Empirical Orthogonal Functions and Normal Modes

GERALD R. NORTH

Goddard Laboratory for Atmospheric Sciences, NASA/Goddard Space Flight Center, Greenbelt, MD 20771

(Manuscript received 9 September 1983, in final form 16 November 1983)

ABSTRACT

An attempt to provide physical insight into the empirical orthogonal function (EOF) representation of data fields by the study of fields generated by linear stochastic models is presented in this paper. In a large class of these models, the EOFs at individual Fourier frequencies coincide with the orthogonal mechanical modes of the system—provided they exist. The precise mathematical criteria for this coincidence are derived and a physical interpretation is provided. A scheme possibly useful in forecasting is formally constructed for representing any stochastic field by a linear Hermitian model forced by noise.

1. Introduction

Empirical orthogonal functions (EOFs) have recently become popular tools in atmospheric science since their introduction several decades ago (Obukhov, 1947; Lorenz, 1956; Kutzbach, 1967). They are equivalent to the "principal components" used in multivariate statistics and are close relatives to the bases used in factor analysis. In stochastic field studies in the mathematical literature, they are called the Karhunen-Loeve basis functions (Loeve, 1978). The EOF basis set defined directly from observational data can be used for representing a climatological field very economically (e.g., Kutzbach, 1967), as a basis for a set of predictors (Barnett and Hasselmann, 1979) or as a means of physically interpreting the data (e.g., Wallace and Gutzler, 1981). Hence, there is considerable interest in the theory of these widely used functions.

If one thinks of the snapshots of geophysical data maps as realizations of random fields generated by some kind of stochastic process, it is possible to construct various second moment statistics linking one point and another in the map. The resulting covariance matrix is real and symmetric and therefore possesses a set of orthogonal eigenvectors with positive eigenvalues. The map associated with each eigenvector represents a pattern which is statistically independent of the others and spatially orthogonal to them. The eigenvalue indicates the amount of variance accounted for by the pattern. The eigenvector patterns contain information about the multidimensional probability distribution which constitutes the climate and are, therefore, of theoretical interest.

Although EOFs have been used for some time, several questions persist concerning their value in practical situations. For example, it is well-known that the EOF patterns depend sensitively upon the boundaries chosen for defining their domain—a sensitivity not unexpected since the patterns have to preserve the integral orthogonality constraint over the domain (Buell, 1979; North *et al.,* 1982). Another difficulty encountered is the sampling stability of the patterns; it appears that many realizations are required for the sample EOFs to approach the true EOFs (Karhunen–Loeve functions), particularly if the eigenvalues are closely spaced (North *et al.,* 1982), although in some applications this may not be a serious drawback. When measurement errors are superimposed upon the true geophysical fields, additional sampling effects can lead to distortions in the patterns and errors in the eigenvalue spectrum (Cahalan, 1983). The problem of selecting the statistically significant EOFs from those indistinguishable from either measurement noise or sampling error or a mixture is of considerable interest and has been studied extensively by Preisendorfer *et al.* (1981), who have listed a number of selection rules. Finally, the question of physical interpretation of EOFs remains open although advances have been made by Buell (1971, 1978, 1979), Wallace and Dickinson (1972) and Preisendorfer and his collaborators (1977, 1979, 1981). I have also recently become aware (M. Ghil, personal communication, 1983) of related research in the Soviet literature (Monin and Obukhov, 1967).

The purpose of this paper is to build upon the work cited above in providing a physical interpretation of EOFs by means of linear stochastic models which have solution EOFs that are intimately related to the model dynamics. Preisendorfer (1979) has shown that the EOFs coincide with the normal modes of the mechanical system in many coupled oscillator systems that are driven by noise. The present paper considers the case of continuum linear systems (cf. Preisendorfer *et al.,* 1981) forced by noise and presents general criteria

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

for the EOF–normal mode symmetry to occur. In so doing, a clearer understanding of the phenomenon emerges.

The approach presented here parallels to some extent the development of time series analysis (e.g., Jenkins and Watts, 1968), in that we attempt to understand observational data by first studying the solutions of linear stochastic models forced by white noise. We extend the theory into the spatial dimensions where no such symmetry as stationarity (translation invariance) usually exists, except in cases such as homogeneous turbulence (Batchelor, 1959) or problems where rotational invariance on the sphere can be assumed (Obukhov, 1947; North and Cahalan, 1981). When the translation symmetry is present, the optimum basis functions are uniquely determined [$e^{i\mathbf{k}\cdot\mathbf{r}}$ for homogeneous three-dimensional; $Y_{lm}(\mathbf{r})$ for rotational invariance on the sphere] and an analysis of their relative weights (spectrum) suffices for a complete identification of the model. In climatology there is usually no such translational or rotational symmetry and the Karhunen-Loeve functions must be estimated (EOFs) from the data.

The next section introduces four-dimensional EOFs by finding the eigenvectors of the cross-spectrum matrix—a procedure used by Wallace and Dickinson (1972) but from a rather different point of view. These four-dimensional EOFs are like the spatial patterns obtained in the usual three-dimensional applications, but in this case a complete set is obtained at each frequency. In some cases the shapes are the same for all frequencies, hence, they do not add new information. As we shall see, there are many dynamical systems driven by noise in which the EOFs are not frequency dependent, hence, the three-dimensional EOF formulation is adequate. On the other hand, some interesting geophysical models (e.g., the vorticity equation) exhibit EOFs which are frequency dependent. For this reason, four-dimensional EOFs are introduced first and notation is established. In Section 3, mechanical systems are classified according to whether or not they exhibit the mechanical mode–EOF symmetry when stochastically forced. Section 4 adds some remarks mostly motivated by the rather obvious analogy to univariate time series analysis; in particular, a formal scheme is proposed which is potentially useful in statistical prediction. Conclusions in Section 5 are followed by Appendix A, which lists a number of relevant results from linear analysis, and Appendix B, which provides some details about the noise forcing.

