5

Application of generalised stability theory to deterministic and statistical prediction

Petros J. Ioannou
Department of Physics,
National and Capodistrian University of Athens

Brian F. Farrell
Harvard University, Cambridge

Understanding of the stability of deterministic and stochastic dynamical systems has evolved recently from a traditional grounding in the system's normal modes to a more comprehensive foundation in the system's propagator and especially in an appreciation of the role of non-normality of the dynamical operator in determining the system's stability as revealed through the propagator. This set of ideas, which approach stability analysis from a non-modal perspective, will be referred to as generalised stability theory (GST). Some applications of GST to deterministic and statistical forecast are discussed in this review. Perhaps the most familiar of these applications is identifying initial perturbations resulting in greatest error in deterministic error systems, which is in use for ensemble and targeting applications. But of increasing importance is elucidating the role of temporally distributed forcing along the forecast trajectory and obtaining a more comprehensive understanding of the prediction of statistical quantities beyond the horizon of deterministic prediction. The optimal growth concept can be extended to address error growth in non-autonomous systems in which the fundamental mechanism producing error growth can be identified with the necessary non-normality of the system. The influence of model error in both the forcing and the system is examined using the methods of stochastic dynamical systems theory. In this review deterministic and statistical prediction, i.e. forecast and climate prediction, are separately discussed.

Predictability of Weather and Climate, ed. Tim Palmer and Renate Hagedorn. Published by Cambridge University Press.

P1: FYX/FGC P2: FXS
0521848822c05.xml CUUK341B-Palmer February 14, 2006 9:11

5.1 Introduction

The atmosphere and ocean are constantly evolving and the present state of these systems, while notionally deterministically related to previous states, in practice becomes exponentially more difficult to predict as time advances. This loss of predictability of the deterministic state is described as sensitive dependence on initial conditions and quantified by the asymptotic exponential rate of divergence of initially nearby trajectories in the phase space of the forecast system (Lorenz, 1963) given by the first Lyapunov exponent (Lyapunov, 1907; Oseledets, 1968). Moreover, the optimality of the Kalman filter as a state identification method underscores the essentially statistical nature of the prediction problem (Ghil and Malanotte-Rizzoli, 1991; Berliner, 1996). The initial state is necessarily uncertain but so is the forecast model itself and the system is subject to perturbations from extrinsic and subgrid-scale processes. Given all these uncertainties the notion of a single evolving point in phase space is insufficient as a representation of our knowledge of forecast dynamics, and some measure of the uncertainty of the determination of the system state and the evolution of this uncertainty must be included in a comprehensive forecast system theory (Epstein, 1969; Ehrendorfer, this volume; Palmer, this volume).

The appropriate methods for studying errors in deterministic and statistical forecast are based on the system's propagator and proceed from advances in mathematics (Schmidt, 1906; Mirsky, 1960; Oseledets, 1968) and dynamical theory (Lorenz, 1963, 1965, 1985; Farrell, 1988, 1990; Lacarra and Talagrand, 1988; Molteni and Palmer, 1993; Penland and Magorian, 1993; Buizza and Palmer, 1995; Farrell and Ioannou, 1996a, 1996b; Moore and Kleeman, 1996; Kleeman and Moore, 1997; Palmer, 1999; DelSole and Hou, 1999a, 1999b; Ehrendorfer, this volume; Palmer, this volume; Timmermann and Jin, this volume).

We review recent advances in linear dynamical system and stability theory relevant to deterministic and statistical forecast. We begin with deterministic error dynamics in autonomous and non-autonomous certain systems and then address the problem of prediction of statistical quantities beyond the deterministic time horizon; finally, we study model error in certain and uncertain systems.

5.2 Deterministic predictability of certain systems

The variables in a certain forecast model are specified by the finite dimensional state vector $\mathbf{y}$ which is assumed to evolve according to the deterministic equation

$$\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y}). \tag{5.1}$$

Consider a solution of the forecast equations $\mathbf{y}(t)$ starting from a given initial state. Sufficiently small forecast errors $\mathbf{x} \equiv \delta\mathbf{y}$ are governed in the linear limit by the tangent

P1: FYX/FGC P2: FXS
0521848822c05.xml CUUK341B-Palmer February 14, 2006 9:11

linear equations

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\,\mathbf{x}, \tag{5.2}$$

in which the Jacobian matrix

$$\mathbf{A}(t) \equiv \left.\frac{\partial \mathbf{f}}{\partial \mathbf{y}}\right|_{\mathbf{y(t)}}, \tag{5.3}$$

is evaluated along the known trajectory $\mathbf{y}(t)$ and is considered to be known.

The matrix $\mathbf{A}(t)$ is time dependent and in general its realisations do not commute, i.e. $\mathbf{A}(t_1)\mathbf{A}(t_2) \neq \mathbf{A}(t_2)\mathbf{A}(t_1)$. It follows that the evolution of the error field cannot be determined from analysis of the eigenvalues and eigenfunctions of $\mathbf{A}$, as would be the case for time independent normal matrices, but instead the analysis must be made using the methods of generalized stability theory (GST) (for a review see Farrell and Ioannou, 1996a, 1996b). GST concentrates attention on the behaviour of the propagator $\Phi(t, 0)$, which is the matrix that maps the initial error $\mathbf{x}(0)$ to the error at time t:

$$\mathbf{x}(t) = \Phi(t, 0)\,\mathbf{x}(0). \tag{5.4}$$

Once the matrix $\mathbf{A}(t)$ of the tangent linear system is available, the propagator is readily calculated. Consider a piecewise approximation of the continuous operator $\mathbf{A}(t)$: $\mathbf{A}(t) = \mathbf{A}_i$ where $\mathbf{A}_i$ is the mean of $\mathbf{A}(t)$ over $(i-1)\tau \leq t < i\tau$ for small enough τ. At time $t = n\tau$ the propagator is approximated by the time ordered product

$$\Phi(t, 0) = \prod_{i=1}^{n} e^{\mathbf{A}_i \tau}. \tag{5.5}$$

If $\mathbf{A}$ is autonomous (time independent) the propagator is the matrix exponential

$$\Phi(t, 0) = e^{\mathbf{A}t}. \tag{5.6}$$

Deterministic error growth is bounded by the optimal growth over the interval $[0, t]$:

$$\|\Phi(t, 0)\| \equiv \max_{\mathbf{x}(0)} \frac{\|\mathbf{x}(t)\|}{\|\mathbf{x}(0)\|}. \tag{5.7}$$

This maximisation is over all initial errors $\mathbf{x}(0)$. The optimal growth for each t is the norm of the propagator $\|\Phi(t, 0)\|$. The definition of the optimal implies a choice of norm. In many application $\|\mathbf{x}(t)\|^2$ is chosen to correspond to the total perturbation energy.

We illustrate GST by applying it to the simple autonomous Reynolds[1] matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} -1 & 100 \\ 0 & -2 \end{pmatrix}. \tag{5.8}$$

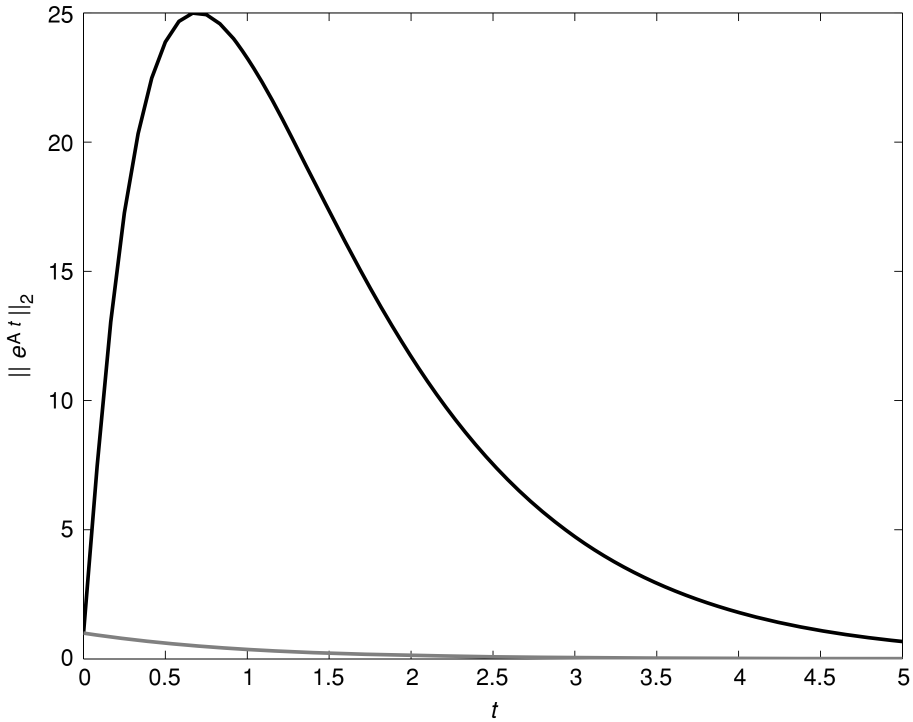

Figure 5.1 The upper curve gives the optimal growth as function of time for the simple matrix (5.8). The optimal growth is given by the norm of the propagator $e^{\mathbf{A}t}$. The lower curve shows the evolution of the amplitude of the least damped eigenmode which decays at a rate of -1.

