A GUIDE TO LIAPUNOV VECTORS

B. Legras & R. Vautard
Laboratoire de Météorologie Dynamique du CNRS
24 rue Lhomond, F-75231 PARIS Cedex 05

published in
Predictability vol I, editor T. Palmer
ECWF Seminar, ECMWF, Reading, UK, pp. 135-146, 1996

1 INTRODUCTION

All along the seminar at the ECMWF, there was a tense and stimulating debate about the techniques for generating perturbations able to issue ensemble forecasts. This debate was animated by two principal actors who, understandably, defended their own methodologies. However, when the meeting closed, there was still a number of open points left. The purpose of our note is to briefly review theoretical issues about dynamical systems, that may help to elucidate these points. It is not intended to present the authors' original results. Most of the results summarized here are detailed in Eckmann & Ruelle (1985) and Goldhirsch, Sulem & Orszag (1987), among others. The cornerstone of the debate was the comparison between "bred grown modes" (BGMs; Toth & Kalnay, 1993, 1996) and "singular vectors" (SVs; Buizza *et al.*, 1993; Buizza & Palmer, 1995), also called "optimal vectors". From a practical point of view, the performance of these two generation methods has been compared within the framework of a simple quasi-geostrophic system (Molteni & Marshall, 1993) by Houtekamer & Derome (1995), but doubt can be cast on the realism of the model in the real-world situation, where convective instabilities exist at much smaller scales and much higher growth rates.

The original problem is primarily technical: Assuming that we roughly know the initial distribution of analysis errors, the -obviously- most correct way of generating random perturbations around the initial analysis is to use a random generator having the same statistics as the error statistics. The major problem is that the space spanned by these errors is probably too large to be sufficiently explored by a small number of selected perturbations such as is possible today. Hence, one has to make compromises.

The BGM method consists in an initial generation of small random perturbations, from which the nonlinear model is integrated a long time before the target time of the forecast. Perturbations are rescaled over short cycles (typically 12 hours) and grown again. At the end of the process, all perturbations are scaled to a desired (albeit arbitrary) amplitude. The SV method consists essentially in projecting random perturbations onto the fastest growing directions over a given finite time interval, which is typically of one day and half in the current application at ECMWF. Both techniques generate perturbations projecting onto unstable features of the atmosphere. Our intention is not to give more credit to one of them, but only to clarify their theoretical foundations based on what is known from the theory of dynamical systems.

An important practical limitation of our discussion is that we focus here on infinitesimal perturbations. Using this assumption, both methods should, in principle, generate perturbations mostly projecting onto the fast convective modes, which are admittedly not responsible for loss of predictability in the large scales at short leads. In the BGM method, which is nonlinear by essence, these instabilities saturate and are not expected to show up in the bred perturbations. In the SV method, which is linear, one way of avoiding this difficulty is simply to employ a large-scale model. The questions we address here are: (i)

What is the definition of Liapunov vectors, and how are the perturbations generated by the two methods related to them ? (ii) Are the two sets of perturbations identical, or spanning the same instability subspace ? (iii) Are the perturbations tangent to the attractor ? (iv) What are the effective growth rates of the perturbations ? (v) How sensitive are the generated sets of perturbations to analysis errors ?

Within the stated assumptions, we will show that SV and BGM methods are related to two different families of vectors, the forward and backward Liapunov vectors, which have specific dynamical properties. We felt necessary to devote a significant part of this article to the presentation of the theoretical background of our discussion, which is, for clarity, introduced using an analytical example of flow trajectory.

2 THE LINEAR TANGENT OPERATOR

2.1 Definitions

Let us assume that the atmosphere's evolution is governed by the nonlinear dynamical system :

$$\frac{d\mathsf{x}(t)}{dt} = \mathsf{F}(\mathsf{x}(t)), \tag{1}$$

where $\mathsf{x}(t)$ is an element of the finite-dimensional phase space $\mathbb{R}^n$, $\mathsf{x}(t) \in \mathbb{R}^n$. Infinitesimal perturbations $\mathsf{y}(t)$ are therefore governed by the tangent linear system:

$$\frac{d\mathsf{y}(t)}{dt} = \frac{\partial \mathsf{F}(\mathsf{x}(t))}{\partial \mathsf{x}}\mathsf{y}(t), \ \mathsf{y}(t) \in \mathbb{R}^n. \tag{2}$$

This system being linear, it has a resolvent $\mathbf{M}(t_1,t_2)$ that yields

$$\mathsf{y}(t_2) = \mathbf{M}(t_1,t_2)\,\mathsf{y}(t_1). \tag{3}$$

What follows is simply a detailed discussion of the properties of the resolvent $\mathbf{M}(t_1,t_2)$. Note that, regardless of the sign of $t_2 - t_1$, one has

$$\mathbf{M}(t_1,t_2) = \mathbf{M}(t_2,t_1)^{-1}. \tag{4}$$

We assume by convention that $t_2 > t_1$ unless specified, and we are particularly interested in the far future limit (i.e. when $t_2 \to \infty$ while t_1 is kept fixed at the present time), and in the far past limit (i.e. when $t_1 \to -\infty$ while t_2 is kept fixed at the present time).

2.2 An analytical example

In order to introduce and illustrate the properties of $\mathbf{M}(t_1,t_2)$, we use a simple example of flow defined in $\mathbb{R}^3$ with three components (x_0,x_1,x_2) such that the basic trajectory is

$$x_0(t) = t, \ x_1(t) = x_2(t) = 0.$$

We assume that perturbations $(y_1(t),y_2(t))$ in the plane (x_1,x_2) transversal to the trajectory do remain in this plane. We neglect any nonlinearity, assuming that the perturbation only evolves according to the linear tangent system (2). We further assume that the instantaneous linear operator $\partial\mathsf{F}/\partial\mathsf{x}$ has one unstable and one stable direction which are orthogonal in the chosen basis, and that these two directions rotate uniformly in time around the basic trajectory with unit angular velocity. Although not very realistic, this example contains sufficiently rich aspects that illustrate the behavior of infinitesimal perturbations along a flow trajectory which undergoes permanently changing stability conditions.

