Available online at www.sciencedirect.com
SCIENCE DIRECT®
Ocean Modelling 7 (2004) 227–258
Ocean Modelling
www.elsevier.com/locate/ocemod

A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model

Andrew M. Moore [a,*], Hernan G. Arango [b], Emanuele Di Lorenzo [c], Bruce D. Cornuelle [c], Arthur J. Miller [c], Douglas J. Neilson [c]

[a] *Program of Atmospheric and Oceanic Sciences, and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Campus Box 311, Boulder, CO 80309-0311, USA*
[b] *Institute of Marine and Coastal Sciences, Rutgers University, 71 Dudley Rd., New Brunswick, NJ 08901-8521, USA*
[c] *Climate Research Division, Scripps Institute of Oceanography, University of California, San Diego, La Jolla, CA 92093-0224, USA*

Received 22 April 2003; received in revised form 16 October 2003; accepted 14 November 2003

Abstract

The regional ocean modelling system (ROMS) is a new generation ocean general circulation model that is rapidly gaining favour in the ocean modelling community. The tangent linear and adjoint versions of ROMS have recently been developed, and a new suite of tools that utilize these models for a variety of applications are now available to the ocean modelling community. In this paper we will describe the tangent linear and adjoint components of ROMS, and present examples from the tools that are currently available to ROMS users. In particular we will consider the finite time eigenmodes and the adjoint finite time eigenmodes of the tangent linear propagator, the singular vectors of the propagator, and its forcing singular vectors and stochastic optimals. The pseudospectra of the tangent linear resolvent matrix are also considered. Examples of each type of calculation will be presented for a time evolving double gyre ocean circulation in a rectangular ocean basin.

Keywords: ROMS; Tangent linear models; Adjoint models; Eigenmodes; Singular vectors; Forcing singular vectors; Stochastic optimals; Pseudospectra

* Corresponding author.
E-mail address: andy@bondi.colorado.edu (A.M. Moore).

1463-5003/$ - see front matter
doi:10.1016/j.ocemod.2003.11.001

1. Introduction

The regional ocean modelling system (ROMS) is a new state-of-the-art model that can be configured for any region of the world ocean ranging from local to basin scale. The numerics of the model are described in detail by Shchepetkin and McWilliams (2003), and the model already has a large user base. ROMS has been used to model the circulation in a variety of different regions of the world ocean (e.g. Haidvogel et al., 2000; Malanotte-Rizzoli et al., 2000; She and Klinck, 2000; Di Lorenzo et al., 2004; MacCready and Geyer, 2001; Penven et al., 2000, 2001a,b; Marchesiello et al., 2002).

This paper describes a suite of powerful new tools now available for ROMS that are based on the tangent linear and adjoint versions of the model. We will describe what essentially amounts to an end-to-end ocean prediction and analysis system that can be used in any region of the global ocean. Before doing this, however, it is useful to briefly review some of the applications of tangent linear and adjoint models to date in meteorology, oceanography and the climate modelling.

It has long been recognized that a tangent linear model and its associated adjoint are powerful tools for analyzing numerous aspects of model sensitivity and the dynamics of flow fields. However, it is only with the fairly recent advent of fast multiprocessor computers that the full potential of these tools has been realized in complex systems and for operational applications.

For the discrete dynamical systems described by numerical models, the tangent linear model provides the Jacobian of the dynamical operators that are tangent linear to a solution trajectory of the nonlinear system. The adjoint of this Jacobian operator provides information about the sensitivity of the system to variations in the model state vector, boundary conditions or model parameters. In fact any linear problem that can be expressed in terms of finding the extrema of a suitably defined scalar function of the model state vector will typically yield the adjoint Jacobian as a necessary component of the solution. It is for this reason that the use of tangent linear and adjoint models in geosciences, particularly in meteorology, has become so wide-spread because of the variety of problems for which they can be used. Some examples include sensitivity analysis, data assimilation, stability analysis, and ensemble prediction.

Some of the first applications of adjoint models in the meteorological literature are described by Hall and Cacuci (1983) who explored the sensitivity of an atmospheric model to variations in all of its free parameters. It can be shown that the adjoint of the tangent linear Jacobian of the model yields information about the gradient of a suitably defined scalar function of the model parameters. In particular, a single integration of the adjoint model equations will yield information about the linear sensitivity (gradient) of the chosen functional at all model grid points and timesteps. If the function is some measure of forecast error, then the adjoint forecast model can be used to determine the sensitivity of the final time forecast error to initial condition error (e.g. Lacarra and Talagrand, 1988; Errico et al., 1993; Vukecevic, 1991; Rabier et al., 1996; Corti and Palmer, 1997). The use of adjoint models for sensitivity analysis is not confined to meteorology, and they have also been used in oceanography (Junge and Haine, 2001; Galanti and Tziperman, 2003) and for problems in climate modelling (Galanti et al., 2002).

Tangent linear and adjoint models have found most wide spread use in meteorology in the field of data assimilation. In this case a scalar function, usually called the cost function (J) is defined as a measure of the squared difference between the forecast model trajectory and the available observations, and other information such as a background field (e.g. climatology or a prior model

forecast). It was shown by Talagrand and Courtier (1987) that the gradient of J with respect to variations in the model initial conditions can be computed from a single integration of the adjoint forecast model. Once this gradient has been computed, the initial conditions can be adjusted to yield a smaller value of J upon subsequent integrations of the forecast model. Thus an iterative algorithm can be designed in which the forecast model and adjoint model are integrated alternately. The forecast model initial conditions are adjusted at the start of each iteration so as to yield a smaller value of J. The iteration process continues until the minimum of J is found which corresponds to the best fit between the model and the observations. For large operational numerical weather prediction models, it is not practical to iterate to the minimum of J, so the procedure is usually terminated after some predefined tolerance or convergence criterion is satisfied.

In its original form, this so-called 4-dimensional variational data assimilation scheme (4DVAR) imposes the model dynamics as a strong constraint. There is a large body of literature on this subject and it is beyond the scope of this paper to provide an exhaustive list of references. One of the great advantages of 4DVAR, however, is that any type of observational data can be assimilated into a model provided that it is a known function of the model state variables. The strong constraint 4DVAR approach to data assimilation has also been used in oceanography (e.g. Long and Thacker, 1989a,b; Moore, 1991; Weaver et al., 2002; Vialard et al., 2002) and in climate modelling (Kleeman et al., 1995).

The strong constraint imposed by the model dynamics in the original formulation of 4DVAR can be relaxed to allow for the inevitable errors and uncertainties in the model equations, forcing fields, and boundary conditions. One very powerful variant of this approach is the so-called method of representers described in detail by Bennett (1992, 2002). For each observation a representer function can be computed that represents a space-time correlation function. The representer functions are found by solving the nonlinear Euler–Lagrange equations iteratively, and the tangent linear forecast model and its adjoint play a key role in this process. The method of representers has been applied to a number of data assimilation problems in oceanography and climate modelling (e.g. Bennett and Thorburn, 1992; Bennett et al., 1993; Bennett et al., 1998; Egbert et al., 1994; Bennett, 2002). Chua and Bennett (2001) describe an inverse ocean model (IOM) system which can be interfaced with the tangent linear and adjoint versions of different models and used for data assimilation using the representer method. An indirect application of the representer method is described by Egbert et al. (1994), Amodei (1995) and Courtier (1997).

Tangent linear and adjoint models can also be used to explore the stability of dynamical systems, and to understand the dynamics of perturbation growth. The traditional approach to stability analysis involves computing the eigenvalues and eigenvectors of the tangent linear model (Pedlosky, 1979, his chapter 7). These eigenvectors represent dominant patterns of variability that might be expected if the associated eigenvalues have a positive real part. Such eigenvector (normal mode) analyses have been used in meteorology to understand various aspects of the large scale circulation (e.g. Charney, 1947; Eady, 1949; Simmons and Hoskins, 1976; Frederiksen, 1982). In general, the eigenvectors of observed and modelled circulation fields do not form an orthonormal basis, in which case the eigenvectors of the associated adjoint model are also of interest since they represent the optimal excitations for the corresponding eigenvectors of the tangent linear model (Branstator, 1985). In general these two sets of eigenvectors have very different structures. Thus the adjoint eigenvectors indicate how a particular eigenvector of the tangent linear system can be forced or excited (Frederiksen, 1997).

The non-orthogonality of the eigenvectors of the atmospheric and oceanic circulations has important consequences for perturbation growth. The fastest growing of all possible perturbations is the fastest growing singular vector of the tangent linear system. The study of the singular vectors of atmospheric flows dates back to the work of Lorenz (1965), and has more recently become a subject of intense research, driven largely by the pioneering work of Farrell (1982a,b, 1984, 1985, 1988a,b, 1989a,b, 1990) and Farrell and Ioannou (1993a,b,c,d,e,f, 1994a,b, 1996a,b,c) who have used these ideas to understand the basic mechanisms of different atmospheric phenomena. The singular vectors of any forecast model circulation can be computed directly using the tangent linear and adjoint forecast models, and a number of studies in meteorology have demonstrated that the predictability of some aspects of the atmospheric circulation can be understood in terms of singular vector growth (e.g. Palmer, 1988). There is a growing body of literature on the singular vectors of the atmospheric general circulation (e.g. Buizza et al., 1993; Buizza, 1995, 1997; Buizza and Palmer, 1995; Molteni et al., 1993, 1996; Hartmann et al., 1995; Joly, 1995) that has considerably advanced our understanding of atmospheric predictability.

Singular vectors have also been computed for the ocean. Moore and Farrell (1993) and Moore and Mariano (1999) have explored the singular vector structure and perturbation growth in the Gulf Stream, and Moore et al. (2002) have examined the singular vectors of a wind-driven double gyre. Singular vectors have also proved valuable for understanding some aspects of the El Nino Southern Oscillation (ENSO) (Moore and Kleeman, 1996, 1997a,b, 1998, 1999a,b, 2001; Penland and Sardeshmukh, 1995a,b; Xue et al., 1994, 1997a,b, 1999; Blumenthal, 1991; Chen et al., 1997; Thompson, 1998; Fan et al., 2000; Zavala-Garay et al., 2003), and other climate variability (DelSole and Hou, 1999).

The singular vectors of the atmospheric general circulation are used at some operational centers for ensemble prediction. Since the singular vectors represent the fastest growing (in the linear limit) of all possible perturbations, they can provide valuable information about the most unstable directions of phase space, and the most likely areas of forecast error growth. Forecast ensembles can be created by perturbing the model initial conditions with different combinations of the fastest growing singular vectors, and the resulting ensemble yields information about the atmospheric probability density function, and the potential predictability of the atmosphere at different times (e.g. Mureau et al., 1993; Molteni et al., 1993, 1996; Betti and Navarra, 1995; Buizza, 1997). These same techniques have been explored by Moore (1999a) for generating ensembles of ocean forecasts, and by Moore and Kleeman (1998) for estimating the predictability of ENSO.

Lastly, we briefly mention an exciting new application of singular vectors, namely that of adaptive observations (Palmer et al., 1998). The singular vectors of the atmospheric general circulation identify regions where perturbations are likely to grow very rapidly. Thus for forecasting it is important to minimize the initial condition errors in these regions. One way to achieve this is by making additional observations in the regions identified by the singular vectors as regions of rapid error growth. Based on this idea of so-called "adaptive" or "targeted" observations, two field experiments have been performed, one in the North Atlantic (FASTEX) and one in the North Pacific (NORPEX) during which the impact on forecast skill of additional observations collected by aircraft in regions of rapid singular vector growth was assessed (Langland et al., 1999; Bishop and Toth, 1999; Szunyogh et al., 1999). Similar types of experiments may prove useful for ocean prediction, particularly on seasonal timescales in the tropics.

From the preceding discussion, it is clear that tangent linear and adjoint model-based techniques hold great promise for physical oceanography, and have a great deal to offer. This was our primary motivation for developing the tangent linear and adjoint versions of the ROMS system which are described in Section 2. A description of the various platforms/drivers that directly use the tangent linear and adjoint codes is given in Section 3. Section 3 also includes a brief review of the theory behind each application for those readers who are unfamiliar with these ideas. Example calculations are presented in Section 4, in which we apply the ideas of Section 3 to the classical baroclinic double gyre circulation in a rectangular ocean basin. We end with a summary in Section 5.

2. Tangent linear and adjoint ROMS

Throughout this, and subsequent sections, we will represent the ROMS equations symbolically. The full equations of the nonlinear, tangent linear and adjoint models can be found in Arango et al. (2003). Here we will denote the ROMS model state vector as $\mathbf{S}$ which represents all of the model prognostic variables at ocean grid points. The ROMS equations of motion can then be represented symbolically as:

$$\frac{\partial \mathbf{S}}{\partial t} = N(\mathbf{S}), \tag{1}$$

where $N(\mathbf{S})$ represents the model dynamical operators, which in general will be nonlinear.The tangent linear form of ROMS was derived by considering small perturbations $\mathbf{s}$ to $\mathbf{S}$, and performing a first-order Taylor expansion of (1) which yields:

$$\frac{\partial \mathbf{s}}{\partial t} = \left.\frac{\partial N(\mathbf{S})}{\partial \mathbf{S}}\right|_{\mathbf{S}_0} \mathbf{s} = \mathbf{A}\mathbf{s}, \tag{2}$$

where $\mathbf{S}_0$ is a solution of (1), and $\mathbf{A} \equiv \frac{\partial N(\mathbf{S})}{\partial \mathbf{S}}|_{\mathbf{S}_0}$ is assumed to be real throughout. Eq. (2) will be referred to as the tangent linear model (TLM). The adjoint model is derived by considering the inner-product of (2) with an arbitrary vector $\mathbf{s}^\dagger$ where the inner-product defines an appropriate norm. For the applications considered in Sections 3 and 4, $\mathbf{s}^\dagger$ can be viewed as a generating vector that yields unique solutions of (2) only in the state dimensions that are activated by the operator $\mathbf{A}$ (Lanczos, 1961). The physical interpretation of $\mathbf{s}^\dagger$ depends on the nature of the problem at hand, and is discussed in more detail in Appendix A. Since we are dealing exclusively with discrete systems, we will use matrix–vector notation throughout, in which case an inner-product will be denoted as $\mathbf{s}^{\dagger T}\mathbf{G}\mathbf{s}$, where $\mathbf{G}$ is symmetric positive definite and defines the norm of interest (e.g. perturbation energy). The adjoint $\mathbf{A}^\dagger$ of $\mathbf{A}$ in (2) is given by the bilinear identity $\mathbf{s}^{\dagger\mathrm{T}}\mathbf{G}\mathbf{A}\mathbf{s} = \mathbf{s}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{G}\mathbf{s}^\dagger = \mathbf{s}^{\mathrm{T}}\mathbf{G}(\mathbf{G}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{G})\mathbf{s} = \mathbf{s}^{\mathrm{T}}\mathbf{G}\mathbf{A}^\dagger\mathbf{s}^\dagger$, therefore $\mathbf{A}^\dagger = \mathbf{G}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{G}$. The inner-product of $\mathbf{s}^\dagger$ with (2) yields $\partial(\mathbf{s}^{\dagger\mathrm{T}}\mathbf{G}\mathbf{s})/\partial t - \mathbf{s}^{\mathrm{T}}\mathbf{G}\partial\mathbf{s}^\dagger/\partial t - \mathbf{s}^{\mathrm{T}}\mathbf{G}\mathbf{A}^\dagger\mathbf{s}^\dagger = 0$. Therefore the adjoint model can be represented symbolically as:

$$-\frac{\partial \mathbf{s}^\dagger}{\partial t} = \mathbf{A}^\dagger\mathbf{s}^\dagger, \tag{3}$$

where the negative sign on the left-hand side of (3) implies integration backwards in time, and $\partial(\mathbf{s}^{\mathrm{T}}\mathbf{G}\mathbf{s}^\dagger)/\partial t = 0$ for all solutions of (2) and (3) and is a direct consequence of the bilinear identity.

