Tellus (2003), 55A, 208–218

TELLUS
ISSN 0280–6495

Coordinate invariance in stochastic dynamical systems

By JEFFREY B. WEISS*, *Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO 80309, USA*

(Manuscript received 27 May 2002; in final form 7 November 2002)

ABSTRACT

Stochastic dynamical systems have been used to model a broad range of atmospheric and oceanic phenomena. Previous work has focused on the stochastic differential equation formulation of these systems, has largely remained in a single coordinate system, and has highlighted the role of non-normality of the deterministic dynamics. Here, the coordinate independent properties of stochastic dynamical systems are studied. The properties previously attributed to non-normality, which can be removed by a coordinate transformation, are more fundamentally seen to be coordinate-dependent manifestations of violations of detailed balance. Systems violating detailed balance can both amplify and rectify the random forcing. New coordinate-invariant measures of noise amplification are introduced and shown to achieve their lower bound when detailed balance is satisfied. Rectification results in a coherent phase space velocity which gives rise to a structured nonzero flux of all physically important quantities such as energy and momentum. The qualitative and quantitative features of these fluxes provide new predictions which can be used to further validate previously proposed stochastic models of geophysical systems.

1. Introduction

The idea of modeling a complex physical system with a random process is several centuries old and remains attractive today. By treating part of a system as random, one can avoid following its complex, detailed behavior and focus instead on its statistics. The result is an enormous simplification. This strategy is most successful when the dynamics can be divided into a physically interesting slow component forced by a fast component with high-dimensional chaotic dynamics. One then considers the fast component as a random perturbation on the deterministic dynamics of the slow component.

The use of stochastic dynamics in atmospheric systems goes back several decades to work by Epstein (1969) and Leith (1974). More recently the stochastic paradigm has been applied to a number of geophysical systems. A partial list includes baroclinic waves (Farrell and Ioannou, 1993; 1994), El Niño dynamics (Penland and Magorian, 1993; Penland and Sardeshmukh, 1995; Moore and Kleeman, 1996), the Gulf Stream (Moore and Farrell, 1993) quasi-geostrophic jets (DelSole and Farrell, 1996), low-frequency variability (Newman et al., 1997), synoptic eddies (Whitaker and Sardeshmukh, 1998), climate statistics (DelSole and Hou, 1999), storm tracks (Zhang and Held, 1999), ocean gyres (Moore, 1999), atmospheric chemistry (Farrell and Ioannou, 2000), and atmospheric angular momentum (Weickmann et al., 2000).

The recent interest in stochastic systems arises in part from the discovery that deterministically stable systems can amplify random noise if the deterministic part of the dynamics is non-normal (Farrell and Ioannou, 1996a,b). Prior to this observation, it was commonly thought that the large variability seen in geophysical systems could only be maintained by either large external forcing, random or otherwise, or internal instabilities of the mean state. Amplification of small perturbations, however, provides another mechanism for large variability. The broad use of stochastic models means that a better understanding of the theory

*e-mail: jeffrey.weiss@colorado.edu

of stochastic dynamical systems will impact a large part of atmospheric and oceanic science.

Despite previous successes, a fundamental question about the role of non-normality remains unanswered. Noise amplification occurs when the linearized deterministic dynamics is non-normal. However, a non-normal matrix can be made normal by an appropriate coordinate change. According to long-standing principles, a fundamental property of a system should not depend on an arbitrary choice of coordinates. A related issue is that analyses of geophysical systems are carried out in a variety of coordinate systems. The diversity of coordinate systems often reflects attempts to focus on different aspects of the same phenomena. It is thus important to understand how stochastic dynamical systems behave under coordinate transformations.

There are two complementary ways to view stochastic dynamical systems. The first view is that of stochastic ordinary differential equations (SDEs). In an SDE, the state of the system is described by a differential equation which contains random forcing, and one naturally writes equations for the moments of the trajectories. In a system with Gaussian statistics, all the information is contained in the first two moments.

The second view of stochastic dynamical systems is called the Fokker–Planck view. Here, one focuses on the probability distribution function (pdf) of the system and obtains a linear deterministic partial differential equation, the Fokker–Planck equation (FPE), for its evolution. This has two advantages: (1) the FPE is deterministic as opposed to the random SDE, and (2) the FPE is linear in the pdf, rather than the possibly nonlinear SDE. The price, however, is that the FPE is a partial differential equation instead of the simpler SDE, which is an ordinary differential equation. These trade-offs mean that some questions are more naturally answered in the SDE view, while others are more naturally answered in the FPE view.

Most previous applications of stochastic dynamics to geophysical systems have taken the SDE view. Here, we consider the FPE view as well. This leads to previously unexplored, coordinate invariant properties of the system, and to new predictions that, for the most part, have not been tested against the phenomena being modeled. It should be noted that in an information-theoretic sense an FPE contains no information not already contained in its associated SDE. However, as we see below, looking at the FPE leads one to investigate issues that are not obvious in the SDE view. The converse is also no doubt true, but, so far, the SDE has received the majority of the attention.

In this paper we focus on the simplest multivariate stochastic dynamical system: linear deterministic dynamics perturbed by additive Gaussian white noise. Throughout the paper we will mostly use vector and matrix notation, where a single symbol describes multidimensional quantities. Occasionally, however, it is more convenient to use index notation, where the components of the vectors and matrices are explicitly identified, and summation over repeated indices is assumed. We will switch between these notations where appropriate without comment.

2. Stochastic differential equations

In this section we review the mathematics of stochastic differential equations (SDEs), also called Langevin equations. More detail can be found in a number of sources (Gardiner, 1983; Farrell and Ioannou, 1996a,b). Although fluid systems such as the atmosphere and ocean are fundamentally described by partial differential equations, it is often convenient (and necessary for numerical modeling) to discretize the system to a finite number of dimensions, resulting in an ordinary differential equation. We take such truncated ordinary differential equations as our starting point.

The SDE for a real stable linear deterministic dynamical system with additive Gaussian white noise is

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \mathbf{A}\,\boldsymbol{x}(t) + \mathbf{F}\boldsymbol{\xi}(t), \quad (1)$$

where $\boldsymbol{x}\ (t)$ is the real N dimensional vector describing the state of the system, $\mathbf{A}$ is the real $N \times N$-dimensional matrix describing the linear deterministic dynamics, $\boldsymbol{\xi}(t)$ is a real N dimensional vector describing the noise process, and $\mathbf{F}$ is the so-called forcing function, a real $N \times N$-dimensional matrix describing how the noise impacts the state vector. For the dynamical system to have stable solutions, $\mathbf{A}$ must be stable, i.e., all of the eigenvalues of $\mathbf{A}$ must have negative real parts. The requirement that $\boldsymbol{\xi}$ be Gaussian white noise implies that it is completely described by its first two moments,

$$\langle \boldsymbol{\xi}(t) \rangle = 0, \quad \langle \boldsymbol{\xi}(t)\boldsymbol{\xi}^T(t') \rangle = \mathbf{I}\delta(t - t'), \quad (2)$$

where $\mathbf{I}$ is the $N \times N$ identity matrix, and the superscript T denotes the transpose. The requirement that the covariance of $\boldsymbol{\xi}$ be proportional to the identity is completely general, since any other covariance can be absorbed into the forcing function $\mathbf{F}$. We restrict ourselves to real SDEs because the associated FPE is simpler. Since an N-dimensional complex SDE can always be transformed into a $2N$-dimensional real SDE the case of real SDEs is completely general.