Explicitely, we have

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{R}_t \begin{pmatrix} \chi_1 & 0 \\ 0 & \chi_2 \end{pmatrix} \mathbf{R}_{-t} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \tag{5}$$

where $\mathbf{R}_t$ is the rotation of angle t with respect to the x_0-axis. We assume in the sequel that $\chi_1 > 0$, $\chi_2 < 0$, and $\delta = \frac{1}{2}(\chi_1 - \chi_2) > 1$. Using the transformed variables

$$\begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} = \mathbf{R}_{-t} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

we obtain a linear system with constant coefficients

$$\frac{dy'_1}{dt} = \chi_1 y'_1 + y'_2, \tag{6}$$

$$\frac{dy'_2}{dt} = -y'_1 + \chi_2 y'_2, \tag{7}$$

whose eigenvalues are $\sigma_{1,2} = \frac{1}{2}(\chi_1 + \chi_2) \pm \sqrt{\delta^2 - 1}$. The associated normal eigenvectors in the rotated space are

$$\mathsf{h}'_1 = \frac{1}{\sqrt{2\delta\alpha}} \begin{pmatrix} \alpha \\ -1 \end{pmatrix}, \tag{8}$$

$$\mathsf{h}'_2 = \frac{1}{\sqrt{2\delta\alpha}} \begin{pmatrix} -1 \\ \alpha \end{pmatrix}, \tag{9}$$

with $\alpha = \delta + \sqrt{\delta^2 - 1}$. Notice that the two eigenvectors are not orthogonal; the cosine of their angle is $-1/\delta$.

It is easy to see that the resolvent $\mathbf{M}(t_1, t_2)$ is

$$\mathbf{M}(t_1, t_2) = \mathbf{R}_{t_2} \mathbf{S} \mathbf{D}_\tau \mathbf{S}^{-1} \mathbf{R}_{-t_1}, \tag{10}$$

$$\text{where } \mathbf{S} = \frac{1}{\sqrt{2\delta\alpha}} \begin{pmatrix} \alpha & -1 \\ -1 & \alpha \end{pmatrix}, \ \mathbf{D}_\tau = \begin{pmatrix} s_2 & 0 \\ 0 & s_1 \end{pmatrix},$$

$$s_1 = e^{\sigma_1 \tau}, \ s_2 = e^{\sigma_2 \tau}, \ \tau = t_2 - t_1.$$

3 EIGENVALUES AND EIGENVECTORS OF $\mathbf{M}(t_1, t_2)$

3.1 Example

The characteristic equation for an eigenvalue ν of $\mathbf{M}(t_1, t_2)$ is

$$\cos\phi \left(1 + \frac{\nu^2}{s_1 s_2}\right) - \nu \left\{ \cos\tau \cos\phi \left(\frac{1}{s_1} + \frac{1}{s_2}\right) + \sin\tau \sin\phi \left(\frac{1}{s_2} - \frac{1}{s_1}\right) \right\} = 0, \tag{11}$$

where $\cot(\phi/2) = \alpha$. For large enough τ, $1/s_2$ is large compared to $1/s_1$, so that we obtain the two eigenvalues

$$\nu_1 \approx s_1 \frac{\cos(\tau - \phi)}{\cos\phi}, \tag{12}$$

$$\nu_2 \approx s_2 \frac{\cos\phi}{\cos(\tau - \phi)}. \tag{13}$$

These eigenvalues are real but oscillate strongly in time. They are not valid when $\cos(\tau - \phi) = 0$. At such times, the eigenvalues are complex:

$$\nu_{1,2} = \pm i \sqrt{s_1 s_2} \left(1 - \frac{s_2}{4 s_1} \sin^2\phi\right)^{1/2} - \varepsilon \frac{s_2}{2} \sin\phi,$$

where $\varepsilon = \sin(\tau - \phi)$. Thus, uniform convergence towards a plain exponential growth or decay, as we may have naively expected, does not occur here. However, the amplitude of the oscillations in ν_1 and ν_2 does vary exponentially in time, being proportional to s_1 for ν_1 and to s_2 for ν_2.

When (12,13) is valid, the two normal eigenvectors are respectively

$$\mathsf{e}_1 = \frac{1}{2\sqrt{\delta\alpha}} \begin{pmatrix} \alpha\cos t_2 + \sin t_2 \\ \alpha\sin t_2 - \cos t_2 \end{pmatrix} \quad \text{for} \quad \nu = \nu_1, \tag{14}$$

$$\mathsf{e}_2 = \frac{1}{2\sqrt{\delta\alpha}} \begin{pmatrix} \alpha\sin t_1 + \cos t_1 \\ \sin t_1 - \alpha\cos t_1 \end{pmatrix} \quad \text{for} \quad \nu = \nu_2. \tag{15}$$

Thus, the eigenvector associated with ν_2 depends only on t_1 while the eigenvector associated with ν_1 depends only on t_2. In other words, the stable eigenvector converges in the far future limit (when $t_2 \to \infty$ and t_1 is kept fixed) while the unstable eigenvector does not. In the far past limit (when $t_1 \to -\infty$ and t_2 is kept fixed), the unstable eigenvector converges while the stable vector does not.

In our example, the eigenvectors swap roles periodically when $\cos(\tau - \phi) = 0$. However, the time interval during which each eigenvector deviates by more than a fixed offset from the directions given by (14,15) decreases exponentially with time.

3.2 General results

The example in the previous subsection demonstrates that we cannot expect a simple asymptotic behavior for eigenvalues ν_i and eigenmodes e_i of $\mathbf{M}(t_1,t_2)$ when one of the times t_1 or t_2 tends to ∞ or $-\infty$. Indeed, for these eigenmodes no mathematical result that is both complete and general exists. However, our example clearly shows the possibility of some asymptotic properties.

Goldhirsh *et al.* (1987) derive such properties under the assumption that the eigenvalues of $\mathbf{M}(t_1,t_2)$ are real and nondegenerate when τ is large enough. They show that there exist limiting exponents $\lambda_i = \lim_{t_2\to\infty} \nu_i(t_1,t_2)$ and that the most stable eigenvector converges to a fixed vector depending only on t_1 in the far future limit. Conversely, the same limiting exponents are obtained in the far past limit but it is now the most unstable eigenvector which converges. The other eigenvectors do not converge in the far future or far past limit but one obtains a family of embedded invariant eigenspaces. We postpone the detailed description of this important property to the next section where it is established using rigorous mathematical arguments.

It suffices to state here that the example studied so far does not satisfy the assumptions of Goldhirsh and colleagues. Nevertheless, the asymptotic properties of the envelop eigensolutions are very similar to those of the asymptotic modes described by these authors.

