Geophysical Fluid Dynamics, Nonautonomous Dynamical Systems, and the Climate Sciences

Michael Ghil and Eric Simonnet

Abstract This contribution introduces the dynamics of shallow and rotating flows that characterizes large-scale motions of the atmosphere and oceans. It then focuses on an important aspect of climate dynamics on interannual and interdecadal scales, namely the wind-driven ocean circulation. Studying the variability of this circulation and slow changes therein is treated as an application of the theory of nonautonomous dynamical systems. The contribution concludes by discussing the relevance of these mathematical concepts and methods for the highly topical issues of climate change and climate sensitivity.
Chapter 1

Effects of Rotation

The first two chapters of this contribution are dedicated to an introductory review of the effects of rotation and shallowness on large-scale planetary flows. The theory of such flows is commonly designated as geophysical fluid dynamics (GFD), and it applies to both atmospheric and oceanic flows, on Earth as well as on other planets. GFD is now covered, at various levels and to various extents, by several books [36, 60, 72, 107, 120, 134, 164].

The virtue, if any, of this presentation is its brevity and, hopefully, clarity. It follows most closely, and updates, Chapters 1 and 2 in [60]. The intended audience includes the increasing number of mathematicians, physicists and statisticians that are becoming interested in the climate sciences, as well as climate scientists from less traditional areas — such as ecology, glaciology, hydrology, and remote sensing — who wish to acquaint themselves with the large-scale dynamics of the atmosphere and oceans.

For tutorial purposes, we start this chapter with the dynamical effects of rotation on large-scale flows in the simplest cases. The Rossby number is defined as a nondimensional measure of the importance of rotation, and the usefulness of a rotating frame of reference is outlined.

Equations of fluid motion are first stated and discussed in an inertial frame of reference. The necessary modifications for passing to a rotating frame are derived. The Coriolis acceleration is shown to be the main term that distinguishes between the equations in a rotating and an inertial frame.

Vorticity is defined and approximate conservation laws for it, namely Kelvin’s theorem and Ertel’s theorem, are derived. The important distinction between barotropic and baroclinic flows is made. Finally, motion at small Rossby number is discussed, the geostrophic approximation is stated, and the Taylor-Proudman theorem for barotropic flow is derived from it. Geostrophic flow is shown to be a good qualitative approximation for atmospheric cyclones.
1 The Rossby Number

Newtonian mechanics recognizes a special class of coordinate systems, called inertial frames, and by rotation we always mean rotation relative to an inertial frame. As we shall see, Newtonian fluid dynamics relative to a rotating frame have a number of unusual features.

For example, an observer mowing with constant velocity in an inertial frame feels no force. If the observer, however, has a constant velocity with respect to a frame attached to a rotating body, a Coriolis force will generally be felt, acting in a direction perpendicular to the velocity. Such “apparent” forces lead to dynamical results which are unexpected and peculiar when compared with conventional dynamics in an inertial frame. We should therefore begin by verifying that this shift of viewpoint to a noninertial frame is really called for by the physical phenomena to be studied, as well as by our individual, direct perception of these flows, as rotating observers.

To examine this question for the dynamics of fluids on the surface of the Earth, we define \( \Omega \) to be the Earth’s angular speed, so \( 2\pi / \Omega \equiv 24 \text{ hr} \). Therefore, at a radius of 6380 km a point fixed to the equator and rotating with the surface moves with speed 1670 km/hr as seen by an (inertial) observer at the center of the Earth. On the other hand, the large-scale motion of the fluid relative to the solid surface never exceeds a speed of about 300 km/hr, which occurs in the atmospheric jet streams. Thus the speeds relative to the surface are but a fraction of inertial speeds, and it will be useful to remove the fixed component of motion due to the Earth’s rotation, by introducing equations of motion relative to rotating axis.

In this rotating frame, we define \( U \) to be a characteristic speed relative to the surface, and define \( L \) as a length over which variations of relative speed of order \( U \) occur. From these parameters we then obtain a characteristic time \( T = L / U \), which is a time scale for the evolution of the fluid structure in question. Since \( \Omega \) is an inverse time, the product \( T \Omega \), or more customarily the Rossby number

\[
\varepsilon = \frac{U}{2\Omega L}
\]  

is a dimensionless measure of the importance of rotation, when assessed over the lifetime or period of the structure.

The nature of this rotational effect will be given a dynamical basis forthwith. For the moment, it is important only to realize that small \( \varepsilon \) means that the effects of rotation are important. With this in mind, we show in Table 1 values of the Rossby number for a few large-scale flows on the Earth and Jupiter.

We conclude that rotation is important in all of these systems and that a rotating coordinate system is probably useful for their study. Minor changes in these estimates are needed to account for the fact that the proper measure of rotation is the local angular velocity of the tangent plane to the planet, as shown in Sec. 2 below.

Also, these estimates only mean that Coriolis forces, proportional to \( 2\Omega U \), are typically larger than the forces needed to produce the acceleration measured relative to the rotating frame, these being of magnitude \( U^2 / L \) (or \( U / T \)) times the fluid
Geophysical Fluid Dynamics, Dynamical Systems, and the Climate Sciences

Table 1

| r          | Feature         | L     | U         | ε      |
|------------|-----------------|-------|-----------|--------|
| Earth      | Gulf Stream     | 100 km| 1 m/sec   | 0.07   |
| Ω = 7.3 · 10⁻⁵ s⁻¹ | Weather system | 1000 km| 20 m/sec  | 0.14   |
|           | Core            | 3000 km| 0.1 cm/sec| 2 · 10⁻⁷|
| Jupiter    | Bands, Red Spot | 10⁴ km| 50 m/sec  | 0.015  |

density. There may be other forces, such as viscous or magnetic forces, which enter into the dynamical balance. In the Earth’s fluid core, for example, magnetic and Coriolis forces are probably comparable, while in the free atmosphere and ocean it is predominantly the “pressure” forces which compete with the Coriolis force in the dominant dynamical balance.

2 Equations of Motion in an Inertial Frame

We consider a fluid of density \( \rho \), moving under the influence of a pressure field \( p(r,t) \) and a body force \( F(r,t) \), where \( r = (x,y,z) = (x₁,x₂,x₃) \) is the position vector and \( t \) is time. The equations describing the local conservation of momentum and mass are

\[
\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla p + \frac{1}{\rho} \mathbf{F},
\]

(2a)

\[
\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0,
\]

(2b)

where \( \mathbf{u}(r,t) \) is the velocity field and \( \nabla = (\partial_x, \partial_y, \partial_z) \) is the gradient operator, with bold type used for vectors. Equation (2b) is often called either the “continuity equation” or the “mass conservation equation.”

For a given \( \mathbf{F} \), (2) comprises four scalar equations for the three velocity components of \( \mathbf{u} \), the density \( \rho \) and the pressure \( p \). Another relation is therefore needed to close the system, e.g., an equation of state connecting \( p \) and \( \rho \).

In many problems of geophysical interest, such as ocean dynamics, the approximate incompressibility of the fluid plays a prominent role. In these cases, \( \mathbf{u} \) is prescribed to be a solenoidal, i.e., divergence-free, vector field, and thus it closes the system. For the moment, we leave the closure unspecified and deal with the incomplete system (2). We also take \( \mathbf{F} \) at first to be a conservative field, such as gravitation, defined in terms of a potential \( \phi \) by

\[
\mathbf{F} = -\rho \nabla \phi.
\]

(3)
We are thus excluding here the possibility of the important contribution of the non-conservative viscous stresses.

The operator

\[ \frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \]

occurring in (2) is called the material derivative or the Lagrangian derivative, since it is the time derivative calculated by an observer moving with the local or material velocity. In the following, we will use the notation \( D/Dt \) rather than \( d/dt \) when appropriate to avoid any confusion. To verify this last claim let \( G(r, t) \) be a field and let \( r = R(t) \) be the position of a fluid particle at time \( t \). Then \( \mathbf{u} = (u, v, w) = (u_1, u_2, u_3) = \frac{dR}{dt} \) is the velocity of the particle, and we have

\[ \frac{d}{dt} G(R(t), t) = \frac{\partial G}{\partial t} + \frac{dR}{dt} \cdot \nabla G = \frac{\partial G}{\partial t} + u_i \frac{\partial G}{\partial x_i} \]

where the summation convention of repeated indices is used. A quantity \( Q \) (scalar, vector or tensor) which satisfies \( DQ/Dt = 0 \) is said to be a material invariant because the quantity \( Q(R(t), t) \) just stays constant along the trajectory defined by \( R \).

2.1 Thermodynamics

Very often, the equation of state does not involve the pressure and density alone. Typically, one often has

\[ \rho_{\text{ocean}} = \rho(S, T, p), \quad \rho_{\text{atmos}} = \rho(T, p), \]

where \( S \) is the salinity and \( T \) the temperature. One needs, therefore, two additional equations for \( T \) (and \( S \)). In the following, we consider the oceanic case, although the cases of both a dry and a humid atmosphere can be treated in an analogous way.

We will suppose later on that the fluid is in a local thermodynamic equilibrium. This assumption implies that the variables entering Eq. (6) below are locally related or that they can be simply expressed in mathematical form:

\[ E = E(\alpha, \eta, S), \]

where \( E \) is the specific internal energy, \( \alpha \equiv 1/\rho \) is called the specific volume, \( \eta \) the specific entropy and \( S \) the salinity. Taking the total differential of (6),

\[ dE = \frac{\partial E}{\partial \alpha} d\alpha + \frac{\partial E}{\partial \eta} d\eta + \frac{\partial E}{\partial S} dS, \]

enables one to interpret it in a illuminating way: The quantity \( -\partial E(\alpha, \eta, S)/\partial \alpha \) is interpreted as the pressure \( p \), while the term \( (\partial E/\partial \alpha)d\alpha \) is the change in energy
due to mechanical compression. Note that \( d\alpha < 0 \) and \( p > 0 \) give \( dE > 0 \), i.e. energy must be supplied to reduce volume if fluid resists compression.

Similarly, \( \partial E(\alpha, \eta, S)/\partial \eta \) is interpreted as the temperature \( T \), and the term \( (\partial E/\partial \eta)d\eta \) is the energy change due to heating. Finally, \( \partial E(\alpha, \eta, S)/\partial S \) is called the chemical potential \( \mu \) of seawater. One has therefore

\[
dE = -pd\left(\frac{1}{\rho}\right) + Td\eta + \mu dS. \tag{8}
\]

It appears that the salinity and specific entropy are conserved along particle motion when there are no heat sources and no molecular diffusion, i.e.

\[
\frac{D\eta}{Dt} = \frac{DS}{Dt} = 0. \tag{9}
\]

In order to have an equation involving \( T \) instead of \( \eta \), one expresses the entropy \( \eta \) as a function of \( p, S \) and \( T \) only. It is possible to eliminate \( \alpha \) between \( \rho(\alpha, \eta, S) \) and \( T(\alpha, \eta, S) \). Using (9), one has

\[
\frac{D\eta}{Dt} = \frac{\partial \eta}{\partial p} \frac{Dp}{Dt} + \frac{\partial \eta}{\partial T} \frac{DT}{Dt} + \frac{\partial \eta}{\partial S} \frac{DS}{Dt} = \frac{\partial \eta}{\partial p} \frac{Dp}{Dt} + \frac{\partial \eta}{\partial T} \frac{DT}{Dt} = 0. \tag{10}
\]

It appears that the partial derivatives are known quantities:

\[
\left(\frac{\partial \eta}{\partial T}\right)_p \equiv C_p T \text{ and } \left(\frac{\partial \eta}{\partial p}\right)_T \equiv -\frac{C_T}{\rho}, \tag{11}
\]

where \( C_p \) is the specific heat at constant pressure and salinity and \( C_T \) the coefficient of thermal expansion. Therefore, one ends up with

\[
\frac{DT}{Dt} = \frac{T}{\rho} \frac{C_T}{C_p} \frac{Dp}{Dt}, \tag{12a}
\]

\[
\frac{DS}{Dt} = 0. \tag{12b}
\]

For the oceans, the right-hand side of the first equation is often neglected, unless pressure variations are very large compared to temperature fluctuations, as could be the case in the deep ocean. Note that alternative formulations to close the system (2) do exist, depending on which variables one considers, instead of the temperature \( T \) addressed here. The overall procedure, however, is always the same and it makes use of the internal energy (6) and of the identities (7) and (9).