The solution of the SDE (1) is easily seen to be

$$\boldsymbol{x}(t) = \mathrm{e}^{\mathbf{A}t}\boldsymbol{x}(0) + \int_0^t \mathrm{d}s \mathrm{e}^{\mathbf{A}(t-s)}\mathbf{F}\boldsymbol{\xi}(s), \tag{3}$$

which gives the positive definite covariance matrix $\mathbf{C} = \langle \boldsymbol{x}\boldsymbol{x}^T \rangle$

$$\mathbf{C}(t) = \mathrm{e}^{\mathbf{A}t}\mathbf{C}(0)\mathrm{e}^{\mathbf{A}^T t} + \int_0^t \mathrm{d}s \mathrm{e}^{\mathbf{A}(t-s)}\mathbf{F}\mathbf{F}^T \mathrm{e}^{\mathbf{A}^T(t-s)}, \tag{4}$$

whose time evolution is given by

$$\frac{\mathrm{d}\mathbf{C}(t)}{\mathrm{d}t} = \mathbf{A}\mathbf{C}(t) + \mathbf{C}(t)\mathbf{A}^T + 2\mathbf{D}, \tag{5}$$

where the diffusion matrix $\mathbf{D}$ is defined to be $\mathbf{D} = \mathbf{F}\mathbf{F}^T/2$. Since the covariance only depends on $\mathbf{F}\mathbf{F}^T$, it is invariant under orthogonal transformations of the noise, $\hat{\boldsymbol{\xi}} = \mathbf{R}\boldsymbol{\xi}$, where $\mathbf{R}$ is an orthogonal matrix, $\mathbf{R}^T = \mathbf{R}^{-1}$.

In the limit $t \to \infty$ the system reaches equilibrium and the moments become independent of time. Since the focus of this paper is on the equilibrium state, we will not use any special notation to indicate equilibrium quantities; any quantity without explicit time dependence will denote the equilibrium value. Since the deterministic equilibrium $\boldsymbol{x} = 0$ is stable the mean equilibrium position is $\langle \boldsymbol{x} \rangle = 0$. From eq. (5) the equilibrium covariance satisfies

$$\mathbf{A}\mathbf{C} + \mathbf{C}\mathbf{A}^T + 2\mathbf{D} = 0. \tag{6}$$

Equation (6) is one of the primary equations in this paper. A matrix equation with this structure is called a Lyapunov equation and can be solved with a variety of numerical techniques (Lu and Wachspress, 1991). Physically, eq. (6) is called the fluctuation–dissipation relation (FDR) because it relates the equilibrium fluctuations ($\mathbf{C}$) and the deterministic dissipation ($\mathbf{A}$). These two names are not synonymous; in one dimension the FDR is not a Lyapunov equation but a simple scalar equation. Because we wish to emphasize the physical content of the equation we will refer to eq. (6) as the FDR.

When $\mathbf{A}$ is stable, all eigenvalues have negative real parts and all eigenvectors decay exponentially. If the eigenvectors are orthogonal, then all initial conditions decay exponentially and no amplification is possible. However, when $\mathbf{A}$ does not commute with its transpose, $\mathbf{A}\mathbf{A}^T - \mathbf{A}^T\mathbf{A} \neq 0$, it is called non-normal and its eigenvectors are not orthogonal (Horn and Johnson, 1985). In this case interference between the exponentially decaying eigenvectors can cause transient amplification. Typically, fluid systems such as the atmosphere and ocean give non-normal operators when linearized about physically relevant mean states, and thus exhibit transient amplification.

In a continuously forced system, such as the SDE studied here, the continuous excitation by random noise results in the system maintaining a finite variance in equilibrium. Previous theoretical work has focused on the case where the diffusion matrix is proportional to the identity matrix, e.g., Ioannou (1995) and Farrell and Ioannou (1996a). In such systems with normal deterministic dynamics, all excitations decay and the equilibrium variance of the system is of the same order as the size of the noise. In such systems with non-normal deterministic dynamics, the continuous amplification of the random noise can result in an equilibrium variance much larger than the size of the noise (Ioannou, 1995).

One difficulty with this view is that a non-normal system can always be made normal by an appropriate choice of norm, or equivalently, by an appropriate coordinate change. Thus, transient amplification and increased equilibrium variance depend solely on an arbitrary choice of coordinates. According to long-standing principles, fundamental properties of a system should be independent of such arbitrary choices and, in this view, non-normality should not influence the basic properties of the dynamics.

3. The Fokker–Planck equations

An ensemble of systems governed by the same SDE can be described by their pdf, $p(\boldsymbol{x}, t)$, where $p(\boldsymbol{x}, t)\,\mathrm{d}^N x$ is the probability of finding a member of the ensemble in a phase space volume $\mathrm{d}^N x$ around the state $\boldsymbol{x}$ at time t. The equation describing the evolution of p, called the Fokker–Planck equation, can be derived from the SDE governing the system (Gardiner, 1983; Risken, 1984; van Kampen, 1992). The Fokker–Planck equation corresponding to the SDE (1) is

$$\frac{\partial p(\boldsymbol{x}, t)}{\partial t} = -\nabla[\mathbf{A}\boldsymbol{x}\, p(\boldsymbol{x}, t)] + \nabla\mathbf{D}\nabla\, p(\boldsymbol{x}, t). \tag{7}$$

In the previous section we saw that the covariance is independent of orthogonal transformations of the noise. Since the FPE depends only on $\mathbf{D}$ and not on $\mathbf{F}$ itself, all properties of the stochastic system are independent of such transformations.