4 SINGULAR VALUES AND VECTORS OF $\mathbf{M}(t_1,t_2)$

4.1 Definitions

Although it seems natural to study the resolvent's eigenstructure, it is well known that the growth of perturbations is not related to it but to the resolvent's singular structure.

The singular values of $\mathbf{M}$ are the eigenvalues of the normal operators $\mathbf{M}^*\mathbf{M}$ or $\mathbf{M}\mathbf{M}^*$ where $\mathbf{M}^*$ is the adjoint of $\mathbf{M}$ with respect to the scalar product $<.,.>$ on $\mathbb{R}^n$, that is $<\mathbf{M}^*\mathsf{x}',\mathsf{x}> \equiv <\mathsf{x}',\mathbf{M}\mathsf{x}>$ for any pair of vectors (x',x) in $\mathbb{R}^n$. If the canonical scalar product associated with the chosen basis is used, the matrix form of $(\mathbf{M}(t_1,t_2))^*$ is obtained by transposing $\mathbf{M}(t_1,t_2)$.

For a given perturbation y defined at $t = t_1$, the quantity

$$\tilde{\nu} = \frac{<\mathbf{M}(t_1,t_2)\mathsf{y},\mathbf{M}(t_1,t_2)\mathsf{y}>}{<\mathsf{y},\mathsf{y}>} = \frac{<\mathsf{y},(\mathbf{M}(t_1,t_2))^*\mathbf{M}(t_1,t_2)\mathsf{y}>}{<\mathsf{y},\mathsf{y}>} \tag{16}$$

is the growth (or decay) of its norm under the action of the flow during the interval $[t_1,t_2]$.

Since $\mathbf{M}^*\mathbf{M}$ is a symmetric positive-definite operator, it has positive real eigenvalues $\mu_i(t_1,t_2)$ and orthogonal eigenvectors $\mathsf{f}_i^+(t_1,t_2)$. Equation (16) implies that the eigenvectors describe the axes of inertia of the error growth. More precisely, if the perturbations are contained at time t_1 within a spherical ball of unit radius, they evolve at time t_2 towards an ellipsoid, with axes along the vectors $\mathbf{M}(t_1,t_2)\,\mathsf{f}_i^+(t_1,t_2)$ and lengths $(\mu_i(t_1,t_2))^{1/2}$.

In turn, the vectors $\mathbf{M}\mathsf{f}_i^+$ are eigenvectors of $\mathbf{M}\mathbf{M}^*$ with the same eigenvalues μ_i. In other words, the perturbations contained within a spherical ball at time t_2 were contained at time t_1 within an ellipsoid with axis along the vectors $\mathbf{M}(t_1,t_2)^{-1}\mathsf{f}_i^-(t_1,t_2)$ and lengths $(\mu_i(t_1,t_2))^{-1/2}$, where $\mathsf{f}_i^-(t_1,t_2)$ is the eigenvector of $\mathbf{M}(t_1,t_2)(\mathbf{M}(t_1,t_2))^*$ associated with μ_i. We have

$$\mathbf{M}\mathsf{f}_i^+ = \sqrt{\mu_i}\,\mathsf{f}_i^-, \quad \mathbf{M}^*\mathsf{f}_i^- = \sqrt{\mu_i}\,\mathsf{f}_i^+.$$

The eigenvectors $\mathsf{f}_i^+(t_1,t_2)$ are the forward singular vectors as defined by Buizza & Palmer (1995) while the eigenvectors $\mathsf{f}_i^-(t_1,t_2)$ are the backward singular vectors. We show below that breeding modes are related to $\mathsf{f}_i^-(t_1,t_2)$ when $t_1 \to -\infty$.

4.2 Example

The normal matrix associated with the example of section 3.1 is

$$\mathbf{M}^*\mathbf{M} = \mathbf{R}_{t_1}\mathbf{S}^{-1}\mathbf{D}_\tau\mathbf{S}^2\mathbf{D}\mathbf{S}^{-1}\mathbf{R}_{-t_1},$$

which depends on t_2 only through $\mathbf{D}$. The singular value equation is

$$(\mu^2 + s_1^2 s_2^2)\cos^2\phi - \mu(s_1^2 + s_2^2 - 2\sin^2\phi) = 0.$$

It has two real roots which are asymptotically, for large τ,

$$\mu_1 = \frac{s_1^2}{\cos^2\phi}, \;\; \mu_2 = s_2^2\cos^2\phi.$$

These singlar values do not exhibit any of the oscillations which are obtained for the eigenvalues of $\mathbf{M}(t_1,t_2)$.

The asymptotic normal eigenvectors of $\mathbf{M}^*\mathbf{M}$ are

$$\mathsf{f}_1^+(t_1) = \frac{1}{\sqrt{2\delta\alpha}}\mathbf{R}_{t_1}\begin{pmatrix}\alpha\\1\end{pmatrix}, \;\; \mathsf{f}_2^+(t_1) = \frac{1}{\sqrt{2\delta\alpha}}\mathbf{R}_{t_1}\begin{pmatrix}-1\\\alpha\end{pmatrix}.$$

We see that there is no dependence on t_2, which has been dropped from the arguments, and that $\mathsf{f}_2^+(t_1)$ is identical to $\mathsf{h}_2(t_1)$.

Similarly, the asymptotic normal eigenvectors of $\mathbf{M}\mathbf{M}^*$ are

$$\mathsf{f}_1^-(t_2) = \frac{1}{\sqrt{2\delta\alpha}}\mathbf{R}_{t_2}\begin{pmatrix}\alpha\\-1\end{pmatrix}, \;\; \mathsf{f}_2^-(t_2) = \frac{1}{\sqrt{2\delta\alpha}}\mathbf{R}_{t_2}\begin{pmatrix}1\\\alpha\end{pmatrix}.$$

We see that in this case there is no dependence on t_1 and that $\mathsf{f}_1^-(t_2)$ is identical to $\mathsf{h}_1(t_2)$.

The key point revealed by this example is that the two families of vectors $\mathsf{f}_i^-(t)$ and $\mathsf{f}_i^+(t)$ are different, that is the "grown" and "yet-to-grow" perturbations have different main axes of inertia. This is one main reason for which BGM method differs from SV method.