Consider the model tangent linear system:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}. \tag{5.9}$$

Traditional stability theory concentrates on the growth associated with the most unstable mode, which in this example gives decay at rate -1, suggesting that the error decays exponentially at this rate. While this is indeed the case for very large times, the optimal error growth, shown by the upper curve in Figure 5.1, is much greater at all times than that predicted by the fastest growing mode (the lower curve in Figure 5.1). The modal prediction fails to capture the error growth because $\mathbf{A}$ is non-normal, i.e. $\mathbf{A}\mathbf{A}^\dagger \neq \mathbf{A}^\dagger\mathbf{A}$ and its eigenfunctions are not orthogonal.

The optimal growth is calculated as follows:

$$G = \frac{\|\mathbf{x}(t)\|^2}{\|\mathbf{x}(0)\|^2} = \frac{\mathbf{x}(t)^\dagger\mathbf{x}(t)}{\mathbf{x}(0)^\dagger\mathbf{x}(0)} = \frac{\mathbf{x}(0)^\dagger e^{\mathbf{A}^\dagger t} e^{\mathbf{A}t}\mathbf{x}(0)}{\mathbf{x}(0)^\dagger\mathbf{x}(0)}. \tag{5.10}$$

This Rayleigh quotient reveals that the maximum eigenvalue of the positive definite matrix $e^{\mathbf{A}^\dagger t}e^{\mathbf{A}t}$ determines the square of the optimal growth at time t. The corresponding eigenvector is the initial perturbation that leads to this growth, called the optimal perturbation (Farrell, 1988). Alternatively, we can proceed with a Schmidt decomposition (singular value decomposition) of the propagator:

$$e^{\mathbf{A}t} = \mathbf{U}\Sigma\mathbf{V}^{\dagger} \tag{5.11}$$

with $\mathbf{U}$ and $\mathbf{V}$ unitary matrices and Σ the diagonal matrix with elements the singular values of the propagator, σ_i, which give the growth achieved at time t by each of the orthogonal columns of $\mathbf{V}$. The largest singular value is the optimal growth and the corresponding column of $\mathbf{V}$ is the optimal perturbation. The orthogonal columns, $\mathbf{v}_i$, of $\mathbf{V}$ are called optimal vectors (or right singular vectors), and the orthogonal columns, $\mathbf{u}_i$, of $\mathbf{U}$ are the evolved optimal vectors (or left singular vectors) because from the Schmidt decomposition we have

$$\sigma_i \mathbf{u}_i = e^{\mathbf{A}t}\mathbf{v}_i. \tag{5.12}$$

The forecast system has typical dimension 10^7 so we cannot calculate the propagator directly as in Eq. (5.5) in order to obtain the optimal growth. Instead we integrate the system

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} \tag{5.13}$$

forward to obtain $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0)$ (or its equivalent in a time dependent system), and then integrate the adjoint system

$$\frac{d\mathbf{x}}{dt} = -\mathbf{A}^{\dagger}\mathbf{x} \tag{5.14}$$

backward in order to obtain $e^{\mathbf{A}^{\dagger}t}\mathbf{x}(t) = e^{\mathbf{A}^{\dagger}t}e^{\mathbf{A}t}\mathbf{x}(0)$. We can then find the optimal vectors (singular vectors) by the power method (Moore and Farrell, 1993; Molteni and Palmer, 1993; Errico, 1997). The leading optimal vectors are useful input for selecting the ensemble members in ensemble forecast (Buizza, this volume; Kalnay *et al.*, this volume) because they span and order in growth the initial error (Gelaro *et al.*, 1998). They also identify sensitive regions that can be targeted for further observation (Thorpe and Petersen, this volume).

We have remarked that optimal growth depends on the norm. The choice of norm is dictated by the physical situation; we are usually interested in growth in energy but other norms can be selected to concentrate on the perturbation growth in other physical measures such as growth in square surface pressure, or in square potential vorticity (for a discussion of the choice of the inner product see Palmer *et al.*, 1998; for a discussion of norms that do not derive from inner products see Farrell and Ioannou, 2000). Formally for autonomous operators there exist ‘normal coordinates’ in which the operator is rendered normal; however, this coordinate system is not usually physical in the sense that the inner product in these coordinates is not usually associated with a physically useful measure. But a more deeply consequential reason why the concept of 'normal coordinates' is not useful is that time dependent operators, such as the tangent linear forecast system, are inherently non-normal, in the sense that there is no transformation of coordinates that renders a general $\mathbf{A}(t)$ normal at all times. It follows that analysis of error growth in time dependent systems necessarily proceeds through analysis of the propagator as outlined above.

The tangent linear forecast system is generally assumed to be asymptotically unstable in the sense that the Lyapunov exponent of the tangent linear system is positive. Lyapunov showed that for a general class of time dependent but bounded matrices $\mathbf{A}(t)$ the perturbations $\mathbf{x}(t)$ grow at most exponentially so that $\| \mathbf{x}(t) \| \propto e^{\lambda t}$ as $t \to \infty$, where λ is the top Lyapunov exponent of the tangent linear system which can be calculated by evaluating the limit

$$\lambda = \overline{\lim}_{t\to\infty} \frac{\ln \| \mathbf{x}(t) \|}{t}. \tag{5.15}$$

This asymptotic measure of error growth is independent of the choice of norm, $\| \cdot \|$.

It is of interest and of practical importance to determine the perturbation subspace that supports this asymptotic exponential growth of errors. Because this subspace has a much smaller dimension than that of the tangent linear system itself, a theory that characterises this subspace can lead to economical truncations of the tangent linear system. Such a truncation could be used in advancing the error covariance of the tangent linear system which is required for optimal state estimation. We now show that the inherent non-normality of time dependent operators is the source of the Lyapunov instability which underlies the exponential increase of forecast errors and that understanding the role of non-normality is key to understanding error growth.

Consider a harmonic oscillator with frequency ω. In normal coordinates (i.e. energy coordinates), $\mathbf{y} = [\omega x, v]^T$, where x is the displacement and $v = \dot{x}$, the system is governed by

$$\frac{d\mathbf{y}}{dt} = \mathbf{A}\mathbf{y}, \tag{5.16}$$

with

$$\mathbf{A} = \omega \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{5.17}$$

This is a normal system $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A}$ and the system trajectory lies on a constant energy surface, which is a circle. In these coordinates the perturbation amplitude is the radius of the circle and there is no growth.

Assume now that the frequency switches periodically between ω_1 and ω_2: at $T_1 = \pi/(2\omega_1)$ units of time the frequency is ω_1 and $T_2 = \pi/(2\omega_2)$ units of time the frequency is ω_2. There is no single transformation of coordinates that renders the matrix $\mathbf{A}$ simultaneously normal when $\omega = \omega_1$ and when $\omega = \omega_2$ so we revert to

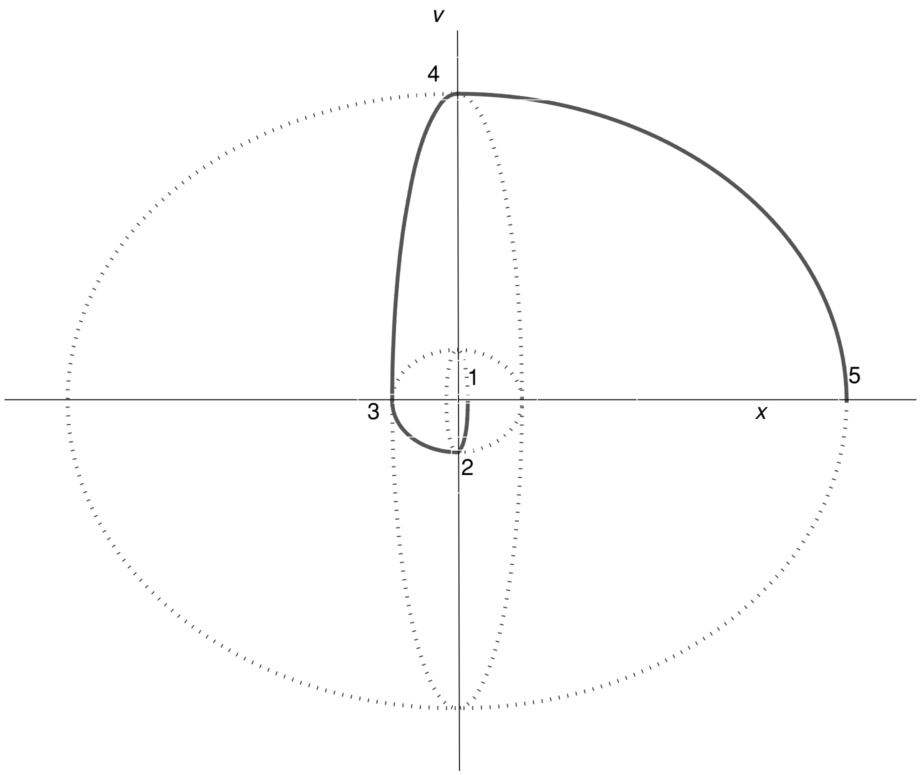

Figure 5.2 The parametric instability of the harmonic oscillator governed by (5.18) is caused by the non-normality of the time dependent operator. So long as there is instantaneous growth and the time dependent operators do not commute, asymptotic exponential growth occurs. For this example $\omega_1 = 1/2$ and $\omega_2 = 3$. See text for explanation of numbers 1 to 5.