**Example.** In the case of the atmosphere, the ideal gas law is a good approximation of the equation of state for dry air.

\[
\rho = \frac{p}{kT}. \tag{13}
\]
The specific entropy $\eta$ is

$$\eta = C_p \log T - R \log p. \quad (14)$$

One obtains $C_T = -\rho \left( \frac{\partial \eta}{\partial p} \right)_T = \frac{1}{T}$. Plugging this expression into (12a) and using (13), one obtains

$$\frac{D T}{D t} = \frac{R}{C_p} \frac{D p}{D t}, \quad (15)$$

or with straightforward algebraic manipulations

$$\frac{D \theta}{D t} = 0, \text{ with } \theta = T \left( \frac{p_0}{p} \right)^{R/C_p}. \quad (16)$$

The quantity $\theta$ is called the potential temperature and $p_0$ a constant pressure reference.

### 3 Equations in a Rotating Frame

We would like to use the form taken by (2) relative to a rotating coordinate system. That is, we would like to work with the velocity as perceived by an observer fixed in the rotating frame, and with time differentiation as would be performed by such an observer. Spatial differentiation, involving a limit process at fixed time, is an operation which is invariant under transformation to a rotating frame. The essential problem is thus to deal properly with time derivatives.

#### 3.1 Absolute and relative derivative

Since differentiation involves a limit process, it is convenient to begin with an infinitesimal rotation of the axis. Consider a constant vector $\mathbf{P}$ fixed rigidly relative to the rotating frame. Let the frame rotates with respect to the inertial frame through a small angle $d\theta$ about the $z$ or $x_3$-axis. If $\mathbf{P}$ has coordinates $(P_1 = a \cos \varphi, P_2 = a \sin \varphi, P_3)$ in the rotated frame, an inertial observer will measure in his frame, after rotation, some coordinates $(P'_1 = a \cos(\varphi + d\theta), P'_2 = a \sin(\varphi + d\theta), P'_3 = P_3)$. Using the smallness of $d\theta$ one easily obtains $(P'_1 = P_1 - P_2 d\theta, P'_2 = P_2 + P_1 d\theta, P'_3 = P_3)$ for $\mathbf{P}$. Thus, under this rotation, we have the invariant vector relation

$$d\mathbf{P} = d\theta \times \mathbf{P}, \quad (17)$$

where the vector $d\theta = (0, 0, d\theta)$. Introducing time $t$, we write (17) as

\footnote{Note that, since only the variables $T$ and $p$ are involved, we could simply derive (14) by integrating (11); in this case, one obtains $TC_T = \text{const.}$}
\[
\left( \frac{dP}{dt} \right)_0 = \frac{d\theta}{dt} \times P \equiv \Omega \times P, 
\]

where we have defined \( \Omega(t) = (0, 0, d\theta/dt) \) as the instantaneous angular velocity of the rotating frame. Equation (18) involves the subscript "0", which will stand for “as determined by an observer in the inertial frame”, sometimes called absolute. Since \( P \) is a constant vector in the rotating frame, (18) simply states that, in this case, the inertial observer sees a vector attached to the origin rotating (about the \( z \)-axis).

Examples of such “rotating, fixed vectors” are a set of orthonormal basis vectors in the rotating frame, \( (i_1, i_2, i_3) \), \( i_k \) being the unit vector along the \( x_k \) axis. Thus, we can recompute (18) as follows. Define

\[
P = P_1 i_1 + P_2 i_2 + P_3 i_3 
\]

as the representation of \( P \) in the rotating basis. Since \( P_1, P_2, \) and \( P_3 \) are independent of time we have, using (18) and the summation convention,

\[
\left( \frac{dP}{dt} \right)_0 = P_k \frac{d i_k}{dt} = P_k (\Omega \times i_k) = \Omega \times P_k = \Omega \times P. 
\]

Thus we see that the basis vectors carry the effect of rotation in their time dependence.

Suppose now that the coordinates \( P_k \) in (19) depend on time, \( P_k = P_k(t) \). Repeating (20) we obtain

\[
\left( \frac{dP}{dt} \right)_0 = \left( \frac{dP}{dt} \right)_r + \Omega \times P, 
\]

where

\[
\left( \frac{dP}{dt} \right)_r = P_k'(t) i_k 
\]

is the time derivative of \( P \) relative to the rotating observer, with \( P_k' = dP_k(t)/dt \). The relation (21) is the basic result that must be used to transform the equations of motion to a rotating frame of reference.

We remark that \( \Omega \) is allowed to depend on time \( t \) in (21). This can be of interest in geophysical fluid dynamics when modeling the precession of the Earth’s axis of rotation. However, this time dependence has negligible effects for most phenomena in the dynamics of the atmosphere, oceans, or fluid core of the earth and will be largely ignored below. We also note that, for the purpose of interpreting a formula such as (18) or (21), it is sometimes convenient to think of the relation as applying at an instant when the two frames of reference coincide.

Actually, we do not need (21) for the continuity equation (23), since it is obvious that the material derivative of a scalar field is independent of the coordinate frame. Indeed, its value is determined by the time dependence observed for a given particle of fluid. To verify the invariance explicitly, let \( G(r,t) \) be a scalar field as observed in the inertial frame.

Then, for partial time derivatives we have \( P = r \) in (18) and thus
\[ \left( \frac{\partial G}{\partial t} \right)_r = \left( \frac{\partial G}{\partial t} \right)_0 + (\Omega \times r) \cdot \nabla G. \]  

(22a)

Now, applying (21) to the vector \( r = R(t) \), which represents the particle position at time \( t \), yields

\[ u_0 = \left( \frac{dR}{dt} \right)_0 = \left( \frac{dR}{dt} \right)_r = u_r + \Omega \times r, \]  

(22b)

where

\[ u_r = \left( \frac{dR}{dt} \right)_r, \]  

(22c)

is the velocity vector relative to the rotating frame. Using (22a, 22b) one obtains

\[ \left( \frac{\partial G}{\partial t} \right)_r + u_r \cdot \nabla G = \left( \frac{\partial G}{\partial t} \right)_0 + u_0 \cdot \nabla G, \]  

which confirms the invariance of the material derivative.

### 3.2 Relative acceleration and Coriolis acceleration

The right-hand side of (22b) is another vector function of \( t \), and we may repeat the process to obtain the acceleration:

\[ \left( \frac{d^2 R}{dt^2} \right)_0 = \frac{d}{dt} \left[ (u_r + \Omega \times r) \right]_r + \Omega \times \left[ u_r + \Omega \times r \right] \]

\[ = \left( \frac{du_r}{dt} \right)_r + 2\Omega \times u_r + \Omega \times r + \Omega \times (\Omega \times r) \]  

(23)

where \( u_0 \) is the velocity perceived by the inertial observer. Using (23) and (2a) and recalling the invariance of spatial derivatives we thus obtain the desired system of equations relative to the rotating frame:

\[ \left[ \left( \frac{\partial u_r}{\partial t} \right)_r + (u_r \cdot \nabla)u_r + 2\Omega \times u_r + \Omega \times r + \Omega \times (\Omega \times r) \right] = -\frac{1}{\rho} \nabla p - \nabla \phi \]  

(24a)

\[ \left( \frac{\partial \rho}{\partial t} \right)_r + \nabla \cdot \rho u_r = 0. \]  

(24b)

If \( \Omega \) is taken to be constant, ie \( \dot{\Omega} = 0 \) and if we use the identity

\[ \Omega \times (\Omega \times r) = -\nabla \left( \frac{|\Omega \times r|^2}{2} \right), \]  

(25)
then Eq. (24a) takes the form

\[ \left( \frac{\partial u_r}{\partial t} \right)_r + \left( u_r \cdot \nabla \right) u_r + 2 \Omega \times u_r = -\frac{1}{\rho} \nabla \rho - \nabla \phi_c \]  

(26a)

\[ \phi_c = \phi - \frac{1}{2} |\Omega \times r|^2, \]  

(26b)

where \( \phi_c \) now absorbs the centripetal acceleration. The latter is only about 1/300 of the gravitational acceleration at the surface of the Earth and so is negligible for most atmospheric and oceanographic purposes. The remaining term which distinguishes Eq. (26) for the relative acceleration from Eq. (2a) for the absolute acceleration is the Coriolis acceleration \( 2 \Omega \times u_r \).

If \( U \) and \( L \) are the characteristic scales of Section 1 defined now in the rotating system, and if \( L/U \) is taken as the characteristic time, then in order of magnitude

\[ \frac{\partial u_r}{\partial t} \sim U^2/L, \quad 2 \Omega \times u_r \sim 2 \Omega U, \]  

(27)

and so the Rossby number has the dynamical meaning of a characteristic ratio of the relative acceleration of a fluid element to the Coriolis acceleration.

For future reference, we now drop the subscript “\( r \)” and rewrite the basic equations (24) as:

\[ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + 2 \Omega \times \mathbf{u} = -\frac{1}{\rho} \nabla \rho - \nabla \phi_c - \ddot{\Omega} \times \mathbf{r} \]  

\[ \frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0. \]  

(28)

4 Vorticity

For any velocity field \( \mathbf{u} \) we define the associated vorticity field by

\[ \omega = \nabla \times \mathbf{u}, \]  

(29)

where we recall the explicit calculation:

\[ \nabla \times \mathbf{u} = \begin{pmatrix} \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \\ \frac{\partial}{\partial x} \end{pmatrix} \times \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{\partial w}{\partial z} - \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial x} - \frac{\partial w}{\partial y} \\ \frac{\partial v}{\partial y} - \frac{\partial u}{\partial x} \end{pmatrix}. \]  

(30)

**Remark:** The vorticity is a divergence-free vector field: \( \nabla \cdot \omega = \nabla \cdot \nabla \times \mathbf{u} = 0 \). In particular, thanks to Stoke’s theorem, for any surface \( S \) enclosing a volume \( V \) in \( \mathbb{R}^3 \) (for instance a sphere), one has
\[ \int_S \omega \cdot dS = \int_V \nabla \cdot \omega \, dV = 0. \] (31)

Define now a vortex line as the integral curve of the vorticity field \( \omega \) meaning that any points on the vortex line is tangent to the vorticity vector field. A vortex tube is composed of vortex lines going through a closed curve (see Fig. 1). Now applying (31) to a portion of this vortex tube determined by two different closed curves \( C_1 \) and \( C_2 \), one obtains the results that the circulation is the same around \( C_1 \) and \( C_2 \) (see Fig. 1). Indeed the lateral surface between \( C_1 \) and \( C_2 \) does not contribute to the circulation since it is tangent to the vector field. One can thus write

\[ \int_{C_1} u \cdot dr = \int_{C_2} u \cdot dr. \] (32)

Note that this is merely a “static” property since the vortex tube is frozen \((t \text{ is fixed}).\) A dynamical view will be provided by Kelvin’s theorem. □

For solid body rotation with angular velocity \( \Omega \), the velocity field is obtained by setting \( u_r = 0 \), so that (cf. 225) \( u = u_0 = \Omega \times r \). The associated vorticity is

\[ \nabla \times (\Omega \times r) = \Omega (\nabla \cdot r) - \Omega \cdot \nabla r = 3\Omega - \Omega = 2\Omega. \] (33)

Applied locally in an arbitrary velocity field, vorticity has the same meaning: it is equal to twice the angular velocity of a fluid element.

This interpretation of vorticity in terms of a local angular velocity is misleading in some respects, since the term “fluid element” is not really appropriate for describing what is basically a point property. More precisely, the rate-of-strain tensor given by the matrix of partial derivatives of the components of \( u \) at any point has a symmetric and an antisymmetric part. Vorticity is associated with the antisymmetric part, while the symmetric part gives rise to a straining field. The effect of the latter is to distort “fluid elements”, by differentially stretching and compressing them. With this caveat, we return to the discussion of vorticity.
Since global solid body rotation can be characterized, according to (33), as a state of uniform vorticity, the mechanics of the fluid relative to a rotating coordinate system can alternatively be regarded as a mechanics of deviations from a state of uniform vorticity. As we have already noted, the practical usefulness of reference to a state of uniform rotation will depend on the magnitude of the deviations, and the Rossby number can aso be regarded as a measure of relative vorticity as a fraction of $2\Omega$. The properties of vorticity relative to the rotating frame are thus clearly "inherited" from the mechanics in the inertial frame, so for the present discussion we return to (2) and omit subscripts.