4.3 The Oseledec theorem and the far future limit

Unlike the eigenvalues and eigenvectors of $\mathbf{M}$, the asymptotic behavior of the eigenvalues and eigenvectors of $\mathbf{M}^*\mathbf{M}$ is governed by one of the most important theorems in dynamical systems theory, the so-called Oseledec (1968) theorem). It states that for almost any trajectory $\mathsf{x}(t)$, which is a solution to (1), and for almost any scalar product $<.,.>$ on $\mathbb{R}^n$, the following three facts hold

1. For any vector e in $\mathbb{R}^n$, there exists an exponent

$$\lambda(\mathsf{e}) = \lim_{t_2\to\infty} \frac{1}{t_2 - t_1} \ln\left(\frac{\|\mathbf{M}(t_1,t_2)\mathsf{e}\|}{\|\mathsf{e}\|}\right),$$

   which is finite, does not depend on t_1 and takes at most n values $\lambda_1 > \lambda_2 > \cdots > \lambda_n$. In practice, we assume that these values are distinct, which is generally the case. Here $\|.\|$ is the norm induced by the scalar product.

2. The limit operator

$$\mathbf{S}_\infty(t_1) = \lim_{t_2\to\infty} [(\mathbf{M}(t_1,t_2))^* \mathbf{M}(t_1,t_2)]^{\frac{1}{2(t_2-t_1)}}$$

   also exists. Notice that the non-integer power of a symmetric positive-definite matrix can easily be defined by diagonalization. This limit operator however depends on t_1, or, more precisely, on the initial point $\mathsf{x}(t_1)$. In fact, only the direction of the eigenvectors of $\mathbf{S}_\infty(t_1)$ depend on $\mathsf{x}(t_1)$, but not the associated eigenvalues μ_i, which are the squares of the exponential of the above λ_i's.

3. There exist a sequence of embedded subspaces

$$F_n^+(t_1) \subset F_{n-1}^+(t_1) \subset \cdots \subset F_1^+(t_1) = \mathbb{R}^n,$$

   such that on the complement $F_i^+(t_1)\backslash F_{i+1}^+(t_1)$ of F_{i+1}^+ in F_i^+ (the set-theoretic difference between F_i^+ and F_{i+1}^+), the exponential growth rate (or decay) is λ_i.

The exponents λ_i are called the *Liapunov exponents* and the eigenvectors $\mathsf{f}_i^+(t_1)$ of $\mathbf{S}_\infty(t_1)$ are usually called the *Liapunov vectors*. A generally overlooked but essential result is that, while the exponents and the embedded subspaces $F_i^+(t_1)$ are independent of the scalar product, all the eigenvectors, but the last one, do depend on it. The last Liapunov vector $\mathsf{f}_n^+(t_1)$ corresponds to the most stable direction, that is to the fastest decaying perturbation. The Liapunov vectors are orthogonal, and the i-th one belongs to $F_i^+(t_1)\backslash F_{i+1}^+(t_1)$. In fact, given an embedded sequence of subspaces as above and a scalar product, there is a unique sequence of such orthonormal vectors which satisfies

$$F_i^+(t_1) = \mathsf{f}_n^+(t_1) \oplus \mathsf{f}_{n-1}^+(t_1) \oplus \cdots \oplus \mathsf{f}_i^+(t_1),$$

where $\mathsf{x}\oplus\mathsf{y}$ stands for the direct sum of the vectors x and y, i.e. for their span $\alpha\mathsf{x}+\beta\mathsf{y}$. For forced dissipative chaotic systems, the first Liapunov exponent is strictly positive, leading to the sensitive dependence on initial data. The sum of the exponents, which is the average trace of the tangent linear operator, is strictly negative because of the dissipativity. Hence the last exponent is strictly negative, and furthermore one intermediate exponent is zero.

There are many corollaries and consequences of this theorem. One is the fact that the growth rate of the surface element spanned initially at $t = t_1$ by any two vectors (e,e') is exponential in time with an exponent $\lambda_1 + \lambda_2$, unless the surface intersects $F_3^+(t_1)$, and more generally the growth rate of k-dimensional volumes is $\lambda_1 + \lambda_2 + ... + \lambda_k$, unless they intersect $F_{k+1}^+(t_1)$. The subspaces $F_i^+(t_1)$ are invariant under the tangent flow, i.e. for any time t_2,

$$\mathbf{M}(t_1,t_2)F_i^+(t_1) = F_i^+(t_2).$$

Such is not the case for the Liapunov vectors themselves, except for $\mathsf{f}_n^+(t_1)$, since the tangent resolvent $\mathbf{M}(t_1,t_2)$ does not generally preserve the orthogonality.

From the Oseledec theorem, one also sees that the Liapunov vectors are the limits when $t_2 \to \infty$ of the forward singular vectors $\mathsf{f}_i^+(t_1,t_2)$. Therefore we have used the same symbol for both (only dropping the dependence on t_2 for the Liapunov vector). It is appropriate therefore to call the $\mathsf{f}_i^+(t_1)$'s *forward Liapunov vectors* as well.

Notice finally that the first forward Liapunov vector is not the direction of the fastest growing perturbation for large future time, since all perturbations contained within $F_1^+(t_1)\backslash F_2^+(t_1)$ eventually grow at the same average rate λ_1. In other words, perturbations starting from, say, the subspace spanned by $\mathsf{f}_1^+(t_1)$ and $\mathsf{f}_2^+(t_1)$ are, in practice, not growing faster than almost any random perturbation. This indicates that, for long-range forecasts, the use of singular vectors for initial perturbations looses its relevance.

4.4 The far past limit

Another important corollary of the Oseledec theorem arises from the time symmetry. Whenever the trajectory lies on the attractor, which is an invariant set of the nonlinear flow (1), the theorem remains valid when the limit is $t_1 \to -\infty$ while t_2 is fixed and $\mathbf{M}(t_1,t_2)\,(\mathbf{M}(t_1,t_2))^*$ is considered instead of $(\mathbf{M}(t_1,t_2))^*\mathbf{M}(t_1,t_2)$. In that case, the Liapunov exponents are the same as for the far future limit but the asymptotic eigenvectors of $\mathbf{M}(t_1,t_2)\,(\mathbf{M}(t_1,t_2))^*$ are now the *backward Liapunov vectors* $\mathsf{f}_i^-(t_2)$. A sequence of embedded subspaces

$$F_1^-(t_2) \subset F_2^-(t_2) \subset \cdots \subset F_n^-(t_2) = \mathbb{R}^n$$

is again defined in such a way that on $F_i^-(t_2)\backslash F_{i-1}^-(t_2)$, initial perturbations (at time t_2) grow like $-\lambda_i$ when reversing time.