The continuous counterpart of the adjoint equations described by (3) are derived by repeated integration of the associated inner-product involving (2), and the adjoint boundary conditions are chosen so as to render zero the resulting boundary integrals (see Lanczos, 1961, Section 4.17).

For many adjoint model applications it is very important to use the adjoint of the discrete model equations rather than the discretized form of the continuous adjoint equations. As noted above $\mathbf{A}^{\dagger} = \mathbf{G}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{G}$, so for the L2-norm ($\mathbf{G} = \mathbf{I}$) the adjoint operator reduces to the matrix transpose. This is the norm that we have used to derive the discrete adjoint of ROMS, although the definition $\mathbf{A}^{\dagger} = \mathbf{G}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{G}$ shows that the adjoint operator for other norms can be derived from the L2-norm adjoint by a simple linear transformation of the input variables by $\mathbf{G}$ and a transformation of the output variables by $\mathbf{G}^{-1}$ (Ehrendorfer and Tribbia, 1997). The identity $\partial(\mathbf{s}^{\dagger\mathrm{T}}\mathbf{A}\mathbf{s})/\partial\mathbf{s} = \mathbf{A}^{\mathrm{T}}\mathbf{s}^{\dagger}$ indicates that the adjoint operator relative to the L2-norm can be computed by multiplying each line of the TLM code by the corresponding adjoint variable, and then differentiating with respect to the TLM variable (e.g. Long and Thacker, 1989a). This practice yields a code that evaluates the transpose of the TLM operators. Giering and Kaminski (1998) provide useful recipes for this operation, and have in fact developed an automatic differentiation and adjoint model compiler. The tangent linear and adjoint versions of ROMS, however, were hand coded using these same recipes. This option was considered preferable to using an automatic compiler since it affords a greater understanding and control of the code structure of the resulting models.

3. ROMS drivers

As discussed in Section 1, there are numerous very powerful applications of tangent linear and adjoint models in meteorology, oceanography and climate modelling. One of the main motivations for developing the tangent linear and adjoint versions of ROMS was to make these applications available to the ocean modelling community by providing a series of easy to use platforms. In this section we will describe these, and some supporting theoretical background for those readers who are unfamiliar with the subject.

3.1. Eigenmodes of the tangent linear model

Following traditional methods, the stability of a circulation described by (1) can be explored by computing the eigenvectors of the tangent linear matrix $\mathbf{A}$ in (2). These are often referred to as "normal modes," and for autonomous systems (i.e. $\mathbf{A}$ is independent of time), the eigensolutions are of the form $\mathbf{s} = \hat{\mathbf{s}}\mathrm{e}^{\sigma t}$, where $\mathbf{A}\hat{\mathbf{s}} = \sigma\hat{\mathbf{s}}$. However, in general the eigenmodes of $\mathbf{A}$ rarely form a normal set, except for particular choices of norm. If the real component of at least one of the eigenvalues σ is positive then the system is asymptotically unstable and the amplitude of all perturbations will grow without bound.

Integral solutions of (2) can be written as $\mathbf{s}(t) = \mathbf{R}(0,t)\mathbf{s}(0)$, where $\mathbf{R}(0,t)$ is called the propagator of (2), and the notation $\mathbf{R}(t_1,t_2)$ refers to an integration forward in time for the interval $[t_1,t_2]$. For autonomous systems the propagator can be written as $\mathbf{R}(0,t) = \mathrm{e}^{\mathbf{A}t}$. In this case, the eigenvectors of $\mathbf{A}$ and $\mathbf{R}$ are identical, while the eigenvalues of $\mathbf{R}$ are given by $\mathrm{e}^{\sigma t}$.

For non-autonomous systems ($\mathbf{A} = \mathbf{A}(t)$), the propagator cannot, in general, be written as a simple exponential function of $\mathbf{A}$. In this case, the eigenvectors of $\mathbf{R}(0,t)$ depend on the time interval t and the time evolution of $\mathbf{S}_0$ over this interval. These eigenvectors are often referred to as the "finite time normal modes." Since they do not form an orthonormal set, to avoid any confusion we will refer to them as finite time eigenmodes (FTEs).

The eigenvectors of the propagator $\mathbf{R}(0,t)$ for either an autonomous or non-autonomous ROMS circulation can be computed iteratively. A single integration of an arbitrary perturbation state vector $\mathbf{u}$ forward in time over the interval $[0,t]$ by the TLM (cf. Eq. (2)) is equivalent to the matrix–vector product $\mathbf{R}(0,t)\mathbf{u}$. By combining forward integrations of the TLM with the Arnoldi algorithm (Golub and van Loan, 1989; Sorensen, 1992), it is possible to compute a portion of the eigenspectrum of $\mathbf{R}(0,t)$. In practice, the public domain software package ARPACK (Lehoucq et al., 1997) can be used to do this via the routine **dnaupd** (see www.caam.rice.edu/software/ARPACK).

3.2. Eigenmodes of the adjoint tangent linear model

The eigenvalues of $\mathbf{A}$ and $\mathbf{A}^\dagger$ are identical and occur in complex conjugate pairs. The eigenvectors of $\mathbf{A}^\dagger$ are of interest because they represent the optimal excitations of the eigenvectors of $\mathbf{A}$ with corresponding complex conjugate eigenvalues. We will denote by $\{\sigma_n, \hat{\mathbf{s}}_n\}$ and $\{\lambda_n, \hat{\mathbf{u}}_n\}$ the {eigenvalue, eigenvector} sets of $\mathbf{A}$ and $\mathbf{A}^\dagger$ respectively, where $\lambda_n = \sigma_n^*$, and * denotes the complex conjugate. It is easy to show that:

$$\hat{\mathbf{u}}_m^{\mathrm{H}}\mathbf{G}\hat{\mathbf{s}}_n(\sigma_n - \lambda_m^*) = \hat{\mathbf{r}}_m^{\mathrm{H}}\hat{\mathbf{s}}_n(\sigma_n - \lambda_m^*) = \hat{\mathbf{r}}_m^{\mathrm{H}}\hat{\mathbf{s}}_n(\sigma_n - \sigma_m) = 0, \tag{4}$$

where $\hat{\mathbf{r}}_m = \mathbf{G}\hat{\mathbf{u}}_m$ are the eigenvectors of $\mathbf{A}^{\mathrm{T}}$, and superscript H denotes the conjugate transpose. For $n = m$ we have $\sigma_n = \sigma_m$ while for $n \neq m$ we require $\hat{\mathbf{r}}_m^{\mathrm{H}}\hat{\mathbf{s}}_n = 0$. Eq. (4) is an expression of the biorthogonality between the eigenmodes of $\mathbf{A}$ and $\mathbf{A}^\dagger$.

To understand the physical relationship between the eigenmodes and adjoint eigenmodes, consider an initial value problem in which an arbitrary perturbation $\mathbf{s}(0)$ is composed of a linear superposition of the eigenmodes $\hat{\mathbf{s}}_n$, namely $\mathbf{s}(0) = \sum_{n=1}^{N} a_n\hat{\mathbf{s}}_n = \mathbf{E}\mathbf{a}$ where $\mathbf{E}$ is the matrix whose columns are the eigenvectors $\hat{\mathbf{s}}_n$, and $\mathbf{a}$ is the vector of amplitudes a_n. Furthermore, assume that the perturbation has unit norm, $\mathbf{s}(0)^{\mathrm{T}}\mathbf{X}\mathbf{s}(0) = 1$ where $\mathbf{X}$ is symmetric positive definite, and $\mathbf{X}$ and $\mathbf{G}$ may define the same or different norms. According to (4), the amplitude a_k of the kth eigenmode is $a_k = \hat{\mathbf{r}}_k^{\mathrm{H}}\mathbf{s}(0)/(\hat{\mathbf{r}}_k^{\mathrm{H}}\hat{\mathbf{s}}_k)$. The unit-norm perturbation $\mathbf{s}(0)$ that maximizes the amplitude a_k of the kth eigenmode can be found by maximizing the Lagrange function L given by

$$L = a_k^* a_k + \chi(\mathbf{s}(0)^{\mathrm{H}}\mathbf{X}\mathbf{s}(0) - 1) = a_k^* a_k + \chi(\mathbf{a}^{\mathrm{H}}\mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{E}\mathbf{a} - 1), \tag{5}$$

where χ is an unknown Lagrange multiplier. At the extrema of L we have $\partial L/\partial \mathbf{a} = 0$ and $\partial L/\partial \chi = 0$ which yields:

$$\chi\mathbf{E}^{\mathrm{H}}\mathbf{X}(\mathbf{E}\mathbf{a}) = -a_k\epsilon_k, \tag{6}$$

where ϵ_k is the unit vector with all elements zero except the kth element. Eq. (6) shows that all elements of the vector on the right-hand side are zero except that corresponding to a_k. The left-hand side of (6) implies then that $\mathbf{X}\mathbf{E}\mathbf{a}$ (which is equivalent to $\mathbf{X}\mathbf{s}(0)$) must be orthogonal to all the

columns of $\mathbf{E}$ (i.e. rows of $\mathbf{E}^{\mathrm{H}}$) except for column k. According to (4) this is a property of the adjoint eigenmodes, so $\mathbf{XEa} = \hat{\mathbf{r}}_k$ and thus $\mathbf{s}(0) = \mathbf{X}^{-1}\hat{\mathbf{r}}_k = \mathbf{X}^{-1}\mathbf{G}\hat{\mathbf{u}}_k$, showing that a linear transformation of the kth adjoint eigenmode is the optimal excitation for the kth eigenmode (i.e. it is the perturbation that has the largest projection on the kth eigenmode). The coefficient $a_k = \hat{\mathbf{r}}_k^{\mathrm{H}}\mathbf{X}^{-1}\hat{\mathbf{r}}_k/(\hat{\mathbf{r}}_k^{\mathrm{H}}\hat{\mathbf{s}}_k)$ shows that the amplitude of the kth eigenmode depends on the projection of $\hat{\mathbf{s}}_k$ on $\hat{\mathbf{r}}_k$, and $\mathbf{X}^{-1}\hat{\mathbf{r}}_k$ is the optimal excitation for $\hat{\mathbf{s}}_k$ in that of all possible norms, it produces the most rapid growth in $\mathbf{s}^{\mathrm{T}}\mathbf{Xs}$ as $\hat{\mathbf{s}}_k$ emerges from the evolving perturbation.

The importance of the adjoint eigenmodes becomes clear when we consider the difference between normal and non-normal systems described by the TLM (2). For a normal system, $[\mathbf{A}, \mathbf{A}^{\dagger}] = \mathbf{AA}^{\dagger} - \mathbf{A}^{\dagger}\mathbf{A} = 0$ and the eigenvectors of $\mathbf{A}$ and $\mathbf{A}^{\dagger}$ are identical. Therefore, in this case the optimal excitation of any eigenmode of $\mathbf{A}$ is a perturbation with the structure of the eigenmode itself. For a non-normal system, $[\mathbf{A}, \mathbf{A}^{\dagger}] \neq 0$ and the eigenvectors of $\mathbf{A}$ and $\mathbf{A}^{\dagger}$ are different. In this case, the optimal excitation of any eigenmode of $\mathbf{A}$ can have a very different structure when compared to the eigenmode itself, depending on the inner-product that defines $\mathbf{A}^{\dagger}$.

It is therefore often desirable to compute the adjoint eigenmodes of the TLM for a circulation of interest. The biorthogonality relation (4) shows that they are required if the projection of a perturbation onto a particular eigenmode of $\mathbf{A}$ is desired. In addition, the adjoint eigenmodes reveal the structures of perturbations that will be most effective for exciting the corresponding eigenmodes of $\mathbf{A}$. In sequel we will assume that the adjoint operator $\mathbf{A}^{\dagger}$ was derived relative to the L2-norm as described in Section 2. Therefore the adjoint of the TLM will be described by the transposed propagator $\mathbf{R}^{\mathrm{T}}(t, 0)$ where the notation $\mathbf{R}^{\mathrm{T}}(t_2, t_1)$ indicates that the adjoint model is integrated backwards in time over the interval $[t_2, t_1]$.

The eigenvectors of the adjoint propagator $\mathbf{R}^{\mathrm{T}}(t, 0)$ can be computed in an analogous way to those of $\mathbf{R}(0, t)$ by noting that a single integration of an arbitrary perturbation state vector $\mathbf{u}$ backward in time over the interval $[t, 0]$ by the adjoint TLM of ROMS (cf. Eq. (3)) is equivalent to the matrix–vector product $\mathbf{R}^{\mathrm{T}}(t, 0)\mathbf{u}$. The ARPACK routine **dnaupd** can be used to compute a subset of the required adjoint eigenmodes in an iterative manner.

3.3. Singular vectors

While the eigenvalues of $\mathbf{A}$ are traditionally viewed as a measure of the asymptotic stability of the system, they are not always the best measure of the ability for perturbations to grow. As noted in Section 3.1, the eigenvectors of $\mathbf{A}$ are typically non-orthogonal for realistic ocean circulations. This is an important characteristic of a non-normal system. The non-orthogonality of the eigenvectors of $\mathbf{A}$ implies that each eigenvector has a non-zero projection on other members of the spectrum. In this case it is possible for linear interference of the eigenvectors to produce transient perturbation growth, even when all of the eigenvectors are themselves asymptotically stable. The fastest growing of all possible perturbations over a given time interval when measured in terms of the growth of a given norm are the singular vectors of the propagator $\mathbf{R}(0, t)$. Specifically, we consider the factor μ by which the squared magnitude of a perturbation $\mathbf{s}$ grows over the time interval $[0, t]$:

$$\mu = \frac{\mathbf{s}^{\mathrm{T}}(t)\mathbf{Xs}(t)}{\mathbf{s}^{\mathrm{T}}(0)\mathbf{Xs}(0)} = \frac{\mathbf{s}^{\mathrm{T}}(0)\mathbf{R}^{\mathrm{T}}(t, 0)\mathbf{XR}(0, t)\mathbf{s}(0)}{\mathbf{s}^{\mathrm{T}}(0)\mathbf{Xs}(0)}, \tag{7}$$

where $\mathbf{X}$ defines the norm used as a measure of perturbation growth. When $\mathbf{A}^{\dagger}$ is defined using the L2-norm ($\mathbf{G} = \mathbf{I}$), $\mathbf{R}^{\mathrm{T}}(t,0)$ represents the propagator of the adjoint equation (3). From the Rayleigh quotient theorem, (7) shows that the eigenvector of $\mathbf{R}^{\mathrm{T}}(t,0)\mathbf{X}\mathbf{R}(0,t)$ with the largest eigenvalue is the fastest growing perturbation of the tangent linear equation (2). This is identical to the right singular vector of $\mathbf{R}(0,t)$ with largest singular value $\mu^{1/2}$. The singular vectors of $\mathbf{R}(0,t)$ are typically a better measure of the stability of the tangent linear equation (2), and for this reason they have been used extensively for assessing the predictability of the atmospheric circulation.