The FPE can be rewritten in terms of a velocity,

$$\frac{\partial p(\boldsymbol{x}, t)}{\partial t} + \nabla \cdot [\boldsymbol{u}(\boldsymbol{x}, t) p(\boldsymbol{x}, t)] = 0, \tag{8}$$

where

$$\boldsymbol{u}(\boldsymbol{x}, t) = \mathbf{A}\,\boldsymbol{x} - \mathbf{D}\nabla \ln p(\boldsymbol{x}, t). \tag{9}$$

The natural logarithm in the above equation arises from factoring a p from the diffusion term in the FPE, resulting in a factor of $(1/p)\,\nabla p$. This continuity equation form of the FPE expresses the fact that probability is neither created nor destroyed, but is advected through phase space by $\boldsymbol{u}$. The phase space velocity $\boldsymbol{u}$ has two parts: the velocity produced by the deterministic dynamics, $\mathbf{A}\boldsymbol{x}$, which is independent of the pdf; and a velocity arising from down-gradient diffusion, $-\mathbf{D}\nabla$ $\ln p(\boldsymbol{x}, t)$, which depends on p. The quantity $\boldsymbol{J}(\boldsymbol{x}, t) = \boldsymbol{u}\,(\boldsymbol{x}, t)\; p(\boldsymbol{x}, t)$ is called the probability current.

In equilibrium the pdf is independent of time, and thus the divergence of the equilibrium probability current is zero. For linear systems with additive Gaussian noise, the equilibrium pdf is itself Gaussian,

$$p(\boldsymbol{x}) = \mathcal{N}(\mathbf{Q})\mathrm{e}^{-\boldsymbol{x}^T \mathbf{Q}\boldsymbol{x}/2}, \tag{10}$$

where $\mathcal{N}(\mathbf{Q})$ is the normalization factor and the positive definite matrix $\mathbf{Q}$ is the inverse of the covariance matrix, $\mathbf{Q} = \mathbf{C}^{-1}$. Gradients of p are given by $\nabla p(\boldsymbol{x}) = -\mathbf{Q}\boldsymbol{x}\,p(\boldsymbol{x})$, and the phase space velocity is then linear in $\boldsymbol{x}$, $\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{\Omega}\boldsymbol{x}$, where $\boldsymbol{\Omega} \equiv (\mathbf{A} + \mathbf{DQ})$ is a matrix of frequencies. The equilibrium FPE requires

$$\mathrm{Trace}(\boldsymbol{\Omega})\,p(\boldsymbol{x}) - \boldsymbol{x}^T \mathbf{Q}\Omega\boldsymbol{x}\,p(\boldsymbol{x}) = 0. \tag{11}$$

Since the two terms in eq. (11) have different powers of $\boldsymbol{x}$, and the equation must hold for all $\boldsymbol{x}$, each term must be zero separately. One can show that the second term is equivalent to the FDR (6), and that the second term being zero implies the first term is also zero. Thus, the equilibrium solution of the FPE is precisely that pdf which satisfies the FDR

4. Coordinate transformations

We now investigate how stochastic dynamical systems transform under different norms, or equivalently, in different coordinate systems. A positive definite matrix $\mathbf{M}$ defines an inner product $(\boldsymbol{x}, \boldsymbol{x}) \equiv \boldsymbol{x}^T\mathbf{M}\boldsymbol{x}$, and an associated vector norm $||\boldsymbol{x}|| = (\boldsymbol{x}^T\mathbf{M}\boldsymbol{x})^{1/2}$. Since $\mathbf{M}$ is positive definite it can be decomposed into $\mathbf{M} = \mathbf{R}^T\mathbf{R}$, where $\mathbf{R}$ is non-singular. This defines a new vector $\hat{\boldsymbol{x}} = \mathbf{R}\boldsymbol{x}$, and a new, Euclidean, norm, $||\hat{\boldsymbol{x}}|| = (\hat{\boldsymbol{x}}^T\hat{\boldsymbol{x}})^{1/2} = (\boldsymbol{x}^T\mathbf{M}\boldsymbol{x})^{1/2} = ||\boldsymbol{x}||$. Thus, considering different norms in one coordinate system is equivalent to considering different coordinate systems with a single norm. Here we use the Euclidean norm and consider the effect of coordinate transformations.

Let $\hat{\boldsymbol{x}} = \mathbf{R}\boldsymbol{x}$ be a new coordinate system obtained by some transformation matrix $\mathbf{R}$. For simplicity, we restrict ourselves to non-singular transformations, which is consistent with associating coordinate transformations and norms. To consider transformations which project onto a subspace of the system, and are associated with semi-norms, one would have to generalize the current work. One can perform coordinate transformations within the FPE Gardiner (1983); Risken (1984), but it is simpler to first transform the SDE and then study the associated FPE. In $\hat{\boldsymbol{x}}$ coordinates, the SDE (1) becomes

$$\frac{\mathrm{d}\hat{\boldsymbol{x}}(t)}{\mathrm{d}t} = \hat{\mathbf{A}}\hat{\boldsymbol{x}}(t) + \hat{\mathbf{F}}\boldsymbol{\xi}(t), \tag{12}$$

where $\hat{\mathbf{A}} = \mathbf{RAR}^{-1}$, and $\hat{\mathbf{F}} = \mathbf{RF}$. The transformed diffusion matrix is then $\hat{\mathbf{D}} = \hat{\mathbf{F}}\hat{\mathbf{F}}^T/2 = \mathbf{RDR}^T$. The transformed covariance matrix, $\hat{\mathbf{C}} = \langle\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^T\rangle = \mathbf{RCR}^T$, is the solution to the transformed FDR, and its inverse is $\hat{\mathbf{Q}}$. In the new coordinates the phase space velocity is $\hat{\boldsymbol{u}} = \hat{\Omega}\hat{\boldsymbol{x}} = \mathbf{R}\boldsymbol{u}$, where $\hat{\Omega} = \hat{\mathbf{A}} + \hat{\mathbf{D}}\hat{\mathbf{Q}} = \mathbf{R}\Omega\mathbf{R}^{-1}$.

The choice of coordinate system for studying a particular physical system is highly subjective. One would thus like to focus on properties that are coordinate invariant and consider their manifestations in various coordinate system. Here we consider several common choices for coordinate systems.

One common analysis tool is to study a system's empirical orthogonal functions (EOFs). The EOFs of a system are the eigenvectors of the covariance matrix $\mathbf{C}$. In EOF coordinates $\mathbf{C}$ and its inverse $\mathbf{Q}$ are diagonal. The surfaces of constant probability in the equilibrium pdf are nested N dimensional ellipsoids whose axes are the EOFS.

EOF coordinates have the disadvantage that the dominant EOF can obscure the behavior of EOFs with smaller variance. It can thus be advantageous to scale each EOF by the square root of its variance, resulting in a coordinate system where each degree of freedom has the same variability. In this coordinate system the variance is unity in all directions, $\mathbf{C} = \mathbf{Q} = \mathbf{I}$, and surfaces of constant probability in the pdf are N

dimensional spheres. We will call this coordinate system identity–covariance coordinates. Here the FDR is $\mathbf{A} + \mathbf{A}^T + 2\mathbf{D} = 0$, or $\mathbf{D} = -\mathbf{A}_{\text{sym}}$, where $\mathbf{A}_{\text{sym}}$ is the symmetric part of $\mathbf{A}$. The frequency matrix which defines the phase space velocity is skew-symmetric, $\Omega = \mathbf{A} + \mathbf{DQ} = \mathbf{A}_{\text{skew}}$, where $\mathbf{A}_{\text{skew}} \equiv (\mathbf{A} - \mathbf{A}^T)/2$ is the skew-symmetric part of $\mathbf{A}$.