the state $\mathbf{y} = [x, v]^T$ with dynamical matrices:

$$\mathbf{A}_{1,2} = \begin{pmatrix} 0 & 1 \\ -\omega_{1,2}^2 & 0 \end{pmatrix}. \tag{5.18}$$

When the frequency is ω_1 the state $\mathbf{y}$ traverses the ellipses of Figure 5.2 that are elongated in the direction of the x axis and when the frequency is ω_2 it traverses the ellipses of Figure 5.2 that are elongated in the v axis (both marked with dots). The dynamics of this system can be understood by considering the evolution of the initial condition $\mathbf{y}(0) = [1, 0]$ marked with 1 in Figure 5.2. Initial condition 1 goes to 2 at time $t = T_2$ under the dynamics of $\mathbf{A}_2$, then the dynamics switch to $\mathbf{A}_1$ taking the system from 2 to 3 at time $t = T_1 + T_2$; reverting back to $\mathbf{A}_2$, the system advances from 3 to 4 at $t = T_1 + 2T_2$, and then under $\mathbf{A}_1$, 4 goes to 5 at time $t = 2(T_1 + T_2)$ with coordinates $\mathbf{y}(2(T_1 + T_2)) = [36, 0]$. As time advances the trajectory clearly grows exponentially as this cycle is repeated despite the neutral stability of the system at each instant of time. How is this possible? The key lies in the inherent non-normality of the operator in time dependent systems. If the operator were time independent and stable, transient growth would necessarily give way to eventual decay. In contrast, a time dependent operator reamplifies the perturbations that would have otherwise decayed. This process of continual rejuvenation producing asymptotic destabilisation is generic and does not depend on the stability properties of the instantaneous operator state (Farrell and Ioannou, 1999).

As a further example consider harmonic perturbations $\Psi(y,t)e^{ikx}$ on a time dependent barotropic mean flow $U(y,t)$ in a β plane channel $-1 \le y \le 1$. The perturbations evolve according to

$$\frac{d\Psi}{dt} = \mathbf{A}(t)\Psi, \tag{5.19}$$

with time dependent operator:

$$\mathbf{A} = \nabla^{-2}\left(-ik\,U(y,t)\nabla^2 - ik\left(\beta - \frac{d^2U(y,t)}{dy^2}\right)\right), \tag{5.20}$$

in which discretised approximations of the operators on the right-hand side is implied. According to Rayleigh's theorem (Drazin and Reid, 1981) this flow cannot sustain growth unless the mean vorticity gradient $Q_y = \beta - U''$ changes sign. Let us consider only flows that are asymptotically stable at all times by Rayleigh theorem, and for simplicity that the mean velocity switches periodically between the two flows shown in the left panels of Figure 5.3. The corresponding mean vorticity gradient is shown in the right panels of the same figure. Despite the asymptotic stability of each instantaneous flow the periodically varying flow is asymptotically unstable. The Lyapunov exponent of the instability as a function of the switching period is shown in Figure 5.4.

This instability arises from sustaining the transient growth of the operator through time dependence. The same process is operative when the flow is varying continuously in time. Then the Lyapunov exponent for given statistically stationary fluctuations in operator structure can be shown to depend on two parameters: the fluctuation amplitude and the autocorrelation time, T_c, of the fluctuations (Farrell and Ioannou, 1999). Snapshots of the perturbation structure revealing the process of accumulation of transient growth by the interaction of the perturbations with the time dependent operator in a continuously varying flow are shown in Figure 5.5. This mechanism of error growth predicts that the perturbation structure should project most strongly on the subspace of the leading optimal (singular) vectors. This is indeed the case, as can be seen in the example in Figure 5.6.

Study of the asymptotic error structure in more realistic tangent linear systems confirms the conclusions presented above (Gelaro *et al.*, 2002). We conclude that error structure in forecast systems projects strongly on the optimal vectors. This

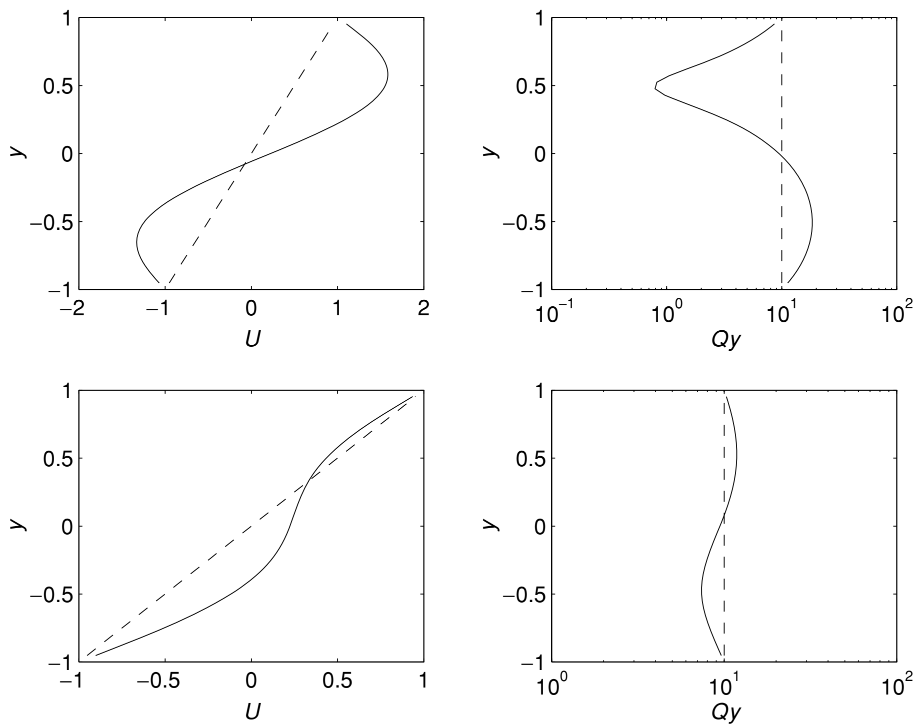

Figure 5.3 (Left panels) the mean flow velocity as a function of latitude for the Rayleigh stable example. (Right panel) the associated mean vorticity gradient, $Q_y = \beta - U''$, with $\beta = 10$.

result is key for dynamical evolution of the error covariance which is required for optimal state estimation (Farrell and Ioannou, 2001).

5.3 Model error in deterministic forecast

We have already discussed methods for determining the impact of uncertainties in the initial state on the forecast. However, as the initialisation of forecast models is improved with the advent of new data sources and the introduction of variational assimilation methods, the medium range forecast will become increasingly affected by uncertainties resulting from incomplete physical parametrisations and numerical approximations in the forecast model, and by the necessarily misrepresented subgrid-scale chaotic processes such as cumulus convection which act as temporally stochastic distributed forcing of the forecast error system. These influences, referred to collectively as model error, conventionally appear as an external forcing in the forecast error system (Allen *et al.*, this volume).

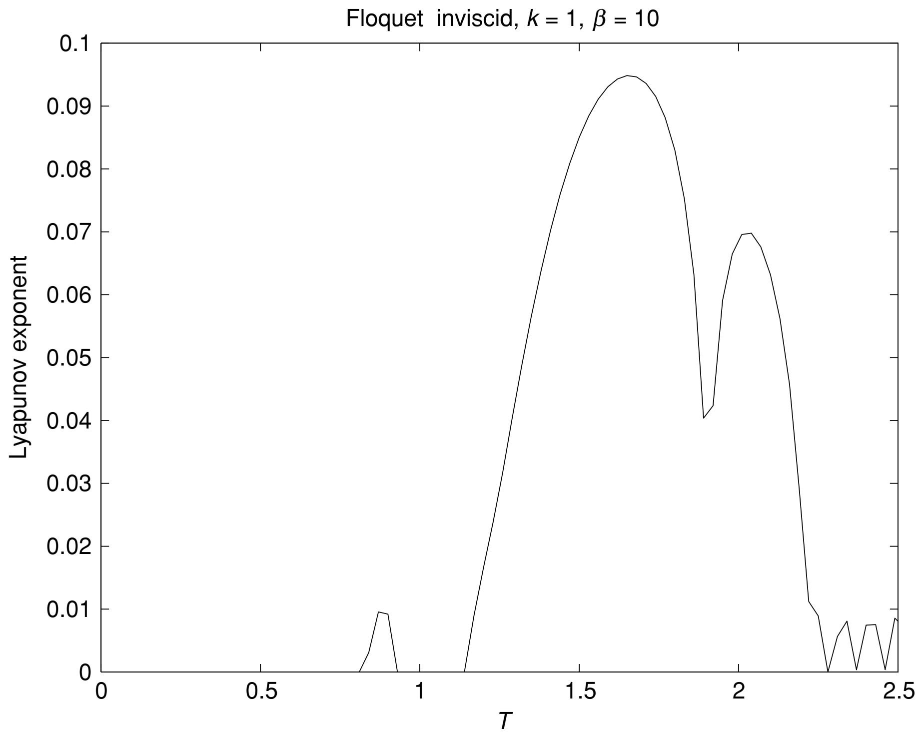

Figure 5.4 Lyapunov exponent as a function of switching period T for the example shown in Figure 5.3. The time dependent flow results from periodic switching every T time units between the Rayleigh stable flow profiles shown in Figure 5.3. The zonal wavenumber is $k = 1$ and $\beta = 10$.