4.1 Kelvin’s theorem

We now use the vector identity (please check by using formula (30))

$$(\mathbf{u} \cdot \nabla)\mathbf{u} = -\mathbf{u} \times \omega + \nabla \left( \frac{1}{2} |\mathbf{u}|^2 \right)$$

so that, (2a) can be reformulated as

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \left( \frac{1}{2} |\mathbf{u}|^2 \right) - \mathbf{u} \times \omega + \frac{1}{\rho} \nabla p = -\nabla \phi.$$  \hspace{1cm} (35)

The particularly simple case in which $\mathbf{u} \times \omega = 0$ is called a Beltrami flow. Taking the curl of (35) we obtain an equation for the vorticity,

$$\frac{\partial \omega}{\partial t} + (\mathbf{u} \cdot \nabla)\omega - (\omega \cdot \nabla)\mathbf{u} + \omega \nabla \cdot \mathbf{u} - \frac{1}{\rho^2} \nabla \rho \times \nabla p = 0.$$ \hspace{1cm} (36a)

Using (2b) to eliminate $(\nabla \cdot \mathbf{u})$ from (36a) we have

$$\frac{\partial \omega}{\partial t} + (\mathbf{u} \cdot \nabla)\omega - (\omega \cdot \nabla)\mathbf{u} - \frac{\omega}{\rho} \frac{\partial \rho}{\partial t} - \frac{1}{\rho^2} \nabla \rho \times \nabla p = 0.$$ \hspace{1cm} (36b)

Multiplying (36b) by $1/\rho$ and rearranging we obtain finally a simpler equation for $\omega/\rho$,

$$\frac{D(\omega/\rho)}{Dt} = \frac{\omega}{\rho} \cdot \nabla \mathbf{u} + \frac{1}{\rho^3} \nabla \rho \times \nabla p.$$ \hspace{1cm} (37)

It shows that the two terms on the right-hand-side will tend to change $\omega/\rho$ as one moves with the fluid. In particular, the vector $\omega/\rho$ is not in general a material invariant.

It would appear that the vorticity vector itself does not provide one with conservation laws having an intuitive appeal, basically because of the already mentioned non-local nature of physical reasoning based upon “fluid elements.” It is advanta-
geous, therefore, to recognize the divergence-free (solenoidal) property of the vorticity vector and to exploit the non-local concept of flux of a divergence-free vector field. This non-local approach is the basis for the fundamental theorem discovered by Lord Kelvin. *Kelvin’s theorem* studies the evolution of the flux of \( \omega \) through a surface \( S \) bounded by a simple closed contour \( C \), when the latter consists of material points, i.e. of points which move with the fluid (see Fig. 2).

**Fig. 2** Kelvin’s theorem: the circulation \( \Gamma_t \) at time \( t \) is the same than the circulation \( \Gamma_{t+dt} \) at time \( t+dt \) provided the fluid is barotropic and the corresponding closed contours consists of material points. Note that the vortex tube is displaced by the flow but might not be a vortex tube anymore (the vortex tube at time \( t+dt \) is shown in thin dashed contours).

By Stoke’s theorem, the flux we study is given by

\[
\int_S \omega \cdot ds = \int_C u \cdot dr, \tag{38a}
\]

where

\[
\Gamma(t) = \int_C u \cdot dr \tag{38b}
\]

is called the *circulation* on \( C \). Let \( C \) be given by parametric equations \( \{r = R(\alpha,t), 0 \leq \alpha < 1\} \); then a given material point in the curve corresponds to a given value of \( \alpha \). The parameter \( \alpha \) therefore is a material of Lagrangian parameter. One may think of \( C \) as a smoke ring for instance. One obtains
\[
\frac{D\Gamma}{Dt} = \frac{D}{Dt} \int_C u \cdot \frac{\partial R}{\partial \alpha} d\alpha
\]
\[
= \int_C \frac{Du}{Dt} \cdot \mathbf{dr} + \int_C u \cdot \left( \frac{D}{Dt} \left( \frac{\partial R}{\partial \alpha} \right) \right) d\alpha
\]
\[
= \int_C \frac{Du}{Dt} \cdot \mathbf{dr} + \int_C u \cdot \frac{\partial u}{\partial \alpha} d\alpha
\]
\[
= \int_C \left[ -\nabla p \rho - \nabla \phi \right] \cdot \mathbf{dr} = -\int_C \frac{\nabla p}{\rho} \cdot \mathbf{dr}, \tag{39}
\]

where the term \( \int_C u \cdot \frac{\partial u}{\partial \alpha} d\alpha = \int_C d(|u|^2/2) = 0 \) vanishes on a closed contour and the material derivative of \( r = R(\alpha, t) \) is just \( u \), i.e. \( u = \frac{Dr}{Dt} \). Note that the time derivative is a material derivative here and since \( C \) is moving with the fluid it is possible to commute integration and time derivative as done in the second line of this calculation letting \( C \) “unchanged”.

From the last form of (39), we see that if \( p \) and \( \rho \) are functionally related, i.e. \( \rho = \rho(p) \), then \( \int_C \frac{\nabla p}{\rho} \cdot \mathbf{dr} = \int_{p_0}^{p} \frac{dp}{\rho(p)} = 0 \) so that the circulation is conserved. Another way to see it, is to apply Stokes theorem on the right-hand-side which becomes
\[
-\int_C \frac{\nabla p}{\rho} \cdot \mathbf{dr} = \int_S \frac{\nabla \rho \times \nabla p}{\rho^2} \cdot \mathbf{ds}, \tag{40}
\]
which is zero if \( \rho \) and \( p \) are functionally dependent. We now state Kelvin’s theorem

**Theorem 1 (Kelvin’s theorem).** If the fluid is barotropic, i.e. surfaces of constant \( p \) coincide with surfaces of constant \( \rho \) then for any closed material contour moving with the fluid
\[
\frac{D\Gamma}{Dt} = 0 \quad \tag{41}
\]

The proof of this theorem, like Eq. (37), emphasizes the role of the pressure and density fields in the evolution of vorticity.

One calls the quantity
\[
B = \frac{\nabla \rho \times \nabla p}{\rho^2}, \tag{42}
\]
the **baroclinic vector**.

Thus, Kelvin’s theorem asserts the conservation of circulation in barotropic, inviscid flows. Flow situations in which \( B \neq 0 \) will be considered in the next chapters.
4.2 Ertel’s theorem

To probe deeper into the connection between Kelvin’s theorem and local variations of vorticity, it is useful to adopt a fully Lagrangian description of the motion, by defining $r(t; r^0)$ to be the position $(x_1, x_2, x_3)$ at time $t$ of a fluid particle initially at $r^0 = (x_1^0, x_2^0, x_3^0)$. Since $r^0$ labels a particle, it is fixed following the motion and the Jacobian tensor $J = \{J_{ij}\} = \left\{\frac{\partial x_i}{\partial x_0^j}\right\}$ satisfies

$$\frac{D}{Dt} J_{ij} = \frac{\partial}{\partial x_0^i} J_{kj} = \frac{\partial}{\partial x_0^i} \frac{\partial u_k}{\partial x_k} J_{kj} = \frac{\partial}{\partial x_k} \frac{\partial x_k}{\partial x_0^i} J_{kj}.$$  \hspace{1cm} (43a)

For any differential $dl$ separating two nearby points $r^0$ and $r^0 + dl$ on a material curve at time $t = 0$, the quantity $J dl$ (with components $J_{ij} dl_j$) can be regarded as the vector connecting the two nearby points on the material curve at all subsequent times. By repeating the calculation (43a) for this latter quantity, we obtain

$$\frac{D}{Dt} J_{ij} dl_j = \frac{\partial}{\partial x_0^i} J_{kj} dl_j = \frac{\partial}{\partial x_k} \frac{\partial x_k}{\partial x_0^i} J_{kj} dl_j.$$

(43b)

These calculations remarkably show that $\omega/\rho$ in (37) and $dr \equiv J dl$ in (43b) satisfy the same equation provided the baroclinic vector $B$ is zero. Now $dr$ is a differential along a material curve. The instantaneous “trajectories” of the vector field $\omega/\rho = (\omega/\rho)(r, t)$ are defined analogously as the integral curves of the system

$$\frac{dr}{d\sigma} = \frac{\omega}{\rho} (r(\sigma), t).$$

(44)

To understand the meaning of this equation, the vector field $\omega/\rho$ is “frozen” ($t$ is fixed), then one starts from an initial position at say $r(\sigma = \sigma_0)$ and follows the trajectory so that it always stays tangent to the vector field. This procedure defines an instantaneous line or trajectory. We can, at any time, select such an instantaneous trajectory and thereafter follow the evolution of this curve as a material curve. A priori, such a curve as it evolves has no reason to be a solution of (44) anymore. However, since $\omega/\rho$ and $dr$ satisfy the same equation, it follows, assuming uniqueness of the solution, that this material curve remains a trajectory of $\omega/\rho$. Therefore, the trajectories of $\omega/\rho$ are material curves in a barotropic flow. One usually refers to these trajectories as vortex lines, and they can be visualized as being carried about by the fluid although here we consider the vector field $\omega/\rho$ instead of the vorticity field alone.
The first term on the right of (37), which might have the appearance of a “source” of vorticity, is in reality part of the statement of the material nature of vortex lines, although \( \omega/\rho \) itself is not a material invariant. The only true source of vorticity in (37) is the barotropic vector \( B \). The material changes of \( \omega/\rho \) which occur when \( B = 0 \) are due to changes in the geometry of the vortex lines, i.e., to the twisting and stretching of small elements threaded by material lines. As a small pencil of vortex lines, commonly called a vortex tube is stretched out by the flow, the cross section diminishes and the local vorticity grows in order to maintain, in accordance with Kelvin’s theorem, a fixed circulation about the tube.

The most concise local statement concerning vorticity is an immediate consequence of the material nature of vortex lines. If \( B = 0 \), the vector \( (\omega/\rho) J^{-1} = (\partial \rho/\partial r) \cdot \nabla r \) is a material invariant.

To see this note first, by differentiating \( \frac{D}{Dt} J^{-1} = I \), that

\[
\frac{D}{Dt} \left( \frac{\omega}{\rho} J^{-1} \right) = \left[ \frac{D}{Dt} \left( \frac{\omega}{\rho} \right) \right] J^{-1} + \frac{\omega}{\rho} \frac{D}{Dt} (J^{-1}) = \left( \frac{\omega}{\rho} \cdot \nabla u \right) J^{-1} - \frac{\omega}{\rho} J^{-1} (J : \nabla u) J^{-1} = 0, \tag{45}
\]

which establishes the invariance. Since \( J \) is initially the identity tensor, we see that the invariant in (45) is just the initial value of \( \omega/\rho \), or

\[
\frac{\omega}{\rho} (r, t) = \frac{\omega}{\rho} (r^0, 0) J(r, t). \tag{46}
\]

Now, suppose that \( \lambda \) is any materially invariant scalar, i.e. \( D\lambda/\rho = 0 \). Then, we may regard \( \lambda \) as a function of the Lagrangian labels \( r^0 \), i.e. \( \lambda (r^0) = \lambda (r) = \text{cst.} \). Then, by remarking that \( \nabla \lambda = \nabla_0 \lambda J^{-1} \) and \( D(\nabla \lambda) /Dt = 0 \), we have

\[
\frac{D}{Dt} \left( \frac{\omega}{\rho} \cdot \nabla \lambda \right) = \nabla_0 \lambda \cdot \frac{D}{Dt} \left( \frac{\omega}{\rho} J^{-1} \right) = 0. \tag{47}
\]

We may now state

**Theorem 2 (Ertel’s theorem).** Let \( \lambda \) be a scalar field which is invariant by the flow \( (D\lambda/\rho = 0) \), then provided the fluid is barotropic or \( \lambda \) is a function of \( \rho \) and \( p \) only, one has

\[
\frac{D}{Dt} \left( \frac{\omega}{\rho} \cdot \nabla \lambda \right) = 0. \tag{48}
\]

\(^2\) or using indices \( (\omega/\rho) J^{-1} = \frac{\partial \lambda_0}{\partial \rho} \frac{\partial \rho}{\partial x_j} \)
The quantity $\frac{\omega}{\rho} \cdot \nabla \lambda$ is called (Ertel) potential vorticity. The case where $\lambda$ is a function of $\rho$ and $p$ only should be clear since $\nabla \lambda = \frac{\lambda}{\rho} \nabla \rho + \frac{\lambda}{p} \nabla p$ is orthogonal to the baroclinic vector $B$. Therefore, in contradistinction to Kelvin’s theorem, Ertel’s theorem can also be generalized to baroclinic flows. In barotropic flows, $B = 0$, and $\lambda$ is unrestricted by being function of $\rho$ and $p$.