If one wishes to focus on the eigenvectors of the deterministic dynamics, it is natural to use coordinates where $\mathbf{A}$ is normal. In normal coordinates the eigenvectors of $\mathbf{A}$ are orthogonal and perturbations strictly decay.

Much of the previous theoretical work considered the coordinate system where the noise forcing function is the identity, $\mathbf{F} = \mathbf{I}$ and $\mathbf{D} = \frac{1}{2}\mathbf{I}$. This coordinate system, which we will call identity-noise coordinates, is the natural choice if one has no knowledge of the structure of the noise. If, however, the noise arises from an underlying deterministic chaotic process, one may want to include the structure of the correlations in the noise. For example, in stochastic El Niño models, where the fast chaotic atmospheric forcing of tropical Pacific sea surface temperatures is modeled as random noise, the coordinates used by previous investigators are EOF coordinates, and the diffusion matrix is not proportional to the identity matrix (Penland and Sardeshmukh, 1995).

5. Detailed Balance

In Section 3 we have seen that equilibrium requires the phase space velocity to be divergence free. This can be achieved in two ways: (1) either $\boldsymbol{u} \equiv 0$, or (2) $\boldsymbol{u} \neq 0$ but $\nabla \cdot \boldsymbol{u} = 0$. The first possibility, $\boldsymbol{u} \equiv 0$, is a necessary and sufficient condition for the physical situation called detailed balance (Gardiner, 1983; Risken, 1984). When a system is in detailed balance, the probability of a transition between any two states $S_1 \rightarrow S_2$ is the same as the probability of the reverse transition $S_2 \rightarrow S_1$. The second possibility, $\boldsymbol{u} \neq 0$, occurs when detailed balance is violated. The two possibilities are drawn schematically in Fig. 1. Considering the two terms that make up $\boldsymbol{u}$ [eq. (9)], detailed balance occurs when the deterministic component of the velocity exactly balances the down-gradient diffusion due to noise. When a system in equilibrium violates detailed balance the random noise is rectified into a nonzero probability current.

We now explore the relation between detailed balance and the matrices defining the dynamical system. When $\boldsymbol{u} = 0$ then, from the definition of $\boldsymbol{u}$ [eq. (9)],

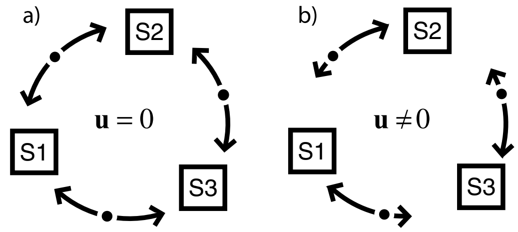

Fig. 1. Schematic diagram of a system with three states in equilibrium indicating the relationship between transition probabilities and detailed balance. The size of the transition probability is indicated by the length of the arrow connecting the two states. (a) Forward and reverse transition probabilities are equal, detailed balance is satisfied, and $\boldsymbol{u} = 0$. (b) Forward and reverse transition probabilities are not equal, detailed balance is violated, and $\boldsymbol{u}$ rotates clockwise.

$$\mathbf{D}^{-1}\mathbf{A}\boldsymbol{x} = \nabla \ln p, \tag{13}$$

which requires that the vector $\mathbf{D}^{-1}\mathbf{A}\boldsymbol{x}$ be the gradient of some scalar function. For a vector $\boldsymbol{v}$ to be a gradient it must obey $\partial_i v_j - \partial_j v_i = 0$, which, in this case, becomes

$$\mathbf{AD} - \mathbf{DA}^T = 0. \tag{14}$$

One can also show that the converse is true, and thus $\mathbf{AD} - \mathbf{DA}^T = 0$ if and only if the system satisfies detailed balance.

The phase space velocity that arises when detailed balance is violated has some interesting properties. First, because $\boldsymbol{u}^T \cdot \nabla p = 0$, the phase space velocity is directed along surfaces of constant probability at all points in phase space. Further, one can show from the FDR that the eigenvalues of $\boldsymbol{\Omega}$ must have zero real part. Thus, $\boldsymbol{u}$ is a pure rotation along surfaces of constant probability.

Now consider the effect of coordinate changes. In coordinates $\hat{\boldsymbol{x}} = \mathbf{R}\boldsymbol{x}$, the condition for detailed balance, eq. (14), becomes

$$\hat{\mathbf{A}}\hat{\mathbf{D}} - \hat{\mathbf{D}}\hat{\mathbf{A}}^T = \mathbf{R}(\mathbf{AD} - \mathbf{DA}^T)\mathbf{R}^T = 0. \tag{15}$$

Thus, whether or not detailed balance is satisfied is coordinate invariant (provided $\mathbf{R}$ is nonsingular). Similarly, since $\hat{\boldsymbol{u}} = \mathbf{R}\boldsymbol{u}$, whether or not the velocity is zero is also coordinate invariant.

Considering coordinate changes allows another interpretation of detailed balance. One can show that satisfying detailed balance is equivalent to the existence of a coordinate transformation which simultaneously diagonalizes $\mathbf{A}$ and $\mathbf{D}$. If $\mathbf{A}$ and $\mathbf{D}$ are both diagonal, then the original N-dimensional system reduces to N uncoupled one-dimensional systems. Thus, systems violating detailed balance are intrinsically multi-dimensional, while systems satisfying detailed balance are essentially a collection of one-dimensional systems.

Many previous studies of stochastic models of geophysical systems, where the emphasis was on the non-normality of $\mathbf{A}$, considered only identity–noise coordinates [e.g., (Farrell and Ioannou, 1996a) and Ioannou (1995)]. In this case, satisfying detailed balance implies that $\mathbf{A}$ is symmetric and hence normal. Thus, the systems previously considered, where $\mathbf{D} = \frac{1}{2}\mathbf{I}$ and $\mathbf{A}$ is non-normal, necessarily violated detailed balance. Here we see that the fundamental coordinate invariant property underlying these systems is not non-normality, but the violation of detailed balance.

It is important to note that, in general, normality of the deterministic component of the dynamics, expressed by the normality of $\mathbf{A}$, and the property of detailed balance are completely independent. By choosing an appropriate noise forcing function $\mathbf{F}$ that gives a diffusion matrix $\mathbf{D}$ which satisfies eq. (14), one can satisfy detailed balance with non-normal $\mathbf{A}$. On the other hand, by transforming a non-normal system in identity–noise coordinates, which violates detailed balance, to normal coordinates, one obtains a system which violates detailed balance but has normal $\mathbf{A}$. Even in identity–noise coordinates detailed balance and normality are not equivalent: here systems satisfying detailed balance must be normal, but systems violating detailed balance can be either normal or non-normal. The important point which makes detailed balance coordinate invariant but non-normality of $\mathbf{A}$ coordinate dependent is that coordinate changes transform both the deterministic dynamics and the noise forcing.