The singular vectors of $\mathbf{R}(0,t)$ can be computed for ROMS by noting that a single integration of a perturbation $\mathbf{u}$ forward in time with the tangent linear model (2) over $[0,t]$, multiplication of the result by $\mathbf{X}$, followed by an integration of the result backwards in time with the adjoint model over $[t,0]$ is all equivalent to the matrix–vector operation $\mathbf{R}^{\mathrm{T}}(t,0)\mathbf{X}\mathbf{R}(0,t)\mathbf{u}$. The implied matrix operator $\mathbf{R}^{\mathrm{T}}(t,0)\mathbf{X}\mathbf{R}(0,t)$ is symmetric, so a Lanczos algorithm (Golub and van Loan, 1989) can be used to find select eigenvectors and eigenvalues. In practice this is accomplished using the ARPACK routine **dsaupd**.

3.4. Stochastic optimals and forcing singular vectors

In general, the ocean circulation is subject to ubiquitous environmental noise forcing arising from numerous sources, including stochastic variations in surface forcing and nonlinear backscatter due to internal processes within the ocean. Perturbations induced in the ocean by stochastic forcing can amplify considerably depending on how non-normal the underlying tangent linear dynamics happen to be. The sources of energy for these perturbations are the perturbation eddy stresses usually associated with classic barotropic and baroclinic instability. However, these sources of energy can be significant even for asymptotically stable systems. In a stochastically forced, asymptotically stable, non-normal system in statistical equilibrium, there is a balance between the input of perturbation energy by the stochastic forcing, the release of perturbation energy from the basic state flow $\mathbf{S}_0$, and dissipation of perturbation energy. The energy sources and sinks due to the basic state flow can also be interpreted geometrically as due to linear interference of the non-orthogonal eigenmodes.

Consider the stochastically forced tangent linear equation:

$$\frac{\partial \mathbf{s}}{\partial t} = \mathbf{A}\mathbf{s} + \mathbf{f}(t), \tag{8}$$

where $\mathbf{f}(t)$ represents the stochastic forcing, and it is assumed that $\mathbf{A}$ is asymptotically stable. We will make two assumptions about the stochastic forcing $\mathbf{f}$: (i) that $\langle \mathbf{f}(t) \rangle = 0$, where $\langle \cdot \rangle$ is the expectation/ensemble average operator, and (ii) that the covariance of $\mathbf{f}$ is separable in space and time as $\langle \mathbf{f}(t)\mathbf{f}^{\mathrm{T}}(t') \rangle = D(t,t')\mathbf{C}$, where $D(t,t')$ describes the temporal correlation of $\mathbf{f}(t)$, and $\mathbf{C}$ is its spatial covariance at zero time lag. Assumption (i) is not as restrictive as it may appear as a non-zero ensemble mean of $\mathbf{f}$ can be absorbed into $\mathbf{A}$ (Gardiner, 1985, section 4.1). Consider an ensemble of solutions $\mathbf{s}(t)$ of (8) generated with different realizations of $\mathbf{f}(t)$ drawn from the same distribution. Under these assumptions, the stochastically induced variance $V_{\mathbf{X}}$ about the ensemble mean at some time t can be expressed as:

$$V_{\mathbf{X}}(t) = \mathrm{tr}\{\mathbf{Z}\mathbf{C}\}, \tag{9}$$

where $\mathbf{Z} = \int_0^t \int_0^t D(t', t'')\mathbf{R}^{\mathrm{T}}(t, t'')\mathbf{X}\mathbf{R}(t', t)\,\mathrm{d}t'\,\mathrm{d}t''$ (Farrell and Ioannou, 1993c). For a stationary ergodic system (as in the case where $\mathbf{R}$ is asymptotically stable), the ensemble mean state will be identical to the time mean of an individual ensemble member (Gardiner, 1985, section 3.7.1), so $V_{\mathbf{X}}$ can also be interpreted as the variance of a single ensemble member about its time mean. The symmetry of $\mathbf{C}$ and $\mathbf{Z}$ allows (9) to be rewritten as:

$$V_{\mathbf{X}}(t) = \sum_{i=1}^{N}\sum_{j=1}^{N} q_i p_j (\mathbf{Q}_i \cdot \mathbf{P}_j)^2, \tag{10}$$

where $\{p_j, \mathbf{P}_j\}$ are the {eigenvalue, eigenvector} pairs of $\mathbf{C}$ (i.e. the EOFs of the stochastic forcing $\mathbf{f}$ at zero time lag), and $\{q_i, \mathbf{Q}_i\}$ the corresponding eigenvalues and eigenvectors of $\mathbf{Z}$.

The eigenvectors of $\mathbf{Z}$ are often referred to as the forcing orthogonal functions (FOFs) (Farrell and Ioannou, 1993c) or stochastic optimals (SOs) (Kleeman and Moore, 1997). They form an orthogonal set of eigenvectors, and like the EOFs of $\mathbf{C}$ they can be ordered according to decreasing magnitude of the eigenvalues q_i. Suppose that the stochastic forcing $\mathbf{f}(t)$ is composed of all the SOs, and that each explains an equal fraction of the *forcing variance*. The SO ($\mathbf{Q}_1$) with largest eigenvalue then represents the spatial pattern of stochastic forcing that accounts for the largest fraction of stochastically induced variability *in the model*, given by $q_1/\mathrm{tr}(\mathbf{Z})$, about the ensemble mean $\langle \mathbf{s}(t) \rangle$. Thus $\mathbf{Q}_1$ represents the pattern of stochastic forcing that is most disruptive to the tangent linear model (8) for a given ocean circulation $\mathbf{S}_0$. The physical meaning of (10) is therefore very clear; the level of stochastically induced variability depends on the projection of the EOFs of $\mathbf{f}$ (characterized by $\mathbf{C}$) onto the SOs of the TLM. Recent theoretical calculations and modelling by Chang et al. (2004a,b) also shows that of all possible stochastic forcing, $\mathbf{Q}_1$ yields the signal with greatest predictability.

The SOs also have physical significance for an asymptotically unstable operator $\mathbf{R}$ since $\langle \int_0^{\mathrm{T}} \mathbf{R}(0,\tau)\mathbf{f}(\tau)d\tau \rangle = 0$ for a white noise forcing $\mathbf{f}(\tau)$ regardless of the stability of $\mathbf{R}$ (Gardiner, 1985, section 4.2.6(e)). In this case, the ensemble mean $\bar{\mathbf{s}}(t) = \langle \mathbf{s}(t) \rangle = \mathbf{R}(0,t)\mathbf{s}(0)$, and $V_{\mathbf{X}}$ represents the variance of the resulting ensemble about $\bar{\mathbf{s}}$ at time t. The SOs provide information about the fraction of ensemble variance that can be explained at this time which is useful for interpreting forecast ensembles and the inherent predictability of the forecast flow field.

The integrand of $\mathbf{Z}$ can be evaluated in much the same way as described in Section 3.3. However, the time integrals in $\mathbf{Z}$ generally render the implicit computation of the action of $\mathbf{Z}$ on an arbitrary vector very expensive computationally. If it is assumed that the stochastic forcing $\mathbf{f}(t)$ is white in time, one of the time integrals in $\mathbf{Z}$ disappears. The remaining time integral can be approximated using the trapezoidal rule, and the action of $\mathbf{Z}$ on a vector becomes computationally tractable. A suitable choice of semi-norm can also greatly simplify the computation of $\mathbf{Z}$ (Moore et al., 2003b). Since $\mathbf{Z}$ represents a symmetric matrix, the SOs can be evaluated using the ARPACK Lanczos algorithm **dsaupd**.

Consider again Eq. (8) but suppose now that the forcing $\mathbf{f}$ is constant in time. The solution of (8) can be written as:

$$\mathbf{s}(t) = \mathbf{R}(0,t)\mathbf{s}(0) + \int_0^{\mathrm{t}} \mathbf{R}(t',t)\mathbf{f}\,\mathrm{d}t'. \tag{11}$$

Suppose also that the initial condition $\mathbf{s}(0) = 0$. Since $\mathbf{f}$ is independent of time, (11) can then be written in this case as $\mathbf{s}(t) = \mathbf{M}(t)\mathbf{f}$, where $\mathbf{M} = \int_0^{\mathrm{t}} \mathbf{R}(t',t)\,\mathrm{d}t'$. Of interest is the structure of the

constant forcing $\mathbf{f}$ that maximizes the value of some norm, such as energy, over the time interval $[0, t]$. Equivalently we would like to find $\mathbf{f}$ that maximizes:

$$\lambda = \frac{\mathbf{s}^{\mathrm{T}}(t)\mathbf{X}\mathbf{s}(t)}{\mathbf{f}^{\mathrm{T}}\mathbf{f}} = \frac{\mathbf{f}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}(t)\mathbf{X}\mathbf{M}(t)\mathbf{f}}{\mathbf{f}^{\mathrm{T}}\mathbf{f}}. \tag{12}$$

The forcing pattern that maximizes λ is clearly the eigenvector of $\mathbf{M}^{\mathrm{T}}(t)\mathbf{X}\mathbf{M}(t)$ with largest eigenvalue. Because of the obvious connection with the singular vectors of Section 3.3, the eigenvectors of $\mathbf{M}^{\mathrm{T}}\mathbf{X}\mathbf{M}$ are called the forcing singular vectors (FSVs) (Barkmeijer et al., 2003). FSVs are of interest in numerical weather prediction because they can be used to generate ensembles of forecasts associated with different possible realizations of systematic errors in model physical parameterizations. In ocean modelling and prediction, FSVs may prove useful for exploring the impact of systematic errors in surface forcing fields.

4. Illustrative examples: an ocean double gyre

In this section we will present example calculations that use the ROMS tangent linear model, adjoint model, and the platforms described in Section 3.

As a general illustrative example, we consider a double gyre ocean circulation in a rectangular basin. This circulation has been studied extensively in the oceanographic literature as a surrogate for the subtropical and subpolar gyres of the major world oceans (see Pedlosky (1996) and Chang et al. (2001) for excellent indepth reviews). In the present example, the model is configured in the form a flat-bottomed, rectangular ocean basin 1000 km in longitude, 2000 km in latitude, and 500 m deep. The ROMS equations were solved on a mid-latitude β-plane centered at 45 °N with a horizontal resolution of 18.5 km. The model consists of four equally spaced vertical levels of thickness 125 m.

The model circulation was forced by (i) a zonally uniform, zonal wind stress of the form $\tau_x = -\tau_0 \cos(2\pi y/L_y)$, where $\tau_0 = 0.05\ \mathrm{N\,m^{-2}}$ and L_y is the meridional extent of the basin, and (ii) weakly relaxing the temperature of the water column to a simple analytical profile which is a function of both latitude and depth, and is given by $T_c(y,z) = h(z)(1 - 3y/(h(z)L_y))$ where $h(z) = T_0 + a\mathrm{e}^{z/100}(10 - 0.4\tanh(z/100))$, $a = 0.95$ °C and $T_0 = 5$ °C. The following ROMS options were used in the calculations presented here: third-order upstream horizontal advection, second-order horizontal diffusion along z-surfaces, linear bottom drag, and second-order vertical viscosity and diffusion. The pertinent model parameters are: coefficient of horizontal eddy diffusivity, $1.28\times10^{3}\ \mathrm{m^2\,s^{-1}}$; coefficients of vertical viscosity and diffusivity, $0.1\ \mathrm{m^2\,s^{-1}}$; coefficient of bottom drag, $8\times10^{-7}\ \mathrm{s^{-1}}$; meridional gradient of the Coriolis parameter $\beta = 2\times10^{-11}\ \mathrm{m^{-1}\,s^{-1}}$. No explicit horizontal viscosity was used in the model since the implicit dissipation inherent in the third-order upstream horizontal advection scheme is sufficient to render well behaved solutions (Shchepetkin and McWilliams, 1998). The nonlinear model equations were solved subject to no-slip boundary conditions on velocity and no-gradient conditions on temperature at the meridional and zonal boundaries. The tangent linear and adjoint model boundary conditions were also no-slip and no-gradient, and in addition, zero tangent linear and adjoint fluxes of heat and momentum were prescribed at the surface and bottom.

Using the aforementioned forcing, the nonlinear model (1) was run from a state of rest for 50 years. Fig. 1 shows a timeseries of the basin integrated kinetic energy from the model, and indicates that after a spin-up period of about 5 years, the model has reached a statistically steady state. Snap-shots of the model velocity and temperature in the upper layer are shown in Fig. 2 during the first 10 days of year 51. The resulting double gyre circulation is unsteady, and during the 10 day period shown two temperature fronts are present near the western boundary. We will refer to these as the northern and southern front regions which are labelled N and S in Fig. 2a

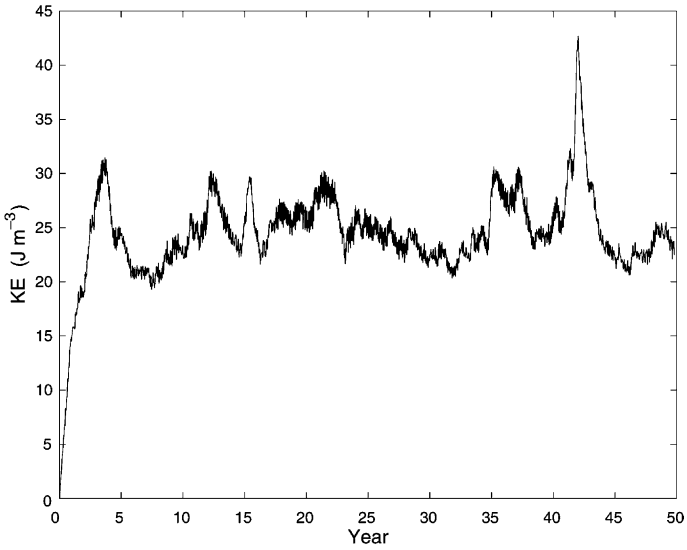

Fig. 1. A timeseries of basin integrated kinetic energy from the model during the 50 year spin-up.