2. Four-dimensional EOFs

In this section we establish notation by introducing a set of basis functions for a random field defined on a spatial domain such that at each point, the field evolves according to a stationary time series. Some geophysical fields are ideally of this type (ignoring seasonal and other periodic stationarity breaking effects). Once these preliminaries are established we are prepared to generate some analytical examples of such behavior in the next section.

Consider a scalar function $T(\mathbf{r}, t)$ defined at a point $\mathbf{r}$ and time t. The function is a random field if some prescription is provided for generating independent *realizations* for the field such that all relevant multivariate probability densities can be constructed. The infinite member ensemble of realizations provides a convenient means of taking averages or of finding various moments of the multidimensional probability distributions.

Since the time dependence is assumed to be that of a stationary stochastic process, we may take the Fourier transform of $T(\mathbf{r}, t)$ which is indicated by

$$T_\omega(\mathbf{r}) = \int_{-\infty}^{\infty} e^{-i\omega t} T(\mathbf{r}, t)dt. \tag{1}$$

The cross spectrum $\gamma_\omega(\mathbf{r}, \mathbf{r}')$ between points $\mathbf{r}$ and $\mathbf{r}'$ can be defined as

$$\langle T_\omega(\mathbf{r}) T^*_{\omega'}(\mathbf{r}')\rangle = 2\pi\delta(\omega - \omega')\gamma_\omega(\mathbf{r}, \mathbf{r}'), \tag{2}$$

where the angular brackets indicate ensemble average and an asterisk denotes complex conjugation. Since it can be shown directly that the cross spectrum is a Hermitian operator [i.e., $\gamma^*_\omega(\mathbf{r}', \mathbf{r}) = \gamma_\omega(\mathbf{r}, \mathbf{r}')$, Jenkins and Watts, 1968; Wallace and Dickinson, 1972; see Appendix A for a discussion of Hermitian operators], we may use $\gamma_\omega(\mathbf{r}, \mathbf{r}')$ as a kernel in the eigenvalue problem

$$\int \gamma_\omega(\mathbf{r}, \mathbf{r}')\psi_\alpha(\mathbf{r}', \omega)d\mathbf{r}' = \lambda_\alpha(\omega)\psi_\alpha(\mathbf{r}, \omega),$$
$$\alpha = 1, 2, 3, \ldots, \tag{3}$$

where the integral runs over the spatial domain defining the system and $\psi_\alpha(\mathbf{r}, \omega)$ is the eigenfunction corresponding to the eigenvalue $\lambda_\alpha(\omega)$. Note that the complex eigenfunctions $\psi_\alpha(\mathbf{r}, \omega)$ and the real eigenvalues $\lambda_\alpha(\omega)$ depend upon (angular) frequency ω.

Following North *et al.* (1982) but for this more general (four-dimensional) case, we can expand the field $T_\omega(\mathbf{r})$ into an infinite series of the $\psi_\alpha(\mathbf{r}, \omega)$:

$$T_\omega(\mathbf{r}) = \sum_\alpha T_{\alpha,\omega}\psi_\alpha(\mathbf{r}, \omega). \tag{4}$$

Similarly, the original evolving field may be written

$$T(\mathbf{r}, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \sum_\alpha T_{\alpha,\omega} e^{i\omega t}\psi_\alpha(\mathbf{r}, \omega)d\omega. \tag{5}$$

The coefficients $T_{\alpha,\omega}$, given by

$$T_{\alpha,\omega} = \iint_{-\infty}^{\infty} e^{-i\omega t}\psi^*_\alpha(\mathbf{r}, \omega)T(\mathbf{r}, t)d\mathbf{r}dt, \tag{6}$$

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

are random variables satisfying the condition

$$\langle T^*_{\alpha,\omega} T_{\alpha',\omega'} \rangle = 2\pi\lambda_\alpha(\omega)\delta_{\alpha\alpha'}\delta(\omega - \omega') \quad (7)$$

and indicate that the random weights associated with the different eigenfunctions are uncorrelated. The variance associated with $\psi_\alpha(\mathbf{r}, \omega)$ is the eigenvalue $\lambda_\alpha(\omega)$. A more detailed discussion can be found in Wallace and Dickinson (1972).

The function $e^{i\omega t}\psi_\alpha(\mathbf{r}, \omega)$ may be thought of as a four-dimensional EOF since $e^{i\omega t}$ is the appropriate EOF basis function for a random stationary (time translation invariant statistics) function. In what follows, we shall call $\psi_\alpha(\mathbf{r}, \omega)$ or occasionally just ψ_α, the EOF at frequency ω. It has the familiar properties of EOFs except that it is complex. By the same convention used with the usual EOFs we can order the labels α such that at the frequency ω,

$$\lambda_1(\omega) > \lambda_2(\omega) > \lambda_3(\omega) > \cdots . \quad (8)$$

This ordering insures that $\psi_1(\mathbf{r}, \omega)$ accounts for the most variance, $\psi_2(\mathbf{r}, \omega)$ the next most, etc., in the integral representing the total variance at ω

$$\int \langle T_\omega(\mathbf{r}) T_{\omega'^*}(\mathbf{r}) \rangle d\mathbf{r} = \sum_\alpha \lambda_\alpha(\omega) 2\pi\delta(\omega - \omega'), \quad (9)$$

which can be derived from (4) using the orthogonality of the $\psi_\alpha(\mathbf{r}, \omega)$.

In practice we make measurements of an evolving field such as $T(\mathbf{r}, t)$ and we can then perform certain statistical manipulations of the data such as computing the sample EOFs depending upon the purpose. One motive might be to gain some physical insight or even to try to identify an associated model of the dynamical processes leading to $T(\mathbf{r}, t)$. In general this will be difficult or even impossible; however, in the next section we shall find a large class of models whose solution EOFs can be studied analytically.