6. Energy

In the SDE view of a stochastic system average quantities are obtained by averaging over realizations of the random noise. In the FPE view averages are obtained by averaging over phase space. Thus, in the SDE view averages combine information about systems at different points in phase space, while in the FPE view the dependence of averages on the state of the system arises naturally.

The generalized energy of a dynamical system is the Euclidean norm of the state vector, $E = \boldsymbol{x}^T \cdot \boldsymbol{x}$. This generalized energy corresponds to the physical energy when $\boldsymbol{x}$ is in the appropriate coordinate system. As noted by previous authors, the energy is given by the trace of the covariance matrix, $E = \text{Trace}\ (\mathbf{C})$. In the SDE view the mean energy is $E = \langle \boldsymbol{x}^T \cdot \boldsymbol{x} \rangle$ where the average is over all realizations of the noise and over systems at different points in phase space. In the FPE view the energy is given by an average over phase space,

$$E = \int \mathrm{d}^N x \boldsymbol{x}^T \cdot \boldsymbol{x}\, p(\boldsymbol{x}, t) \tag{16}$$

which naturally gives rise to the phase space energy density $\mathcal{E}(\boldsymbol{x}, t) \equiv \boldsymbol{x}^T \cdot \boldsymbol{x}\, p(\boldsymbol{x}, t)$.

The time evolution of the energy density can be written using the continuity form of the FPE (8) to obtain

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot (\boldsymbol{u}\mathcal{E}) = \mathcal{E}_\mathrm{s}, \tag{17}$$

where $\mathcal{E}_\mathrm{s} = 2(\boldsymbol{x}^T \cdot \boldsymbol{u})p = 2(\boldsymbol{x}^T \boldsymbol{\Omega} \boldsymbol{x})p$ is the energy density source. The quantity $\boldsymbol{u}\mathcal{E}$ can be interpreted as an energy flux in phase space. In the following we will refer to $\mathcal{E}_\mathrm{s}$ as a source, with the understanding that a negative source is an energy density sink. From the two terms that make up $\boldsymbol{u}$ [eq. (9)] we see that there is a source from the deterministic dynamics, $2(\boldsymbol{x}^T \mathbf{A} \boldsymbol{x})p$, and a source from the noise, $2(\boldsymbol{x}^T \mathbf{DQ}\, \boldsymbol{x})p$. Note that even though the deterministic dynamics is stable and the eigenvalues of $\mathbf{A}$ have negative real part, its contribution to the energy source can be either positive or negative. Similarly, even though $\mathbf{D}$ and $\mathbf{Q}$ are both positive definite, the noise induced energy source can also be either a source or a sink. Thus, eq. (17) says that the energy density is advected through phase space by the phase space velocity and modified by a deterministic energy source and a noise induced energy source.

When the system is in equilibrium the energy density must be constant in time and $\nabla \cdot (\boldsymbol{u}\mathcal{E}) = 2\boldsymbol{x}^T \cdot \boldsymbol{u}\, p$. Thus, a system which satisfies detailed balance in equilibrium, i.e., $\boldsymbol{u} = 0$, has no energy flux and no energy sources. A system that violates detailed balance necessarily has a nonzero energy flux in phase space whose divergence is balanced by energy sources. Because the integral over phase space of the flux must be zero, locations with sources must balance locations with sinks. If the energy sources are identically zero there is a nonzero divergence-free energy flux.

The quantitative value of the energy flux $\boldsymbol{u}\mathcal{E}$ will depend on the coordinate system, but its presence or absence, like the detailed balance condition, is invariant under coordinate transformations. Thus, a coordinate invariant property of the previously studied non-normal systems is that they have a nonzero phase space energy flux. The presence of sources and sinks, on the other hand, is coordinate dependent. In identity–covariance coordinates $\boldsymbol{\Omega}$ is skew-symmetric and $\mathcal{E}_s = 0$.

One can find the locations in phase space where the energy density and its sources are maximized or minimized. Considering $\boldsymbol{x}^T \cdot \nabla\mathcal{E} = 0$ shows that the energy density maxima are on the two standard deviation probability surface where $\boldsymbol{x}^T \mathbf{Q}\boldsymbol{x} = 2$. Similarly, the extrema of the energy density sources also occur on the two standard deviation probability surface. Thus, these extrema occur when the system is in a relatively unlikely state corresponding to a two standard deviation event.

More generally, any quantity that is obtained by an average over phase space will display similar properties. Consider some property $B = \int d^N x b(\boldsymbol{x})\ p(\boldsymbol{x}, t)$ with density $\mathcal{B} \equiv b(\boldsymbol{x})p(\boldsymbol{x}, t)$. Then the density is governed by an equation $\partial_t \mathcal{B} + \nabla \cdot (\boldsymbol{u}\mathcal{B}) = \mathcal{B}_s$, where the source is given by $\mathcal{B}_s = (\boldsymbol{u}^T \cdot \nabla b)p$. In a system where detailed balance is satisfied the flux $\boldsymbol{u}\mathcal{B}$ and the source $\mathcal{B}_s$ are zero, while when detailed balance is violated the system supports a flux and possible sources. Further, if $b \sim x^n$, then the extrema of the density and the density sources occur on the probability surface corresponding to n standard deviations, i.e., at locations $\boldsymbol{x}$ which satisfy $\boldsymbol{x}^T \mathbf{Q}\boldsymbol{x} = n$. Thus, a system violating detailed balance must have nonzero phase space fluxes of all quantities for which one can define a phase space density: energy, momentum, heat, etc.

The nonzero fluxes resulting from violations of detailed balance provide new, mostly unexplored, predictions of previously studied stochastic models of geophysical systems. While Farrell and Ioannou (1994), for example, did study the heat flux in stochastically excited baroclinic waves, it has not been previously noted how general these fluxes are and they have not been studied in most applications. It should be emphasized that no additional assumptions are needed to obtain these fluxes: they are required in any system violating detailed balance. Testing the predicted flux of energy and other quantities against observations will provide new evidence for the success or failure of stochastic models of atmospheric and oceanic phenomena.

7. Noise amplification

In previous work (Ioannou, 1995) it has been shown that stochastic dynamical systems in identity–noise coordinates with non-normal deterministic components can amplify the random noise and, as a result, the system can sustain a relatively large variance. Here we investigate the relationship between noise amplification in equilibrium and coordinate transformations.