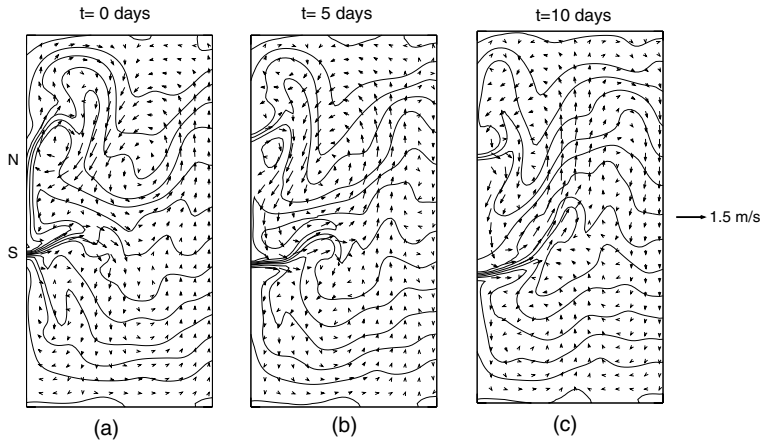

Fig. 2. Snap-shots of the ocean surface temperature (contours) and surface current (vectors) every 5 days during the first 10 days of year 51 of the model integration. The contour interval is 0.2 °C and the vector scale is indicated.

respectively. A strong jet is associated with front S and flows north-eastwards away from the western boundary. As the circulation evolves the two fronts merge near the central basin latitude around day 25 (not shown). The 10 day circulation shown in Fig. 2 will be used as the time evolving basic state in all of the calculations described below.

In addition to the ubiquitous published literature on the double gyre circulation, another primary motivation for considering this circulation in the present context is that similar calculations have been presented by Moore (1999b) and Moore et al. (2002) using a barotropic quasi-geostrophic model. Thus some direct comparisons of the quasi-geostrophic and primitive equation calculations can be made.

4.1. Finite time eigenmodes

As a first application of the tangent linear version of ROMS we will consider the FTEs of the circulation depicted in Fig. 2. Recall from Section 3.1 that the FTEs are the eigenvectors of the propagator $\mathbf{R}(0,\tau)$ linearized about a time evolving circulation. For the calculations presented here, we have chosen $\tau = 10$ days, and have used the time evolving flow in Fig. 2a–c. The eigenvalues of the first 200 members of the FTE spectrum with largest amplitude in this case are shown in Fig. 3a. Recall that since $\mathbf{R}(0,\tau)$ represents a real matrix, the eigenvalues occur in complex conjugate pairs. The eigenvalues with magnitudes greater than 1 represent eigenvectors which grow in amplitude over the 10 day period, and in Fig. 3 these fall outside the shaded circle. Eigenvectors with eigenvalues that fall within the shaded circle all decay over the interval τ. In all there are 23,328 members of the entire FTE spectrum for the current configuration of ROMS, and the remaining 23,128 would populate the central region of the shaded circle in Fig. 3a.

The sea level associated with the real and imaginary components of the first three fastest growing FTEs is shown in Fig. 4 which is an excellent surrogate for barotropic stream function. The temperature and velocity structures of the FTEs exhibit similar features (not shown). The

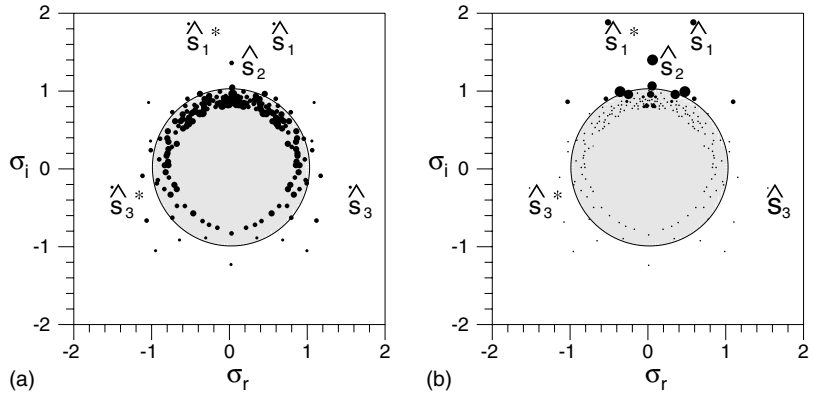

Fig. 3. The complex plane showing the eigenvalues $\sigma = \sigma_r + \mathrm{i}\sigma_i$ of the first 200 members of the FTE spectrum with largest eigenvalue amplitude. The size of each dot is directly proportional to $\log_{10}(|v|)$ in (a) and to $|a_k|$ in (b) where v and a_k are defined in Sections 4.2 and 4.3 respectively.

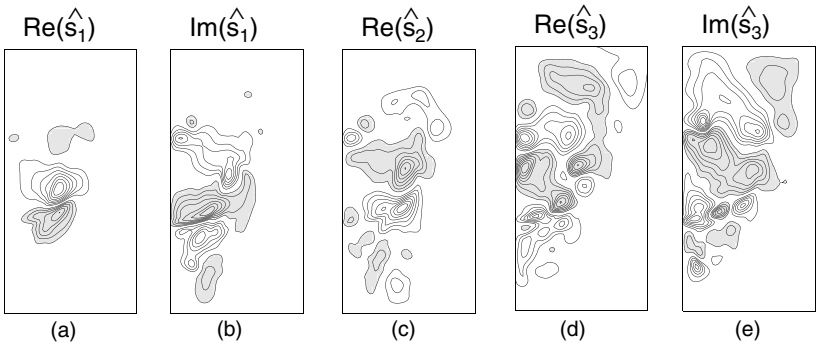

Fig. 4. Contour plots of perturbation sea level for the real and imaginary components of the first three members of the FTE spectrum with largest eigenvalue modulus. In each panel the contour interval is arbitrary, and shaded and unshaded regions are of opposite sign.

eigenvalues of these FTEs are identified in Fig. 3a. The structures of the FTEs suggest that growth is achieved through barotropic processes along the northern and southern fronts and jet that are present in the basic state flow in Fig. 2a–c. The vertical structure of the FTEs (not shown) reveals that growth occurs through baroclinic processes also. The structures of the FTEs shown in Fig. 4 are fairly typical of those possessed by many other members of the spectrum. The growth of the FTEs is fairly modest however, and the amplitude of the fastest growing, $\hat{\mathbf{s}}_1$, increases by only a factor of 1.4 over the 10 day interval τ. Recall that for FTEs, there is no formal requirement for this growth to occur exponentially in time as in the case of an autonomous system, so we cannot assign an exponential doubling time to them.

4.2. Adjoint finite time eigenmodes

As noted in Section 3.2, the eigenvectors of the adjoint propagator $\mathbf{R}^{\mathrm{T}}(\tau,0)$ are also of interest for several reasons. First, courtesy of the biorthogonality relation (4) they can be used to compute the projection of any arbitrary perturbation on the eigenvectors of $\mathbf{R}(0,\tau)$. Second, the eigenvectors of $\mathbf{R}^{\mathrm{T}}(\tau,0)$ and $\mathbf{R}(0,\tau)$ can be used to quantify the non-normality of a dynamical system. Third, the eigenvectors of $\mathbf{R}^{\mathrm{T}}(\tau,0)$ represent the perturbations that have the largest projection on the corresponding eigenvectors of $\mathbf{R}(\tau,0)$ and at the same time maximize the growth of the L2-norm (see the discussion surrounding Eqs. (5) and (6)). As noted in Section 2, the adjoint version of ROMS was derived using an L2-norm. However, it is easy to show that the same adjoint model can be used to derive the adjoint relative to other norms of interest by way of simple linear transformations of the input and output vectors (Ehrendorfer and Tribbia, 1997) as noted in Section 2.

As a first application of the adjoint tangent linear version of ROMS we will consider the adjoint finite time eigenmodes (AFTEs) of the circulation depicted in Fig. 2. As in Section 4.1 we will restrict our attention to the first 10 days of the time evolving basic state circulation shown in Fig. 2a–c. The eigenvalue spectrum of $\mathbf{R}^{\mathrm{T}}(\tau,0)$ is identical to that of $\mathbf{R}(0,\tau)$ shown in Fig. 3a. Fig. 5 shows the sea level associated with the real and imaginary components of the first three fastest growing AFTEs. These three AFTEs have eigenvalues that are the complex conjugates of the

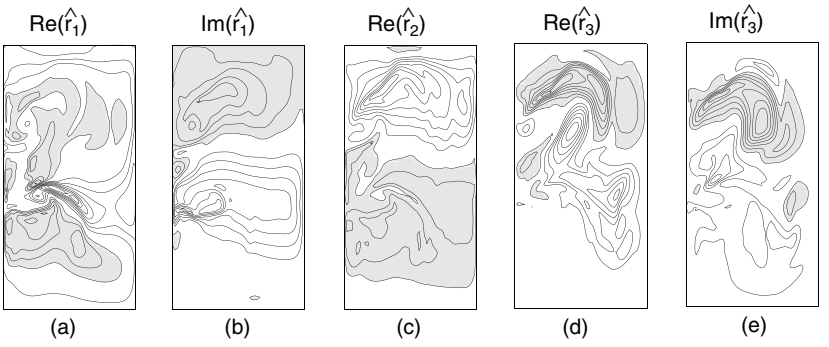

Fig. 5. Same as Fig. 4 except for the AFTEs.

FTEs shown in Fig. 4. Inspection of Figs. 4 and 5 reveals that the FTEs and their optimal excitations, the AFTEs, have very different spatial structures. This is an important feature of non-normal (and non-self-adjoint) systems. Thus the optimal way to excite a particular FTE with large amplitude is to perturb the flow with a disturbance that has a structure very dissimilar to that of the FTE. The AFTEs have sea level structures that are generally large-scale over much of the basin, and complicated in the vicinity of the ocean jets and fronts. The large-scale components of the first two AFTEs are reminiscent of low wavenumber adjoint barotropic basin modes, in agreement with related quasi-geostrophic calculations described by Moore (1999b). The AFTEs in Fig. 5 would excite the corresponding FTEs shown in Fig. 4 and at the same time maximize the growth of the L2-norm over the 10 day interval τ. The AFTEs that maximize the growth of perturbation energy over the same time interval have very similar spatial structures (not shown).

The dissimilarity in the structure of the FTEs and corresponding AFTEs is due to the non-normal nature of the tangent linear propagator $\mathbf{R}(0,\tau)$ of (2) linearized about the time evolving basic state circulation in Fig. 2. It is the inhomogeneities in the basic state circulation that render the system non-normal, and in fact, except under very special circumstances, non-autonomous systems will always be non-normal (Farrell and Ioannou, 1999). Following Farrell and Ioannou (1999) the degree of non-normality of the tangent linear system can be quantified by considering the dot product of the FTEs and AFTEs. Recall that for two real vectors, **a** and **b**, the dot product is defined as $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$ where θ is the angle subtended by the two vectors. Now suppose that **a** is an FTE and **b** is the corresponding AFTE for the L2-norm. As a consequence of the biorthogonality relation (4), the AFTE **b** will be orthogonal to all other FTEs except **a**. Thus the angle subtended by **b** and all other members of the FTE eigenspectrum will be $\pi/2$. Non-normality is related to the non-orthogonality of the FTEs, or equivalently their degree of linear dependence. The closer the value of $|\theta|$ to $\pi/2$, the more non-normal the FTE is considered to be ecause as $\theta \to \pm\pi/2$, the more parallel or antiparallel (i.e. the more linearly dependent) **a** becomes to the other members of the FTE spectrum. In general, the FTEs and AFTEs will be complex vectors, so θ will also be complex. However, the degree of non-normality can still be conveniently quantified using the dot product of the vectors. Following Farrell and Ioannou (1999) we define a measure of non-normality for a particular FTE as:

$$v_k = |\mathbf{U}^{-1}\hat{\mathbf{r}}_k||\mathbf{U}^{\mathbf{T}}\hat{\mathbf{s}}_k|/\hat{\mathbf{r}}_k^{\mathrm{H}}\hat{\mathbf{s}}_k, \tag{13}$$

where $\mathbf{G} = \mathbf{U}\mathbf{U}^{\mathrm{T}}$ introduced in Section 2 has now been factorized, and the notation of Sections 3.1 and 3.2 has been used.

As an example of this type of calculation, we will restrict our attention to the perturbation energy norm. The perturbation energy equation and the form of $\mathbf{U}$ in this case are given in Moore et al. (2003a). Return now to the eigenvalue spectrum depicted in Fig. 3a. The size of each point in this figure is directly proportional to $\log_{10}(|\nu_k|)$ for each of the corresponding eigenmodes. Fig. 3a shows that the most non-normal FTEs characterized by large values of $|\nu_k|$ are typically the damped modes of the tangent linear system in this case. We will return to this point in Section 4.3.

4.3. Singular vectors

Recall from Section 3.3 that while the AFTEs are optimal for exciting the corresponding FTE, they are not the fastest growing of all possible perturbations. This latter class of perturbations are the singular vectors (SVs) of the tangent linear propagator $\mathbf{R}(0,\tau)$. According to (7), the SVs are the eigenvectors of $\mathbf{R}^{\mathrm{T}}(\tau,0)\mathbf{X}\mathbf{R}(0,\tau)$, and these computations require the use of the tangent linear and adjoint versions of ROMS in tandem. In all of the SV calculations documented here, $\mathbf{X}$ was chosen to define the perturbation energy norm.

The influence of the optimal growth time τ on the SVs will be considered first. Fig. 6a shows perturbation energy growth factors, μ in (7), for the first 10 members of the SV spectrum for $\tau = 3$, 10 and 30 days. The first τ days of the time evolving basic state shown in Fig. 2 were used in each case (however only the first 10 days of the basic state trajectory are shown in Fig. 2). Fig. 6a indicates that while μ increases with τ, the shape of the spectrum is similar in each case. In general, the SV spectrum is dominated by a few members as illustrated in Fig. 6b which shows μ for the first 100 members for the case $\tau = 10$ days. Beyond SV number 20 or so, the spectrum is relatively flat, and the energy growth factors are an order of magnitude smaller than that of the fastest growing SV.

The perturbation sea level at initial time and time τ is shown in Fig. 7 for the fastest growing SV when $\tau = 10$ days. The initial SV structure is generally large-scale, and has a more complex structure where the southern most front and jet in Fig. 2 separate from the coast. The first three members of the SV spectrum (not shown) typically favour the southern front region, while SV number 4 (not shown) favours the northern front region. SV number 5 and many others (not shown) favour both regions for growth. The horizontal and vertical structures (not shown) of the SVs as they evolve reveal that perturbation energy growth occurs via barotropic and baroclinic processes, the source of perturbation energy being the shearing and straining flows associated with the jet circulation. The SVs for other optimization times τ up to 30 days (not shown) are qualitatively similar to those discussed above. The tendency for the initial SV structures to favour the western boundary agrees qualitatively with the related quasi-geostrophic calculations of Moore et al. (2002) who showed that this tendency was due to the interaction of short barotropic Rossby waves with the western boundary currents.