3. Noise-forced linear models

In this section we consider a class of stochastically forced linear systems which can be solved by formal decomposition into mechanical modes. The class may be defined by the governing equation

$$H\left(\frac{\partial}{\partial t}, \nabla, \mathbf{r}\right)\Psi(\mathbf{r}, t) = f(\mathbf{r}, t), \quad (10)$$

along with homogeneous boundary conditions (either Ψ or its normal derivative vanishes on the boundary). In (10), $f(\mathbf{r}, t)$ is a stochastic field stationary in time and "white" in space, i.e.,

$$\langle f(\mathbf{r}, t) f(\mathbf{r}', t') \rangle = \sigma(|t - t'|)\delta(\mathbf{r} - \mathbf{r}'), \quad (11)$$

where $\delta(\mathbf{r} - \mathbf{r}')$ is the Dirac function and $\sigma(|t - t'|)$ is some given lag-autocovariance function. For most purposes it is convenient to think of this forcing as a randomly located point source with its strength modulated by a stationary process. Some examples are provided in Appendix B.

We begin by Fourier analyzing (10) according to (1):

$$H_\omega \Psi_\omega(\mathbf{r}) = f_\omega(\mathbf{r}), \quad \text{with} \quad H_\omega = H(-i\omega, \nabla, \mathbf{r}). \quad (12)$$

Following are three examples of physical systems with H_ω exhibiting different characteristic behaviors.

(Case I) The wave equation for a drumhead clamped at its edges:

$$H^w_\omega = -\omega^2 - c^2\nabla^2,$$

$$\Psi_\omega(\mathbf{r}_b) = 0, \ \mathbf{r}_b \quad \text{a boundary point.} \quad (13)$$

The forcing here is analogous to an erratic drummer providing statistically stationary impulsive, localized blows to the drum surface.

(Case II) The diffusion equation:

$$H^D_\omega = -i\omega - \nabla \cdot D\nabla, \quad (14)$$

with, e.g., absorbing boundaries $\Psi_\omega(\mathbf{r}_b) = 0$. The forcing here is a statistically steady supply of polluting drops which diffuse away to the boundary.

(Case III) The barotropic vorticity equation for the streamfunction (e.g., Longuet-Higgins, 1964):

$$H^v_\omega = -i\omega\nabla^2 + \beta(y)\frac{\partial}{\partial x}, \quad (15)$$

with basic state at rest and the boundary condition for a channel $\hat{\mathbf{n}} \cdot \nabla\Psi_\omega(r_b) = 0$, where $\hat{\mathbf{n}}$ is a unit vector along the channel boundary. The forcing here is analogous to randomly located point sources of vorticity whose strengths are modulated by a stationary time series of arbitrary spectrum σ_ω. Although not treated here explicitly, the corresponding linearized multicomponent Boussinesq model also has essentially identical properties (Leith, 1980).

In each of these physically interesting cases, the state $\Psi_\omega(\mathbf{r})$ is a random field and the prescription for constructing the ensemble is provided by the random forcing. It is interesting to note that many of the results to follow are not restricted to the form of (10), (11). The most important element is that a prescription needs to be provided for assembling an infinite number of realizations which are equivalent to those generated by (10), (11). The crucial feature is that the prescription allots, on the average, equal forcing variance to all normal modes of the system. This is easily accomplished by $\delta(\mathbf{r} - \mathbf{r}_0)$, $\mathbf{r}_0$ uniformly distributed because it equipartitions energy (average amplitude squared) into any orthonormal basis set. This can be seen as follows:

$$\delta(\mathbf{r} - \mathbf{r}_0) = \sum_\alpha \phi_\alpha(\mathbf{r})\phi^*_\alpha(\mathbf{r}_0).$$

The amplitude of $\phi_\alpha(\mathbf{r})$ is the random variable $\phi^*_\alpha(\mathbf{r}_0)$ and $|\phi_\alpha(\mathbf{r}_0)|^2$ averages to unity over the volume if $\mathbf{r}_0$ is uniformly distributed (orthonormality). It is important to note that this result will hold for any orthonormal

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

basis set ϕ_α. As an example of an alternative prescription, the ensemble could be constructed for these same systems by setting the forcing to zero and starting with random initial conditions (e.g., $\Phi(\mathbf{r}, 0) = \Phi_0\delta(\mathbf{r} - \mathbf{r}_0)$ or

$$\left.\frac{\partial\Phi}{\partial t}\right|_{t=0} = A\delta(\mathbf{r} - \mathbf{r}_0),$$

plucking or striking), $\mathbf{r}_0$ random, and examining the system a fixed time t_0 later. Simmons *et al.* (1983) have recently studied the evolution of Rossby waves on the sphere excited by localized randomly located disturbances. The ensemble of solutions they produce is analogous to this latter procedure. In some of the nondecaying systems, we could use a single initial condition and observe the system with snapshots at random intervals uniformly distributed over some long time (much greater than any internal period). We conclude that these prescriptions, though not general, are broad enough to be physically interesting. Remarks about relaxing the point source forcing restriction are included in the next section.

We now return to the system (12) and analyze its mechanical properties. A very important special case arises when H_ω is a Hermitian operator as in case I. Hermitian operators and some relevant properties are reviewed in Appendix A. In the Hermitian H_ω case we have a complete orthonormal set of eigenfunctions ϕ_α defined by

$$H\phi_\alpha(\mathbf{r}) = \mu_\alpha\phi_\alpha(\mathbf{r}), \quad (16)$$

subject to the boundary conditions and where μ_α are the real eigenvalues, $\alpha = 1, 2, \cdots$ and the ω dependence of ϕ_α, μ_α, H are suppressed in this notation for simplicity. Using

$$\Psi_\omega(\mathbf{r}) = \sum_\alpha \psi_{\omega,\alpha}\phi_\alpha(\mathbf{r}) \quad (17)$$

and orthonormality of the ϕ_α in (12)