In a one-dimensional system it is straightforward to define the gain g as the ratio of the maintained variance to the diffusion coefficient, $g = C/D = -1/A$, where all quantities are scalars. For multi-dimensional systems the relevant quantities are matrices and there is no single obvious definition of the gain. Ioannou (1995) used the energy, $E = \text{Trace}\ (\mathbf{C})$, which is a reasonable choice when the noise is fixed. Here we consider coordinate transformations which can possibly rescale the noise. The FDR (6) shows that for a fixed deterministic dynamics the covariance matrix scales with the diffusion matrix. This simply reflects the fact that larger noise gives a proportionally larger response. Thus, any measure of gain must, in some sense, divide by the diffusion matrix. One could define the gain as the ratio of the system energy to the noise energy, Trace $(\mathbf{C})$/ Trace $(\mathbf{D})$, but this quantity is only invariant under orthogonal coordinate transformations.

We choose to define the gain matrix $\mathbf{G} \equiv \mathbf{C}\mathbf{D}^{-1}$, which is one obvious generalization of the gain in one-dimensional systems. The eigenvalues of $\mathbf{G}$ are all positive, but it is not necessarily symmetric, and hence not necessarily positive definite. Another possible generalization is $\mathbf{D}^{-1}\mathbf{C} = \mathbf{G}^T$. We will focus on scalar measures of noise amplification which give the same result for $\mathbf{G}$ and $\mathbf{G}^T$ and are coordinate invariant. The two obvious candidates are $g^{tr} \equiv$ Trace $(\mathbf{G})$ and $g^{det} \equiv$ Det $(\mathbf{G})$. Under coordinate transformations $\hat{\mathbf{G}} = \mathbf{R}\mathbf{G}\mathbf{R}^{-1}$ and both g^{tr} and g^{det} are coordinate invariant. At this point, there is no compelling reason to choose one definition over another. In specific applications the concept of noise amplification will have concrete physical manifestations which will inform the choice of the amplification measure.

The gain is directly related to the violation of detailed balance. By analogy with the one-dimensional case, we might guess that the gain can be expressed in terms of $\mathbf{G}_0 \equiv -\mathbf{A}^{-1}$. Since $\mathbf{A}$ is stable, $g_0^{tr} \equiv \text{Trace}(\mathbf{G}_0)$ and $g_0^{det} \equiv \text{Det}(\mathbf{G}_0)$ are positive. As in the one-dimensional case, as the deterministic dynamics gets more stable both g_0^{tr} and g_0^{det} decrease. When detailed balance is satisfied, $\boldsymbol{\Omega}$, which can be written as $\boldsymbol{\Omega} = \mathbf{A} + \mathbf{G}^{-1}$, is zero and thus $\mathbf{G} = \mathbf{G}_0$.

One can show that the g_0's are lower bounds on the respective gains. Since g^{tr} and g^{det} are coordinate invariant, one can pick any convenient coordinate system to perform the analysis. In identity-noise coordinates, where $\mathbf{D} = \frac{1}{2}\mathbf{I}$, one can show that

$$g^{tr} = 2\mathrm{Trace}(\mathbf{C}) \geq -2\sum_{i=1}^{N} \frac{1}{\lambda_i + \lambda_i^*} \geq -\mathrm{Trace}(\mathbf{A}^{-1}), \tag{18}$$

where λ_i are the eigenvalues of $\mathbf{A}$, the first inequality is that of Ioannou (1995), and second inequality follows from the fact that the eigenvalues of $\mathbf{A}^{-1}$ are $1/\lambda_i$ and the λ_i's are either real or complex conjugate pairs. Thus, $g^{tr} \geq g_0^{tr}$. For the determinant measure, in identity-covariance coordinates $\mathbf{D} = -\mathbf{A}_{\mathrm{sym}}$ and the so-called Ostrowski–Taussky inequality (Lütkepohl, 1996) gives $g^{det} \geq g_0^{det}$. Thus, the coordinate invariant gains defined here are minimized when the system satisfies detailed balance and the minimum value is the obvious generalization of the one-dimensional case. When detailed balance is violated, the gain can be significantly larger. Noise amplification is only possible when detailed balance is violated.

The determinant defined gain g^{det} has an interpretation in terms of the energetics of the system. In EOF coordinates each eigenvalue of $\mathbf{C}$ represents the energy of the EOF. Thus, $e_{\mathbf{C}} = (\mathrm{Det}(\mathbf{C}))^{1/N}$ can be interpreted as the geometric mean energy per degree of freedom. Due to the transformation properties of $\mathbf{C}$ this quantity, like the energy itself, is not coordinate invariant. However, $(g^{det})^{1/N} = e_{\mathbf{C}}/e_{\mathbf{D}}$, the ratio of the geometric mean energies per degree of freedom of the response and the noise, is coordinate invariant. Violation of detailed balance thus can cause an amplification of the system's energy over that predicted by g_0^{det} in this specific sense.

8. Examples

In this section we explore the behavior of simple two-dimensional dynamical systems to illustrate the theoretical ideas presented above. Although these dynamical systems are not models of any particular physical system, the low dimensionality allows one to easily visualize the properties of the system.

Consider the stochastic dynamical system $\mathrm{d}\boldsymbol{x}(t)/\mathrm{d}t = \mathbf{A}\boldsymbol{x}(t) + \mathbf{F}\boldsymbol{\xi}(t)$ [eq. (1)] with matrices

$$\mathbf{A} = \begin{pmatrix} -1 & -\cot\theta \\ 0 & -2 \end{pmatrix}, \quad \mathbf{F} = \mathbf{I}, \tag{19}$$

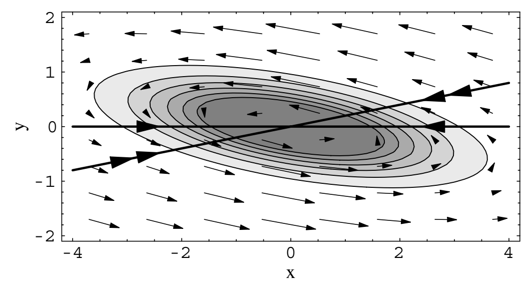

Fig. 2. Properties of the system described by eq. (19) with $\cot(\theta) = 5$, which violates detailed balance. The equilibrium pdf is contoured in gray with darker shades corresponding to larger probabilities. The eigenvectors of $\mathbf{A}$ are indicated by the thick lines where the double arrows indicate the more stable eigenvector. The equilibrium phase space velocity field $\boldsymbol{u}/|\boldsymbol{x}|$ is indicated by the arrows, where the velocity has been scaled by $|\boldsymbol{x}|$ to remove its linear structure.

which is the example of Farrell and Ioannou (1996a). When $\cot(\theta) \neq 0$ the deterministic matrix is non-normal and the system violates detailed balance. The FDR (6) can be solved numerically (Lu and Wachspress, 1991).