The initial large-scale structure of the SV for $\tau = 10$ days is qualitatively quite similar in some aspects to the AFTEs shown in Fig. 5. The super exponential growth of the SVs cannot be sustained beyond a finite time, so to is of interest to compute the projection of the SVs on the FTEs to see which structures are likely to emerge at later times. The projection of any perturbation on each member of the FTE spectrum can be computed by exploiting the biorthogonality

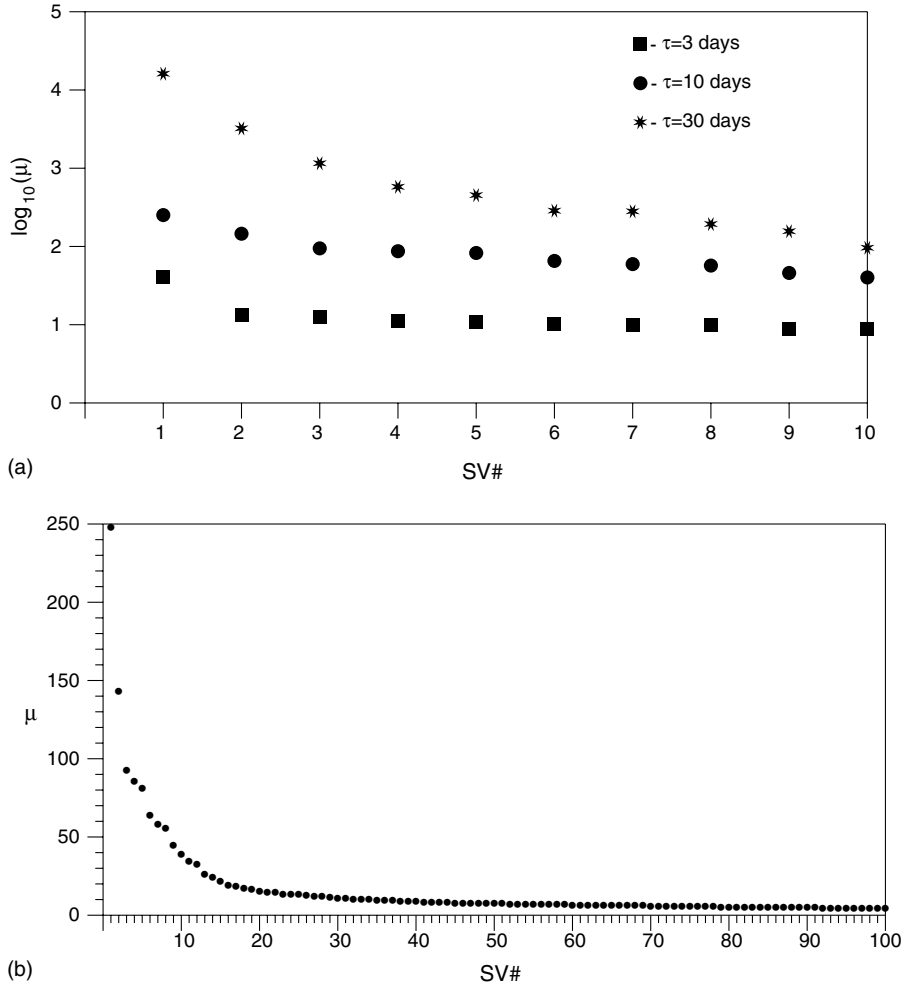

Fig. 6. (a) A plot of $\log_{10}\mu$ vs. SV number for the 10 fastest growing SVs when the optimal growth time $\tau = 3$, 10 and 30 days, and μ is the perturbation energy growth factor of each SV over the appropriate interval τ. (b) A plot of μ vs. SV number for the 100 fastest growing SVs when $\tau = 10$ days.

relation (4). If we denote an SV as $\mathbf{p}$, then $\mathbf{p}$ can be represented as a linear combination of the FTEs $\hat{\mathbf{s}}$, so that $\mathbf{p} = \sum_i a_i \hat{\mathbf{s}}_i$, where in general the amplitude coefficients a_i will be complex. Using (4), the projection of $\mathbf{p}$ on any FTE, say $\hat{\mathbf{s}}_k$, as measured by the magnitude of the amplitude coefficient is given by $a_k = \hat{\mathbf{r}}_k^{\mathrm{H}}\mathbf{p}/(\hat{\mathbf{r}}_k^{\mathrm{H}}\hat{\mathbf{s}}_k)$. Fig. 3b shows a different representation of the FTE spectrum. In this case the size of each point is directly proportional to the modulus of the amplitude coefficient of each FTE computed from the fastest growing SV for $\tau = 10$ days (cf Fig. 7). Fig. 3b indicates that this SV projects most on the FTE shown in Fig. 4c, namely $\hat{\mathbf{s}}_2$. However, unlike the FTE $\hat{\mathbf{s}}_2$ which has an energy growth factor of only 1.45 over the interval τ, the SV of Fig. 7 grows in energy by a factor of almost 250.

The ability of SVs to achieve and sustain such rapid transient super exponential growth is a consequence of the non-normality of the tangent linear propagator $\mathbf{R}(0,\tau)$. The basic idea is as

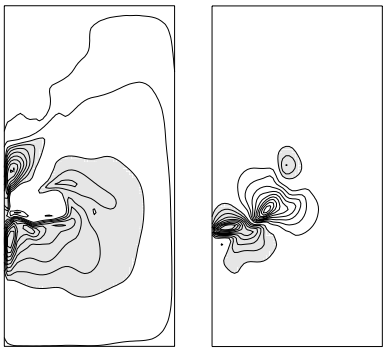

Fig. 7. Contour plots of perturbation sea level at the initial and final time τ for the fastest growing SV when $\tau = 10$ days. In each panel the contour interval is arbitrary, and shaded and unshaded regions are of opposite sign.

follows. Since the FTEs are non-orthogonal, they have a non-zero projection on each other. Thus it is possible to take the FTE $\hat{\mathbf{s}}_2$ in Fig. 4c with a large amplitude, and conceal it with a linear superposition of the other members of the FTE spectrum. Obviously the more non-normal, or linearly dependent, the individual FTEs happen to be, the easier it will be to conceal a particular FTE in this way. Fig. 3b shows that the SV of Fig. 7 also has a large projection on other damped members of the FTE spectrum, and a comparison of Fig. 3b with Fig. 3a reveals that these FTE are relatively non-normal as measured by their values of $|\nu|$. Thus it appears that the FTE $\hat{\mathbf{s}}_2$ can be concealed with a large amplitude by a linear superposition of other relatively non-normal, damped FTEs. This superposition constitutes the SV of Fig. 7a. As the SV evolves, these latter FTEs decay in amplitude and change their phase relative to $\hat{\mathbf{s}}_2$, and $\hat{\mathbf{s}}_2$ is revealed with large amplitude.

4.4. Forcing singular vectors and stochastic optimals

As discussed in Section 3.4, the notion of singular vectors as the fastest growing perturbations from the point of view of an initial value problem can be generalized to perturbation variance maintained by forcing. Two cases were considered, namely, where the forcing was held constant over some time interval τ, and where the forcing was assumed to be stochastic in time with a zero ensemble mean value. The former leads to the idea of forcing singular vectors (FSVs), while the latter leads to stochastic optimals (SOs). Examples of both FSVs and SOs will considered in this section.

The FSVs of the time evolving basic state circulation in Fig. 2 will be considered first, assuming a time interval $\tau = 10$ days. The eigenvalues (λ in (12)) of the first 10 members of the FSV spectrum in this case are shown in Fig. 8. The values of λ can be interpreted as the perturbation energy that the tangent linear model would possess after 10 days if subjected to a constant forcing with unit norm with the structure of the associated FSV, and assuming a zero initial perturbation. The shape of the FSV spectrum is similar to that of the SVs and appears to be dominated by just a few members.

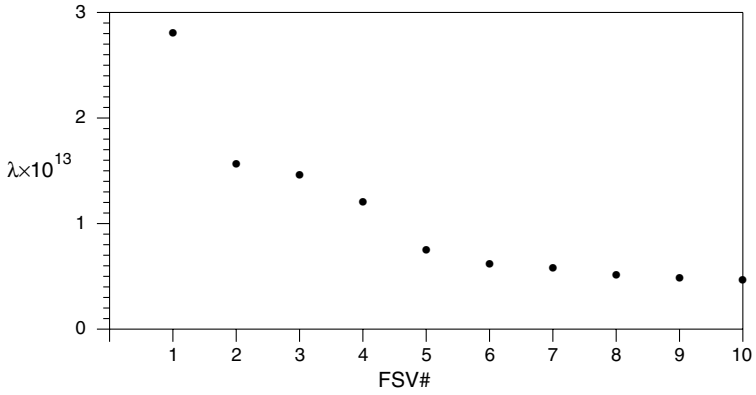

Fig. 8. A plot of λ vs. FSV number for the first 10 members of the FSV spectrum, where λ is defined in (12).

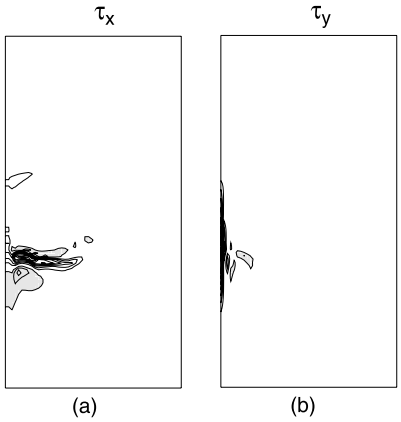

Fig. 9. Contour plots of the zonal (τ_x) and meridional (τ_y) wind stress components of the first member of the FSV spectrum. In each panel the contour interval is arbitrary, and shaded and unshaded regions are of opposite sign.

The structure of the zonal and meridional wind stress component of the first member of the FSV spectrum is shown in Fig. 9. The FSV has a structure that favours the southern front region which is the area of growth of the first three members of the SV spectrum shown in Fig. 6b. Other members of the FSV spectrum (not shown) favour the northern front region where SV number 5 and others are located. Thus it would appear that the FSVs are able to maximize energy growth by continuously exciting perturbations in the regions where large sources of perturbation energy exist by virtue of the basic state flow structure. The tendency of the FSVs to favour the western boundary is in qualitative agreement with the findings of Moore (1999b) and Moore et al. (2002).

We will consider now the SOs of the basic state circulation shown in Fig. 2, assuming as before a time interval $\tau = 10$ days. In the examples shown here, we have assumed that the stochastic forcing is white in time, in which case the temporal correlation function in $\mathbf{Z}$ of (9) is a δ-function, and one of the time integrals disappears. Recall that the eigenvalues q_i of $\mathbf{Z}$ divided by $\mathrm{tr}(\mathbf{Z})$ represent the fraction of stochastically induced variance that the corresponding stochastic optimal

$\mathbf{Q}_i$ would account for in the system if each SO was present in the forcing $\mathbf{f}(t)$ and explained equal forcing variance. Since the matrix $\mathbf{Z}$ is not explicitly computed, the randomized trace estimation technique described by Bai et al. (1995) was used to estimate $\mathrm{tr}(\mathbf{Z})$. Fifty such estimates of $\mathrm{tr}(\mathbf{Z})$ were computed using this method, and the mean trace (denoted $\overline{\mathrm{tr}(\mathbf{Z})}$) and standard deviation of the trace estimates (denoted σ_{tr}) were computed. Fig. 10 shows the percentage explained stochastically induced variance for each of the first 10 members of the SO spectrum. The vertical bars associated with each point show the range of explained variance for $\overline{\mathrm{tr}(\mathbf{Z})} \pm \sigma_{\mathrm{tr}}$. Under the assumption that all SOs are present in the stochastic forcing and explain equal forcing variance, Fig. 10 indicates that $\mathbf{Q}_1$ would account for about 5% of the stochastically induced variance, and that collectively the first 10 SOs would account for about 18% of the stochastically induced variance.

The structure of the zonal and meridional wind stress components of the first member of the SO spectrum is shown in Fig. 11. As for the SVs and FSVs, the SOs favour the western boundary. Qualitatively similar results were obtained for the FSVs and SOs for other time intervals up to 30 days (not shown).

Recall from Fig. 3 that the basic state circulation is unstable. In this case the SOs in Fig. 11 represent the patterns of stochastic forcing that would account for different fractions of the variance at day 10 in an ensemble of stochastically forced flow realizations. If these realizations were, for example, members of an ocean forecast ensemble, then Fig. 11 would indicate which regions of the ocean circulation are particularly sensitive to forcing errors. This will be discussed further in Section 5.

4.5. Quantifying non-normality of autonomous systems

The properties of the various perturbations and forcing fields presented in previous sections depend on the non-normal nature of the tangent linear operator $\mathbf{A}$ in (2). In Section 4.2 we

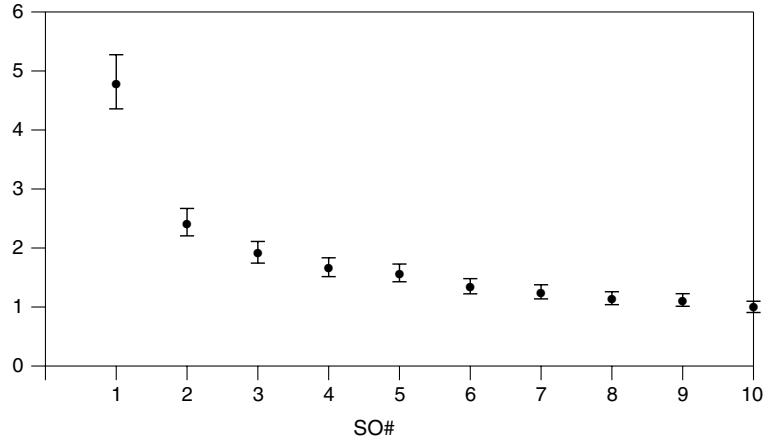

Fig. 10. A plot of $q_i/\overline{\mathrm{tr}(\mathbf{Z})}$ vs. SO number for the first 10 members of the SO spectrum. The vertical bars represent the range of values for $\mathrm{tr}(\mathbf{Z}) \pm \sigma_{\mathrm{tr}}$.

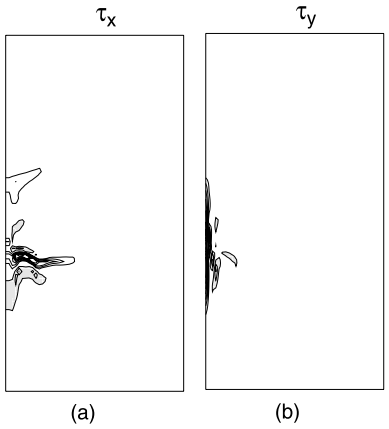

Fig. 11. Contour plots of the zonal (τ_x) and meridional (τ_y) wind stress components of the first member of the SO spectrum. In each panel the contour interval is arbitrary, and shaded and unshaded regions are of opposite sign.

discussed one means of quantifying the non-normality of the system by way of (13). In this section we present an alternative means of quantifying non-normality which can also yield important information about the response of the tangent linear system to forcing at different frequencies, and about the reliability of eigenmode calculations such as those in Sections 4.1 and 4.2.

Consider again the forced tangent linear equation (8) but where now the forcing is at a single frequency, $\mathbf{f}(t) = \mathbf{h}e^{\omega t}$ where ω is in general complex, and $\mathbf{h}$ describes the three-dimensional structure of the forcing field. Assume also that the system is autonomous, and asymptotically stable, in which case the component of the solution of (8) due to the initial condition will decay to zero. Thus after a sufficiently long time $e^{\mathbf{A}t} \to 0$, and solutions of (8) can be written as $\mathbf{s}(t) = e^{\omega t}(\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{h}$. Of interest is the ratio of the norm of the response to the norm of the forcing which is given by $\|\mathbf{s}(t)\|/\|e^{\omega t}\mathbf{h}\| \leqslant \|R(\omega)\|$ where $R(\omega) = (\omega\mathbf{I} - \mathbf{A})^{-1}$ is called the resolvent. Clearly the norm of the resolvent governs the relative amplification of the system response compared to the size of the forcing imposed on the system. As we have seen this amplification is associated with the non-normal character of the system, so $\|R(\omega)\|$ is a useful measure of this property. Contour plots of $\|R(\omega)\|$ vs. ω for ω in the complex plane are often referred to as pseudospectra. The interested reader should consult Trefethen (1996) and Trefethen et al. (1993) for more precise definitions of the pseudospectra and important theorems relating to them. The transient growth of perturbations, such as singular vectors, due to the interference of non-normal eigenmodes can be anticipated from certain properties of the pseudospectra (see Reddy et al. (1993) for more details).