As for the far future limit case, the embedded subspaces do not depend on the scalar product, but the backward Liapunov vector do, except the first one $\mathsf{f}_1^-(t_2)$. The backward Liapunov vectors are the only orthonormal basis satisfying

$$F_i^-(t_2) = \mathsf{f}_1^-(t_2) \oplus \mathsf{f}_2^-(t_2) \oplus \cdots \oplus \mathsf{f}_i^-(t_2)\,.$$

To construct a complete basis of vectors that is independent of the scalar product, and has desirable topological properties, one can (Eckmann and Ruelle, 1985) simply intersect the subspaces $F_i^-(t)$ and $F_i^+(t)$ at time t. When all exponents are distinct, this intersection is of dimension 1, unless there are *homoclinic tangencies*, which are points where the stable and unstable manifolds (see below) are tangent. The discussion of this pathological case which may occur generically is beyond the scope of this article (but see Guckenheimer & Holmes, 1983, pp. 331-40). We restrict ourselves to the case whe a set of vectors $\mathsf{g}_i(t)$ can be defined consistently from these intersections, which are independent of the scalar product, and have the following property:

$$\lim_{|t'|\to\infty} \frac{1}{t'} \ln \|\mathbf{M}(t,t')\mathsf{g}_i(t)\| = \lambda_i\,. \tag{17}$$

This means that these vectors grow exponentially with exponent λ_i in the far future and with exponent $-\lambda_i$ in the far past. To our knowledge, the relevance of these *characteristic vectors* for ensemble prediction has never been investigated.

4.5 Stable and unstable manifold

We define the stable manifold at t_1 as the set of perturbations which are damped as $t_2 \to \infty$. In the analytical example, this set is generated by h_2' in the rotated space (cf section 2.2) and by $\mathsf{h}_2(t_1) = \mathbf{R}_{t_1}\mathsf{h}_2'$ at location $(t_1,0,0)$ in the original space. We define conversely the unstable manifold at t_2 as the set of

perturbations which are damped as $t_1 \to -\infty$. In the analytical example, this set is generated by h_1' in the rotated space and by $\mathsf{h}_1(t_2) = \mathbf{R}_{t_2}\mathsf{h}_1'$ at location $(t_2,0,0)$ in the original space.

From the above discussion we see that the stable manifold is generated by $\mathsf{F}^+_{j^-}$ where j^- is the index of the algebraically largest negative Liapunov exponent and the unstable manifold is generated by $\mathsf{F}^-_{j^+}$ where j^+ is the index of the algebraically smallest positive Liapunov exponent (usually $j^- = j^+ + 2$).

In the presence of nonlinearities and chaos, these manifolds are generally much more complicated and exhibit a fine foliated structure. The very structure of the attractor is strongly associated with that of the unstable manifolds. More precisely, an ergodic attractor is contained within the unstable manifold of any of its points (e.g., Arrowsmith and Place, 1990).

5 STANDARD CALCULATION OF LIAPUNOV EXPONENTS AND VECTORS

From a practical viewpoint, it is not possible to calculate numerically $\mathbf{M}(t_1,t_2)$ for large $t_2 - t_1$ and to diagonalize $(\mathbf{M}(t_1,t_2))^*\mathbf{M}(t_1,t_2)$, since this matrix contains exponentially diverging terms that lead to numerical overflow and underflow. The infinite-time Liapunov vectors cannot be estimated by any diagonalization method. It is, however, possible to overcome this difficulty by an indirect consequence of the Osedelec theorem. It can be shown easily that as $t_2 - t_1 \to \infty$, any random perturbation $\mathsf{e}(t_1)$ starting from time t_1 converges to $\mathsf{f}_1^-(t_2)$. The resolvent operator maps almost all the tangent space at time t_1 onto a small cone around the first backward Liapunov vector at time t_2. In the same way, almost all surfaces are mapped onto the subspace $\mathsf{F}_2^-(t_2)$ spanned by $\mathsf{f}_1^-(t_2)$ and $\mathsf{f}_2^-(t_2)$, and almost all k-dimensional volumes are mapped onto the subspace $\mathsf{F}_k^-(t_2)$. This leads to consider a Schmidt orthogonalization procedure in order to construct numerically the Liapunov vectors and the Liapunov exponents.

Let $\mathbf{E}(t_1)$ be the $n \times k$ matrix having as its columns k perturbations initialized at time t_1. Then at time t_2, we have

$$\mathbf{E}(t_2) = \mathbf{M}(t_1,t_2)\,\mathbf{E}(t_1)\,.$$

The Schmidt orthonormalization consists in performing a polar decomposition of $\mathbf{E}(t_2)$

$$\mathbf{E}(t_2) = \mathbf{Q}(t_2)\mathbf{T}(t_2)\,, \tag{18}$$

where the columns of $\mathbf{Q}$ contain the k orthonormal vectors, and $\mathbf{T}$ is a $k \times k$ upper triangular matrix with positive diagonal coefficients. For almost all initial sets of perturbations, the columns of the matrix $\mathbf{Q}(t_2)$ converge to the Liapunov vectors when $t_2 - t_1$ is large enough while the diagonal elements of $\mathbf{T}$ give the k algebraically largest Liapunov exponents. In practice, the orthonormalization is applied to $\mathbf{E}$ at regular intervals to avoid numerical problems, and the integration restarts from the intermediate Schmidt vectors. The final result (18) is insensitive to the frequency of orthonormalization up to roundoff errors (see Goldhirsh *et. al* for a detailed description of the algorithm). This algorithm is much faster than any diagonalization, its cost being $O(k \times n)$. At finite $t_2 - t_1$, the columns of $\mathbf{Q}$ differs from the singular vectors but it can be shown that the convergence of the i-th column to the Liapunov vector $\mathsf{f}_i^-(t_2)$ (provided it is smoothly varying along the trajectory) is exponential, with error $O(\max[e^{(\lambda_i-\lambda_{i-1})(t_2-t_1)}, e^{(\lambda_{i+1}-\lambda_i)(t_2-t_1)}])$.