The results for $\cot(\theta) = 5$ are shown in Fig. 2. The non-normality of $\mathbf{A}$ is seen by the lack of orthogonality of its eigenvectors. The equilibrium pdf has elongated surfaces of constant probability. As the deterministic non-normality and violation of detailed balance grow (keeping the diffusion matrix constant) the eigenvectors of $\mathbf{A}$ become more parallel and the elongation of the pdf becomes more extreme. Since the phase space velocity $\boldsymbol{u} = \boldsymbol{\Omega x}$ is linear in $\boldsymbol{x}$ its structure is best seen by plotting $\boldsymbol{u}/|\boldsymbol{x}|$. The direction of $\boldsymbol{u}$ follows surfaces of constant probability, and the magnitude is large along the flanks of the pdf, and smaller at the turning points. Although the speed varies around the pdf, the divergence of $\boldsymbol{u}$ is zero, and the acceleration is thus due to changes in spacing of the probability contours.

The energetics are shown in Fig. 3. The energy density has peaks where the major axis of the pdf intersects the two standard deviation probability surface. The energy density source field has maxima and minima on opposite sides of each energy density peak. As one follows the phase space velocity $\boldsymbol{u}$ around the pdf, the energy flux and energy density increase as one goes through the source regions, reach a peak, and then decreases as one goes through the sink regions.

The energetics of the system are perhaps simplest in identity–covariance coordinates, where there

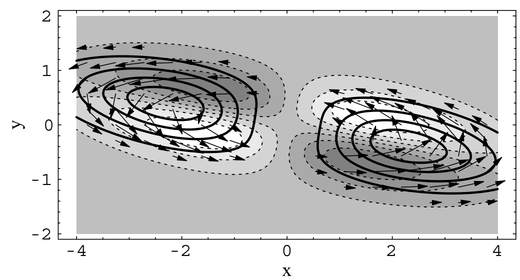

Fig. 3. Energetics for the system described by eq. (19) with $\cot(\theta) = 5$. Contours of the energy density are indicated by solid lines, with equal peaks inside the concentric curves and values near zero outside the curves. The energy density source is contoured with a grayscale where dark shades are positive, and hence are energy density sources, and light shades are negative and energy density sinks. The energy flux is indicated by the arrows, where only vectors with magnitudes above a small threshold are drawn.

are no energy density sources and the energy flux is divergence-free. Since $\mathbf{C} = \mathbf{I}$ the pdf is simply $p(r) = (1/2\pi)\mathrm{e}^{-r^2/2}$ and the probability surfaces are concentric circles. The energy density and the energy flux are axisymmetric, and there is a uniform azimuthal energy flux around the phase space (Fig. 4). The energy density peaks at the radius of the two standard deviation probability surface, $r = \sqrt{2}$, and the magnitude of the energy flux peaks at a slightly larger radius.

We next investigate the noise amplification focusing on the trace definition of gain. Similar results are obtained for the determinant definition. When $\theta = \pi/2$ the system satisfies detailed balance, has a normal deterministic matrix, and thus $g^{tr} = g_0^{tr}$. As θ deviates from $\pi/2$ detailed balance is increasingly violated, the deterministic matrix becomes increasingly non-normal, and g^{tr} grows (Fig. 5). Now consider a system with the same deterministic dynamics but with the diffusion matrix

$$\mathbf{D} = \frac{1}{2}\begin{pmatrix} 1 + 2\cot^2(\theta) & \cot(\theta) \\ \cot(\theta) & 1 \end{pmatrix}. \tag{20}$$

For this system, detailed balance is satisfied regardless of the value of θ and the resulting non-normality of the deterministic dynamics. Now the gain is constant at g_0^{tr} (Fig. 5), demonstrating that the amplification of the noise depends on the property of detailed balance and not on the normality of the deterministic dynamics.

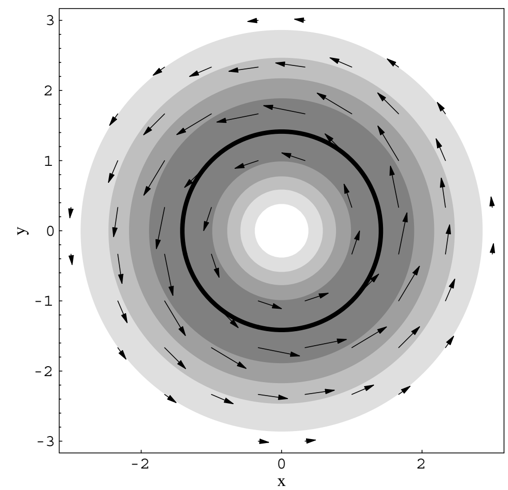

Fig. 4. Properties of the system described by eq. (19) with $\cot(\theta) = 5$ in identity-covariance coordinates. The two standard deviation probability surface is shown as a solid line, the energy density is indicated by the grayscale, where larger values are darker, and the energy flux is indicated by the arrows, where only vectors with magnitudes above a small threshold are drawn.

9. Discussion

In this paper we have focused on the equilibrium state of linear stochastic dynamical systems with additive Gaussian white noise. By viewing the stochastic dynamical system in terms of its Fokker–Planck

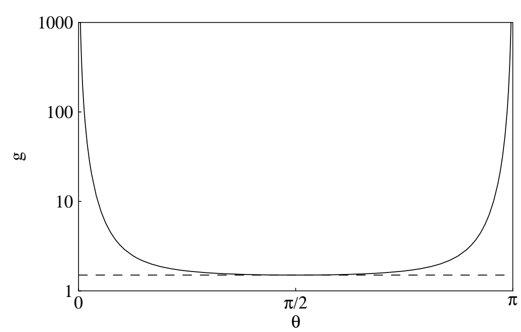

Fig. 5. Gain g^{tr} as a function of θ for a system defined by eq. (19) which violates detailed balance when $\theta \neq \pi/2$ (solid line), and a system with the same deterministic dynamics but with the diffusion matrix given by eq. (20) which satisfies detailed balance for all θ (dashed line).

equation (FPE), the coordinate invariant property of detailed balance is seen to be the fundamental property determining the system's behavior. When a system violates detailed balance, transitions have a preferred direction and the equilibrium state sustains a nonzero phase space velocity. In this case, the random noise is rectified into a flow with a preferred direction. The details of this rectified velocity are coordinate dependent, but its existence is invariant under nonsingular coordinate transformations.

Previous studies focused on the coordinate dependent property of non-normality of the deterministic dynamics. These previously studied systems also violated detailed balance. Here, we take the view that the coordinate invariant nature of detailed balance makes it the preferred concept for interpreting the dynamics. The resulting phenomena have different manifestations in different coordinate systems. In the identity–noise coordinates used in many previous studies, the phenomena are related to the non-normality of the deterministic operator and its ability to transiently amplify perturbations. In other coordinate systems the same underlying phenomena have different manifestations.