Improving understanding of model error and specifically identifying forcings that lead to the greatest forecast errors are centrally important in predictability studies. In analogy with the optimal perturbations that lead to the greatest forecast error in the case of initial condition error, these continuous error sources will be called optimal distributed forcings. In an approach to this problem D'Andrea and Vautard (2000) obtained approximate optimal temporally distributed deterministic forcings of the forecast error system (which they refer to as forcing singular vectors) and Barkmeijer *et al.* (2003) obtained the optimal temporally distributed deterministic forcing of the forecast error system over fixed spatial structures. We here describe the method for determining the general optimal forcing in both the forecast and assimilation systems.

The underlying theory for determining the optimal forcing in the deterministic case is based on analysis of the dynamical system as a mapping from the space of input forcings to the space of states at later time. We seek the deterministic forcing $\mathbf{f}(t)$ of unit norm on $t \in [0, T]$ producing the greatest state norm at time T, i.e. that

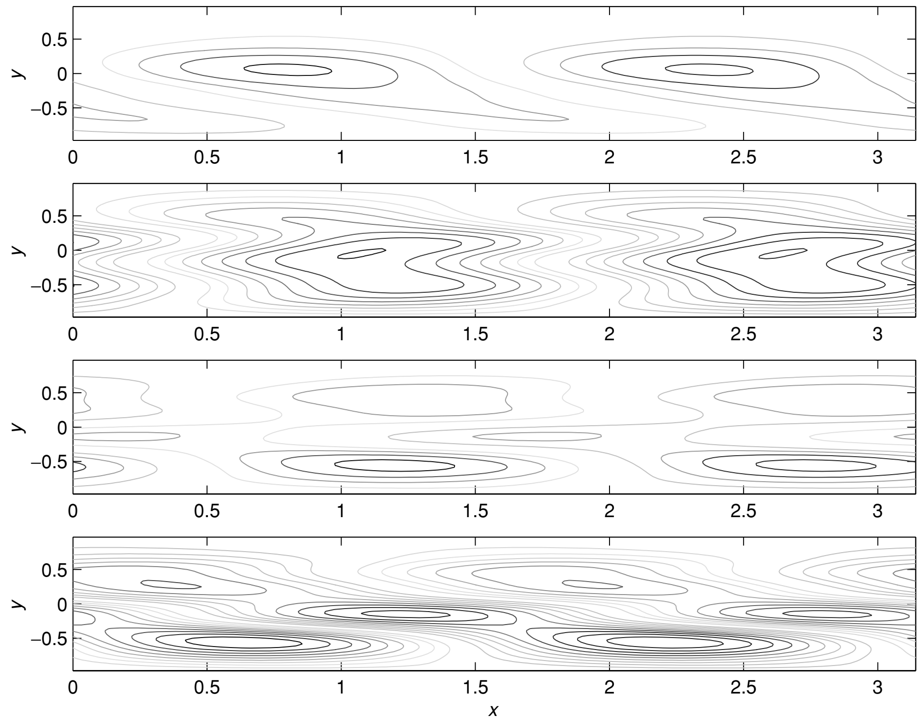

Figure 5.5 For a continuously varying barotropic flow, the structure of the state $\Psi(y, t)e^{ikx}$ in the zonal (x), meridional (y) plane at four consecutive times separated by an autocorrelation time T_c. The rms velocity fluctuation is 0.16 and the noise autocorrelation time is $T_c = 1$. The zonal wavenumber is $k = 2$, $\beta = 0$, and the Reynolds number is $Re = 800$. The Lyapunov exponent is $\lambda = 0.2$. At first (top panel) the Lyapunov vector is configured to grow, producing an increase over T_c of 1.7; in the next period the Lyapunov vector has assumed a decay configuration (second panel from top) and suffers a decrease of 0.7; subsequently (third panel from top) it enjoys a slight growth of 1.1; and finally (bottom panel) a growth by 1.8. Further details can be found in Farrell and Ioannou (1999).

maximises the square norm of the state $\|\mathbf{x}(T)\|^2$, assuming the state is initially zero, $\mathbf{x}(0) = 0$, and that $\mathbf{x}$ obeys the tangent linear forecast equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x} + \mathbf{f}(t). \tag{5.21}$$

The forcing $\mathbf{f}(t)$ over the interval $[0, T]$ is measured in the square integral norm

$$\|\mathbf{f}\|^2_{L_2} = \int_0^T \mathbf{f}^\dagger(t)\mathbf{f}(t)dt, \tag{5.22}$$

while the state $\mathbf{x}$ is measured in the vector square norm

$$\|\mathbf{x}\|^2 = \mathbf{x}^\dagger\mathbf{x}. \tag{5.23}$$

The use of alternative inner products can be easily accommodated.

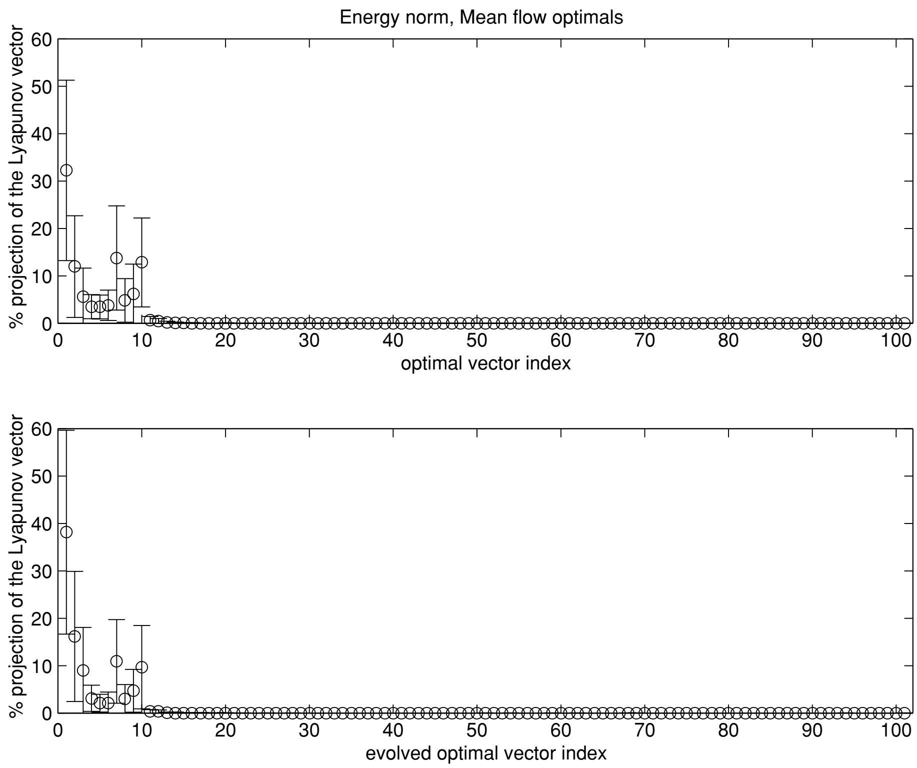

Figure 5.6 (Top panel) Mean and standard deviation of the projection of the Lyapunov vector on the optimal vectors of the mean flow calculated for a time interval equal to T_c in the energy norm. (Bottom panel) The mean projection and standard deviation of the Lyapunov vector on the T_c evolved optimal vectors of the mean flow in the energy norm.