$$\Psi_\omega(\mathbf{r}) = \sum_\alpha \phi_\alpha(\mathbf{r}) \int \frac{\phi_\alpha(\mathbf{r}')f_\omega(\mathbf{r}')d\mathbf{r}}{\mu_\alpha}, \quad (18)$$

we can now form the cross spectrum as in (2)

$$\langle\Psi_\omega(\mathbf{r})\Psi_\omega^*(\mathbf{r}')\rangle = \sum_{\alpha,\beta} \frac{\phi_\alpha(\mathbf{r})\phi_\beta^*(\mathbf{r}')}{\mu_\alpha\mu_\beta}$$

$$\times \left\langle \iint \phi_\alpha(\mathbf{r}_1)f_\omega(\mathbf{r}_1)\phi_\beta^*(\mathbf{r}_2)f_\omega(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \right\rangle, \quad (19)$$

and using (11) and orthonormality of the ϕ_α,

$$\gamma_\omega(\mathbf{r}, \mathbf{r}') = \sigma_\omega \sum_\alpha \frac{\phi_\alpha(\mathbf{r})\phi_\alpha^*(\mathbf{r}')}{\mu_\alpha^2}. \quad (20)$$

It can now be seen directly by insertion into (3) that the $\phi_\alpha(\mathbf{r})$ are the EOFs of the field at frequency ω with eigenvalues $\phi_\omega\mu_\alpha^{-2}$. It follows that Hermitian mechanical systems forced according to (10), (11) have their EOFs coinciding with the normal mechanical modes, defined as the eigenfunctions of H_ω.

The vorticity equation operator (15) is an example of a purely anti-Hermitian operator ($H^+ = -H$). By writing $L = iH$, we see that L is Hermitian and therefore has a complete set of normal modes. A repeat of the steps (17)–(20) reveals that this case also has its EOFs and normal mechanical modes coinciding. Hence, in this important case in geophysical fluid dynamics we do find the EOF-mechanical mode symmetry.

In general, the operator H will not be pure Hermitian or pure anti-Hermitian, although any operator can be decomposed into a sum of Hermitian and anti-Hermitian parts,

$$H_\omega = \frac{1}{2}(H_\omega + H_\omega^+) + \frac{1}{2}(H_\omega - H_\omega^+) = H_R + iH_I, \quad (21)$$

where H_R and H_I are Hermitian by construction. Ordinarily a complex operator such as (21) with $H_I \neq 0$ does not have orthogonal eigenfunctions. Rather, the eigenfunctions are oblique and one makes use of a dual basis set, the eigenfunctions of H^+, each member of which is orthogonal to the corresponding eigenfunction of H (Halmos, 1958). It suffices here to note that in cases such as this, the EOFs, which are manifestly orthogonal, cannot possibly coincide with these oblique mechanical modes. In some interesting cases, however, the oblique modes are accidently orthogonal and we wish to examine this possibility now.

In Appendix A it is noted that two operators have simultaneously the same eigenfunctions if, and only if, they commute. This well-known theorem is very important in the present context since it is the crux of having simultaneous eigenfunctions of γ_ω and H. But before examining that general question, consider case II, the diffusive system which has

$$H_R = -\nabla\cdot D(\mathbf{r})\nabla, \quad (22)$$

$$H_I = -\omega. \quad (23)$$

Note that in this special case H_I is a scalar operator and trivially commutes with the differential operator H_R. Therefore, H_R and H_I can have simultaneous eigenfunctions (those of H_R) and they can be shown directly to be the EOFs of the system by the methods already described.

In contrast, consider the case

$$H_\omega^{DW} = -\omega^2 - iB\omega - c^2\nabla^2, \quad (24)$$

which is that of a damped wave system. If B is constant or any Hermitian operator commuting with ∇^2 such as $\hat{\mathbf{n}}\cdot\nabla$, $\hat{\mathbf{n}}$ a fixed vector, the EOF–normal mode symmetry will hold; however, if B is a function of position for example, H_R and H_I will not commute and the orthogonality of the eigenfunctions will be broken. Clearly, in this situation the eigenfunctions of H_R and the eigenfunctions of H_I do not coincide and the resulting mechanical modes are not orthogonal.

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

One way of interpreting linear mechanical modes is to consider the unforced system and insert energy initially into a single mode. Subsequent evolution of the system will not leak the energy into other modes (subspaces are invariants of the motion). This will be true for a pure vibration or a pure decay system. When the two are mixed this preservation property will ordinarily be lost, except in the accidental case that the two mode patterns coincide (H_R and H_I commute).

As an important GFD example of this failure, consider the barotropic vorticity equation with the basic state having a latitude dependent zonal wind, a system which has been studied extensively recently because of its importance to planetary wave studies (e.g., Grose and Hoskins, 1979). The mechanical system is characterized by the operator

$$H'_v = \left(-i\omega + \bar{u}(y)\frac{\partial}{\partial x}\right)\nabla^2 + \beta(y)\frac{\partial}{\partial x}. \qquad (25)$$

It can be seen directly that this operator is not pure H_I or H_R because the y-dependence in $\bar{u}(y)$ does not permit $u(y)$ and ∇^2 to commute, which is required in the proof that H'_v be anti-Hermitian. Such a system will not have an orthogonal basis set nor will it necessarily have a set of real eigenvalues, implying a possible instability. This latter is not surprising since barotropic instability is known to occur when $\bar{u}(y)$ satisfies certain conditions (Kuo, 1949). Hence, we cannot expect the EOFs to coincide with the normal modes in this case. Simmons *et al.* (1983) have, in fact, argued that the response of a system similar to our prototype (25) has EOFs that look more like the instability modes rather than the standing wave.