One of the objections to using stochastic dynamical systems to model many geophysical phenomena is that they do not seem all that random. While phenomena like ENSO and baroclinic life cycles do have irregular aspects in their behavior, they also have a definite organization that seems distinctly nonrandom. The heuristic image of a random process such as the stereotypical drunkard's walk appears at odds with such organized behavior. However, a drunkard's walk satisfies detailed balance: the drunkard is just as likely to stumble one way as another. Violation of detailed balance necessarily produces a rectification of the noise into an organized phase space velocity which can model the organization seen in geophysical phenomena.

One new result of this paper is that systems that violate detailed balance necessarily have nonzero fluxes of all physical quantities that can be written in terms of a phase space density. The detailed manifestation of these fluxes is, like the phase space velocity, coordinate dependent, but their existence is coordinate invariant. Further, the fluxes have a very specific structure which is imposed by dynamics, and they thus provide detailed predictions which can be used to validate stochastic models.

An important property of previously studied non-normal stochastic systems is they can amplify a small random forcing. Thus, the large variability seen in geophysical systems can be produced by stable non-normal deterministic dynamics perturbed by small noise, rather than requiring instabilities or large forcing. Here, by defining noise amplification in a coordinate-invariant manner, we see that, again, it is the violation of detailed balance rather than non-normality that gives rise to noise amplification. Once again, the specific manifestation of the noise-amplification is coordinate-dependent.

The application of these ideas to the many specific geophysical systems which have already been modeled by stochastic dynamical systems should yield important new tests of the validity of these models. For example a stochastic model of ENSO with sea-surface temperature and thermocline depth as the state variables would provide detailed predictions about the heat stored in the mixed layer, its fluxes, and the heating and cooling of the mixed layer by the random and deterministic components, often associated with the atmosphere and ocean, respectively.

This paper has focused on the equilibrium state. Stochastic models are often used for predictions from a given initial condition, as is done, for example, for tropical sea-surface temperature (Penland and Magorian, 1993). In this case, the appropriate FPE describes the time-dependent decay from a pdf representing the uncertainty in the initial state. Extending the ideas presented here to such time-dependent cases could provide predictions of new quantities similar to the fluxes seen in equilibrium.

Both the deterministic and stochastic components of the dynamics explored here are as simple as possible. Improved stochastic models of geophysical systems may entail further complexity. One can add nonlinearity to the deterministic dynamics, and perturb the dynamics with non-Gaussian, colored or multiplicative noise. All of these complexities can and should be studied from both the SDE and FPE perspective in a variety of coordinate systems.

10. Acknowledgments

The author thanks Judith Berner, Grant Branstator, Keith Kearnes, Andy Moore, Cecile Penland and Bill Young for conversations that contributed to this work. This work was partially supported by NSF OCE-0002370.

REFERENCES

DelSole, T. and Farrell, B. F. 1996. The quasi-linear equilibration of a thermally maintained stochastically excited jet in a quasigeostrophic model. *J. Atmos. Sci.* **53**, 1781–1797.

DelSole, T. and Hou, A. Y. 1999. Empirical stochastic models for the dominant climate statistics of a general circulation model. *J. Atmos. Sci.* **56**, 3436–3456.

Epstein, E. 1969. Stochastic dynamic prediction. *Tellus* **21**, 739–759.

Farrell, B. F. and Ioannou, P. J. 1993. Stochastic dynamics of baroclinic waves. *J. Atmos. Sci.* **50**, 4044–4057.

Farrell, B. F. and Ioannou, P. J. 1994. A theory for the statistical equilibrium energy spectrum and heat flux produced by transient baroclinic waves. *J. Atmos. Sci.* **51**, 2685–2698.

Farrell, B. F. and Ioannou, P. J. 1996a. Generalized instability theory. Part I: Autonomous operators. *J. Atmos. Sci.* **53**, 2025–2040.

Farrell, B. F. and Ioannou, P. J. 1996b. Generalized instability theory. Part II: Nonautonomous operators. *J. Atmos. Sci.* **53**, 2041–2053.

Farrell, B. F. and Ioannou, P. J. 2000. Perturbation dynamics in atmospheric chemistry. *J. Geophys. Res.* **105**, 9303–9320.

Gardiner, C. 1983. *Handbook of stochastic methods*. Springer-Verlag, Berlin, 442 pp.

Horn, R. A. and Johnson, C. R. 1985. *Matrix analysis*. Cambridge University Press, Cambridge, 561 pp.

Ioannou, P. J. 1995. Nonnormality increases variance. *J. Atmos. Sci.* **52**, 1155–1158.

Leith, C. 1974. Theoretical skill of monte carlo forecasts. *Mon. Wea. Rev.* **102**, 409–418.

Lu, A. and Wachspress, E. L. 1991. Solution of Lyapunov equations by alternating direction implicit iteration. *Comput. Math. Appl.* **21**, 43–58.

Lütkepohl, H. 1996. *Handbook of matrices*. John Wiley, New York, 304 pp.

Moore, A. M. 1999. Wind-induced variability of ocean gyres. *Dyn. Atmos. Oceans* **29**, 335–364.

Moore, A. M. and Farrell, B. F. 1993. Rapid perturbation growth on spatially and temporally varying oceanic flows determined using an adjoint method: Application to the Gulf Stream. *J. Phys. Oceanogr.* **23**, 1682–1702.

Moore, A. M. and Kleeman, R. 1996. The dynamics of error growth and predictability in a coupled model of ENSO. *Q. J. R. Meteorol. Soc.* **122**, 1405–1446.

Newman, M., Sardeshmukh, P. D. and Penland, C. 1997. Stochastic forcing of the wintertime extratropical flow. *J. Atmos. Sci.* **54**, 435–455.

Penland, C. and Magorian, T. 1993. Prediction of Niño 3 sea surface temperaturs using linear inverse modeling. *J. Climate* **6**, 1067–1076.

Penland, C. and Sardeshmukh, P. D. 1995. The optimal growth of tropical sea surface temperature anomalies. *J. Climate* **8**, 1999–2024.

Risken, H. 1984. *The Fokker–Planck equation: Methods of solution and applications*. Springer-Verlag, Berlin, 454 pp.

van Kampen, N. 1992. *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam, 465 pp.

Weickmann, K. M., Robinson, W. A. and Penland, C. 2000. Stochastic and oscillatory forcing of global atmospheric angular momentum. *J. Geophys. Res.* **105**, 15543–15557.

Whitaker, J. S. and Sardeshmukh, P. D. 1998. A linear theory of extratropical synoptic eddy statistics. *J. Atmos. Sci.* **55**, 237–258.

Zhang, Y. and Held, I. M. 1999. A linear stochastic model of a GCM's midlatitude storm tracks. *J. Atmos. Sci.* **56**, 3416–3435.