Let us now apply Kelvin’s theorem to a small contour $C$ lying on a surface of constant $\lambda$, where $\lambda$ is again a scalar invariant (see Fig. 3).

![Fig. 3 Kelvin’s theorem and Ertel’s theorem for a material fluid element of cylindrical shape. Note that the closed contour $C$ always stays on the surface $\lambda = \lambda_0$.](image)

Kelvin’s theorem asserts the constancy of $\omega \cdot dS$ where $dS = dSn$ is the area enclosed by $C$, while Ertel’s theorem asserts the constancy of $(\omega/\rho) \cdot (d\lambda/dl)n$. Since the mass $\rho dS dl$ is conserved, the two statements become identical in this local setting.

Intuitively, we see that as a tube of vortex lines is laterally compressed, the intensity of the field will increase, by Kelvin’s theorem applied to the cross-section of the tube. The presence of $\rho$ in Ertel’s local invariant can be traced to the need to simultaneously satisfy Kelvin’s theorem and conservation of mass in a small element undergoing deformation. In this sense, Kelvin’s theorem is the more fundamental characterization of vorticity. Nevertheless, Ertel’s theorem will be useful to us, since in certain cases it is possible to derive an invariant $\lambda$ and to cast the dynamical problem into a form equivalent to Ertel’s equation. This is the case, for instance, in the next chapter.

Our vorticity computations so far have been in the inertial frame, but in view of the correspondence

$$\omega_0 = 2\Omega + \omega_r,$$

that follows from (22) and (33), one can write Ertel’s equation relative to the rotating frame as
\[
\frac{d}{dt} \left( \frac{2\Omega + \omega_r}{\rho} \cdot \nabla \lambda \right) = 0.
\] (49)

5 The Geostrophic Approximation

The geostrophic approximation governs flows at small Rossby number, i.e., at relatively rapid rotation. From now on, unless otherwise specified, the equations of motion will be assumed to be relative to a rotating frame, and given by (28).

To study motion at small Rossby number, we shall neglect the inertial terms \( \partial_t u + u \cdot \nabla u \) in (28) relative to the Coriolis force. Generally, this requires that both \( U/2\Omega L \) and \( 1/2\Omega T \) be small, where \( U, L, \) and \( T \) are scales of the velocity field.

5.1 The Taylor-Proudman theorem

The geostrophic approximation to (28) is then given by

\[
2\Omega \times u + \frac{1}{\rho} \nabla p = -\nabla \phi_c.
\] (50)

The curl of (50) yields

\[
-2(\Omega \cdot \nabla) u + 2\Omega (\nabla \cdot u) = B.
\] (51)

If the baroclinic vector (42) vanishes, then with \( \Omega = (0, 0, \Omega) \) and \( u = (u, v, w) \) the components of (51) reduce to

\[
2\Omega \frac{\partial u}{\partial z} = 0, \quad 2\Omega \frac{\partial v}{\partial z} = 0, \quad 2\Omega \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0.
\] (52)

If, in addition, the fluid can be regarded as incompressible, the constraint \( \nabla \cdot u = 0 \) allows us to conclude from (52) that

\[
\frac{\partial u}{\partial z} = 0.
\] (53)

This is the Taylor-Proudman theorem, stating that the velocity field in the geostrophic approximation is invariant along the axis of rotation for a barotropic, inviscid flow. It highlights a striking feature of rotating flows, which was demonstrated experimentally by G. I. Taylor (Fig. 4).

A cylindrical tank of water was rotated with constant speed, and on the bottom of the tank a small cylindrical obstacle was moved horizontally with very small relative speed. In ordinary inviscid fluid dynamics, we would expect the fluid to move over and around the obstacle, the velocity vector being roughly tangent of the
boundary. But, on the top of the tank $w$ must vanish everywhere. If fluid particles can only move horizontally, with a velocity independent of $z$, then the flow field is determined by its structure in a horizontal plane through the cylinder. Indeed, in the experiment, the obstacle was observed to carry an otherwise stagnant column of fluid with it. This “Taylor columns” is an example of the nonlocal effect which boundary conditions can produce in a rapidly rotating fluid.

5.2 Geostrophic motion

Suppose now that we deal with a rotating layer of fluid of constant density with $\phi = g z$, and with $w = 0$ rather than $B = 0$. Then the components of (50) give

$$-2\Omega v = -\frac{1}{\rho} \frac{\partial p}{\partial x}, \quad 2\Omega u = -\frac{1}{\rho} \frac{\partial p}{\partial y}, \quad \frac{\partial p}{\partial z} = -\rho g. \quad (54)$$

This is an example of a hydrostatic as well as geostrophic flow, wherein the Coriolis force is in equilibrium with the horizontal pressure gradient.

Looking down on the layer, the flow in the vicinity of a pressure minimum rotates in the same direction as the large scale rotation. In meteorology, this would roughly approximate the wind field associated with a low in the pressure field, and is called a cyclonic circulation. Anticyclonic circulations correspond to local highs.

Figure 5 is correct as drawn in the Northern Hemisphere, with the cyclonic wind blowing counterclockwise around the low. In the Southern Hemisphere, the cyclonic wind around a low blows clockwise around it. This simple rule is stated by meteorologists as the Buys-Ballot law: in the Northern Hemisphere, if a person stands with the back to the wind, atmospheric pressure is low to the left and high to the
right. A more entertaining way of stating it is that, if lost, stand with the back to wind — if the high pressure is to your left, you're in the Southern Hemisphere.

Fig. 5 Geostrophic flow in a Northern Hemisphere cyclone. Circles are streamlines, and arrows indicate the sense of rotation.

The geostrophic approximation highlights the unusual features of the dynamics of a fluid viewed in a rotating coordinate system. To an observer in the inertial frame, geostrophic flows are predominantly solid body rotation. In order to understand the physical meaning of the dynamical balance represented in Figure 5, this background solid body rotation must be kept in mind. However, it must also be remembered that a spatial pattern evolving on a moderate time scale relative to a rapidly rotating frame will be seen by the inertial observer as a slow modulation of the basic time dependence due to rotation of the pattern.

It is therefore helpful to deal directly with the relative motion in developing one’s intuition concerning the effects of rotation. For example, in interpreting Fig. 5, the Coriolis acceleration associated with the eddy is seen to be directed towards its center, and this acceleration is in equilibrium with minus the pressure gradient.
Chapter 2
Effects of Shallowness

The most important geometrical approximation in geophysical fluid dynamics stems from the effective shallowness of the fluid layers relative to global horizontal scales of atmospheric and oceanic flow. In the present chapter, we shall study this approximation in the simplest setting of an incompressible inviscid fluid of constant density, with uniform body force, namely gravitation, acting vertically downwards. The resulting approximate description of motion in a frame which rotates about a vertical axis is usually referred to as rotating shallow-water theory.

The equations of motion for a constant rotation rate $\Omega$ are subjected to a systematic scale analysis, where $H$ is the vertical scale, $L$ is the horizontal scale and the aspect ratio $\delta \equiv H/L$ is small. The shallow-water equations with rotation are thus derived, and shown to satisfy a particular form of Ertel’s theorem. To wit, potential vorticity $(f_0 + \zeta)/(h - h_0)$ is conserved, where $\zeta$ is the fluid’s relative vorticity, $f_0 = 2\Omega$ the background vorticity, and $h - h_0$ the height of a fluid column.

Small-amplitude solutions of these equations are studied in a basin with prescribed depth, using linear theory. Gravity waves are obtained as natural modes when $f_0 = 0$, and their modification when $f_0 \neq 0$ is investigated, yielding Poincaré’s inertia-gravity waves. The “edge waves” that are related to the Poincaré waves, and called Kelvin waves, are also derived.

Finally, we study low-frequency modes that satisfy an approximate geostrophic balance between Coriolis force and pressure-gradient force. Stationary geostrophic modes, geostrophic contours of a basin and the infinite multiplicity of solutions they determine are analyzed. Rossby waves over sloping bottom topography are shown to remove this geostrophic degeneracy, and we examine their relationship with the high-frequency Poincaré and Kelvin waves, along with their physical structure.

6 Derivation of the Equations for Shallow Water

For constant $\Omega = (0, 0, \Omega)$ and $\rho$, we write the governing equations in the form
\[ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + 2 \Omega \times \mathbf{u} = -\frac{1}{\rho} \nabla p - g i_3, \quad (55a) \]

\[ \nabla \cdot \mathbf{u} = 0, \quad (55b) \]

where \( 2\Omega = (0, 0, f_0) \), \( f_0 \) is the Coriolis parameter, and the centripetal acceleration is neglected. We shall think of the fluid layer of constant density \( \rho \), which might represent an ocean basin, as occupying the region \( h_0(x,y) \leq z \leq h(x,y,t) \). The upper surface \( z = h(x,y,t) \) will be assumed to be a constant pressure \( p_0 \) (see Fig. 6). If \( H \) is a typical value of \( h \) and \( h_0 \), and \( L \) is a typical horizontal scale of changes in \( h \) in \( p \) and in the velocity components \( (u,v,w) \), the layer is considered shallow if

\[ \delta \equiv \frac{H}{L} \ll 1. \quad (56) \]

Since there is no preferred horizontal direction, we take \( u \) and \( v \) to be of possibly comparable size and of order \( U \).

It is convenient to introduce the horizontal velocity \( \mathbf{u}_h = (u,v,0) \) as well as the horizontal gradient \( \nabla_h = (\partial_x, \partial_y, 0) \) and write (55b) in the form

\[ \nabla_h \cdot \mathbf{u}_h + \frac{\partial w}{\partial z} = 0. \quad (57) \]

The first term on the left is then of order \( U/L \). If \( w \) varies by an amount of order \( W \) over the layer, the second term is of order \( W/H \). These two terms must be comparable in order to account for conservation of matter in the presence of a divergence of horizontal velocity. Thus we take \( W \) to be of order \( \delta U \). This small vertical component of velocity is consistent with small inclination of both surfaces in Fig. 6.

With this scaling, the material derivative is \( O(U^2/L) \) and contains no negligible terms, provided that the time scale is taken as \( L/U \). The horizontal pressure forces, which drive the horizontal motions, must be of order comparable to the acceleration. Taking the Coriolis and inertial accelerations to be also comparable to each other, so that the Rossby number \( U/2\Omega L \) is \( O(1) \), we have \( p \sim \rho U^2 \). Then the horizontal momentum balance is given by

\[ \frac{\partial \mathbf{u}_h}{\partial t} + (\mathbf{u}_h \cdot \nabla_h) \mathbf{u}_h + w \frac{\partial \mathbf{u}_h}{\partial z} + 2 \Omega \times \mathbf{u}_h = -\frac{1}{\rho} \nabla_h p. \quad (58) \]

On the other hand, the vertical momentum balance is

\[ \frac{1}{\rho} \frac{\partial}{\partial z} (p + \rho g z) = O\left(\delta U^2/L\right). \quad (59a) \]

yielding the hydrostatic approximation

\[ p = -\rho g z + P(x,y) + O\left(\delta^2 \rho U^2\right). \quad (59b) \]

For a shallow layer this simplifies therefore to
where $P(x, y, t)$ is an arbitrary function. The latter is fixed by the dynamic condition $p(x, y, h, t) = p_0 = \text{cst}$ on the free surface. Thus

$$p = \rho g (h - z) + p_0.$$  

(61)

We find that the pressure is also of order $\rho g H \sim \rho U^2$, yielding the characteristic horizontal velocity $U \sim \sqrt{gH}$.