The connection between the normality of the operator H_ω and the stability of the free system can be understood in the following way. First, consider H_ω to be Hermitian ($H_\omega = H_R$). The free system corresponds to the eigenfunctions of H_ω corresponding to vanishing eigenvalue (null space). For example, in Case (I), the drumhead, above this means $-\omega^2 + \lambda_\alpha = 0$, where λ_α is the positive eigenvalue of $-\nabla^2$ (subject to the boundary conditions). The free system will be stable if the solution to this constraint has all roots for ω real. Ordinarily, this will be the case for systems like (1). Now suppose an infinitesimal part $i\epsilon H_I$ is added to H_R. The null space of $H_R + i\epsilon H_I$ corresponds (to first order in ϵ) to the condition $\mu_\alpha + i\epsilon C_\alpha = 0$ where μ_α is the eigenvalue of H_R and C_α [$\approx(\Psi_\alpha, H_I\Psi_\alpha)$] is a number of the order of unity. The solution of this constraint for ω is almost sure to lead to a complex root and therefore, a nonoscillatory motion. Of course, as noted above, if the operator is normal H_I and H_R will commute and the corresponding eigenfunctions will coincide. If these two operators do not commute the exponential modes and the standing wave modes will not coincide. For ϵ small as in this example, the eigenfunctions of H_ω will correspond to those of H_R except for corrections of the order of ϵ. When ϵ is not small, the eigenfunctions of H_ω will not bear any simple relation to those of H_R or H_I. Of some interest is the case where ϵ is very large, since this case may occur naturally as in the Simmons *et al.* (1983) study.

Now we return to the general question of EOF-mechanical mode symmetry. In certain cases we have seen that (20) holds. In light of the commuting operator arguments just presented, it is interesting to ask about whether the kernel $\gamma_\omega(r, r')$, thought of as an operator Γ_ω and H commute. Formally, we can write for the Hermitian H_ω case

$$\Gamma_\omega = \sigma_\omega H_\omega^{-2}, \qquad (26)$$

since σ_ω is a scalar (number) and since H commutes with any power of itself, the operators Γ_ω and H_ω can have simultaneous eigenfunctions. The analog of (26) when H is not Hermitian is

$$\Gamma_\omega = \sigma_\omega(H_\omega^+ H_\omega)^{-1}, \qquad (27)$$

which is itself Hermitian, since

$$[(H_\omega^+ H_\omega)^{-1}]^+ = (H_\omega^+ H_\omega)^{-1}, \qquad (28)$$

using $(AB)^+ = B^+A^+$ and the analogous rule for inverses. However, in this case (27) Γ_ω and H do not commute since H and H^+ will not commute if H_R and H_I do not. We thus find that if H is a normal operator ($H_\omega H_\omega^+ = H_\omega^+ H_\omega$) the EOFs and the normal modes for the system (10), (11) will coincide.

Having analyzed cases I, II and III for their EOF–normal mode symmetry, it is perhaps helpful to reexamine the frequency dependence of the corresponding four-dimensional EOFs. Note that in both cases I and II the ω dependence simply appears as an additive scalar in H_ω. This structure leads to the property that the EOF shapes do not depend upon ω; instead, all the ω dependence will occur in the eigenvalues $\lambda_\alpha(\omega)$. It is obvious that this result can be generalized to all systems of the form $H = \partial/\partial t - K_1(\mathbf{r}, \nabla)$ (fluid-type systems) or $H = \partial^2/\partial t^2 - K_2(\mathbf{r}, \nabla)$ (Hamiltonian-type systems). A three-dimensional analysis would have sufficed in these cases. On the other hand, when the time derivative occurs in H as it does in the vorticity equation (III) (also the Boussinesq version), the model EOF shapes must depend upon ω. Whether empirically derived EOFs taken from real data depend upon ω or not could be a useful clue in model identification studies.

4. Additional remarks

In pursuing the analogy wth univariate time series analysis we wish to bring up and discuss several extensions of the theory just presented.

a. Smoothed forcing

Nature is not likely to provide many systems with white noise in space as a forcing. We should inquire

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

about the analogy to moving averages (MA) of white spatial noise as a forcing. It is not appropriate to simply "smear" $f_\omega(\mathbf{r})$ in space by a convolution since the convolution is special to systems with translation invariance. Consider the result of an integral smearing operator S acting upon the noise forcing f_ω:

$$f'_\omega(\mathbf{r}) = \int S(\mathbf{r}, \mathbf{r}')f_\omega(\mathbf{r}')d\mathbf{r}'. \tag{29}$$

As an example consider the operator ∇^{-2}, the inverse of the Laplacian with appropriate boundary conditions understood. If $f_\omega(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)$, $\mathbf{r}_0$ random, we can find the result of (29) by knowing the Green's function for the Laplacian. For example, in the infinite three-dimensional domain

$$\nabla^{-2}\delta(\mathbf{r} - \mathbf{r}_0) = \frac{1}{4\pi}\frac{1}{|\mathbf{r} - \mathbf{r}_0|}. \tag{30}$$

The latter is equivalent to Coulomb's Law for a point charge. Applying more powers of ∇^{-2} will continue the smearing process to any desired level. By applying a polynomial in ∇^{-2} one can build up a moving spatial average of arbitrary structure.

A moment's reflection suggests that the smearing operator will interfere with the EOF–normal mode symmetry because it will no longer produce equipartition on the average. The problem can be analyzed by multiplying the resulting mechanical equation through by S^{-1} (assuming S^{-1} exists):

$$S^{-1}H_\omega\Psi_\omega = f_\omega, \tag{31}$$

with the result equivalent to (10), (11) except that H_ω is replaced by $S^{-1}H_\omega$. The EOF–normal mode symmetry will occur if $S^{-1}H_\omega$ is a normal operator or if each is individually a normal operator and they commute with each other. In this case statistical equipartitioning will be preserved. As we might have expected, the smearing operator has to be very special otherwise we can now have three distinct sets of modes: vibration, unstable and forcing.