The optimal forcing, $\mathbf{f}(t)$, is the forcing that maximises the final state at time T, i.e. that maximises the quotient

$$R_d = \frac{\| \mathbf{x}(T) \|^2}{\| \mathbf{f} \|_{L_2}^2}. \tag{5.24}$$

It can be shown (Dullerud and Paganini, 2000; Farrell and Ioannou, 2005) that this intractable maximisation over functions, $\mathbf{f}(t)$, can be transformed to a tractable maximisation over states, $\mathbf{x}(T)$. Specifically the following equality is true:

$$\max_{\mathbf{f}(t)} \frac{\| \mathbf{x}(T) \|^2}{\| \mathbf{f}(t) \|_{L_2}^2} = \max_{\mathbf{x}(T)} \frac{\| \mathbf{x}(T) \|^2}{\mathbf{x}(T)^\dagger \mathbf{C}^{-1} \mathbf{x}(T)}, \tag{5.25}$$

where $\mathbf{C}$ is the finite time state covariance matrix at time T under the assumption of temporally white noise forcing with unit covariance $\mathbf{I}$. The covariance can be obtained by integrating from $t = 0$ to $t = T$ the Lyapunov equation

$$\frac{d\mathbf{C}}{dt} = \mathbf{A}(t)\mathbf{C} + \mathbf{C}\mathbf{A}^{\dagger}(t) + \mathbf{I}, \tag{5.26}$$

with the initial condition $\mathbf{C}(0) = 0$. Note that the form of the second optimisation in Eq. (5.25) is reminiscent of the covariance or Mahalanobis metric used in predictability analysis (Palmer *et al.*, 1998), suggesting the interpretation of optimals weighted by the Mahalanobis metric as structures that are most easily forced.

Quotient (5.25) is maximised for unit forcing by the state

$$\mathbf{x}_{opt}(T) = \sqrt{\lambda_1}\mathbf{v}_1, \tag{5.27}$$

where λ_1 is the maximum singular value of $\mathbf{C}$ and $\mathbf{v}_1$ is the corresponding singular vector of $\mathbf{C}$ ($\mathbf{v}_1$ is conventionally called the top empirical orthogonal function (EOF) of $\mathbf{C}$). It can be shown (Farrell and Ioannou, 2005) that the optimal forcing and the associated state of the system can be obtained simultaneously by integrating the following coupled forward and adjoint systems backwards over the finite interval from time $t = T$ to the initial time $t = 0$:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x} + \mathbf{f}$$
$$\frac{d\mathbf{f}}{dt} = -\mathbf{A}^{\dagger}(t)\mathbf{f}, \tag{5.28}$$

with $\mathbf{x}(T) = \sqrt{\lambda_1}\mathbf{v}_1$ and $\mathbf{f}(T) = \mathbf{v}_1/\sqrt{\lambda_1}$. The initial state $\mathbf{x}_{opt}(0) = 0$ is recovered as a consistency check.

5.4 Prediction of statistics of certain systems

Beyond the limit of deterministic forecast it is still possible to predict the statistical properties which constitute the climate of a system. Consider the perturbation structure, $\mathbf{x}$, produced by the forced equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{n}(t). \tag{5.29}$$

Here $\mathbf{A}$ may be the deterministic linear operator governing evolution of large-scale perturbations about the mean midlatitude flow, and $\mathbf{F}\mathbf{n}(t)$ an additive stochastic forcing with spatial structure $\mathbf{F}$, representing neglected non-linear terms. For simplicity we assume that the components of $\mathbf{n}(t)$ are white noise with zero mean and unit variance. We wish to determine the perturbation covariance matrix (or density matrix)

$$\mathbf{C}(t) = \langle \mathbf{x}\mathbf{x}^{\dagger} \rangle, \tag{5.30}$$

where $\langle\cdot\rangle$ denotes the ensemble average over the realisations of the forcing $\mathbf{F}\mathbf{n}(t)$. If a steady state is reached, $\langle\cdot\rangle$ is also the time mean covariance. We argue

(Farrell and Ioannou, 1993; DelSole, 1996, 1999, 2001, 2004)[2] that the midlatitude jet climatology can be obtained in this way because the transient climatology in the midlatitudes is the statistical average state resulting from random events of cyclogenesis. Because cyclogenesis is a rapid transient growth process primarily associated with the non-normality of **A**, its statistics are well approximated by the structure of the linear operator **A**. The diagonal elements of the steady state covariance **C** are the climatological variance of **x**, and they locate the storm track regions. All mean quadratic fluxes are also derivable from **C**, from which the observed climatological fluxes of heat and momentum can be obtained. In this way we obtain a theory for the climate and can address systematically statistical predictability questions such as how to determine the sensitivity of the climate, that is of **C**, to changes in the boundary conditions and physical parameters which are reflected in changes in the mean operator **A** and the forcing structure matrix **F**.

If $\mathbf{n}(t)$ is a white noise process it can be shown (Farrell and Ioannou, 1996a) that

$$\mathbf{C}(t) = \int_0^t e^{\mathbf{A}s}\mathbf{Q}e^{\mathbf{A}^\dagger s}\, ds, \tag{5.31}$$

where

$$\mathbf{Q} = \mathbf{F}\mathbf{F}^\dagger \tag{5.32}$$

is the covariance of the forcing. It can also be shown that the ensemble mean covariance evolves according to the deterministic equation

$$\frac{d\mathbf{C}}{dt} = \mathbf{AC} + \mathbf{CA}^\dagger + \mathbf{Q} \equiv \mathbf{HC} + \mathbf{Q}, \tag{5.33}$$

where **H** is a $n^2 \times n^2$ matrix if **C** is a $n \times n$ covariance matrix. It should be noted that the above equation is also valid for non-autonomous $\mathbf{A}(t)$. If **A** is time independent the solution of the above equation is

$$\begin{aligned}\mathbf{C}(t) &= e^{\mathbf{H}t}\,\mathbf{C}(0) + \left(\int_0^t e^{\mathbf{H}(t-s)} ds\right)\mathbf{Q}\\ &= e^{\mathbf{H}t}\,\mathbf{C}(0) + \mathbf{H}^{-1}(e^{\mathbf{H}t} - \mathbf{I})\mathbf{Q}.\end{aligned} \tag{5.34}$$

As $t \to \infty$ and assuming the operator **A** is stable a steady-state is reached, which satisfies the steady-state Lyapunov equation

$$\mathbf{AC}^\infty + \mathbf{C}^\infty\mathbf{A}^\dagger = -\mathbf{Q}. \tag{5.35}$$

This equation can be readily solved for $\mathbf{C}^\infty$, from which ensemble mean quadratic flux quantities can be derived.

Interpretation of **C** requires care. The asymptotic steady-state ensemble average, $\mathbf{C}^\infty$, is the same as the time averaged covariance and can be obtained from a single realisation of $\mathbf{x}(t)$ by averaging the covariance over a sufficient long interval. However, the time dependent $\mathbf{C}(t)$ cannot be associated with a time average[3] but rather is necessarily an ensemble average. With this consideration in mind, $\mathbf{C}(t)$ from Eq. (5.34) is appropriate for evolving the error covariance in ensemble prediction as will be discussed in the next section. In this section we consider a time independent and stable $\mathbf{A}$ and interpret the steady state $\mathbf{C}^\infty$ as the climatological covariance.

It has been demonstrated that such a formulation accurately models the midlatitude climatology (Farrell and Ioannou, 1994, 1995; DelSole, 1996, 1999, 2001, 2004; Whitaker and Sardeshmukh, 1998; Zhang and Held, 1999) and reproduces the climatological heat and momentum fluxes. The asymptotic covariance captures the distribution of the geopotential height variance of the midlatitude atmosphere as well as the distribution of heat and momentum flux in the extratropics.

The algebraic equation (5.35) gives $\mathbf{C}^\infty$ as an explicit functional of the forcing covariance $\mathbf{Q}$ and the mean operator $\mathbf{A}$, which is in turn a function of the mean flow and the physical process parameters. This formulation permits systematic investigation of the sensitivity of the climate to changes in the forcing and structure of the mean flow and parameters.

We first address the sensitivity of the climate to changes in the forcing under the assumption that the mean state is fixed.

We determine the forcing structure, $\mathbf{f}$, given by a column vector, that contributes most to the ensemble average variance $\langle E(t)\rangle$. This structure is the stochastic optimal (Farrell and Ioannou, 1996a; Kleeman and Moore, 1997; Timmermann and Jin, this volume).

The ensemble average variance produced by stochastically forcing this structure (i.e. introducing the forcing $\mathbf{f}n(t)$ in the right-hand side of Eq. (5.29) can be shown to be

$$\langle E(t)\rangle = \langle \mathbf{x}^\dagger \mathbf{x}\rangle = \mathbf{f}^\dagger\, \mathbf{B}(t)\, \mathbf{f}, \tag{5.36}$$

where $\mathbf{B}(t)$ is the stochastic optimal matrix

$$\mathbf{B}(t) = \int_0^t e^{\mathbf{A}^\dagger s} e^{\mathbf{A}s}\, ds. \tag{5.37}$$

The stochastic optimal matrix satisfies the time dependent back Lyapunov equation, analogous to Eq. (5.33):

$$\frac{d\mathbf{B}}{dt} = \mathbf{BA} + \mathbf{A}^\dagger\mathbf{B} + \mathbf{I}. \tag{5.38}$$

If $\mathbf{A}$ is stable the statistical steady state $\mathbf{B}^\infty$ satisfies the algebraic equation

$$\mathbf{B}^\infty\mathbf{A} + \mathbf{A}^\dagger\mathbf{B}^\infty = -\mathbf{I}, \tag{5.39}$$

which can be readily solved for $\mathbf{B}^\infty$.