According to (61), $\nabla_h p$ is independent of $z$. Hence we expect from (58) that $u_h$ would remain independent of $z$ if initially so. Assuming that this is the case, Eq. (58) simplifies to

$$\frac{D u_h}{Dt} + 2\Omega \times u_h + g\nabla_h h = 0,$$  

(62a)

where, for simplicity, the material derivative in a horizontal plane is denoted by

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u_h \cdot \nabla_h.$$  

(62b)

Equation (57) may be integrated to give

$$w \equiv \frac{Dz}{Dt} = -z\nabla_h \cdot u_h + w_0(x, y, t),$$  

(63)
involving a second arbitrary function. To determine it, one imposes the kinematic conditions that the upper and lower fluid boundaries be material surfaces. If a boundary is defined implicitly by \( S(x, y, z, t) = 0 \), the latter condition is that \( dS/dt = 0 \) on \( S = 0 \). This gives using (63) applied to \( z = h \) and \( z = h_0 \)

\[
0 = \frac{D}{Dt} (z - h)_{z=h} = -h \nabla_h \cdot u_h + w_0 - \frac{Dh}{Dt}, \tag{64a}
\]

\[
0 = \frac{D}{Dt} (z - h_0)_{z=h_0} = -h_0 \nabla_h \cdot u_h + w_0 - \frac{Dh_0}{Dt} \tag{64b}
\]

\[
= -h_0 \nabla_h \cdot u_h + w_0 - u_h \cdot \nabla_h h_0.
\]

Thus (64) determines \( w_0 \) and provides one equation to supplement (62a). If we substract one equation from the other, the supplemental equation is obtained in the form

\[
\frac{D}{Dt} (h - h_0) = - (h - h_0) \nabla_h \cdot u_h. \tag{65}
\]

Note that (65) has, in horizontal coordinates, the form of a continuity equation for a “density” \( h - h_0 \) (cf. Eq. (2b)). Physically, such an equation describes how a vertical fluid column changes in height in response to changes in horizontal divergence of the velocity field. Note, in this connection, that the absence of any vertical variation of \( u_h \) implies that vertical columns remain vertical.

The linear variation of \( w \) with \( z \) in fact implies that the fluid column is stretched uniformly and gives rise to an interesting material invariant of the shallow-water equations. Taking the exact material derivative of \( z - h_0 \) and using (63) and (64b), we obtain

\[
\frac{D}{Dt} (z - h_0) = w - u_h \cdot \nabla_h h_0 = -z (\nabla_h \cdot u_h) + w_0 - u_h \cdot \nabla_h h_0 \tag{66}
\]

\[
= -(z - h_0) \nabla_h \cdot u_h,
\]

we may thus combine (65) and (66) to obtain

\[
\frac{D}{Dt} \lambda = 0, \quad \lambda = \frac{z - h_0}{h - h_0}. \tag{67}
\]

Thus, a particle of fluid retains its position in a column as a given fraction of the height of the column.

We are now in a position to apply Ertel’s theorem (Section 4.2) to the material invariant \( \lambda \). To do this, we need the shallow-water approximation to the inertial or absolute vorticity \( \omega_0 \). Taking into account that \( u_h \) is independent of \( z \) and \( \left( \partial_z w, \partial_z w \right) = O(\delta U/H) \) we have

\[
\omega_0 = (\partial_y w - \partial_x v, \partial_x u - \partial_y w, f_0 + \zeta) = (0, 0, f_0 + \zeta) + O(\delta \zeta), \tag{68}
\]
where $\zeta = \partial_x v - \partial_y u$ is the relative vorticity of the flow. Actually, to justify (68) it must be shown that there is no term of order $\delta$ in the equation for $u_h$, which could contribute $O(1)$ terms through the $z$-derivatives. But this follows from the estimate (59), which indicates that the shallow-water equations determine the leading terms in an expansion in $\delta^2$. Thus, the contributions of $\partial_z u$ and $\partial_z v$ to (68) are in fact of order $\delta$ as stated.

Consequently, in the shallow-water approximation, Ertel’s theorem, cf. Eq. (49), yields a potential vorticity equation in the form

$$\frac{D}{Dt} \left( f_0 + \zeta \right) = 0,$$

(69a)

where $\nabla \lambda = (\nabla_h \lambda, \frac{1}{h-h_0})$ and the term $\nabla_h \lambda$ does not contribute to the scaled gradient $\omega_0 \nabla \lambda$ up to order $O(\delta \zeta)$.

We can interpret this equation quite easily in physical terms: since $h-h_0$ is inversely proportional to the horizontal cross-sectional area of a thin column, (69a) is in fact the shallow-water version of a local form of Kelvin’s theorem in a rotating frame.

One obtains some simple consequences of (69a) by expanding it as

$$\frac{1}{f_0 + \zeta} \frac{D \zeta}{Dt} = \frac{1}{h-h_0} \frac{D(h-h_0)}{Dt},$$

(69b)

where it is assumed that $f_0$ does not depend on the horizontal coordinates, i.e. $Df_0/Dt = 0$. Since planetary vorticity $f_0$ is positive, $f_0 > 0$ in the Northern Hemisphere, it follows from (69b) that relative vorticity $\zeta$ and layer depth $h-h_0$ increase and decrease together, at least when $\zeta$ is cyclonic, $\zeta > 0$. This monotonic dependence of $\zeta$ on $h-h_0$ will also hold for anticyclonic $\zeta$, provided $f_0 \gg |\zeta|$. The latter is typically the case for large-scale geophysical flows. The same conclusions apply in the Southern Hemisphere, where $f_0 < 0$, for cyclonic flows of arbitrary strength and for weak anticyclonic flows; notice that cyclonic flows in the Southern Hemisphere are characterized by $\zeta < 0$, cf. Fig. 1.3 and accompanying remarks.

Of course, we do not know what the changes in $\zeta$ are until we find $h(x,y,t)$. In Chapter 10 we shall see that, for small Rossby number, relations connecting $u_h$ (and therefore $\zeta$) to $h$ become quite simple. In this case, (69) becomes a nonlinear partial differential equation in $h$, which completely determines the shallow-water dynamics.

The shallow-water theory described by Eqs. (62a, 65) has a number of interesting features, both with and without the Coriolis force term. We study here some aspects of the relevant linear theory, and take up in the next chapter some aspects of the nonlinear equations. In other words, the next two sections will address the small-amplitude limit of shallow-water theory. Results about the theory’s limit for small Rossby number can be found in several of the books cited at the beginning of Chapter 1 and they are applied here in Chapter 3.
Remarks. It is possible to derive Eqs. (62a, 65) directly from (55), given the hydrostatic approximation (59b) and the fact that the horizontal velocities are independent of \( z \). It suffices to integrate (55) with respect to \( z \), from \( h_0 \) to \( h \). The vertically integrated advection term \( \int_{h_0}^{h} \omega \partial_z u_h \, dz \) is zero. To see this, one uses

\[
\int_{h_0}^{h} \omega \partial_z u_h \, dz = \left[ \omega u_h \right]_{h_0}^{h} - \int_{h_0}^{h} u_h \partial_z \omega \, dz = u_h \frac{D(h-h_0)}{Dt} + \int_{h_0}^{h} u_h \nabla_h \cdot u_h \, dz.
\]

Integrating the divergence-free equation (55b), one has \( w(z = h) - w(z = h_0) = - (h-h_0) \nabla_h \cdot u_h \), which yields directly Eq. (65). Plugging this expression into the expression above gives

\[
\int_{h_0}^{h} \omega \partial_z u_h \, dz = -u_h (h-h_0) \nabla_h \cdot u_h + (h-h_0) u_h \nabla_h \cdot u_h = 0.
\]

Hence, the horizontal pressure terms, when integrated over the depth of the layer and using (61), yield \( \int_{h_0}^{h} (1/\rho) \nabla_h p \, dz = g(h-h_0) \nabla h \). Therefore, one obtains (62a) multiplied everywhere by the factor \( h-h_0 \).

Equation (69a) can also be obtained directly from (62a) and (65) by taking the curl to yield

\[
\frac{D \zeta}{D t} + f_0 \nabla_h \cdot u_h = \frac{D \zeta}{D t} - f_0 \frac{1}{h-h_0} \frac{D(h-h_0)}{D t} = (h-h_0) \frac{D}{D t} \left( \frac{f_0 + \zeta}{h-h_0} \right) = 0.
\]

7 Reduced-gravity shallow-water models

In this section, we would like to obtain the so-called \((N+1/2)\)-layer shallow-water model that extends the result above in the direction of a fully 3-D model. The terminology involving a "1/2-layer" is shorthand for the assumption of a deep, motionless layer above which the active flow can be approximated by \( N \) layers with nonzero velocities. In this situation, the effective gravity felt in the active layers is smaller than \( g \), and it is referred to as reduced gravity. As we shall see at the end of this section, the atmospheric case with \( N+1 \) fully active layers is quite similar, except for the need of using a reduced gravity.

More precisely, in the \((N+1/2)\)-layer setting, the oceans are modeled by \( N \) active layers of finite depth and constant density \( \rho_n \) and another quiescent layer of infinite depth and of larger density \( \rho_{N+1} \). Roughly speaking, the active layers model the ocean above the thermocline and the last layer represents the fluid below it.

The term "reduced gravity" is used because the interfaces below the surface are in slow motion due to the fact that the force acting on them is essentially the gravity field multiplied by a factor that is proportional to density differences between the layers. This factor is in general small, so that the gravity is reduced by an order of
magnitude. A simple experiment with a layer of oil above one of water in a large cup can help illustrate this concept: the interface between the two layers, when moving the cup a bit initially, is slowly damped compared to the fast gravity waves at the surface.

In the shallow-water context, the main hypothesis is the validity of the hydrostatic approximation, with constant pressure $P_0$ at the surface. Since the horizontal gradient of the pressure becomes independent of $z$, the velocities are also independent of $z$ at all times, if they are initially so; see Sec. 6. We assume that the depth associated with each layer is $h_k > 0$, and the vertical coordinate $z$ is taken to be equal to $h_0$ at the surface, cf. Fig. 7.

Note that the derivatives in the vertical direction must be kept although it is assumed that the velocities $u_k = (u_k, v_k)$ within each layer $k$ are independent of $z$. This is due to the fact that the complete velocity field $(u_1, \cdots, u_N)$ obviously depends on $z$, although it does so in a discontinuous way, from layer to layer. In particular, when interpreting the derivatives in discretized form, terms like $\frac{\partial u_k}{\partial z}$ at the interface between layer $k$ and $k+1$, say, would be proportional to the velocity difference in a first-order linear approximation:

$$A_v \frac{\partial u_k}{\partial z} \bigg|_{z=h_0-h_1-\cdots-h_k} \approx A_v \frac{u_k - u_{k+1}}{h_k + h_{k+1}}. \quad (70)$$

One must be aware that estimating these derivatives is more a matter of parametrization of subgrid-scale processes, and it is difficult to assess what they should be in reality. For oceanic applications, these vertical derivatives are either neglected, except at the surface, or take the form of Rayleigh damping between the layers, in view of (70).