One climate model of this type has been studied by North and Cahalan (1981). The model was a diffusive climate model with constant coefficients

$$H_\omega^{\mathrm{NC}} = -i\omega C - D\nabla^2 + B, \tag{32}$$

which is a rotationally invariant normal operator having the spherical harmonics as eigenfunctions. The system was forced by noise which was rotationally invariant on the sphere (homogenous noise). Such a forcing field can be decomposed into a Karhunen-Loeve basis of spherical harmonics. Hence, the symmetry in the mechanical operator and the same symmetry in the forcing causes the EOF–normal mode coincidence even though the forcing is not white in space.

b. Forecasting

Suppose we have infinitely many perfectly measured realizations of the field $\psi(\mathbf{r}, t)$, which might even have nonlinear effects included. Is it possible to construct a linear empirical model from the four-dimensional EOFs and their spectra that might be useful for forecasting? The analog in time series analysis is a theorem by Wiener (1947). We can sketch how such a model can be constructed in principle. For simplicity we consider a Hermitian system forced by white noise in space and time ($\sigma_\omega = 1$). From (26) we may write

$$H_\omega^2 = \Gamma_\omega^{-1}. \tag{33}$$

The right-hand side of (33) can be represented by its eigenfunction decomposition

$$\Gamma_\omega^{-1} = \sum_\alpha \frac{\psi_\alpha(\mathbf{r})\psi_\alpha^*(\mathbf{r}')}{\lambda_\alpha}; \tag{34}$$

a solution for H_ω is

$$H_\omega = \sum_\alpha \frac{\psi_\alpha(\mathbf{r})\psi_\alpha^*(\mathbf{r}')}{\lambda_\alpha^{1/2}}, \tag{35}$$

and in principle we find H by Fourier inversion. Implicit is the assumption that these operators exist ($\lambda_\alpha > 0$). Ignoring this last, it is conceivable that a program based upon (35) might be useful in forecasting, although experience with time series analysis suggests that things are never so simple. Clearly, a number of problems are likely to arise early on. For example, the $\lambda_\alpha^{1/2}$ in the denominator signals trouble since the EOFs contributing least variance will be emphasized most; these will also be the ones least well estimated from the data. (On the other hand, we actually seek ψ_ω which is proportional to H_ω^{-1}.) Even with these caveats the scheme could prove useful in some cases. It is also conceivable that clues about the true mechanical system can be found by studying the empirical system H_ω. In any case, this class of problems seems worthy of further study.

c. Multicomponents

The problems we have discussed so far are in no way restricted to unicomponent fields. We could have discussed multicomponent fields $\psi_i(r, t)$ from the beginning with H^{ij} a matrix operator. In this way we can generalize the EOFs even further. The analog of this class of models in time series analysis is the bivariate autoregressive systems of Jenkins and Watts (1968). An example of a Hermitian multicomponent system is the linearized Boussinesq approximation to a fluid whose basic state is at rest or in solid rotation (Leith, 1980).

5. Conclusion

In this paper we have seen that a large class of linear stochastically forced mechanical systems have their EOFs coincident with the mechanical free modes of the system. These cases are the ones where the mechanical system is governed by a normal operator with

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

the forcing being white in space and stationary in time. The EOF–mechanical mode symmetry is broken when the mechanical modes are not orthogonal, a situation which occurs if the system has decay and/or growth modes not coinciding with its vibration or wave modes. Understanding how stochastic models behave provides some insight into the behavior of natural systems as has been the case in the history of time series analysis. It is conceivable that phenomenological models can be developed from the space–time statistics of the data that can even be useful in the forecasting problem in analogy to the ARMA models from univariate time series analysis. But as the history of univariate spectral analysis and model identification has shown, much research will be needed to understand how to implement and evaluate such procedures in view of our limited data samples and their contamination by measurement and data reduction (analysis) errors.

Acknowledgments. Several colleagues including J. M. Wallace, M. Ghil, W. K. M. Lau, P. B. James and especially R. F. Cahalan and T. L. Bell have patiently listened to my expositions of this work and they have offered several helpful suggestions.

APPENDIX A

Some Relevant Results from Linear Analysis

A very short but neat summary of linear analysis is contained in the Appendix of Leith (1980). Longer discussions can be found in the excellent books by Halmos (1957, 1958) or the more traditional book by Courant and Hilbert (1953). In the applications encountered in this paper we may think of a space of functions, defined on a spatial domain, with some good behavior property such as being square integrable and having certain boundary properties at the domain edge such as being zero or having zero normal derivative. There exist certain infinite sequences of functions in the space called basis sets which can be used as the expansion terms in an infinite series. A complete basis set is one that can be used to represent any function in the space by such an infinite series with appropriate weights.

The inner product of two functions in the space can be defined as the complex number

$$\int \phi^*(\mathbf{r})\psi(\mathbf{r})d\mathbf{r} \equiv (\phi, \psi), \tag{A1}$$

where the integral runs over the whole spatial domain. Linear operators are linear transformations carrying (mapping) one function in the space to another. The action of an operator H acting upon an element ϕ of the function space yielding the result ψ is denoted

$$\psi = H\phi, \tag{A2}$$

which may be written in the function representation as an integral transform

$$\psi(\mathbf{r}) = \int H(\mathbf{r}, \mathbf{r}')\phi(\mathbf{r}')d\mathbf{r}'. \tag{A3}$$

A differential operator such as the familiar ∇^2 can also be written in this form by use of the Dirac delta function

$$\nabla^2\psi(\mathbf{r}) = \int \delta(\mathbf{r} - \mathbf{r}')\nabla'^2\psi(\mathbf{r}')d\mathbf{r}'. \tag{A4}$$

The Hermitian adjoint of an operator A is defined as the operator $A+$ which gives the same complex number for the inner products

$$(\phi, A\psi) = (A^+\phi, \psi) \tag{A5}$$

for any ϕ and ψ in the space. An operator is Hermitian if it equals its Hermitian adjoint.