In order to obtain the forward Liapunov vectors, we use the property that almost any perturbation $\mathsf{e}^*(t_2)$ at time t_2 has an image at time t_1 by the adjoint $(\mathbf{M}(t_1,t_2))^*$, that is

$$\mathsf{e}^*(t_1) = (\mathbf{M}(t_1,t_2))^*\,\mathsf{e}^*(t_2)\,,$$

which converges to the first forward Liapunov vector $\mathsf{f}_1^+(t_1)$ as $t_2 - t_1$ is large. In the same way almost all k-dimensional volumes are mapped onto the subspace spanned by $\mathsf{f}_1^+(t_1) \oplus \mathsf{f}_2^+(t_1) \oplus \cdots \oplus \mathsf{f}_k^+(t_1)$. Therefore the algorithm is now to integrate backward in time the linear equation adjoint to (1), from t_2 to t_1, that is

$$\mathbf{E}^*(t_1) = (\mathbf{M}(t_1,t_2))^*\,\mathbf{E}^*(t_2)\,,$$

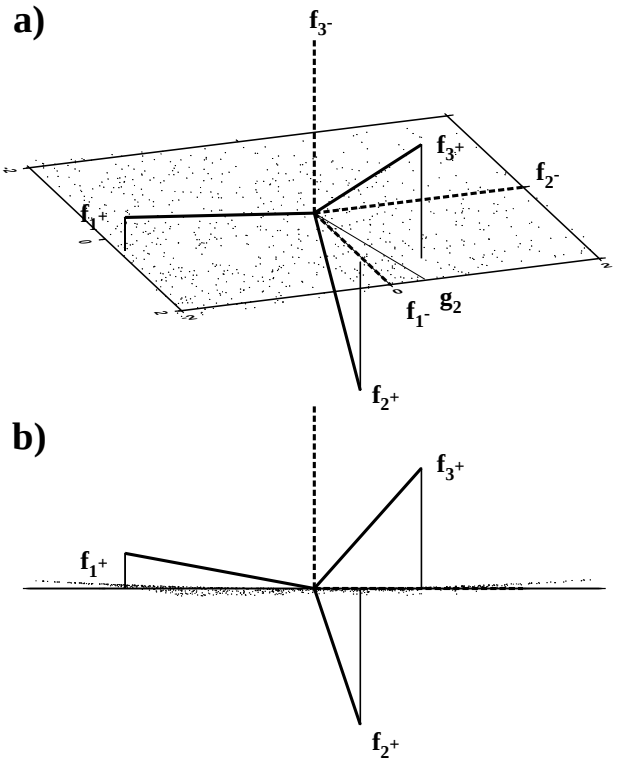

Figure 1: Liapunov $\mathsf{f}^{\pm}$ and characteristic g vectors. See text for details

and to orthonormalize $\mathbf{E}^*(t_1)$ in the same way as $\mathbf{E}(t_2)$:

$$\mathbf{E}^*(t_1) = \mathbf{Q}^*(t_1)\mathbf{T}^*(t_1)\,. \tag{19}$$

Again, the orthonormalization process can be performed at regular intervals during the backward integration of the adjoint and the integration restarts from the intermediate Schmidt vectors. The diagonal elements of $\mathbf{T}^*$ now converge to the k algebraically smallest Liapunov exponents, while the computational cost is still $O(k \times n)$. This procedure is, however, only valid when one wants to estimate the Liapunov vectors, not the finite-time singular vectors.

Finally, it is also not difficult to prove that if we replace $\mathbf{M}^*$ by $\mathbf{M}(t_1,t_2)^{-1}$ in (19), starting from a random set of perturbations at time t_2, and using the Schmidt orthonormalization procedure as above, integrating backward in time from t_2 to t_1, one would also obtain the forward Liapunov vectors, but they would come up in reverse order.

6 NUMERICAL EXAMPLE

We use the classical 3-variable Lorenz system (1963) to illustrate numerically the previous sections in the case of chaotic dynamics. The Lorenz equations are initially integrated 500 time units in order to remove transient behavior. The forward Liapunov vectors are then calculated using the backward Schmidt procedure described in section 5 with the adjoint equations, and a relaxation period of 10 time units. The backward Liapunov vectors are calculated in the same way, but with the Schmidt orthonormalization applied to the direct tangent linear system (2). Average growth rates are calculated over 1000 independent points on the attractor.

Figure 1a shows the backward f^- and forward f^+ Liapunov vectors, and the characteristic g vectors at an arbitrarily chosen point. The three-dimensional representation uses the backward Liapunov vectors (dashed lines) as a reference basis, with the chosen point as the origin. The forward Liapunov vectors

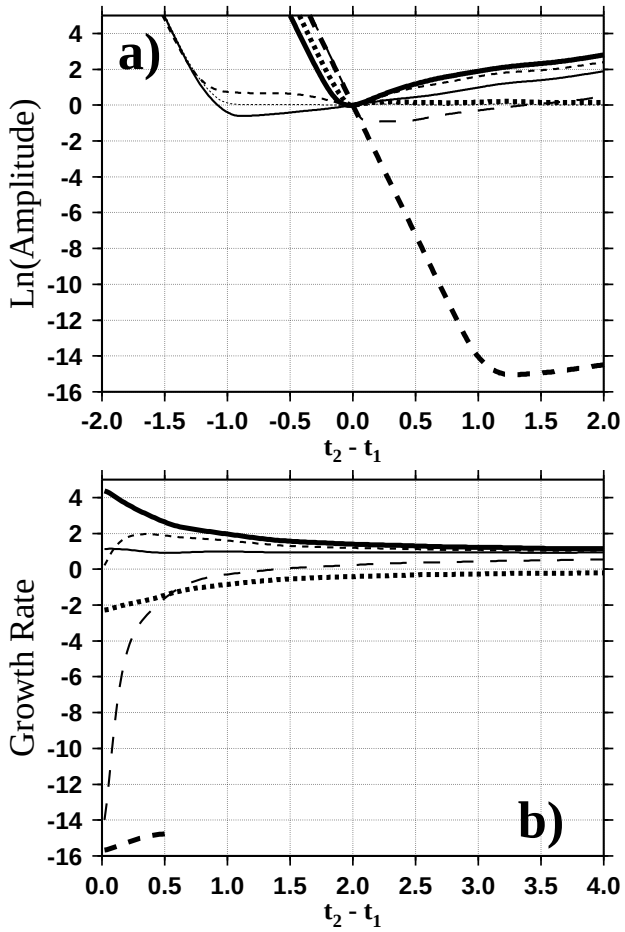

Figure 2: (a) Growth of the perturbation amplitude, when starting at time t_1 from the forward Liapunov vectors $\mathsf{f}_i^+(t_1)$(heavy curves), the backward Liapunov vectors $\mathsf{f}_i^-(t_1)$(light curves) and the second characteristic vector $\mathsf{g}_2(t_1)$. (light dotted curve). The first vectors in each case are displayed as solid curves, the second vectors as the dashed curves with short dashes, and the third vectors as the long-dashed curves. (b) Same as (a) but for the average forward growth rates over finite times for perturbations equal initially to the finite-time singular vectors (heavy curves) and to the backward Liapunov vectors [as in (a)] (light curves).