Following Trefethen (1996), it can be shown that $\|R(\omega)\|$ obeys the following inequality:

$$\frac{1}{\operatorname{dist}(\omega, \Lambda(\mathbf{A}))} \leqslant \|R(\omega)\| \leqslant \frac{\kappa(\mathbf{E})}{\operatorname{dist}(\omega, \Lambda(\mathbf{A}))}, \tag{14}$$

where $\Lambda(\mathbf{A})$ denotes the eigenspectrum of $\mathbf{A}$; $\operatorname{dist}(\omega, \Lambda(\mathbf{A}))$ is the distance function and defines the shortest distance in the complex plane between the forcing frequency ω and the eigenspectrum $\Lambda(\mathbf{A})$; $\mathbf{E}$ is the matrix introduced in Section 3.2 whose columns are the eigenvectors of $\mathbf{A}$; and $\kappa(\mathbf{E}) = \|\mathbf{E}\|\|\mathbf{E}^{-1}\|$ is the condition number of $\mathbf{E}$. If the system is normal, the columns of $\mathbf{E}$ are

orthogonal, and for the L2-norm, $\kappa = 1$. In this case (14) becomes an equality, and $\|R(\omega)\|$ simply varies inversely as the distance of the forcing frequency from the eigenspectrum. This describes the well known phenomena of resonance. If, on the other hand, the system is non-normal, the columns of $\mathbf{E}$ will no longer be linearly dependent, and κ may become very large. In this case the response of the system is bounded by $\kappa(\mathbf{E})/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$ which may be very large, even when $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$ is small, by virtue of $\kappa(\mathbf{E})$ being large. The possibility of a large response to forcing at frequencies distant from the eigenspectrum is referred to as pseudoresonance (Trefethen et al., 1993) and some interesting oceanographic examples are presented in Moore et al. (2002) and Aiken et al. (2002).

The ability of a system to pseudoresonate at frequencies far removed from an eigenfrequency is clearly a direct consequence of the non-normality of the system. However, (14) shows that pseudospectra also provide direct information about $\kappa(\mathbf{E})$. This is also of interest because pseudospectra indicate the sensitivity of the eigenvalues of $\mathbf{A}$ to small changes in the elements of $\mathbf{A}$ (Trefethen, 1991). Very large values of $\|R(\omega)\|$ in the vicinity of the eigenspectrum are indicative of ill-conditioning of $\mathbf{E}$ and the associated eigenvalues and eigenvectors should be treated as unreliable.

Pseudospectra can be computed using the tangent linear version of ROMS as a by-product of the eigenmode calculations described in Sections 3.1 and 4.1 using the method described by Wright and Trefethen (2001). The eigenmode calculations of ARPACK proceed using an Arnoldi iteration technique, which itself is achieved by using the Krylov subspace of the associated power method to reduce the original matrix $\mathbf{A}$ to upper Hessenberg form. The pseudospectra of the associated upper Hessenberg matrix $\mathbf{H}$ are a subset of those of $\mathbf{A}$ for the converged eigenvalues. If the L2-norm is used, the pseudospectra of $\mathbf{A}$ (given by $\|R(\omega)\|_2$) are simply the inverse of the smallest singular values of the QR factorization of $\mathbf{H}$. Interested readers should consult Wright and Trefethen (2001) for further details.

An example of a pseudospectra using ROMS and the basic state circulation of Fig. 2 will be presented next. In this case we have used the circulation of Fig. 2a to define an autonomous operator $\mathbf{A}$. The eigenfrequencies of the first 200 members of the eigenspectrum for the resulting autonomous operator $\mathbf{A}$ are shown as points in Fig. 12a. Note that several of the eigenmodes have positive real eigenfrequency components which means that they grow exponentially. Thus we are violating the assumption of asymptotic stability introduced above. Nonetheless, the resulting pseudospectra reveal information about the non-normality of the system, and about the reliability of the eigenfrequencies (and eigenmodes). Contours of $\|R(\omega)\|_2$ are also shown in Fig. 12a and represent the pseudospectra of the autonomous propagator defined by the circulation in Fig. 2a. Notice that the contour interval is not uniform, and that large values are found in the vicinity of the eigenfrequencies as expected. The unshaded region near the origin represents the eigenfrequencies of eigenvectors that are considered to be highly unreliable because of the very large values of $\|R(\omega)\|_2$ encountered in this region.

Recall, that if the system was normal, $\|R(\omega)\|_2$ would simply be described by the inverse distance function. Fig. 12b shows contours of $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$ for the eigenspectrum of $\mathbf{A}$. Compared to Fig. 12b, Fig. 12a indicates that the non-normality of $\mathbf{A}$ leads to a considerable distortion of the pseudospectra contours compared to $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$. The ratio of $\|R(\omega)\|_2$ to $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$ is plotted in Fig. 12c and gives a direct indication of the degree of non-normality exhibited by the system since this ratio is everywhere larger than unity.

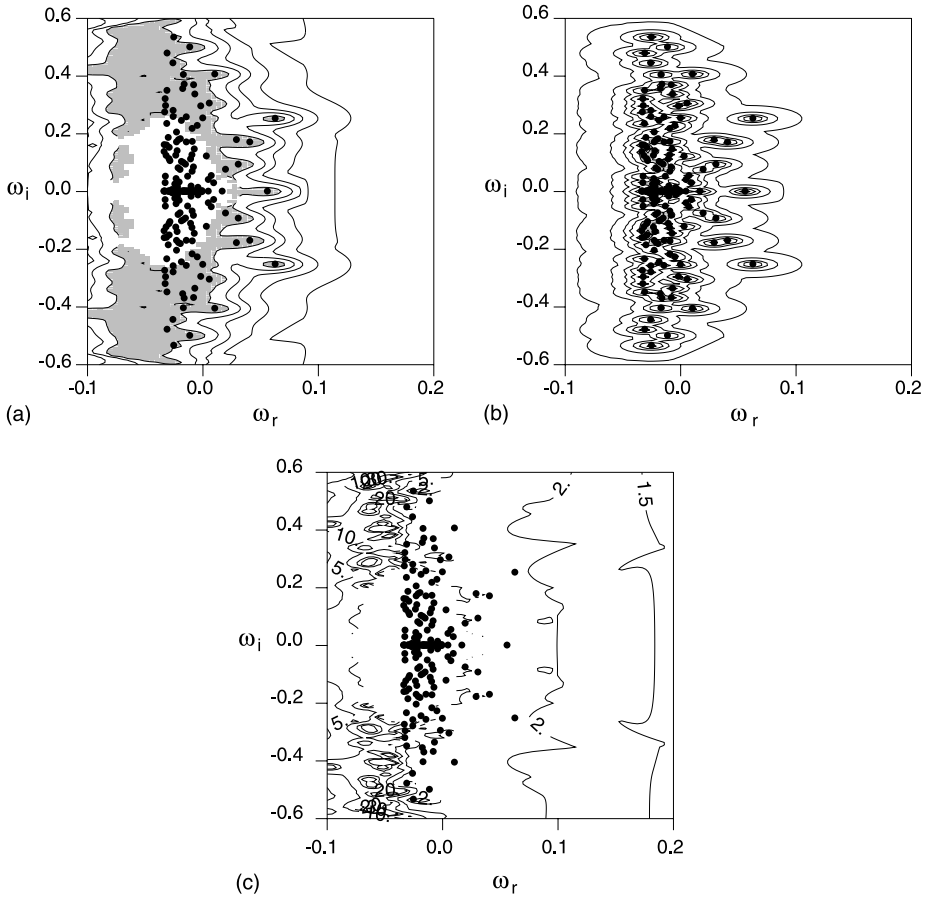

Fig. 12. Contour plots of: (a) $\|R(\omega)\|_2$ in the complex frequency plane where $\omega = \omega_r + i\omega_i$. The points represent the eigenfrequencies of the first 200 members of the eigenspectrum of **A**; (b) $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$, where Λ represents the eigenspectrum of **A** represented by the points; (c) $\|R(\omega)\|\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$. The frequency axes are labelled in units of days^{-1}. In (a) and (b) contours values plotted are 10, 20, 30, 50 and 100, and values greater than 100 are shaded. The unshaded region in (a) near the origin represents values very much greater than 100 and is indicative that the eigenvalues and eigenvectors in this region are unreliable. In (c) contour values plotted are 1.5, 2, 5, 10, 20 and 30.

An example of a pseudospectra for an asymptotically stable circulation is shown in Fig. 13. In this case a steady barotropic double gyre circulation was considered as illustrated in Fig. 13a. The first 200 members of the eigenspectrum of the tangent linear matrix **A** of (2) linearized about this circulation are shown in the lower panel of Fig. 13b. In this case all of the eigenmodes have eigenfrequencies with negative real parts, although there is one eigenfrequency very close to the imaginary axis $\omega_r = 0$. The upper panel of Fig. 13b shows $\|R(\omega)\|_2$ vs. ω for purely oscillatory forcing frequencies ($\omega_r = 0$), which is the case of most geophysical interest. For comparison, Fig. 13b also shows a plot of $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$ vs. ω for the same forcing frequencies. Due to the non-normal character of the eigenmodes of the double gyre circulation, Fig. 13b shows that the system response to any forcing frequency according to $\|R(\omega)\|_2$ is significantly larger than would be expected from the conventional idea of resonance (given by $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$). This is particularly

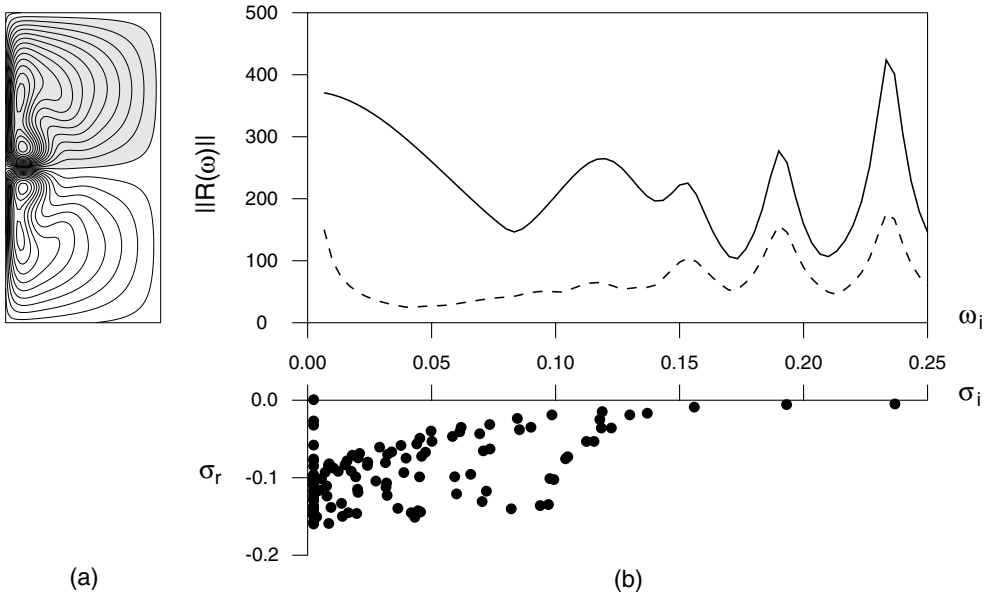

Fig. 13. (a) Contour plot of sea level for the steady barotropic double gyre circulation. Shaded regions represent depressed sea level. (b) A plot of $\|R(\omega)\|_2$ (solid line) and $1/\mathrm{dist}(\omega, \Lambda(\mathbf{A}))$ (dashed line) vs. $\omega = \omega_\mathrm{r} + \mathrm{i}\omega_\mathrm{i}$ when $\omega_r = 0$. The lower panel shows the eigenfrequencies ($\sigma = \sigma_\mathrm{r} + \mathrm{i}\sigma_\mathrm{i}$) associated with the first 200 members of the eigenspectrum of the circulation in (a). Both ω and σ are in units of days^{-1}.

true for forcing with periods between 20 and 50 days ($\omega_i = 0.01$–0.05). In this case $\|R(\omega)\|_2$ is closely related to the energy norm. For a resting ocean the eigenmodes of system (i.e. the "basin modes") are orthogonal with respect to the energy norm and the system dynamics are normal. The straining flow of the double gyre circulation in Fig. 13a destroys the orthogonality of the eigenmodes, and it is straining flow that renders **A** non-normal.

5. Summary

The aim of this paper was to provide a partial review of some of the many powerful applications of tangent linear and adjoint models in oceanography, and to describe a new suite of ocean modelling tools based on the tangent linear and adjoint versions of ROMS that are available to the ocean modelling community at large.

In Section 4 we presented examples of the various drivers that are currently available for ROMS that utilize the tangent linear and adjoint models. As noted in the Section 1, each of these drivers can provide valuable information about various aspects of the ocean circulation. The FTEs described in Section 4.1 represent the dynamical modes of variability of the perturbed circulation, and are one measure of the stability of the basic state flow. The AFTEs of Section 4.2 are of interest because not only do they allow one to compute the projection of any arbitrary perturbation onto the FTEs by virtue of the biorthogonality relation, but they also represent the

optimal excitations for a particular FTE. The FTEs and AFTEs can also be used to quantify the degree of non-normality of the linearized dynamical system.

In Section 4.3 we explored the SVs of the system. In the linear limit, these are the fastest growing perturbations that can exist in the system for the chosen norm and optimal growth time. The growth of SVs is typically initially much faster than suggested by the most unstable FTE, and can be understood in terms of the linear interference of the non-normal FTEs. An extension of these ideas to systems subject to stochastic forcing, or uncertainties in forcing, yields the SOs and FSVs described in Section 4.4. The FSVs are the patterns of forcing that are constant in time that yield the largest growth (in the linear limit) of the chosen norm over the chosen time interval. If the forcing is stochastic in time, the SOs represent the patterns of forcing that account for different fractions of the stochastically induced variance in the system.

The pseudospectra described in Section 4.5 provide information not only about the associated non-normality of the eigenspectrum of the circulation under consideration, but also information about the reliability of FTE and AFTE computations. In addition, pseudospectra reveal information about the response of stable, autonomous systems to forcing at different frequencies, and allow pseudoresonances to be identified.

The aforementioned calculations were performed with (1) linearized about an unsteady, baroclinic, double gyre circulation. The gravest SVs, FSVs, SOs and AFTEs indicate that this circulation is particularly sensitive to perturbations near the western boundary, and in regions characterized by straining and shearing flows. The latter observation is dynamically consistent with what we understand about sources of perturbation energy and energy release from the basic state via barotropic and baroclinic processes. The former observation is consistent with the findings of related calculations using quasi-geostrophic models (Moore, 1999b; Moore et al., 2002).