The product of the two operators A and B is denoted AB and when acting upon an element ϕ we mean

$$AB\phi = A(B\phi) = (AB)\phi \tag{A6}$$

which in function representation is

$$\iint A(\mathbf{r}, \mathbf{r}')B(\mathbf{r}', \mathbf{r}'')\phi(\mathbf{r}'')d\mathbf{r}'d\mathbf{r}''. \tag{A7}$$

Two operators A, B are said to commute if $AB = BA$ when acting on any element in the space. Ordinarily operators do not commute. An operator is normal if it commutes with its adjoint. Hermitian operators are normal operators. The adjoint of a product AB is B^+A^+.

The eigenvectors of an operator H are the solutions ϕ_α, $\alpha = 1, 2, \ldots,$ of

$$H\phi_\alpha = \lambda_\alpha\phi_\alpha, \tag{A8}$$

where the numbers λ_α are the eigenvalues. The eigenvalues of a Hermitian operator are real. The eigenvectors of a normal operator are orthogonal and can be normalized such that

$$(\phi_\alpha, \phi_\beta) = \delta_{\alpha\beta} = \int \phi_\alpha^*(\mathbf{r})\phi_\beta(\mathbf{r})d\mathbf{r}, \tag{A9}$$

where ϕ_α and ϕ_β correspond to distinctly different eigenvalues λ_α and λ_β. Usually it is possible to show that for well-behaved operators acting upon a bounded domain, the eigenvectors form a complete basis set, which may be expressed

$$\sum_\alpha \phi_\alpha(\mathbf{r})\phi_\alpha^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'). \tag{A10}$$

When this holds, any function in the space can be written in the infinite series representation

$$\psi(\mathbf{r}) = \sum_\alpha \psi_\alpha\phi_\alpha(\mathbf{r}). \tag{A11}$$

A normal operator can be decomposed into an expansion of its eigenvectors and eigenvalues

$$H(\mathbf{r}, \mathbf{r}') = \sum_\alpha \lambda_\alpha\phi_\alpha(\mathbf{r})\phi_\alpha(\mathbf{r}'). \tag{A12}$$

Two operators can have simultaneously the same eigenvectors if, and only if, they commute.

The inverse of an operator A, denoted A^{-1}, ordinarily exists if A has no vanishing eigenvalues. For a product $(AB)^{-1} = B^{-1}A^{-1}$. An operator commutes with itself and any power of itself including its inverse.

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

APPENDIX B

Noise Forcing

The linear models described in this paper are all forced by a noise function $f(\mathbf{r}, t)$ introduced in Eqs. (10) and (11), cf. also Balgovind *et al.* (1983). In this section, some elaboration on such forcing functions is provided. Some insight can be gained by providing some actual examples of stochastic fields $f(\mathbf{r}, t)$ which satisfy (11).

Example 1: Consider the stochastic function

$$f_\nu(\mathbf{r}, t) = g[T\delta(t - t_\nu) - 1][V\delta(\mathbf{r} - \mathbf{r}_\nu) - 1], \quad \text{(B1)}$$

where the subscript ν denotes an index labeling the realization, T is a (long) time interval suitable for averaging over and V is the volume of the domain. The field $f_\nu(\mathbf{r}, t)$ is a random field because t_ν and r_ν are random variables uniformly distributed over their domains, i.e., the probability density function, $P(t_\nu, \mathbf{r}_\nu) = (TV)^{-1}$ is constant. Using this probability density we can compute

$$\langle f(\mathbf{r}, t)\rangle = 0, \quad \text{(B2)}$$

$$\langle f(\mathbf{r}, t)f(\mathbf{r}', t')\rangle = g^2T^2V^2\delta(t - t')\delta(\mathbf{r} - \mathbf{r}'), \quad \text{(B3)}$$

for large T and V. Hence, $f_\nu(\mathbf{r}, t)$ defined by (B1) is a special case of (11) where $\sigma(|t - t'|)$ is given by $\delta(t - t')$ (white noise). Any more general form of $\sigma(|t - t'|)$ can be construted by forming the temporal convolution with an appropriate averaging function. The interpretation of $f_\nu(\mathbf{r}, t)$ is that each realization in space–time is a constant negative function (uniform sink) except at the randomly located space–time coordinate $(\mathbf{r}_\nu, t_\nu)$, where it is locally infinite. The negative uniform sink merely assures that zero average force is applied. If this latter were not enforced, standing wave systems forced by (B1) would continue to energize indefinitely. In this example, either factor in brackets in (B1) could be replaced by any white noise function.

Example 2:

$$f_\nu(\mathbf{r}, t) = g\{V\delta[\mathbf{r} - \mathbf{r}_\nu(t)] - 1\}, \quad \text{(B4)}$$

where $\mathbf{r}_\nu(t)$ is a stationary white noise time series with uniform probability density in space. Clearly, (B2) holds by direct computation. To proceed with the rest of the proof it is convenient to discretize in time so that the sequence (dropping the realization label ν)

$$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \cdots \quad \text{(B5)}$$

represents the sequence of (vector) values assumed by $\mathbf{r}(t)$ for the different times. The multivariate probability distribution for $\mathbf{r}_i$ is

$$P(\mathbf{r}_1, \mathbf{r}_2, \cdots) = \prod p(\mathbf{r}_i), \quad \text{(B6)}$$

since each element of the time sequence is statistically independent of the others. The expectation value of (B4) becomes

$$\langle f(\mathbf{r}, i)\rangle = \int \prod_j p(\mathbf{r}_j)g[V\delta(\mathbf{r} - \mathbf{r}_i) - 1]d\mathbf{r}_j. \quad \text{(B7)}$$

All of the factors integrate to unity except the one for which the times are equal, i.e., $i = j$, whence

$$\langle f(\mathbf{r}, i)\rangle = g[Vp(\mathbf{r}) - 1], \quad \text{(B8)}$$

which vanishes since for a uniform distribution of $\mathbf{r}_i$ we have $p(\mathbf{r}) = 1/V$. To find the space–time lagged covariance we must form

$$\langle f(\mathbf{r}', j)f(\mathbf{r}, i)\rangle = \int \prod_\kappa p(\mathbf{r}_\kappa)g^2(V\delta(\mathbf{r}' - \mathbf{r}_j) - 1) \times (V\delta(\mathbf{r} - \mathbf{r}_i) - 1)d\mathbf{r}_\kappa. \quad \text{(B9)}$$