are drawn as the solid heavy lines. All calculated points belonging to the attractor and contained within the cube of size $4\times4\times4$ around the reference point are plotted (light points) in this three-dimensional representation. The surface drawn is $F_2^- = \mathsf{f}_1^- \oplus \mathsf{f}_2^-$. Figure 1b shows the same data but seen from a point which has the same azimuth as figure 1a but with a zero elevation, i.e. it is contained within F_2^-. The Lorenz attractor is now seen as a thin sheet (actually there are infinitely many sheets) to which F_2^- is tangent. The second characteristic vector g_2 is displayed in figure 1a (thin solid line). The first one, g_1, is identical to f_1^-, and the third one, g_3, is identical to f_3^+. In fact, g_2 is the direction of the flow given by (1) at time t while $F_2^+ = \mathsf{f}_3^+ \oplus \mathsf{g}_2$ and $F_2^- = \mathsf{f}_1^- \oplus \mathsf{g}_2$ are respectively the tangent planes to the stable and the unstable manifolds. Notice that the stable manifold is transversal but not orthogonal to the attractor. The main point is that f_1^- and f_2^- are directions within the attractor, while f_1^+ and f_2^+ are not.

Figure 2a shows the average of $\phi(t_2) = \ln\|\mathbf{M}(t_1,t_2)\,\mathsf{e}(t_1)\|$, as a function of $\tau = t_2 - t_1$, over 1000 independent initial states $\mathsf{x}(t_1)$, where $\mathsf{e}(t_1)$ is equal in turn to $\mathsf{f}_1^-(t_1)$, $\mathsf{f}_3^-(t)$, $\mathsf{f}_1^+(t_1)$, $\mathsf{f}_2^+(t_1)$, $\mathsf{f}_3^+(t_1)$ and $\mathsf{g}_2(t_1)$. The linear system (2) is integrated both forward and backward in time. As expected, the perturbation initiated from forward Liapunov vectors exhibit average growth rates for positive times which are the Liapunov exponents, $\lambda_1 = 1$, $\lambda_2 = 0$ and $\lambda_3 = -14$. By ergodicity, one can show that the average

growth rates at $t_2 - t_1 = 0$ are also the Liapunov exponents (for both forward and backward vectors). Notice that the perturbation started along the third forward Liapunov vector decays to zero within the single-precision of our calculations by the time that $t_2 - t_1 \approx 1$. At later times, round-off errors kick the perturbation off the stable manifold, and the growth rate rapidly reaches the value of the first exponent. For negative times, the average growth rates of the forward Liapunov vectors are all close to $+14$. Similarly, for positive times, the backward Liapunov vectors rapidly reach a growth rate equal to the first Liapunov exponent. For negative times, they grow as the opposite of the Liapunov exponents. The first backward Liapunov vector remains numerically on the unstable manifold for a time interval of 1 unit, and is then kicked off by round-off errors. Notice that the decay of the first backward Liapunov vector is much less impressive than for the stable manifold and positive time, owing to the large ratio between $|\lambda_3|$ and $|\lambda_1|$.

Figure 2b shows the average of $\phi(t_2)/\tau$, as a function of τ, for perturbations initialized with (i) the backward Liapunov vectors $\mathsf{f}_i^-(t_1)$ and (ii) the singular vectors $\mathsf{f}_i(t_1,t_2)$ obtained by a direct diagonalization of the matrix $(\mathbf{M}(t_1,t_2))^*\mathbf{M}(t_1,t_2)$. It is clear that the average initial growth rates of the backward Liapunov vectors are the Liapunov exponents. However, when τ is large, the growth rates all tend to the value of the first Liapunov exponent. This demonstrates that backward vectors do not enjoy any special property in the far future. For large τ, the growth rates of singular vectors are the Liapunov exponents. Nevertheless, $\mathbf{MM}^*$ is so ill-conditioned that, in most cases, the third eigenvalue and the third eigenvector cannot be estimated with accuracy for times larger than about 0.5 units. These results contrast with the numerical efficiency of the Schmidt decomposition technique proposed in section 5 in order to compute forward Liapunov vectors. An important point illustrated by figure 2b is that at moderate "optimization times", i.e. for moderate values of τ, the spectrum of the singular values is much broader than that of the growth rates of backward Liapunov vectors. This explains one of the figures presented by T. Palmer which raised considerable controversy. Our results are corroborated by Trevisan & Pancotti (1996) who study the Liapunov vectors associated with unstable periodic orbits of the Lorenz model. These authors recover transient and long-term behavior for singular and Liapunov vectors which is the same as that shown in figure 2b.

In a practical situation, the initial state $\mathsf{x}(t)$ is not known exactly, and the calculation of BGMs and SVs are based on erroneous trajectories of the model. This may have a significant impact since, for the estimation of Liapunov vectors, the large-time limit has to be reached. Hence it is necessary to check how sensitive the calculated Liapunov vectors are to errors in the estimate of the initial state. There is, to our knowledge, no theorem which guarantees the continuity of Liapunov vectors. However, continuity seems to hold for the Lorenz system. To demonstrate this, we perturb the trajectory by adding some random Gaussian noise of standard deviation σ to each variable. Forward Liapunov vectors are then calculated by the adjoint Schmidt technique described above, using the same parameters, but the reference trajectory is now the trajectory starting from the perturbed initial state. Backward Liapunov vectors are calculated by forward Schmidt orthonormalization, refreshing the perturbed reference trajectory every 0.1 time units. Except for the fact that perturbations are still grown from the linear tangent model, this method is similar to BGM, with a breeding cycle of 0.1 time units.