We thus start from $N$ 2-D coupled Navier-Stokes equations in dimensional form, that is

$$\frac{\partial u_k}{\partial t} + (u_k \cdot \nabla) u_k + w_k \frac{\partial u_k}{\partial z} + 2f \times u_k = -\frac{1}{\rho_k} \nabla p_k + A_H \Delta u_k + A_v \frac{\partial^2 u_k}{\partial z^2},$$

$$\nabla \cdot u_k + \frac{\partial w_k}{\partial z} = 0. \quad (71)$$

Using the hydrostatic hypothesis, and the continuity of the pressure at each interface, one obtains the pressure within each layer:

$$p_k(z) = P_0 + \sum_{j=1}^{k-1} g \rho_j h_j + g \rho_k \left( h_0 - z - \sum_{j=1}^{k-1} h_j \right), \quad 1 \leq k \leq N + 1. \quad (72)$$

In order to proceed, we reformulate the nonlinear terms in (73) in a more convenient conservative form using the divergence-free equation, to obtain
Fig. 7 Reduced-gravity shallow-water model with $N$ active layers of density $\rho_k$ and thickness $h_k > 0$, while the bottom layer $N + 1$ is of infinite depth and at rest. The surface deviation is $z = h_0$ and serves as the reference for the vertical coordinate. The pressure $P_0$ at the surface is constant.

\[
\begin{align*}
\frac{\partial u_k}{\partial t} + \nabla \cdot (u_k u_k) + \frac{\partial (u_k w_k)}{\partial z} - 2f v_k &= -\frac{1}{\rho_k} \frac{\partial p_k}{\partial x} + A_H \Delta u_k + A_v \frac{\partial^2 u_k}{\partial z^2}, \\
\frac{\partial v_k}{\partial t} + \nabla \cdot (v_k u_k) + \frac{\partial (v_k w_k)}{\partial z} + 2f u_k &= -\frac{1}{\rho_k} \frac{\partial p_k}{\partial y} + A_H \Delta v_k + A_v \frac{\partial^2 v_k}{\partial z^2}, \\
\nabla \cdot u_k + \frac{\partial w_k}{\partial z} &= 0.
\end{align*}
\] (73)

In the following, we will extensively use the formula

\[
\frac{d}{dx} \int_{a(x)}^{b(x)} f(u,x) \, du = \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}(u,x) \, du + \frac{db}{dx} f(b(x),x) - \frac{da}{dx} f(a(x),x).
\]

We now integrate (73) over the layer depth, that is between $z^- = h_0 - \sum_{j=1}^{k} h_j$ and $z^+ = z^- + h_k$, where $z^-$ and $z^+$ of course depends on $(x,y,t)$. We proceed with each term and, again assuming that $u_k$ does not depend on $z$, obtain:
implies that there are no pressure variations, that is, the relation is readily provided by the motionless horizontal stresses acting on the top and bottom of the layer. The terms on the right-hand side of the last equation above correspond to the horizontal discontinuities of the horizontal pressure force over the layer thicknesses. The requisite relation at the surface is the wind stress, see Fig. 15 and (101) in Chapter 3.

To complete the derivation, one integrates the vertical dissipative terms:

\[ \int_{z^-}^{z^+} \frac{\partial (u_k w_k)}{\partial z} \, dz = u_k (w_k(z^+) - w_k(z^-)) = u_k \left( \frac{Dz^+}{Dt} - \frac{Dz^-}{Dt} \right) = u_k \frac{Dh_k}{Dt}, \]

using the relation (72). The terms on the right-hand side of the last equation above correspond to the horizontal stresses acting on the top and bottom of the layer \( k \). In particular, the stress at the surface is the wind stress, see Fig. 15 and (101) in Chapter 3.

In order to conclude, one must relate \( h_0 \) to the layer thicknesses. The requisite relation is readily provided by the motionless \( N + 1 \)-st layer. Absence of motion implies that there are no pressure variations, that is

\[ \nabla p_{N+1} = 0. \quad (74) \]

In the following sections, we will use the notation

\[ g_j = g \frac{\rho_j - \rho_i}{\rho_j}. \quad (75) \]

We thus find, using (72) and (74), that

\[ g \nabla h_0 = \sum_{j=1}^{N} g_j^{i} \nabla h_j, \quad (76) \]
We are now ready to write the reduced-gravity shallow-water model with $N$ active and one passive layer, using the mass flux vector notation $U_k := u_k h_k$ and the fact that $\frac{Dh_k}{Dt} = \frac{dh_k}{dt} + (u_k \cdot \nabla)h_k$.

\[
\frac{\partial U_k}{\partial t} + \nabla \cdot (U_k \otimes u_k) + f \times u_k = -h_k \mathcal{G} \cdot \nabla \mathcal{H} + h_k (\mathcal{R} \nabla \mathcal{H})_k + \mathcal{D}_H + \mathcal{T}_k
\]
\[
\frac{\partial h_k}{\partial t} + \nabla \cdot U_k = 0,
\]

Here $\nabla \cdot (U_k \otimes u_k)$ is a convenient notation for $(\nabla \cdot (U_k u_k)), \nabla \cdot (V_k u_k))$.

\[
\mathcal{R} = \begin{pmatrix}
0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
g_2^2 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
g_3^2 & g_2^2 & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
g_k^2 & \cdots & g_k^{k-2} & g_k^{k-1} & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
g_N^2 & \cdots & \cdots & \cdots & \cdots & g_N^{N-1} & 0
\end{pmatrix}
\]

and the surface deviation in the densities that corresponds to $h_0$ in vector form is $\mathcal{G} = (g_{N+1}^1, \cdots, g_{N+1}^N)$. The interfacial stress $\mathcal{T}_k$ for $2 \leq k \leq N - 1$ is often assumed to be zero to retain only the surface and bottom stress. Since we assume the velocity vanishes in the $N + 1$-st layer, one expects to get a boundary layer of Ekman type.

Herewith we just obtained the shallow-water equations in a conservative mass-flux form that is quite convenient for numerical discretization. There are two other ways to express (77), one is the momentum form and the other is the vorticity form.

In momentum form, the equations are indeed straightforward to obtain since it suffices to divide by $h_k$ and use the mass-conservation equation:

\[
\frac{\partial u_k}{\partial t} + (u_k \cdot \nabla)u_k + f \times u_k = -\mathcal{G} \cdot \nabla \mathcal{H} + (\mathcal{R} \nabla \mathcal{H})_k + \frac{\mathcal{D}_H}{h_k} + \frac{\mathcal{T}_k}{h_k}
\]
\[
\frac{\partial h_k}{\partial t} + \nabla \cdot U_k = 0.
\]

The vorticity formulation is found by cross-differentiating (79) and defining the vorticities, layer-by-layer, as $\zeta_k = \frac{\partial v_k}{\partial x} - \frac{\partial u_k}{\partial y}$. Mass conservation implies that $\nabla \cdot u_k = -\frac{1}{h_k} \frac{Dh_k}{Dt}$ and the cross-differentiation of the left-hand side of (79) yields the expression $\frac{Dh_k}{Dt} (\zeta_k + f) + (f + \zeta_k) \nabla \cdot u_k$. Thus, the final result is

\[
h_k \frac{Dh_k}{Dt} \left( \frac{\zeta_k + f}{h_k} \right) = \nabla \times \frac{\mathcal{D}_H}{h_k} + \nabla \times \frac{\mathcal{T}_k}{h_k}.
\]
Equation (80) neatly generalizes the expression (69).

Note that the potential vorticity \( \frac{\zeta_k + f}{h_k} \) in each layer is a conservative quantity in the absence of lateral friction and vertical stresses. Of course, one needs to relate \( h_k \) with \( \zeta_k \) to obtain a complete description of the flow. It therefore amounts to solve the shallow-water equations in their primitive form, that is (77) or (79). In the quasi-geostrophic approximation, one can nevertheless find a closed equation for the vorticity–streamfunction variables. Again, we refer, for instance, to Chapter 3 in [60] for the derivation and to Chapter 5 for an application to the wind-driven ocean circulation.

Remark. The case of a genuine \( N+1 \)-layer model is slightly different. In such a model, the relation (72) for the pressure in each layer is the same but instead of having (76) due to \( \nabla p_{N+1} = 0 \), one has \( g \nabla h_0 = g \nabla h_1 + \cdots + g \nabla h_{N+1} \), which involves an additional layer \( N+1 \).

8 Small-Amplitude Motions in a Basin

From now on, we will omit the subscript ‘h’, namely \( \nabla \) and \( u \) are understood as \( \nabla_h \) and \( u_h \). Suppose that the fluid lies in a basin occupying a closed set \( \mathcal{D} \) in the \((x, y)\)-plane, cf. Fig. 8, and that when at rest, i.e. with \( h = \text{const.} \), the depth of the fluid is positive, \( H(x, y) > 0 \), over \( \mathcal{D} \). We then set

\[
h - h_0 = H(x, y) + \eta(x, y, t),
\]

where \( \eta \) and hence \( u = (u, v) \) are small perturbations. Since nonlinear terms in the perturbations are neglected, (62a) and (65) reduce to

\[
\frac{\partial u}{\partial t} - f_0 v + g \frac{\partial \eta}{\partial x} = 0,
\]

\[
\frac{\partial v}{\partial t} + f_0 u + g \frac{\partial \eta}{\partial y} = 0,
\]

\[
\frac{\partial \eta}{\partial t} + \nabla \cdot (H u) = 0.
\]

When \( f_0 = 0 \), we obtain the equations for small-amplitude oscillations in a non-rotating basin. In this case (82) can be combined to give

\[
\frac{\partial^2 \eta}{\partial t^2} - g \nabla \cdot (H \nabla \eta) = 0
\]

If the contour \( \partial \mathcal{D} \) is an impenetrable barrier, the appropriate boundary condition for system (82) is \( u \cdot n = 0 \), where \( n \) is the normal to \( \partial \mathcal{D} \). In this case, (82a,82b) with \( f_0 = 0 \) yields

\[
\nabla \eta \cdot n = 0
\]

as the corresponding boundary condition for (83a).
Fig. 8 A shallow-water basin \( \mathcal{D} \), with layer thickness \( h - h_0 = H(x,y) + \eta(x,y,t) \). The deviation \( \eta \) is assumed to be small compared with \( H \).

If we look for eigenfunctions \( \eta(x,y,t) = e^{i\omega t}N(x,y) \) of (83), where \( N \) has zero integral over \( \mathcal{D} \), then the resulting equation for \( N \) is equivalent to the Euler-Lagrange equation for the following variational problem:

\[
\min_N \int_{\mathcal{D}} H|\nabla N|^2 dxdy, \quad \int_{\mathcal{D}} N^2 dxdy = 1.
\]

Here the minimum is to be taken over all \( N \) which satisfy \( \partial N / \partial n = 0 \) on the boundary and have zero integral over \( \mathcal{D} \). The Lagrange multiplier in the problem may then be identified with \( \omega^2 / g \). In this way, the minimum frequency \( \omega_0 \) is determined variationally by

\[
\omega_0^2 = g \min_N \frac{\int_{\mathcal{D}} H|\nabla N|^2 dxdy}{\int_{\mathcal{D}} N^2 dxdy}.
\]

We note in passing that, for given \( H(x,y) \), \( \omega_0 \) is strictly positive and the Rayleigh-Ritz procedure provides an effective means for constructing approximate eigenfunctions. These eigenfunctions represent standing gravity waves in a shallow basin.

The relative simplicity of this theory is a consequence of the self-adjoint property of the boundary-value problem (83). Writing (83a) as

\[
\frac{\partial^2 \eta}{\partial t^2} = L\eta,
\]

it is easily seen that \( L \) is formally self-adjoint,

\[
\int_{\mathcal{D}} \phi L\psi dxdy = \int_{\mathcal{D}} \psi L\phi dxdy,
\]

for all functions \( \phi, \psi \) satisfying the boundary condition (83b).
We turn now to the effect of rotation on small-amplitude motions. It can be shown from (82) that, if $f_0 \neq 0$, the associated linear operator is no longer self-adjoint. To exhibit this operator, it is convenient to introduce the quantities

$$
\sigma = \nabla \cdot (Hu), \quad s = \nabla \times (Hu)
$$

which may be thought of as representing the divergence and vorticity of the integrated mass flux across the layer. From (82a,82b), we obtain by differentiation

$$
\frac{\partial \sigma}{\partial t} - f_0 s + g \nabla \cdot (H \nabla \eta) = 0, \quad (85a)
$$

$$
\frac{\partial s}{\partial t} + f_0 \sigma + g H \nabla \times \nabla \eta = 0 \quad (85b)
$$

Eliminating $s$ and using $\sigma = -\frac{\partial \eta}{\partial t}$ from (84) and (82c), the equation for $\eta$ becomes

$$
\frac{\partial}{\partial t} \left[ \frac{\partial^2 \eta}{\partial t^2} + f_0^2 \eta - g \nabla \cdot (H \nabla \eta) \right] = f_0 g H \nabla \times \nabla \eta \quad (86)
$$

We consider the simplest case, namely $H = \text{constant}$, so that the last two terms of the above equation vanish. One substitutes $\eta = e^{i(\omega t + k \cdot x)}$ into (86) to obtain a dispersion equation for traveling plane waves in an unbounded domain:

$$
i \omega [f_0^2 - \omega^2] = -i \omega g H |k|^2, \quad (87)
$$

where the wavenumber $k$ has two components $(k_x, k_y)$ and $|k|^2 = k_x^2 + k_y^2$. Disregarding for the moment the wave corresponding to $\omega = 0$, these waves are known as Poincaré or inertia-gravity waves, and represent the generalization to the rotating case of familiar gravity waves. For every angular frequency $\omega \neq 0$, there are two waves with with phase speed along the direction $k$

$$
|c_k| = -\omega / |k| = \pm \left( \frac{f_0^2}{|k|^2} + g H \right)^{1/2}. \quad (88)
$$

One propagates in the positive, the other in the negative direction relative to the wave vector $k$. They are observed in the atmosphere as well as in large bodies of water.