We must consider two cases: when the times i and j are equal and when they are not. The result is

$$g^2\delta_{ij}\{V^2\delta(\mathbf{r} - \mathbf{r}')p(\mathbf{r}) - V[p(\mathbf{r}) + p(\mathbf{r}')] + 1\} + g^2(1 - \delta_{ij})\{V^2p(\mathbf{r})p(\mathbf{r}') - V[p(\mathbf{r}) + p(\mathbf{r}')] + 1\}, \quad \text{(B10)}$$

and using $p(\mathbf{r}) = 1/V$, we have for large V

$$\langle f(\mathbf{r}', j)\rangle f(\mathbf{r}, i)\rangle = g^2V\delta(\mathbf{r} - \mathbf{r}')\delta_{ij} + O(1), \quad \text{(B11)}$$

which is equivalent to (B3) when the δ_{ij} is interpreted as $T\delta(t_i - t_j)$. The physical interpretation is again like that of example 1); the forcing is localized at a point whose location in the domain is determined by a white noise process.

The above examples can be made even more general by summing a finite number of statistically independent terms that are of the same form. These examples have been described partly to show that the class of forcings (10)–(11) is not empty and to show how we can relate this type of forcing to experience.

REFERENCES

Balgovind, R., A. Dalcher, M. Ghil and E. Kalnay, 1983: A Stochastic-dynamic model for spatial structure of forecast error statistics. *Mon. Wea. Rev.*, **111**, 701–722.
Barnett, T. P., and K. Hasselmann, 1979: Techniques of linear prediction with application to oceanic and atmospheric fields in the tropical pacific. *Rev. Geophys. and Space Phys.*, **178**, 949–968.
Batchelor, G. K., 1959: *The Theory of Homogeneous Turbulence.* Cambridge University Press, London, 197 pp.
Buell, C. E., 1971: Integral equation representation for factor analysis. *J. Atmos. Sci.*, **28**, 1502–1505.
——, 1978: The number of significant proper functions of two-dimensional fields. *J. Appl. Meteor.*, **17**, 717–722.
——, 1979: On the physical interpretation of empirical orthogonal functions. *Preprints 6th Conf. on Probability and Statistics in Atmospheric Sciences*, Banff, Amer. Meteor. Soc., 112–117.
Cahalan, R. F., 1983: EOF spectral estimation in climate analysis. *Proc. Second Int. Meeting on Statistical Climatology*, Lisbon, Amer. Meteor. Soc., 451–457.
Courant, R., and D. Hilbert, 1953: *Methods of Mathematical Physics, Vol. I*, Interscience Publishers, 560 pp.
Grose, W. L., and B. J. Hoskins, 1979: On the influence of orography on large-scale atmospheric flow. *J. Atmos. Sci.*, **36**, 223–234.

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC

Halmos, P. R., 1957: *Introduction to Hilbert Space,* 2nd ed., Chelsea, 114 pp.

——, 1958: *Finite-Dimensional Vector Spaces.* Van Nostrand, 200 pp.

Jenkins, G. M., and D. G. Watts, 1968. *Spectral Analysis and Its Applications.* Holden-Day, 525 pp.

Kuo, H. L., 1949: Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *J. Meteor.,* **6,** 105–122.

Kutzbach, J., 1967: Empirical eigenvectors of sea level pressure, surface temperature, and precipitation complexes over North America. *J. Appl. Meteor.,* **6,** 791–802.

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

Loeve, M., 1978: *Probability Theory, Vol. II,* 4th ed., Springer-Verlag, 413 pp.

Longuet-Higgins, M. S., 1964: Planetary waves on a rotating sphere. I. *Proc. Roy. Soc. London,* **A279,** 446–473.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction, Rep. 1, Statist. Forecasting Project., MIT.

Monin, A. S., and A. M. Obukhov, 1967: On statistical investigation of large-scale atmospheric processes, Nauka, Moscow, 195–200 (Russian with English summary). [Available from Amer. Inst. Aeronaut. Astronant., New York, NY 10019.]

North, G. R., and R. F. Cahalan, 1981: Predictability in a solvable stochastic climate model. *J. Atmos. Sci.,* **38,** 504–513.

——, T. Bell, F. Modeng, and R. F. Cahalan, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.,* **110,** 699–706.

Obukhov, A. M., 1947: Statistically homogeneous fields on a sphere. *Usp. Mat. Navk.,* **2,** 196–198.

Preisendorfer, R. W., 1979: Principle components and the motions of simple dynamical systems. SIO Ref. Ser. 70-11. Scripps Institution of Oceanography, 46 pp.

——, and T. P. Barnett, 1977: Significance tests for empirical orthogonal functions, *Proc. Fifth Conf. on Probability and Statistics,* Las Vegas, Amer. Meteor. Soc., 169.

——, F. W. Zwiers and T. P. Barnett, 1981: Foundations of principal component selection rules. SIO Ref. Ser. 81-4. Scripps Institution of Oceanography, 192 pp.

Simmons, A. J., J. M. Wallace and G. W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.,* **40,** 1363–1392.

Wallace, J. M., and R. Dickinson, 1972: Empirical orthogonal representation of time series in the frequency domain. Part I: Theoretical considerations. *J. Appl. Meteor.,* **11,** 887–892.

——, and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.,* **109,** 785–812.

Weiner, N., 1947: *Extrapolation, Interpolation and Smoothing of Stationary Time Series.* Technology Press, Cambridge and Wiley, New York, 306 pp.

Unauthenticated | Downloaded 12/06/23 05:46 AM UTC