Figure 3 shows the average cosine (over 100 independent cases) of the angles between the perturbed and the unperturbed grown vectors as a function of σ. For comparison, a standard deviation of $\sigma = 1$ is about one tenth of the variables' standard deviation. It is clear that the Liapunov vectors are quite robust to perturbations of the initial data. One notices, however, that some vectors are more robust than others. The third backward Liapunov vector is surprisingly more robust than the first two. The same result holds for forward Liapunov vectors. We do not have any explanation at hand. Another important feature is that forward Liapunov vectors are apparently more robust than the backward ones. Should we conclude that SVs are more robust than BGMs ? This would certainly be stretching the interpretation of results obtained here only for the simple Lorenz system.

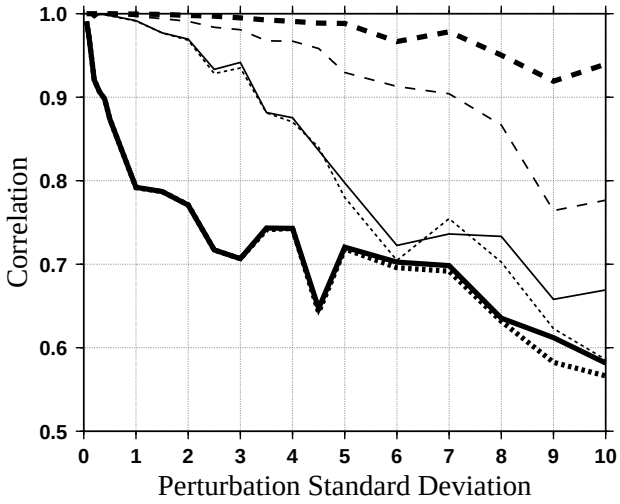

Figure 3: Average cosine of the angle between perturbed Liapunov vectors and the exact Liapunov vectors, as a function of the standard deviation of the initial condition error. Curve patterns as in figure 2a.

7 DISCUSSION

Neither the BGMs of Toth and Kalnay (1994, 1996), nor the SVs of Buizza *et al.* (1993) correspond rigorously to the Liapunov vectors discussed in the previous sections. The BGMs are indeed "grown" from a long past time interval, but take into account the nonlinear saturation of small-scale convective modes. If there truly were a clear physical and energy gap between convectively and baroclinically unstable subsystems of the climate, BGMs would be simply in the linear regime at large scale and amplitude-saturated at convective scales. However, such a gap has never been clearly identified. Therefore, one expects a rather progressive similarity of the BGMs to the backward Liapunov vectors as spatial scale becomes larger.

By contrast, the SVs are by definition linear perturbations, but are "grown" over finite forward times. It is the large time limit of the SVs which equals the forward Liapunov vectors. Despite these reservations, the mathematical theory presented in the previous sections clearly shows that there is no reason to believe that BGMs and SVs bear, in general, any resemblance. BGMs have no particular near- or far-future growth properties. They all grow, as lead time goes to infinity, as the first Liapunov exponent, even when orthonormalized. However, they are more likely to represent the actual initial error, since the BGM methodology is designed to simulate the initial analysis errors. Their correspondence to the leading backward Liapunov vectors leads them to project essentially onto the unstable manifold, which is consistent with the theoretical result of Pires *et al.* (1995) for the errors of assimilated trajectories in chaotic dynamical systems. Possibly, one source of confusion arises from that "instability" refers to past behavior: the unstable manifold is the set of perturbations that have grown fastest in the past (see section 4.5). On the contrary, the design of SVs relate them to future error growth, but does not relate them to initial analysis error.

From our analysis, one is also led to conclude that forward singular or Liapunov vectors are not tangent to the attractor, while the leading backward Liapunov vectors are tangent to the attractor. This property results from the fact that the unstable manifold contains the attractor itself. If BGMs were constructed in a linear fashion, they would therefore characterize perturbations tangent to the attractor.

The fact that BGMs do not, after a certain time, project mostly onto one direction, which would be that of the first Liapunov vector, is a sign that nonlinearity acts significantly, and therefore one should not expect the tangency either. However, nonlinearity may not be the only factor responsible for the lack of directional convergence. In the BGM method, perturbations are recycled by addition to successive anal-

yses which themselves contain errors, while Liapunov vectors are obtained by adding the perturbation to the reference "true" dynamical system trajectory. The sensitivity of the computed Liapunov vectors to initial analysis errors has been examined numerically in section 6 where it has been shown that the last backward Liapunov vector is surprisingly more robust than the other ones. However, the leading forward singular or Liapunov vectors turn out to be more robust than the corresponding backward vectors.

Acknowledgment We thank M. Ghil and I. A. Mogensen for numerous improvements of the manuscript.

References

[] ARROWSMITH, D. K., AND PLACE, C. M. *An Introduction to Dynamical Systems*. Cambridge University Press, 1990.

[] BUIZZA, R., AND PALMER, T., 1995. The singular vector structure of the atmospheric general circulation. *J. Atmos. Sci.* **52**, 1434–1456.

[] BUIZZA, R., TRIBBIA, J., MOLTENI, F., AND PALMER, T., 1993. Computation of optimal unstable structures for a numerical weather prediction model. *Tellus* **45A**, 388–407.

[] ECKMANN, J.-P., AND RUELLE, D., 1985. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57**, 617–656.

[] GOLDHIRSCH, I., SULEM, P.-L., AND ORSZAG, S. A., 1987. Stability and Lyapunov stability of dynamical systems: a differential approach and a numerical method. *Physica* **27D**, 311–337.

[] GUCKENHEIMER, J., AND HOLMES, P. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*. Springer-Verlag, 1983.

[] HOUTEKAMER, P. L., AND DERÔME, J., 1995. Methods for ensemble prediction. *Mon. Wea. Rev.* **123**, 2181–2196.

[] LORENZ, E., 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141.

[] MARSHALL, J., AND MOLTENI, F., 1993. Toward a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.* **50**, 1792–1818.

[] OSELEDEC, V., 1968. A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.* **19**, 197.

[] PIRES, C., VAUTARD, R., AND TALAGRAND, O., 1996. On extending the limits of variational assimilation in chaotic systems. *Tellus* **48A**, 96–121.

[] TOTH, Z., AND KALNAY, E., 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Met. Soc.* **74**, 2317–2330.

[] TOTH, Z., AND KALNAY, E., 1996. Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, submitted.

[] TREVISAN, A. AND PANCOTTI, F., 1996. Periodic orbits, Lyapunov vectors and singular vectors in the Lorenz system *J. Atmos. Sci.*, submitted.