Seen from above, particle paths in the nonrotating case appear as line segments parallel to the direction of propagation of the wave; see panel (a) in Fig. 9. With rotation, cf. panel (b), the segments become ellipses whose eccentricity decreases with rotation. These facts can be most easily derived from a Lagrangian description.

For Poincaré waves in a finite basin, one finds that, in certain simple basin shapes, the boundary condition selects a discrete set of eigenfunctions. Each of the latter exhibits a dependence of frequency upon a discrete ordering parameter that is similar to (87).

Using (87), we write the dispersion relation for inertia-gravity waves in the form
\[
\left( \frac{\omega}{f_0} \right)^2 = 1 + L_R^2 |k|^2 ,
\]
where
\[
L_R = \frac{\sqrt{gH}}{f_0} \quad \text{(90)}
\]
has the dimensions of a length, and is called the external Rossby radius of deformation. The physical importance of this length will become clearer in studying the near-surface ocean circulation in Ch. 3.

In the present context, \(2\pi L_R\) is a horizontal scale below which the Poincaré waves become essentially gravity waves with small Coriolis effect. To take a typical case for the Earth’s atmosphere, let \(H = 10\) km, giving \(L_R \approx 3 \times 10^3\) km. Thus, for horizontal scales of \(2\pi/k < 3 \times 10^3\) km, we note that the gravity contribution to (89) dominates.

Notice that the frequency of these waves increases with wavenumber \(k\). It is always larger than \(f_0\), i.e., their period is smaller than the period of planetary rotation. A typical period of the inertia-gravity waves at length scales \(L_R = O(10^3\) km\) is about 2 hours. For most geophysical problems on such horizontal scales, this determines a short time scale. An important feature, therefore, of models for large-scale atmospheric and oceanic dynamics is an appropriate mechanism for filtering out the possible Poincaré-wave components of small period. Inertia-gravity waves are observed to have, in fact, small amplitudes over these length scales.
8.1 Kelvin waves

Another possible component of oscillations of a basin is the family of Kelvin waves. These waves can be generated near a bounding contour and oscillate at a frequency which can be close to that associated with gravity waves. To take the simplest case, suppose that the basin is in the upper half-plane \( y > 0 \).

Eliminating \( u \) from (82a, 82b) we obtain

\[
\frac{\partial^2 v}{\partial t^2} + f_0^2 v + g \frac{\partial^2 \eta}{\partial y \partial t} - f_0 g \frac{\partial \eta}{\partial x} = 0.
\]  

(91)

if \( v \) vanishes on the boundary we must also have

\[
\frac{\partial^2 \eta}{\partial y \partial t} - f_0 \frac{\partial \eta}{\partial x} = 0, \quad y = 0.
\]  

(92)

With \( \eta = N(y)e^{i(\omega t + k_x x)} \), \( k_x > 0 \), Eq. (86) with \( H = \text{cst} \) yields

\[
gH \frac{d^2 N}{dy^2} - (f_0^2 - \omega^2 + gH k_x^2) N = 0, \quad y > 0
\]

(93a)

\[
\omega \frac{dN}{dy} - k_x f_0 N = 0, \quad y = 0.
\]

(93b)

A new result can be obtained if we assume \( f_0^2 - \omega^2 + gH k_x^2 > 0 \), since then the \( y \)-dependence involves real exponentials rather than real trigonometric functions. To obtain a solution bounded in the upper half plane we must select the decaying exponential. Then (93b) requires \( N = N_0 e^{i k_x x / \omega y} \), \( \omega < 0 \), in which case (93a) gives

\[
(\omega^2 - gH k_x^2) (f_0^2 - 1) = 0.
\]

(94)

Taking \( \omega = -k_x \sqrt{gH} \), one obtains a wave with positive phase velocity identical to that of a pure gravity wave.

It can be checked from (82) that \( v \) vanishes identically for this wave, and that wave crests are parallel to the \( y \)-axis. On each line \( y = \text{cst} \) the wave propagates as an ordinary gravity wave, but the \( x \)-motion is in geostrophic balance with the wave height \( \eta \); the latter decreases in concert with \( u \) as \( y \) increases, at an exponential rate determined by the Rossby radius \( L_R \).

Because of their exponential decrease away from the boundary, Kelvin waves can be regarded as “edge waves”. They have vanishing frequency as \( k_x \to 0 \), a property which separates them from inertial-gravity waves; see Fig. 12 in Sec. below. But for moderate length scales in the \( x \)-direction they fall into the category of “fast” waves. Kelvin waves are particularly important in tropical meteorology and oceanography, where the equator plays the role of the “wall”.

9 Geostrophic Degeneracy and Rossby waves

The time scales associated with large-scale atmospheric phenomena are of the order of days, while those for the ocean are of the order of months. If these time scales are to be reflected in the small-amplitude solutions of the shallow-water equations, the relevant physics must be associated with the modes corresponding to the root $\omega = 0$ of the dispersion relation \[ (87) \] in the constant-depth case.

9.1 Stationary geostrophic modes

Let us consider then the stationary modes in the small-amplitude theory. Eqs. \([82]\), reduce for steady solutions to

\[
\begin{align*}
2\Omega \times u_h + g\nabla_h \eta &= 0 \quad \text{(95a)} \\
\nabla_h \cdot (H u_h) &= 0 \quad \text{(95b)}
\end{align*}
\]

This is a special kind of geostrophic flow (Section 1), and we refer to all solutions of \([95]\) compatible with the condition on the boundary of the basin as stationary geostrophic modes. Eq. \((95a)\) implies that \((g/f_0) \eta\) is a stream function for the flow’s velocity field and that lines \(\eta = \text{cst}\) are streamlines of the flow. In particular, the boundary has to be an isoline of \(\eta\).

For the constant-depth problem we see that arbitrary divergence-free horizontal motions compatible with the condition that the boundary be a streamline are allowed. Eqs. \((95)\) simply tells us what the fluid level must be to bring the system into geostrophic balance. From \((86)\) and \((95)\) we conclude that in the constant-depth case, with \(\eta\) proportional to \(e^{i\omega t}\) in separated variables, the eigenvalue \(\omega = 0\) is infinitely degenerate. Given any boundary contour, there are an infinite number of functions \(\eta\) which satisfy the boundary conditions.

If \(H\) depends upon \(x\) and \(y\), the geostrophic modes may still be infinitely degenerate, since \((95b)\) then implies that the streamfunction \(\eta\) may be an arbitrary function of \(H\) alone. Thus lines of constant depth and streamlines coincide. Although contours of constant \(H\) that intersect the boundary \(\partial \mathcal{D}\) must carry zero velocity, closed contours that do not touch the boundary — or nested islands of such contours — do allow the construction of geostrophic modes, as shown in Fig. 10. The latter type of contours, along which stationary geostrophic flow may occur, are often called geostrophic contours.
9.2 Sloping bottom topography

The question now arises as to what happens to the "disappearing" geostrophic modes when the depth function is perturbed, in such a way that the geostrophic degeneracy is partially or perhaps completely removed. To take a particularly simple case, consider a large basin whose depth function is locally linear

\[ H = H_0 (1 - \beta y), \quad \beta > 0. \] (96)

It is helpful to consider the parameter \( \beta \) as small, \( \beta = O(\varepsilon) \), to retain (96) over a sizeable region, and to disregard boundary effects.

It is customary in dealing with large-scale flow of the atmosphere and oceans to choose the \( x \)-coordinate pointing eastward and the \( y \)-coordinate pointing northward. Such a geometry is able to simulate the important effect of latitudinal variation in the effective local rate of rotation, as shown, for instance by Weeks and colleagues [169, and references therein].

With (96), the geostrophic contours are the lines \( y = \text{const.} \), so the degeneracy of the constant-depth case has been significantly altered. To force recovery of the missing modes in (82) as a perturbation, we suppress completely the remaining geostrophic modes by requiring that \( \eta \) be of the form

\[ \eta = \tilde{\eta} e^{i(\omega t + k_x x)}, \] (97a)

with similar expressions for \( u \) and \( v \), and with small \( \omega / f_0 \).
The smallness of $\omega$ guarantees slow changes in all the variables; in particular, we shall have small $\partial u / \partial t$ and $\partial \eta / \partial t$. Equations (82a, 82c) then imply an approximate geostrophic balance. Since $\partial \eta / \partial y = 0$, (82b) also states that $u$ itself is small. The solutions we consider have therefore predominantly $y$-directed fluid velocity. That is, particle motion of the slowly-varying flow is essentially perpendicular to the geostrophic contours of the stationary flow discussed before, while the slow wave propagation is parallel to these contours. Moreover, the free surface $\eta$ and velocity $(u, v)$ of the flow (97a) are still in a nearly geostrophic, slowly shifting balance.

Substituting $\eta$ from (97a) and $H$ from (96) into (86) yields
\[
i \omega \left( f_0^2 - \omega^2 \right) = -gH_0 i \omega (1 - \beta y) k_x^2 + igH_0 \beta f_0 g k_x, \tag{97b}
\]
which involves $y$ explicitly in the first term on the right. This appears to be inconsistent with the assumed form of $\eta$. However, we are dealing only with small $\omega$, and one can thus neglect the term $\beta y$ in order to obtain a dispersion relation that is still correct to $O(\varepsilon)$. Rearranging terms, the latter takes the form
\[
f_0^2 + gH_0 k_x^2 - \omega^2 = \frac{gH_0 \beta f_0 g k_x}{\omega}. \tag{98}
\]

The solutions of (98) are plotted in Fig. 11, which shows that, for small $\beta$, there are two “perturbed” inertia-gravity modes, together with a third root of order $\beta$. As a matter of fact, letting the left-hand side (LHS) vanish gives directly (89). Since
this third root is small, an approximate dispersion relation for it is

$$\omega = \frac{gH_0f_0k_x}{f_0^2 + gH_0k_x^2} \beta.$$  \hfill (99)

The waves associated with (99) are called Rossby waves and, in the present case of positive $\beta$, their phase velocity $c = -\omega/k_x$ is westward. The frequency of these waves is always smaller than $f_0$, provided $\beta$ is not too large. Rossby waves are further distinguished from both Poincaré and Kelvin waves by the fact that, for short waves, $\omega$ decreases with $k_x$, as seen in Fig. 12.

![Fig. 12 Comparison of dispersion relations for small-amplitude shallow-water waves with rotation, namely $\omega_{\text{poincare}} = \pm f_0(1 + x^2)^{1/2}$ and $\omega_{\text{kelvin}} = \pm f_0x$, and for Rossby waves, $\omega_{\text{rossby}} = \beta(gH)^{1/2}x/(1 + x^2)$, where $x = kL_R$. The ordinate corresponds to the absolute value $|\omega|$ of the frequency and the abscissa is in units of the Rossby radius $L_R$, cf. (90).](image)

We might think of these waves as replacing, among the infinity of geostrophic modes in a basin of constant depth, the family of modes with $y =$ const. as streamlines. In fact, to $O(\varepsilon)$, the flow is still nondivergent and in geostrophic balance,

$$u = -\frac{\partial \psi}{\partial y}, \quad v = \frac{\partial \psi}{\partial x}, \hfill (100a)$$

where $\psi = (g/f_0)\eta$. This balance also implies that, to $O(\varepsilon)$, the relative vorticity in the wave, cf. (68), is given by

$$\zeta = \nabla^2 \psi. \hfill (100b)$$
But it is precisely the small deviation from exact geostrophy — apparent to $O(1)$ only in the vorticity form of the equations, cf. (86) — that eliminates the geostrophic degeneracy and gives rise to the waves.