Having obtained $\mathbf{B}^\infty$ from (5.36) we obtain the stochastic optimal as the eigenfunction of $\mathbf{B}^\infty$ with the largest eigenvalue. The stochastic optimal determines the forcing structure, $\mathbf{f}$, that is most effective in producing variance. Forcings will have

P1: FYX/FGC P2: FXS
0521848822c05.xml CUUK341B-Palmer February 14, 2006 9:11

impact on the variance according to the forcing's projection on the top stochastic optimals (the top eigenfunctions of $\mathbf{B}^{\infty}$).

As another application the sensitivity of perturbation statistics to variations in the mean state can be obtained. Assume, for example, that the mean atmospheric flow U is changed by δU inducing the change $\delta\mathbf{A}$ in the mean operator. The statistical equilibrium that results satisfies the Lyapunov equation

$$(\mathbf{A}+\delta\mathbf{A})(\mathbf{C}^{\infty}+\delta\mathbf{C}^{\infty})+(\mathbf{C}^{\infty}+\delta\mathbf{C}^{\infty})+(\mathbf{A}+\delta\mathbf{A})^{\dagger}=-\mathbf{Q}\,, \tag{5.40}$$

under the assumption that the forcing covariance $\mathbf{Q}$ has remained the same. Because $\mathbf{C}^{\infty}$ satisfies the equilibrium (5.35) the first order correction $\delta\mathbf{C}^{\infty}$ is determined from

$$\mathbf{A}\delta\mathbf{C}^{\infty}+\delta\mathbf{C}^{\infty}\mathbf{A}^{\dagger}\;=\;-(\delta\mathbf{A}\mathbf{C}^{\infty}+\mathbf{C}^{\infty}\delta\mathbf{A}^{\dagger})\,. \tag{5.41}$$

From this one can determine a bound on the sensitivity of the climate by determining the change in the mean operator that will result in the largest change $\delta\mathbf{C}^{\infty}$. This operator change leading to maximum increase in a specified quadratic quantity is called the optimal structural change. Farrell and Ioannou (2004) show that a single operator change fully characterises any chosen quadratic quantity tendency, in the sense that, if an arbitrary operator change is performed, the quadratic tendency, $\delta\mathbf{C}^{\infty}$, is immediately obtained by projecting the operator change on this single optimal structure change.

In this way the sensitivity of quadratic quantities such as variance, energy, and fluxes of heat and momentum, to change in the mean operator can be found. The mean operator change could include jet velocity, dissipation and other dynamical variables, and these jet structure changes, as well as the region over which the response is optimised, can be localised in the jet. The unique jet structure change producing the greatest change in a chosen quadratic quantity also completely characterises the sensitivity of the quadratic quantity to jet change in the sense that an arbitrary jet change increases the quadratic quantity in proportion to its projection on this optimal structure change. This result provides an explanation for observations that substantial differences in quadratic storm track quantities such as variance occur in response to apparently similar influences such as comparable sea surface temperature changes, and moreover provides a method for obtaining the optimal structural change.

5.5 Prediction of statistics of uncertain systems

The sensitivity of forecasts to various aspects of the model can be determined by performing parallel computations of the forecast system in which the uncertain aspects of the model are varied. These integrations produce an ensemble of forecasts (Palmer, Kalnay *et al.*, and Buizza, this volume). The ensemble mean of these predictions is for many systems of interest a best estimate of the future state (Gauss, 1809; Leith, 1974). These ensemble integrations also provide estimates of the probability density

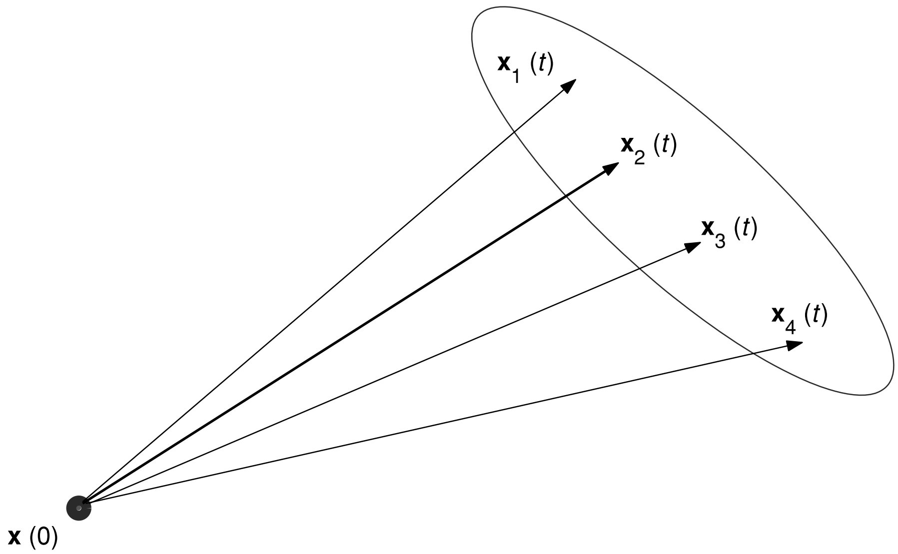

Figure 5.7 Schematic evolution of a sure initial condition $\mathbf{x}(0)$ in an uncertain system. After time t the evolved states $\mathbf{x}(t)$ lie in the region shown. Initially the covariance matrix $\mathbf{C}(0) = \mathbf{x}(0)\mathbf{x}^\dagger(0)$ is rank 1, but at time t the covariance matrix has rank greater than 1. For example if the final states were $\mathbf{x}_i(t)$ $(i = 1, \cdots, 4)$ with equal probability, the covariance at time t: $\mathbf{C}(t) = \frac{1}{4}\sum_{i=1}^{4} \mathbf{x}_i(t)\mathbf{x}_i^\dagger(t)$ would be rank 4, representing an entangled state. In contrast, in certain systems the degree of entanglement is invariant and a pure state evolves to a pure state.

function of the prediction. The covariance matrix of the predicted states $\mathbf{C} = \langle \mathbf{x}\mathbf{x}^\dagger \rangle$, where $\langle\cdot\rangle$ signifies the ensemble mean, provides the second moments of the probability density of the predictions and characterises the sensitivity of the prediction to variation in the model. We wish to determine bounds on the error covariance matrix resulting from such model uncertainties.

5.5.1 The case of additive uncertainty

Consider first a tangent linear system with additive model error. With the assumption that the model error can be treated as a stochastic forcing of the tangent linear system the errors evolve according to

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x} + \mathbf{F}\mathbf{n}(t), \tag{5.42}$$

where $\mathbf{A}(t)$ is the tangent linear operator which is considered certain, $\mathbf{F}$ the structure matrix of the uncertainty which is assumed to be well described by zero mean and unit covariance white noise processes $\mathbf{n}(t)$. Such systems are uncertain and as a result a single initial state maps to a variety of states at a later time depending on the realisation of the stochastic process $\mathbf{n}(t)$. This is shown schematically in Figure 5.7 for the case of four integrations of the model.

Let us assume initially that the ensemble had model error covariance $\mathbf{C}(0)$. At a time t later the error covariance is given by Eq. (5.34). The homogeneous part of (5.34) is the covariance resulting from the deterministic evolution of the initial $\mathbf{C}(0)$ and represents error growth associated with uncertainty in specification of the initial state. Predictability studies traditionally concentrate on this source of error growth. The inhomogeneous part of (5.34) represents the contribution of additive model error.

The deterministic part of the growth of the error covariance at any time is bounded by the growth produced by the optimal perturbation. The forced error growth at any time, by contrast, is bounded by the error covariance at time t forced by the quite different stochastic optimal that is determined as the eigenfunction with largest eigenvalue of the stochastic optimal matrix

$$\mathbf{B}(t) = \int_0^t e^{\mathbf{A}^\dagger s} e^{\mathbf{A}s}\, ds \tag{5.43}$$

at time t. Given an initial error covariance, $\mathbf{C}(0)$, and a forcing covariance, $\mathbf{Q}$, it is of interest to determine the time at which the accumulated covariance from the model error exceeds the error produced by uncertainty in the initial conditions. As an example consider the simple system (5.8). Assume that initially the state has error such that $\mathrm{trace}(\mathbf{C}(0)) = 1$ and that the additive model error has covariance $\mathrm{trace}(\mathbf{Q}) = 1$. The growth of errors due to uncertainty in the initial conditions is plotted as a function of time in Figure 5.8. After approximately unit time the error covariance is dominated by the accumulated error from model uncertainty.

From this simple example it is realised that in both stable and unstable systems as the initial state is more accurately determined error growth will inevitably be dominated by model error. At present the success of the deterministic forecasts and increase in forecast accuracy obtained by decreasing initial state error suggest that improvements in forecast accuracy are still being achieved by reducing the uncertainty in the initial state.

5.5.2 The case of multiplicative uncertainty

Consider now a forecast system with uncertain parametrisations (Palmer, 1999; Sardeshmukh *et al.*, 2001, 2003) so that the tangent linear system operator itself is uncertain and for simplicity takes the form

$$\mathbf{A}(t) = \mathbf{A} + \varepsilon \mathbf{B}\eta(t), \tag{5.44}$$

where $\eta(t)$ is a scalar stochastic process with zero mean and unit variance and $\mathbf{B}$ is a fixed matrix characterising the structure of the operator uncertainty, and ε is a scalar amplitude factor.