All of the calculations described in this paper form the foundation of generalized stability thoery (GST) as described by Farrell and Ioannou (1996a,b). GST is a powerful analysis tool since it allows us to examine the stability of a wide class of flows, including time dependent aperiodic flows, analyses which are precluded by traditional techniques based on modes alone.

The calculations presented here also have important practical applications. Some of the perturbations and forcing fields described in Section 3 have been used in operational numerical weather prediction for generating forecast ensembles. Ensemble members may be generated by either perturbing the forecast initial conditions, boundary conditions, or model physics, or a combination of all three. One approach to ensemble generation is to perturb the model in the directions which are known a priori to yield the largest error growth. This can be achieved by perturbing the forecast initial conditions with SVs (Molteni et al., 1993) or by perturbing the model forcing fields with FSVs or SOs. This is the approach adopted at ECMWF where SVs are used to perturb the model initial conditions, and FSVs are used to perturb the model physics. Efforts are also underway to utilize SOs at ECMWF for generating ensembles of seasonal forecasts using coupled models. Forecast ensembles provide information not only about the most likely future state of the system, but also the associated uncertainty in the state, or its predictability. Clearly the same approaches will prove useful for ocean forecasting and for assessing the predictability of the ocean. A ROMS-based ocean ensemble prediction system is currently being developed.

There are several other important uses for tangent linear and adjoint ocean models that were mentioned in Section 1, but which are still under development for ROMS. Specifically several data

assimilation modules will be available for ROMS in the near future that are based on strong and weak constraint 4DVAR. These applications of ROMS will be described in future publications.

In summary, the adjoint of an ocean model essentially represents a general sensitivity operator. Any problem that can be expressed in terms of finding the extrema of a function of the model state variables can typically be solved iteratively using tangent linear and adjoint models. Therefore, we anticipate that the ROMS suite of tangent linear and adjoint tools will prove useful for investigating a wide range and class of scientific problems of interest to the ocean modelling community.

Acknowledgements

This research was supported by grants from the Office of Naval Research (University of Colorado, award number N00014-01-1-0209; Rutgers University N00014-00-1-0227 and N00014-02-1-1034; Scripps Institute of Oceanography N00014-99-1-0045) and the NSF Division of Ocean Sciences (University of Colorado, OCE0121176; Rutgers University, OCE0121506; Scripps Institute of Oceanography OCE0121332).

Appendix A. Adjoint variables and generating functions

In this appendix, the mathematical and physical significance of the adjoint operator and associated adjoint variable will be discussed. This is by no means an exhaustive discussion, and the interested reader should consult other texts, such as Lanczos (1961) for an in-depth exposé. In the following, we will restrict our attention to the L2-norm adjoint which is identical to the matrix transpose for vector spaces.

A.1. Vector space

Consider the $n \times m$ matrix $\mathbf{A}$ where $n < m$. There are two spaces associated with $\mathbf{A}$, one of dimension n, referred to as the N-space, the other of dimension m, referred to as the M-space. Associated with these two spaces are two sets of eigenvectors $\mathbf{u}$ and $\mathbf{v}$ which are respectively the left and right singular vectors of $\mathbf{A}$ and are related to each other according to $\mathbf{A}\mathbf{v} = \lambda\mathbf{u}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{u} = \lambda\mathbf{v}$, where λ are the associated singular values. Suppose that $\mathbf{A}$ has only p non-zero singular values (i.e. $\lambda_i \neq 0$ for $i = 1, 2, \ldots, p$), in which case $\mathbf{A}$ is said to be activated only in the subspace described by the associated p dimensions of N-space and M-space.

Consider now the under-determined linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$. Unique solutions of this system can be found by letting $\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{y}$ where $\mathbf{y}$ is called the generating vector. In this case we are now reduced to solving the linear system $\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{y} = \mathbf{b}$. This is a completely determined system, and from $\mathbf{y}$ we can compute a unique solution $\mathbf{x}$ using the above relation. The nature of this solution can be explored by noting that $\mathbf{A}^{\mathrm{T}} = \mathbf{V}_p\mathbf{\Lambda}_p\mathbf{U}_p^{\mathrm{T}}$ where $\mathbf{U}_p$ and $\mathbf{V}_p$ are the matrices of the p left and right singular vectors that define the p-dimensional subspace, and Λ_p is a diagonal $p \times p$ matrix composed of the non-zero singular values. The condition $\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{y}$ can be expressed as $\mathbf{x} = \mathbf{V}_p\mathbf{q}$ where $\mathbf{q} = \mathbf{\Lambda}_p\mathbf{U}_p^{\mathrm{T}}\mathbf{y}$. The action of $\mathbf{V}_p$ on $\mathbf{q}$ shows that the solution $\mathbf{x}$ lies completely within the activated p-dimensional subspace of $\mathbf{A}$. The unique solution $\mathbf{x}$ found in this way from the generating vector $\mathbf{y}$

is referred to as the natural solution, and is orthogonal to the non-activated dimensions of **A**. This property of the solution arises from the action of the adjoint matrix $\mathbf{A}^{\mathrm{T}}$ on the generating function **y**.

For most problems of physical interest, one might think that **A** will always be a square matrix, representing a uniquely determined linear system. However, this is not always the case when **A** represents a discretized linear operator. We will therefore consider the extension of the above results to function space and differential operators.

A.2. Function space

All of the results described in Section A.1 have a corresponding analogue in continuous space where we consider functions and operators instead of vectors and matrices. In this case we can use generating functions, the analogues of generating vectors, to solve under-determined systems of differential equations. To illustrate this idea, we will consider the linear barotropic vorticity equation as a specific example of general oceanographic interest. This will also illustrate the physical significance that can be assigned to the generating function in this case.

We denote by $\mathbf{u} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$ the three-dimensional velocity vector where (u, v, w) represent continuous, differentiable functions, and it is understood that $w = 0$ given the nature of the system. The vertical component of relative vorticity can be expressed as $q = \mathbf{k} \cdot \operatorname{curl} \mathbf{u}$. The linear barotropic vorticity equation can then be written as:

$$\partial(\mathbf{k} \cdot \operatorname{curl} \mathbf{u})/\partial t + \beta \mathbf{j} \cdot \mathbf{u} = 0, \tag{A.1}$$

where β is the meridional gradient of the Coriolis parameter. Our aim is to solve (A.1) for **u** subject to appropriate boundary conditions, but as it stands this equation is under-determined since q is a scalar quantity and **u** is a vector. However, following Section A.1, we can find a unique solution for **u** by introducing a scalar generating function ϕ. We will define the inner-product $\{f(x,y), g(x,y)\}$ of two arbitrary, continuous, twice differentiable functions as $\int_0^{L_y} \int_0^{L_x} fg \, \mathrm{d}x \, \mathrm{d}y$ where L_x and L_y define the spatial domain. Consider the inner-product of ϕ with (A.1). Using the continuous form of the bilinear identity leading to (3) (the so-called Green's identity), the adjoint of $\mathbf{k} \cdot \operatorname{curl} \mathbf{u}$ is given by $-\mathbf{k} \times \nabla\phi$, in which case we write $\mathbf{u} = \partial\phi/\partial y\mathbf{i} - \partial\phi/\partial x\mathbf{j}$ in order to solve (A.1). Identifying $-\phi$ with a new scalar function ψ, it is clear that the velocity vector must be written in terms of the familiar geostrophic stream function in order to find unique solutions of (A.1). A consequence of this is that $\nabla \cdot \mathbf{u} = 0$ and $q = \nabla^2\psi$, in which case we are reduced to solving:

$$\partial\nabla^2\psi/\partial t + \beta\partial\psi/\partial x = 0, \tag{A.2}$$

which is the familiar form of the barotropic vorticity equation.

The derivation of (A.2) in this way leads to a deeper understanding of the relationship between stream function, velocity and the vertical component of relative vorticity. Following Section A.1, it is now clear that $-\psi$ is the generating function for **u** which leads to unique solutions of (A.1) only in those dimensions that are activated by the vertical component of the curl operator.

Consider now $\nabla^2 \times$(A.2) which yields $\partial\nabla^2 q/\partial t + \beta\partial q/\partial x = 0$. The adjoint of this equation following the continuous analogue of the steps leading to (3) is $\partial\nabla^2 q^\dagger/\partial t + \beta\partial q^\dagger/\partial x = 0$ which is

identical in form to (A.2). Therefore the stream function ψ plays the role of $q^\dagger$, the adjoint vertical component of relative vorticity. Thus ψ, which itself is directly proportional to pressure, corresponds to $q^\dagger$, and is the generating function for $\mathbf{u}$ that leads to unique solution of (A.1) in only those dimensions that are activated by q. So in this example the adjoint variable $q^\dagger \equiv \psi$ has clear physical significance.

References

Aiken, C.M., Moore, A.M., Middleton, J.H., 2002. The non-normal nature of recirculating flows in the coastal zone. J. Phys. Oceanogr. 32, 2955–2974.

Amodei, L., 1995. Solution approchée pour un problème d'assimilation de donnée météorologie avec prise en compte de l'erreur de modèle. Comptes Rendus de l'Académie des Sciences 321 (Série IIa), 1087–1094.

Arango, H.G., Moore, A.M., Miller, A.J., Cornuelle, B.D., Di Lorenzo, E., Neilson, D.J., 2003. The ROMS tangent linear and adjoint models: a comprehensive ocean prediction and analysis system. IMCS, Rutgers University Technical Report No. 2003-28, available from http://marine.rutgers.edu/po/models/roms/documentation/adjoint.php.

Bai, Z., Fahey, M., Golub, G., 1995. Some large matrix computation problems. J. Comput. Appl. Math. 74, 71–89.

Barkmeijer, J., Iversen, T., Palmer, T.N., 2003. Forcing singular vectors and other sensitive model structure. Quart. J. Roy. Meteor. Soc. 129, 2401–2424.

Bennett, A.F., 1992. Inverse Methods in Physical Oceanography. Cambridge University Press. p. 346.

Bennett, A.F., 2002. Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press. p. 244.

Bennett, A.F., Thorburn, M.A., 1992. The generalized inverse of a nonlinear quasigeostrophic ocean circulation model. J. Phys. Oceanogr. 22, 213–230.

Bennett, A.F., Leslie, L.M., Hagelberg, C.R., Powers, P.E., 1993. Tropical cyclone prediction using a barotropic model initialized by a generalized inverse model. Mon. Wea. Rev. 121, 1714–1729.

Bennett, A.F., Chua, B.S., Harrison, D.E., McPhaden, M.J., 1998. Generalized inversion of tropical atmosphere-ocean (TAO) data and a coupled model of the tropical Pacific. J. Climate 11, 1768–1792.

Betti, A., Navarra, A., 1995. Prediction of the evolution of the variance in a barotropic model. J. Atmos. Sci. 52, 367–379.

Bishop, C.H., Toth, Z., 1999. Ensemble transformation and adaptive observations. J. Atmos. Sci. 56, 1748–1765.

Blumenthal, M.B., 1991. Predictability of a coupled ocean-atmosphere model. J. Climate 4, 766–784.

Branstator, G., 1985. Analysis of general circulation model sea-surface temperature anomaly simulations using a linear model. Part II: Eigenanalysis. J. Atmos. Sci. 42, 2242–2254.

Buizza, R., 1995. Optimal perturbation time evolution and sensitivity of ensemble prediction to perturbation amplitude. Quart. J. Roy. Meteor. Soc. 121, 1705–1738.

Buizza, R., 1997. Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. Mon. Wea. Rev. 125, 99–119.

Buizza, R., Palmer, T.N., 1995. The singular-vector structure of the atmospheric global circulation. J. Atmos. Sci. 52, 1434–1456.

Buizza, R., Tribbia, J., Molteni, F., Palmer, T., 1993. Computation of optimal unstable structures for a numerical weather prediction model. Tellus 45A, 388–407.

Chang, K.-I., Ghil, M., Ide, K., Lai, C.-C., 2001. Transitions to aperiodic variability in a wind-driven double-gyre circulation model. J. Phys. Oceanogr. 31, 1260–1286.

Chang, P., Saravanan, R., DelSole, T., Wang, F., 2004a. Predictability of linear coupled systems. Part I: Theoretical analyses. J. Climate, in press.

Chang, P., Saravanan, R., Wang, F., Ji L., 2004b. Predictability of linear coupled systems. Part II: An application to a simple model of tropical Atlantic variability. J. Climate, in press.

Charney, J.G., 1947. The dynamics of long waves in a baroclinic westerly current. J. Meteor. 4, 135–163.

Chen, Y.-Q., Battisti, D.S., Palmer, T.N., Barsugli, J., Sarachik, E.S., 1997. A study of the predictability of tropical Pacific SST in a coupled atmosphere/ocean model using singular vector analysis: the role of the annual cycle and the ENSO cycle. Mon. Wea. Rev. 125, 831–845.

Chua, B.S., Bennett, A.F., 2001. An inverse ocean modeling system. Ocean Modell. 3, 137–165.

Corti, S., Palmer, T.N., 1997. Sensitivity analysis of atmospheric low-frequency variability. Q. J. R. Meteor. Soc. 123, 2425–2447.

Courtier, P., 1997. Dual formulation of four-dimensional assimilation. Quart. J. R. Meteor. Soc. 123, 2449–2461.

DelSole, T.M., Hou, A.Y., 1999. Empirical stochastic models for the dominant climate statistics of a general circulation model. Mon. Wea. Rev. 127, 2533–2545.

Di Lorenzo, E., Miller, A.J., Neilson, D.J., Cornuelle, B.D., Moisan, J.R., 2004. Modeling observed California current mesoscale eddies and the ecosystem response. Int. J. Remote Sensing, in press.

Eady, E.T., 1949. Long waves and cyclone waves. Tellus 1, 33–52.

Egbert, G.D., Bennett, A.F., Foreman, M.G.G., 1994. TOPEX/POSEIDON tides estimated using a global inverse method. J. Geophys. Res. 99, 24281–24852.

Ehrendorfer, M., Tribbia, J.J., 1997. Optimal prediction of forecast error covariance through singular vectors. J. Atmos. Sci. 54, 286–313.

Errico, R.M., Vukićecić, T., Raeder, K., 1993. Examination of the accuracy of a tangent linear model. Tellus 45A, 462–477.

Fan, Y., Allen, M.R., Anderson, D.L.T., Balmaseda, M.A., 2000. How predictability depends on the nature of uncertainty in initial conditions in a coupled model of ENSO. J. Climate 13, 3298–3313.

Farrell, B.F., 1982a. The initial growth of disturbances in baroclinic flow. J. Atmos. Sci. 39, 1663–1686.

Farrell, B.F., 1982b. Pulse asymptotics of the Charney baroclinic instability problem. J. Atmos. Sci. 39, 507–517.

Farrell, B.F., 1984. Modal and nonmodal baroclinic waves. J. Atmos. Sci. 41, 668–673.

Farrell, B.F., 1985. Transient growth of damped baroclinic waves. J. Atmos. Sci. 42, 2718–2727.

Farrell, B.F., 1988a. Optimal excitation of neutral Rossby waves. J. Atmos. Sci. 45, 163–172.

Farrell, B.F., 1988b. Optimal excitation of perturbations in viscous shear flow. Phys. Fluids 31, 2093–2102.

