor for the unit vector $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ at time T we obtain:

$$\|M_T(e_2)\|^2 = (\exp(T) - 1)^2 + (\exp(T))^2.$$

This coefficient is larger than the amplification coefficient $(\exp(\lambda_2 T))^2$.

The reason for which unitary vectors can grow faster than the eigenvectors is that, since the eigenvectors are not orthogonal, there are unitary vectors, such as e_2, whose component along the dominant eigenvector v_2 is larger than 1 ($\sqrt{2}$ in our case). Since these vectors have non-zero component along the other eigenvector v_1, the norm of their image by M_T can be larger than the norm $\|M_T(v_2)\|$.

In the present case the matrix S_T whose dominant eigenvector maximizes the amplification factor is:

$$S_T = M_T^* M_T = \begin{pmatrix} 1 & \exp(T) - 1 \\ \exp(T) - 1 & 2\exp(2T) - 2\exp(T) + 1 \end{pmatrix}.$$

As time goes to infinity, S_T is equivalent to:

$$\exp(2T) \times \begin{pmatrix} O(1) & O(1) \\ O(1) & 2 \end{pmatrix},$$

which shows that the growth rate $(\exp(T))$ of the dominant eigenvector of A is asymptotically recovered.

Remark. The fact that A is singular in the above example has no importance since it is only with the resolvent M_T, which is non-singular, that we are dealing. The important condition is that the eigenvectors of A must be non-orthogonal for the scalar product chosen in the definition of the amplification factor. We thus see that the fastest growing modes over a finite time interval, are not, contrary to the eigenvectors, independent of the norm.

In the context of Section 7, the fact that the largest amplification coefficient, associated with some small scale perturbation, is initially different from the eigenvalue which asymptotically appears in the exponential behaviour of perturbations for large time scales can be interpreted as follows: these small scale perturbations, when decomposed along the eigenvectors of the constant coefficient tangent linear system, have relatively large components along the large-scale Rossby eigenvectors which dominate the time evolution of the tangent linear system.

REFERENCES

Balgovind, R., Dalcher, A., Ghil, M. and Kalnay, E. 1983. *Mon. Wea. Rev. 111*, 701–722.

Basdevant, C., Legras, B., Sadourny, R. and Béland, M. 1981. A study of barotropic model flows: Intermittency, waves and predictability. *J. Atmos. Sci. 38*, 2305–2326.

Brauer, F. and Nohel, J. A. 1969. *Qualitative theory of ordinary differential equations*. W. A. Benjamin, New York.

Daley, R. 1980. On the optimal specification of the initial state for deterministic forecasting. *Mon. Wea. Rev. 108*, 1719–1735.

Daley, R. 1981. Normal mode initialization. In: *Dynamic meteorology: data assimilation methods*, eds. Bengtsson, L., Ghil, M. and Källen, E. Springer-Verlag, New York, 77–109.

Eckmann, J.-P. and Ruelle, D. 1985. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys. 57*, Part I, 617–656.

Errico, R. M. 1982. Normal mode initialization and the generation of gravity waves by quasi-geostrophic forcing. *J. Atmos. Sci. 39*, 573–586.

Farge, M. and Sadourny, R. 1986. Inhibition of the two-dimensional turbulence by rotation (in French). *Comptes Rendus de l'Académie des Science*, Serie *12*. Turbulence.

Ghil, M., Cohn, S., Tavantzis, J., Bube, K. and Isaacson, E. 1981. Application of estimation theory to numerical weather prediction. In: *Dynamic meteorology: data assimilation methods*, eds. Bengtsson, L., Ghill, M. and Källen, E. Springer-Verlag, New York, 139–224.

Ghil, M., Cohn, S. E. and Dalcher, A. 1982. Sequential estimation, data assimilation and initialization. In: The interaction between objective analysis and initialization. *Proceedings of the 14th Stanstead seminar*. Publications in Meteorology 127, Department of Meteorology, McGill University, Montreal, ed. Williamson, D., 193 pp.

Hoffman, R. N. and Kalnay, E. 1983. Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus 35A*, 100–118.

Jazwinski, A. H. 1970. *Stochastic processes and filtering theory*. Academic Press, New York.

Le Dimet, F. X. and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus 38A*, 97–110.

Leith, C. E. 1974. Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev. 102*, 409–418.

Leith, C. E. 1980. Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci. 37*, 958–968.

Lorenz, E. N. 1982. Atmospheric predictability experiments with a large numerical model. *Tellus 34*, 505–513.

Machenhauer, B. 1977. On the dynamics of gravity oscillations in a shallow-water model with application to normal mode initialization. *Beitr. Phys. Atmos. 50*, 253–271.

Parlett, P. 1980. *The symmetric eigenvalue problem*. Series in computational mathematics, Prentice Hall, Englewood Cliffs, New Jersey.

Pedlosky, J. 1979. *Geophysical fluid dynamics*. Springer-Verlag, New York.

Talagrand, O. and Courtier, P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Q. J. R. Meteorol. Soc.* in press.

Thaning, L. 1984. On the existence of solutions to Machenhauer's nonlinear mode initialization. *Tellus 36A*, 30–41.

Thompson, P. D. 1986. A simple approximate method of stochastic-dynamic prediction for small initial errors and short range. *Mon. Wea. Rev. 114*, 1709–1715.

Urban, B. 1985. Error maximum growth in simple meteorological models (in French). (Météorologie Nationale Internal report.)

Vautard, R. and Legras, B. 1986. Invariant manifolds, quasi-geostrophy and initialization. *J. Atmos. Sci. 43*, 565–584.