An important property of these multiplicative uncertain systems is that different realisations produce highly disparate growths. Fix the inner product with which the perturbation magnitude is measured and concentrate on calculation of the error

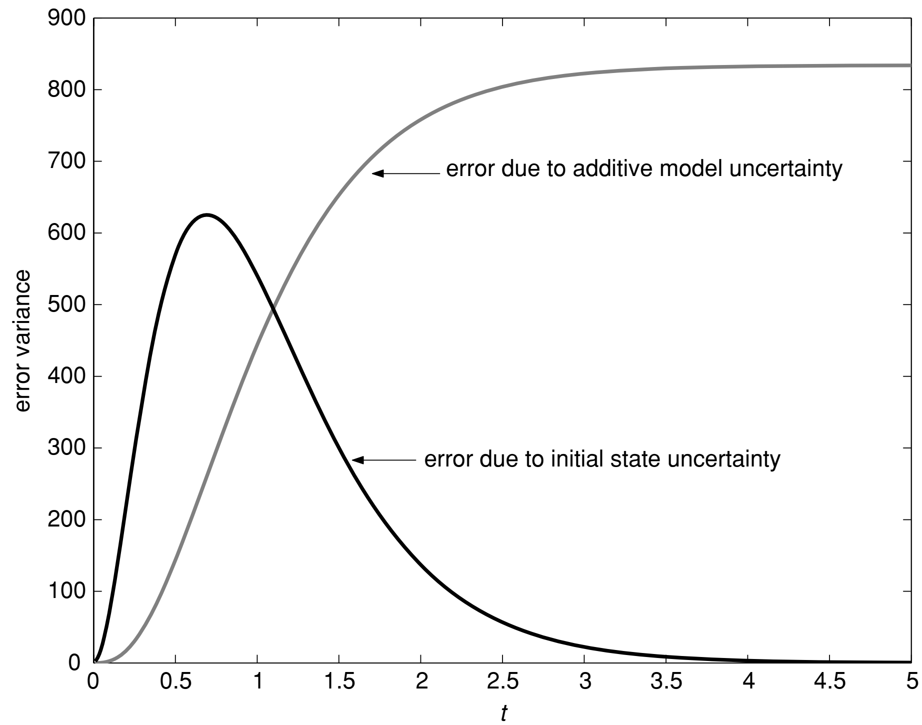

Figure 5.8 Error variance resulting from free evolution of the optimal unit variance initial covariance and evolution of error variance forced by additive uncertainty with unit forcing variance. The operator **A** is the simple 2×2 Reynolds operator given in (5.8).

growth. Because of the uncertainty in the operator, different realisations, η, will result in different perturbation magnitudes, and because the probability density function of η is known, perturbation amplitudes can be ascribed a probability. A measure of error growth is the expectation of the growths:

$$\langle g \rangle = \int P(\omega))\, g(\omega) d\omega, \tag{5.45}$$

where ω is a realisation of η, $P(\omega)$ is the probability of its occurrence, and $g(\omega)$ is the error growth associated with this realisation of the operator. Because of the convexity of the expectation the root-mean-square second moment error growth exceeds the amplitude error growth, i.e.

$$\sqrt{\langle g^2 \rangle} \geq \langle g \rangle. \tag{5.46}$$

It follows that in uncertain systems different moments generally have different growth rates and the Lyapunov exponent of an uncertain system may be negative while higher moments are unstable. This emphasises the fact that rare or extreme events that are weighted more by the higher moment measure, while difficult to predict, are potentially highly consequential.

As an example consider two classes of trajectories which with equal probability give growth in unit time of $g = 2$ or $g = 1/2$. What is the expected growth in unit time? While the Lyapunov growth rate is 0 because

$$\lambda = \frac{\langle \log g \rangle}{t} = \frac{\log 2}{2} + \frac{\log(1/2)}{2} = 0, \tag{5.47}$$

the second moment growth rate is positive:

$$\lambda_2 = \frac{\log \langle g^2 \rangle}{t} = \log\left(\frac{1}{2} \times 4 + \frac{1}{2} \times \frac{1}{4}\right) = 0.75, \tag{5.48}$$

proving that the error covariance increases exponentially fast and showing that uncertain systems may be Lyapunov (sample) stable, while higher order moments are unstable. The second moment measures include the energy, so that this example demonstrates that multiplicative uncertain systems can be Lyapunov stable while expected energy grows exponentially with time. In fact if the uncertainty is Gaussian there is always a higher moment that grows exponentially (Farrell and Ioannou, 2002a).

One implication of this property is that optimal error growth in multiplicative uncertain systems is not derivable from the norm of the ensemble mean propagator. To obtain the optimal growth it is necessary to first determine the evolution of the covariance $\mathbf{C} = \langle \mathbf{x}\mathbf{x}^\dagger \rangle$ under the uncertain dynamics and then determine the optimal $\mathbf{C}(0)$ of unit trace that leads to greatest trace($\mathbf{C}(t)$) at later times.

Consider the multiplicative uncertain tangent linear system:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \varepsilon n(t)\mathbf{B}\mathbf{x}, \tag{5.49}$$

where $\mathbf{A}$ is the sure mean operator and $\mathbf{B}$ is the structure of the uncertainty in the operator and $n(t)$ is its time dependence. Take $n(t)$ to be a Gaussian random variable with zero mean, unit variance and autocorrelation time t_c. Define $\Phi(t, 0)$ to be the propagator for a realisation of the operator $\mathbf{A} + \varepsilon\eta(t)\mathbf{B}$. For that realisation the square error at time t is

$$\mathbf{x}(t)^\dagger \mathbf{x}(t) = \mathbf{x}(0)^\dagger \Phi^\dagger (t, 0)\Phi(t, 0)\, \mathbf{x}(0) \tag{5.50}$$

where $\mathbf{x}(0)$ is the initial error. The optimal initial error, i.e. the initial error that leads to the greatest variance at time t, for this realisation is the eigenvector of

$$\mathbf{H}(t) = \Phi^\dagger (t, 0)\Phi(t, 0) \tag{5.51}$$

with largest eigenvalue.

For uncertain dynamics we seek the greatest expected variance at t by determining the ensemble average

$$\langle \mathbf{H}(t) \rangle = \left\langle \Phi^\dagger(t, 0)\Phi(t, 0) \right\rangle. \tag{5.52}$$

The optimal initial error is identified as the eigenvector of $\langle \mathbf{H}(t) \rangle$ with largest eigenvalue. This eigenvalue determines the optimal ensemble expected error growth. This gives constructive proof of the remarkable fact that there is a single sure initial error that maximises expected error growth in an uncertain tangent linear system; that is, the greatest ensemble error growth is produced when all ensemble integrations are initialised with the same state.

To quantify the procedure one needs to obtain an explicit form of the ensemble average $\langle \mathbf{H}(t) \rangle$ in terms of the statistics of the uncertainty. It turns out that this is possible for Gaussian fluctuations (Farrell and Ioannou, 2002b, 2002c), in which case $\langle \mathbf{H}(t) \rangle$ evolves according to the exact equation

$$\frac{d \langle \mathbf{H} \rangle}{dt} = (\mathbf{A} + \varepsilon^2 \mathbf{E}(t)\mathbf{B})^\dagger \langle \mathbf{H}(t) \rangle + \langle \mathbf{H}(t) \rangle (\mathbf{A} + \varepsilon^2 \mathbf{E}(t)\mathbf{B}) \tag{5.53}$$

$$+ \varepsilon^2 (\mathbf{E}^\dagger(t) \langle \mathbf{H} \rangle \mathbf{B} + \mathbf{B}^\dagger \langle \mathbf{H} \rangle \mathbf{E}(t)) \tag{5.54}$$

where

$$\mathbf{E}(t) = \int_0^t e^{-\mathbf{A}s} \mathbf{B} e^{\mathbf{A}s} e^{-\nu s} \, ds. \tag{5.55}$$

For autocorrelation times of the fluctuations which are small compared with the time scales of $\mathbf{A}$, the above equation reduces to

$$\frac{d \langle \mathbf{H} \rangle}{dt} = \left(\mathbf{A} + \frac{\varepsilon^2}{\nu} \mathbf{B}^2 \right)^\dagger \langle \mathbf{H}(t) \rangle + \langle \mathbf{H}(t) \rangle \left(\mathbf{A} + \frac{\varepsilon^2}{\nu} \mathbf{B}^2 \right) + \frac{2\varepsilon^2}{\nu} \mathbf{B}^\dagger \langle \mathbf{H} \rangle \mathbf{B}. \tag{5.56}$$

As an example application of this result, the ensemble for an uncertain tangent linear system arising from a forecast system with Gaussian statistical distribution of parameter value variation could be constructed from this basis of optimals, i.e. the optimals of $\langle \mathbf{H} \rangle$.