Farrell, B.F., 1989a. Optimal excitation of baroclinic waves. J. Atmos. Sci. 46, 1193–1206.

Farrell, B.F., 1989b. Transient development in confluent and diffluent flow. J. Atmos. Sci. 46, 3279–3288.

Farrell, B.F., 1990. Small error dynamics and the predictability of atmospheric flows. J. Atmos. Sci. 47, 2409–2416.

Farrell, B.F., Ioannou, P.J., 1993a. Stochastic forcing of perturbation variance in unbounded shear and deformation flows. J. Atmos. Sci. 50, 200–211.

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

Farrell, B.F., Ioannou, P.J., 1993c. Stochastic forcing of the linearized Navier–Stokes equations. Phys. Fluids A 5, 2600–2609.

Farrell, B.F., Ioannou, P.J., 1993d. Optimal excitation of three dimensional perturbations in viscous constant shear flow. Phys. Fluids A 5, 1390–1400.

Farrell, B.F., Ioannou, P.J., 1993e. Perturbation growth in shear flow exhibits universality. Phys. Fluids A 5, 2298–2300.

Farrell, B.F., Ioannou, P.J., 1993f. Transient development of perturbations in stratified shear flow. J. Atmos. Sci. 50, 2201–2214.

Farrell, B.F., Ioannou, P.J., 1994a. Variance maintained by stochastic forcing of nonnormal dynamical systems associated with linearly stable shear flows. Phys. Rev. Lett. 72, 1188–1191.

Farrell, B.F., Ioannou, P.J., 1994b. 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., Ioannou, P.J., 1996a. Generalized stability theory. Part I: Autonomous operators. J. Atmos. Sci. 53, 2025–2040.

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

Farrell, B.F., Ioannou, P.J., 1996c. Turbulence suppression by active control. Phys. Fluids A 8, 1257–1268.

Farrell, B.F., Ioannou, P.J., 1999. Perturbation growth and structure in time dependent flows. J. Atmos. Sci. 56, 3622–3639.

Frederiksen, J.S., 1982. A unified three-dimensional instability theory for the onset of blocking and cyclogenesis. J. Atmos. Sci. 39, 969–982.

Frederiksen, J.S., 1997. Adjoint sensitivity and finite-time normal mode disturbances during blocking. J. Atmos. Sci. 54, 1144–1165.

Galanti, E., Tziperman, E., 2003. A midlatitude ENSO teleconnection mechanism via baroclinically unstable long Rossby waves. J. Phys. Oceanogr. 33, 1877–1888.

Galanti, E., Tziperman, E., Harrison, M., Rosati, A., Giering, R., Sirkes, Z., 2002. The equatorial thermocline outcropping—a seasonal control on the tropical Pacific ocean-atmosphere instability strength. J. Climate 15, 2721–2739.

Gardiner, C.W., 1985. Handbook of Stochastic Methods, second ed. Springer. p. 444.

Giering, R., Kaminski, T., 1998. Recipes for adjoint code construction. ACM Trans. Math. Software 24, 437–474.

Golub, G.H., van Loan, C.F., 1989. Matrix computations. Johns Hopkins. p. 642.

Haidvogel, D.B., Arango, H.G., Hedstrom, K., Beckmann, A., Malanotte-Rizzoli, P., Shchepetkin, A.F., 2000. Model evaluation experiments in the North Atlantic basin: simulations in nonlinear terrain-following coordinates. Dyn. Atmos. Oceans 32, 239–281.

Hall, M.C.G., Cacuci, D.G., 1983. Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. J. Atmos. Sci. 40, 2537–2546.

Hartmann, D.L., Buizza, R., Palmer, T.N., 1995. Singular vectors: the effect of spatial scale on linear growth of disturbances. J. Atmos. Sci. 52, 3885–3894.

Joly, A., 1995. The stability of steady fronts and the adjoint method: non-modal frontal waves. J. Atmos. Sci. 52, 3082–3108.

Junge, M.M., Haine, T.W.N., 2001. Mechanisms of North Atlantic wintertime sea surface temperature anomalies. J. Climate 14, 4560–4572.

Kleeman, R., Moore, A.M., 1997. A theory for the limitations of ENSO predictability due to stochastic atmospheric transients. J. Atmos. Sci. 54, 753–767.

Kleeman, R., Moore, A.M., Smith, N.R., 1995. Assimilation of sub-surface thermal data into an intermediate tropical coupled ocean-atmosphere model. Mon. Wea. Rev. 123, 3103–3113.

Lacarra, J.-F., Talagrand, O., 1988. Short-range evolution of small perturbations in a barotropic model. Tellus 40A, 81–95.

Lanczos, C., 1961. Linear Differential Operators. D. Van Nostrand Company Ltd, London and New York. p. 564.

Langland, R., Toth, Z., Gelaro, R., Szunyogh, I., Shapiro, M., Majomdar, S., Morss, R., Rohaly, G.D., Velden, C., Bond, N., Bishop, C., 1999. The north Pacific experiments (NORPEX-98): targeted observations for improved north Americal weather forecasts. Bull. Am. Met. Soc. 80, 1363–1384.

Lehoucq, R.B., Sorensen, D.C., Yang, C., 1997. ARPACK users' guide: solution of large scale eigenvalue problems with implicitly restarted Arnoldi methods. Rice University. p. 140.

Long, R.B., Thacker, W.C., 1989a. Data assimilation into a numerical ocean model. I. The model and the assimilation algorithm. Dynam. Atmos. Oceans 13, 379–412.

Long, R.B., Thacker, W.C., 1989b. Data assimilation into a numerical ocean model. II. Assimilation experiments. Dynam. Atmos. Oceans 13, 413–440.

Lorenz, E.N., 1965. A study of the predictability of a 28-variable atmospheric model. Tellus 17, 321–333.

MacCready, P., Geyer, G.R., 2001. Estuarie salt flux through an isohaline surface. J. Geophys. Res. 106, 11629–11637.

Malanotte-Rizzoli, P., Hedstrom, K., Arango, H.G., Haidvogel, D.B., 2000. Water mass pathways between the subtropical and tropical ocean in a climatological simulation of the North Atlantic ocean circulation. Dyn. Atmos. Oceans 32, 331–371.

Marchesiello, P., McWilliams, J.C., Shchepetkin, A.F., 2002. Equilibrium structure and dynamics of the California current system. J. Phys. Oceanogr., submitted for publication.

Molteni, F., Mureau, R., Palmer, T.N., 1993. Predictability and finite time instability of the northern winter circulation. Quart. J. Roy. Meteor. Soc. 119, 269–298.

Molteni, F., Buizza, R., Palmer, T.N., Petroliagis, T., 1996. The ECMWF ensemble prediction system: methodology and validation. Quart. J. Roy. Meteor. Soc. 122, 73–120.

Moore, A.M., 1991. Data assimilation in a quasigeostrophic open ocean model of the Gulf Stream region using the adjoint method. J. Phys. Oceanogr. 21, 398–427.
Moore, A.M., 1999a. The dynamics of error growth and predictability in a model of the Gulf Stream. II: Ensemble prediction. J. Phys. Oceanogr. 29, 762–778.
Moore, A.M., 1999b. Wind-induced variability of ocean gyres. Dyn. Atmos. Oceans 29, 335–364.
Moore, A.M., 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., Kleeman, R., 1996. The dynamics of error growth and predictability in a coupled model of ENSO. Q. J. R. Meteorol. Soc. 122, 1405–1446.
Moore, A.M., Kleeman, R., 1997a. The singular vectors of a coupled ocean-atmosphere model of ENSO. Part I: Thermodynamics, energetics and error growth. Q. J. R. Meteorol. Soc. 123, 953–981.
Moore, A.M., Kleeman, R., 1997b. The singular vectors of a coupled ocean-atmosphere model of ENSO. Part II: Sensitivity studies and dynamical significance. Q. J. R. Meteorol. Soc. 123, 983–1006.
Moore, A.M., Kleeman, R., 1998. Skill assessment for ENSO using ensemble prediction. Q. J. R. Meteorol. Soc. 124, 557–584.
Moore, A.M., Kleeman, R., 1999a. Stochastic forcing of ENSO by the intraseasonal oscillation. J. Climate 12, 1199–1220.
Moore, A.M., Kleeman, R., 1999b. The non-normal nature of El Nino and intraseasonal variability. J. Climate 12, 2965–2982.
Moore, A.M., Kleeman, R., 2001. On the differences between the optimal perturbations of coupled models of ENSO. J. Climate 14, 138–163.
Moore, A.M., Mariano, A.J., 1999. The dynamics of error growth and predictability in a model of the Gulf Stream. I: Singular vector analysis. J. Phys. Oceanogr. 29, 158–176.
Moore, A.M., Perez, C.L., Zavala-Garay, J., 2002. A non-normal view of the wind-driven ocean circulation. J. Phys. Oceanogr. 32, 2681–2705.
Moore, A.M., Vialard, J., Weaver, A., Anderson, D.L.T., Kleeman, R., Johnson, J.R., 2003a. The role of air-sea interaction in controlling the optimal perturbations of low-frequency tropical coupled ocean-atmosphere modes. J. Climate 16, 951–968.
Moore, A.M., Weaver, A., Vialard, J., Tang, Y., Zavala-Garay, J., Sahami, K., Kleeman, R., Anderson, D.L.T., 2003b. Stochastic optimals, singular vectors and forcing singular vectors of coupled models of ENSO, in preparation.
Mureau, R., Molteni, F., Palmer, T.N., 1993. Ensemble prediction using dynamically conditioned perturbations. Quart. J. Roy. Meteor. Soc. 119, 299–323.
Palmer, T.N., 1988. Medium and extended range predictability of the Pacific-North America mode. Q. J. R. Meteor. Soc. 116, 31–67.
Palmer, T.N., Gelaro, R., Barkmeijer, J., Buizza, R., 1998. Singular vectors, metrics, and adaptive observations. J. Atmos. Sci. 55, 633–653.
Pedlosky, J., 1979. Geophysical Fluid Dynamics. Springer-Verlag. p. 710.
Pedlosky, J., 1996. Ocean Circulation Theory. Springer. p. 453.
Penland, C., Sardeshmukh, P.D., 1995a. Error and sensitivity analysis of geophysical eigensystems. J. Climate 8, 1988–1998.
Penland, C., Sardeshmukh, P.D., 1995b. The optimal growth of tropical sea surface temperature anomalies. J. Climate 8, 1999–2024.
Penven, P., Roy, C., Colin de Verdiere, A., Largier, J., 2000. Simulation and quantification of a coastal jet retention process using a barotropic model. Oceanol. Acta 23, 615–634.
Penven, P., Roy, C., Lutjeharms, J.R.E., Colin de Verdiere, A., Johnson, A., Shillington, F., Freon, P., Brundrit, G., 2001a. A regional hydrodynamic model of the Southern Benguala. S. Afr. J. Sci. 97, 472–476.
Penven, P., Lutjeharms, J.R.E., Marchesiello, P., Roy, C., Weeks, S.J., 2001b. Generation of cyclonic eddies by the Algulhas Current in the lee of the Agulhas Bank. Geophys. Res. Let. 27, 1055–1058.
Rabier, F., Klinker, E., Courtier, P., Hollingsworth, A., 1996. Sensitivity of forecast error to initial conditions. Q. J. R. Meteor. Soc. 122, 121–150.

Reddy, S.C., Schmid, P.J., Henningson, D.S., 1993. Pseudospectra of the Orr-Sommerfeld operator. SIAM J. Appl. Math. 53, 15–47.

Shchepetkin, A.F., McWilliams, J.C., 1998. Quasi-monotone advection schemes based on explicit locally adaptive dissipation. Mon. Wea. Rev. 126, 1541–1580.

Shchepetkin, A.F., McWilliams, J.C., 2003. The regional ocean modeling system: a split-explicit, free-surface, topography-following coordinate ocean model. J. Geophys. Res., submitted for publication.

She, J., Klinck, J.M., 2000. Flow near submarine canyons driven by constant wind. J. Geophys. Res. 105, 28671–28694.

Simmons, A.J., Hoskins, B.J., 1976. Baroclinic instability on the sphere: normal modes of the primitive and quasigeostrophic equations. J. Atmos. Sci. 33, 1454–1477.

Sorensen, D.C., 1992. Implicit application of polynomial filters in a *k*-step Arnoldi method. SIAM J. Matrix Anal. Appl. 13, 357–385.

Szunyogh, I., Toth, Z., Emanuel, K.A., Bishop, C., Woolen, J., Marchok, T., Morss, R., Snyder, C., 1999. Ensemble-based targeting experiments during FASTEX: the impact of dropsonde data from the LEAR jet. Q. J. R. Meteor. Soc. 125, 3189–3217.

Talagrand, O., Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation I. Theor. Quart. J. Roy. Meteor. Soc. 113, 1311–1328.

Thompson, C.J., 1998. Initial conditions for optimal growth in a coupled ocean-atmosphere model of ENSO. J. Atmos. Sci. 55, 537–557.

Trefethen, L.N., 1991. Pseudospectra of matrices. In: Griffiths, D.F., Watson, G.A. (Eds.), Numerical Analysis. Longman Scientific and Technical, Harrow, UK, pp. 234–266.

Trefethen, L.N., 1996. Pseudospectra of linear operators. SIAM Rev. 39, 383–406.

Trefethen, L.N., Trefethen, A.E., Reddy, S.C., Driscoll, T.A., 1993. Hydrodynamic stability without eigenvalues. Science 261, 578–584.

Vialard, J., Weaver, A.T., Delecluse, P., 2002. Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part 2: Physical validation. Mon. Wea. Rev., submitted for publication.

Vukecevic, T., 1991. Nonlinear and linear evolution of initial forecast errors. Mon. Wea. Rev. 119, 1603–1611.

Weaver, A.T., Vialard, J., Anderson, D.L.T., 2002. Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics and consistency checks. Mon. Wea. Rev., submitted for publication.

Wright, T.G., Trefethen, L.N., 2001. Large-scale computation of pseudospectra using ARPACK and EIGS. SIAM J. Sci. Comput. 23, 591–605.

Xue, Y., Cane, M.A., Zebiak, S.E., Blumenthal, M.B., 1994. On the prediction of ENSO: a study with a low order Markov model. Tellus 46A, 512–528.

Xue, Y., Cane, M.A., Zebiak, S.E., 1997a. Predictability of a coupled model of ENSO using singular vector analysis. Part I: Optimal growth in seasonal background and ENSO cycles. Mon. Wea. Rev. 125, 2043–2056.

Xue, Y., Cane, M.A., Zebiak, S.E., Palmer, T.N., 1997b. Predictability of a coupled model of ENSO using singular vector analysis. Part II: Optimal growth and forecast skill. Mon. Wea. Rev. 125, 2057–2073.

Xue, Y., Leetmaa, A., Ji, M., 1999. Impact of sea level on predictability of ENSO: a study with Markov model. In: Second Hayes Symposium on Seasonal to Interannual Climate Variability—The 1997/1998 ENSO Cycle. American Meteorological Society, pp. 81–84.

Zavala-Garay, J., Moore, A.M., Perez, C.L., Kleeman, R., 2003. The response of a coupled model of ENSO to observed estimates of stochastic forcing. J. Climate 16, 2827–2842.