5.6 Conclusions

Generalised stability theory (GST) is required for a comprehensive understanding of error growth in deterministic autonomous and non-autonomous systems. In contrast to the approach based on normal modes in GST applied to deterministic systems, attention is concentrated on the optimal perturbations obtained by singular value analysis of the propagator or equivalently by repeated forward integration of the system followed by backward integration of the adjoint system. The optimal perturbations are used to understand and predict error growth and structure and for such tasks as building ensembles for use in ensemble forecast. In addition this approach provides theoretical insight into the process of error growth in both autonomous and non-autonomous systems. In the case of non-autonomous systems the process of error growth is identified with the intrinsic non-normality of a time dependent system and the unstable error is shown to lie in the subspace of the leading optimal (singular) vectors.

Beyond the deterministic time horizon, GST can be used to address questions of predictability of statistics and of sensitivity of statistics to changes in the forcing and changes in the system operator. As an example of the power of these methods, the sensitivity of a statistical quantity is found to be related to a single structured change in the mean operator.

Finally, we have seen how GST addresses error growth in the presence of both additive and multiplicative model error. In the case of multiplicative model error the role of rare trajectories is found to be important for the stability of higher statistical moments, including quadratic moments which relate sample stability to stability in energy.

Acknowledgements

This work was supported by NSF ATM-0123389 and by ONR N00014-99-1-0018.

Notes

1. It is called the Reynolds matrix because it captures the emergence of rolls in three-dimensional boundary layers that are responsible for transition to turbulence.
2. See also the related linear inverse model perspective (Penland and Magorian, 1993; Penland and Sardeshmukh, 1995).
3. Under certain conditions it can be associated with a zonal mean; for discussion and physical application of this interpretation see Farrell and Ioannou (2003).

References

Barkmeijer, J., T. Iversen and T. N. Palmer (2003). Forcing singular vectors and other sensitive model structures. *Quart. J. Roy. Meteor. Soc.*, **129**, 2401–24.

Berliner, L. M. (1996). Hierarchical Baysian time series models. In *Maximum Entropy and Bayesian Methods*, ed. K. M. Hanson and R. N. Silver, pp. 15–22. Kluwer Academic Publishers.

Buizza, R. and T. N. Palmer (1995). The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–56.

D'Andrea, F. and R. Vautard (2000). Reducing systematic errors by empirically correcting model errors. *Tellus*, **52A**, 21–41.

DelSole, T. (1996). Can quasigeostrophic turbulence be modelled stochastically? *J. Atmos. Sci.*, **53**, 1617–33.

(1999). Stochastic models of shear-flow turbulence with enstrophy transfer to subgrid scales. *J. Atmos. Sci.*, **56**, 3692–703.

(2001). A theory for the forcing and dissipation in stochastic turbulence models. *J. Atmos. Sci.*, **58**, 3762–75.

(2004). Stochastic models of quasigeostrophic turbulence. *Surv. Geophys.*, **25**, 107–49.

DelSole, T. and A. Y. Hou (1999a). Empirical stochastic models for the dominant climate statistics of a general circulation model. *J. Atmos. Sci.*, **56**, 3436–56.

(1999b). Empirical correction of a dynamical model. I: Fundamental issues. *Mon. Weather Rev.*, **127**, 2533–45.

Drazin, P. G. and W. H. Reid (1981). *Hydrodynamic Stability*. Cambridge University Press.

Dullerud, G. E. and F. Paganini (2000). *A Course in Robust Control Theory: a Convex Approach*. Springer Verlag.

Epstein, E. S. (1969). Stochastic-dynamic prediction. *Tellus*, **21**, 739–59.

Errico, R. M. (1997). What is an adjoint mode? *Bull. Am. Meteorol. Soc.*, **78**, 2577–91.

Farrell, B. F. (1988). Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.*, **45**, 163–72.

(1990). Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.*, **47**, 2409–16.

Farrell, B. F. and P. J. Ioannou (1993). Stochastic dynamics of baroclinic waves. *J. Atmos. Sci.*, **50**, 4044–57.

(1994). A theory for the statistical equilibrium energy and heat flux produced by transient baroclinic waves. *J. Atmos. Sci.*, **51**, 2685–98.

(1995). Stochastic dynamics of the mid-latitude atmospheric jet. *J. Atmos. Sci.*, **52**, 1642–56.

(1996a). Generalized stability. I: Autonomous operators. *J. Atmos. Sci.*, **53**, 2025–40.

(1996b). Generalized stability. II: Non-autonomous operators. *J. Atmos. Sci.*, **53**, 2041–53.

(1999). Perturbation growth and structure in time dependent flows. *J. Atmos. Sci.*, **56**, 3622–39.

(2000). Perturbation dynamics in atmospheric chemistry. *J. Geophy. Res.*, **105**, 9303–20.

(2001). State estimation using reduced order Kalman filter. *J. Atmos. Sci.*, **58**, 3666–80.

(2002a). Perturbation growth and structure in uncertain flows: Part I. *J. Atmos. Sci.*, **59**, 2629–46.

(2002b). Perturbation growth and structure in uncertain flows: Part II. *J. Atmos. Sci.*, **59**, 2647–64.

(2002c). Optimal perturbation of uncertain systems. *Stoch. Dynam.*, **2**, 395–402.

(2003). Structural stability of turbulent jets. *J. Atmos. Sci.*, **60**, 2101–18.

(2004). Sensitivity of perturbation variance and fluxes in turbulent jets to changes in the mean jet. *J. Atmos. Sci.*, **61**, 2645–52.

(2005). Optimal excitation of linear dynamical systems by distributed forcing. *J. Atmos. Sci.*, **62**, 460–75.

Gauss, K. F. (1809). *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*. C. H. Davis, Translation, republished by Dover Publications, 1963.

Gelaro, R., R. Buizza, T. N. Palmer and E. Klinker (1998). Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.*, **55**, 1012–37.

Gelaro, R., C. Reynolds and R. M. Errico (2002). Transient and asymptotic error growth in a simple model. *Quart J. Roy. Meteor. Soc.*, **128**, 205–28.

Ghil, M. and P. Malanotte-Rizzoli (1991). Data assimilation in meteorology and oceanography. *Adv. Geophys*., **33**, 141–266.

Kleeman, R. and A. M. Moore (1997). A theory for the limitations of ENSO predictability due to stochastic atmospheric transients. *J. Atmos. Sci.*, **54**, 753–67.

Lacarra, J. and O. Talagrand (1988). Short range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81–95.

Leith, C. E. (1974). Theoretical skill of Monte Carlo forecasts. *Mon. Weather Rev.*, **102**, 409–18.

Lorenz, E. N. (1963). Deterministic non-periodic flow. *J. Atmos. Sci.*, **20**, 130–41.

(1965). A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–33.

(1985). The growth of errors in prediction. In *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, ed. M. Ghil, pp. 243–65. North-Holland.

Lyapunov, M. A. (1907). Problème général de la stabilité du mouvement. Annales Fac. Sciences Toulouse 9. Reprinted by Princeton University Press, 1949.

Mirsky, L. (1960). Symmetric gage functions and unitarily invariant norms. *Q. J. Math*, **11**, 50–9.

Moore, A. M. and R. Kleeman (1996). The dynamics of error growth and predictability in a coupled model of ENSO. *Quart. J. Roy. Meteor. Soc.*, **122**, 1405–46.

Molteni, F. and T. N. Palmer (1993). Predictability and finite time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.*, **119**, 269–98.

Moore, A. M. and B. F. Farrell (1993). Rapid perturbation growth on spatially and temporally varying oceanic flows determined using an adjoint method: application to the Gulf Stream. *J. Phys. Ocean*., **23**, 1682–702.

Oseledets, V. I. (1968). A multiplicative ergodic theorem. Lyapunov characteristic exponents. *Trans. Moscow Math. Soc.*, **19**, 197–231.

Palmer, T. N. (1999). A nonlinear dynamical perspective on climate prediction. *J. Climate*, **12**, 575–91.

Palmer, T. N., R. Gelaro, J. Barkmeijer and R. Buizza (1998). Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, **55**, 633–53.

Penland, C. and T. Magorian (1993). Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Climate*, **6**, 1067–76.

Penland, C. and P. D. Sardeshmukh (1995). The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999–2004.

Sardeshmukh, P. D., C. Penland and M. Newman (2001). Rossby waves in a stochastically fluctuating medium. *Prog. Probability*, **49**, 369–84.

(2003). Drifts induced by multiplicative red noise with application to climate. *Europhys. Lett.*, **63**, 498–504.

Schmidt, E. (1906). Zur Theorie der linearen und nichtlinearen Intgralgleighungen. *Math. Annalen* **63**, 433–476.

Whitaker, J. S. and P. D. Sardeshmukh (1998). A linear theory of extratropical synoptic eddy statistics. *J. Atmos. Sci.*, **55**, 237–58.

Zhang, Y. and I. Held (1999). A linear stochastic model of a GCM's midlatitude storm tracks. *J. Atmos. Sci.*, **56**, 3416–35.