To understand the physical structure of Rossby waves, it is useful to consider their propagation in terms of vorticity balance. According to (69), vorticity and depth along a material trajectory increase and decrease together. Examining the structure of the Rossby wave (97a) at any given time, we observe that $\zeta$ and $v$ have the appropriate variations shown in Fig. 13, namely if $\zeta(x) = \sin(kx)$ then $v(x) = kx \cos(kx)$. Since $v$ advects $\zeta$, wherever $v$ is positive the relative vorticity at that point must be instantaneously decreasing with time, since advection would carry the fluid element from a deeper region into a shallower region ($\beta > 0$). In order for this to happen, the wave pattern must drift westward.

**Fig. 13** Linear Rossby waves must propagate westward when $H = H_0(1 - \beta y)$, with $\beta > 0$, in order to conserve potential vorticity according to (97a). Thus, an initial vorticity wave given by (97a), plotted in the figure as the heavy blue undulating contour, has its vorticity decreasing — as given by the vertical red arrow that indicates the vorticity contribution $-\zeta$ in this case — if the particle moves northward. To the contrary, southward particle motion makes a positive contribution $+\zeta$ to the wave’s vorticity.
Chapter 3

The Wind-driven Ocean Circulation

10 Introduction and motivation

Having introduced the reader to some of the fundamentals of planetary-scale flows in Chapters 1 and 2, we are ready to turn to current topics of the climate sciences and of the applications of dynamical systems theory to them.

Since the early 1970s, the importance of the climate sciences has been rapidly growing, as a field of scholarship as well as a source of information for stewardship of planet Earth [116, 73, 74, 75, 76]. Likewise, the usefulness of dynamical systems and of their successive bifurcations in navigating across the hierarchy of climate models [37, 55, 56, 60, 136] — while far from universally accepted — has found a wide community of users and produced a substantial literature.

An important recent development in this literature was caused by the realization that the theory of differentiable dynamical systems (DDS) — in its classical form dealing with autonomous DDSs — was fairly well adapted to applications in which neither the forcing nor the coefficients were time dependent. To a good approximation, this could be said to be the case in numerical weather prediction [81, 159], but much less so in climate problems in which the seasonal forcing [20, 63, 68, 79, 80, 135, 163, 162] or anthropogenic effects [18, 116, 73, 74, 75] play a major role.

To cope with the latter type of problems, applications of the theory of nonautonomous and random dynamical systems (NDS and RDS) have started to appear in the theoretical climate dynamics literature [59, 27, 10, 43]. The purpose of the present paper is to give some flavor of the systematic application of DDS, NDS and RDS theory across the hierarchy of climate models, from the simplest, conceptual ones all the way to the most elaborate ones [39, 55, 56, 63].

In mathematical terms, this means considering, at the lower end of the hierarchy, small systems of ordinary differential equations (ODEs) and at the higher end full-blown systems of nonlinear partial differential equations (PDEs). In the climate literature, the former are sometimes called “toy models,” while the latter are re-
ferred to as general circulation models (GCMs) or, more recently, as global climate models, which preserves the GCM acronym.

An interesting problem to which these ideas and results will be applied herein is that of the wind-driven circulation in mid-latitude ocean basins [37, 39], subject at first to time-constant and then to purely periodic and more general forms of time-dependent wind stress. The wind-driven circulation is obviously strongest near the surface of the oceans, and its variability lies mostly in the subannual and interannual range. In many simplified, theoretical treatments of the oceans, this circulation is treated separately from the so-called buoyancy-driven, thermohaline or overturning circulation. The latter is driven by mass fluxes at the ocean–atmosphere interface; it involves the density distribution of water masses — which depends on both temperature and salinity, hence its characterization as thermohaline — and it reaches all the way to the bottom of the oceans. The variability of this overturning circulation lies largely in the interdecadal and centennial range [156, 152, 54, 117, 121, 37, 39, 56]. It will only be mentioned here in passing in Section 13.

Key features of the wind-driven circulation are described in Section 11 below. The influence of strong thermal fronts — like the Gulf Stream in the North Atlantic or the Kuroshio in the North Pacific — on the mid-latitude atmosphere above is severely underestimated, thus contributing to the current range of uncertainties in climate simulations over the 21st century. Typical spatial resolutions in recent century-scale GCM simulations [1, 73, 74, 75, 128, 46] are still of the order of 100 km at best, whereas resolutions of 50 km and less would be needed to really capture the strong mid-latitude ocean-atmosphere coupling just above the oceanic fronts [51, 52, 110, 148].

An important additional source of uncertainty comes from the difficulty to correctly parametrize global and regional effects of clouds and their highly complex small-scale physics. This difficulty is particularly critical in the tropics, where large-scale features such as the El-Niño/Southern Oscillation (ENSO) and the Madden-Julian Oscillation are strongly coupled with convective phenomena [103, 114, 63].

The outline of this chapter is the following. First, we describe in Section 11 the most recent theoretical results regarding the internal variability of the mid-latitude wind-driven circulation, viewed as a problem in nonlinear fluid mechanics. These results rely to a large extent on autonomous DDS theory [5, 70]. Next, we summarize in Section 12 the key concepts and methods of NDS [137] and RDS [4] theory.

With the results of Sections 11 and 12 in hand, we address in Section 13 the changes in this purely oceanic variability as a result of time-dependent forcing, on the one hand, and truly coupled ocean–atmosphere variability, on the other. Much of the material in this section is new [50, 125, 167]. A summary and an outlook on future work follow in Section 14. Rigorous mathematical results appear in Appendices A, B and C, which follow closely [59] (Appendix A) and [27] (Appendices B and C).
11 Internal variability of the wind-driven ocean circulation

11.1 Observations

To a first approximation, the main near-surface currents in the oceans are driven by the mean effect of the winds. The trade winds near the equator blow mainly from east to west and are also called the tropical easterlies. In mid-latitudes, the dominant winds are the prevailing westerlies, and towards the poles the winds are easterly again. Thus an idealized version of wind effects on the ocean must involve zonally oriented winds that blow westward at low and high latitudes, and eastward in mid-latitudes.

Five of the strongest near-surface, mid-and-high-latitude currents are the Antarctic Circumpolar Current, the Gulf Stream and the Labrador Current in the North Atlantic, and the Kuroshio and Oyashio off Japan. These currents are all clearly visible in Fig. 14.

![Map of Ocean Currents](image)

Fig. 14 A map of the main oceanic currents: warm currents in red and cold ones in blue. Reproduced from [59], with permission from Elsevier.

The Gulf Stream [153] is an oceanic jet with a strong influence on the climate of eastern North America and of western Europe. From Mexico’s Yucatan Peninsula, the Gulf Stream flows north through the Florida Straits and along the East Coast of the United States. Near Cape Hatteras, it detaches from the coast and begins to drift off into the North Atlantic towards the Grand Banks near Newfoundland. Actually, the Gulf Stream is part of a larger, gyre-like current system, which includes the North Atlantic Drift, the Canary Current and the North Equatorial Current. It is also coupled with the pole-to-pole overturning circulation.

The Coriolis force is responsible for the so-called Ekman transport, which deflects water masses orthogonally to the near-surface wind direction and to the right
In the North Atlantic, this Ekman transport creates a divergence and a convergence of near-surface water masses, respectively, resulting in the formation of two oceanic gyres: a smaller, cyclonic one in subpolar latitudes, the other larger and anticyclonic in the subtropics. This type of *double-gyre* circulation characterizes all mid-latitude ocean basins, including the South Atlantic, as well as the North and South Pacific.

The double-gyre circulation is intensified as the currents approach the East Coast of North America due to the $\beta$-effect. This effect arises primarily from the variation of the Coriolis force with latitude, while the oceans’ bottom topography also contributes to it. The former, planetary $\beta$-effect is of crucial importance in geophysical flows and induces free Rossby waves propagating westward.

The currents along the western shores of the North Atlantic and of the other mid-latitude ocean basins exhibit boundary-layer characteristics and are commonly called western boundary currents. The northward-flowing Gulf Stream and the southward-flowing Labrador Current extension meet near Cape Hatteras and give rise to a strong, mainly eastward jet that eventually reaches Scandinavia. In all four mid-latitude ocean basins — i.e., the North and South Atlantic, as well as the North and South Pacific — the confluence of a warm, poleward current with a cold, equatorward one produces an intense eastward jet. The formation of this jet and of the associated recirculation vortices near the western boundary, to either side of the jet, is mostly driven by internal, nonlinear effects.

Figure 15 illustrates how these large-scale wind-driven oceanic flows self-organize, as well as the resulting eastward jet. Different spatial and time scales contribute to this self-organization: so-called mesoscale eddies in the ocean play the role of the synoptic-scale weather systems in the atmosphere.

Two cold rings and a warm one are clearly visible in Fig. 15. Such warm and cold eddies last for several months up to a year and have a diameter of about 100 km vs. the much larger horizontal dimensions of atmospheric weather systems, of the order of 1000 km. In fact, the scale of roughly 100 km that characterizes these eddies in the ocean is given by the Rossby radius of deformation that was introduced in Ch. 2. This scale is much smaller than that of atmospheric weather systems, although the basic physics is the same. The difference in scale of the oceanic and atmospheric Rossby radius is due to the different stratification of the two fluid media.

The characteristic scale of the eastward oceanic jet that detaches from Cape Hatteras and of the two opposite gyres on either side of it — sketched as the recurving white arrows in the figure — is of several thousand kilometers. These larger-scale oceanic features exhibit their own intrinsic dynamics on time scales of several years to possibly one or two decades.

A striking feature of the wind-driven circulation is the existence of two well-known North Atlantic Oscillations, with a period of about 7-8 and 14-15 years, respectively. Data analysis of various climatic variables, such as sea surface temperature (SST) over the North Atlantic or sea level pressure (SLP) over western Europe, and local surface air temperatures in Central England, as well as of proxy records, such as tree rings in Britain, travertine concretions in southeastern France, and Nile floods over the last millennium or so, all
exhibit strikingly robust oscillatory behavior with a 7-8-yr period and, to a lesser extent, with a 14-yr period. Variations in the path and intensity of the Gulf Stream are most likely to exert a major influence on the climate in this part of the world \[153\]. These climatic implications add to the intrinsic interest of theoretical studies of the low-frequency variability of the oceans’ double-gyre circulation.

Given the complexity of the processes involved, climate studies have been most successful when using not just a single model but a full hierarchy of models, from the simplest to models to the most detailed GCMs \[55, 63\]. In the following, we describe one of the simplest models of the hierarchy used in studying this problem, which still retains many of its essential features.

### 11.2 A simple model of the double-gyre circulation

The simplest model that includes many of the mechanisms described above is governed by the barotropic quasi-geostrophic (QG) equations. The term geostrophic refers to the fact that large-scale rotating flows tend to run parallel to, rather than perpendicular to constant-pressure contours; in the oceans, these contours are associated with the deviation from rest of the surfaces of equal water mass, due to Ekman pumping. Geostrophic balance implies in particular that the flow is divergence-free. The term barotropic, as opposed to baroclinic, has a slightly different meaning in geophysical fluid dynamics than in engineering fluid mechanics: it means that the model describes a single fluid layer of constant density and therefore the solutions do not depend on depth \[69, 60, 120\].

We consider an idealized, rectangular basin geometry and simplified forcing that mimics the distribution of vorticity due to the wind stress, as sketched to the right of Fig. 15. In our idealized model, the amounts of subpolar and subtropical vorticity injected into the basin are equal and the rectangular domain \[\Omega = (0, L_x) \times (0, L_y)\] is symmetric about the axis of zero wind stress curl.

The barotropic two-dimensional (2-D) QG equations in this idealized setting are:
\begin{align}
q_t + J(\psi, q) - \nu \Delta^2 \psi + \mu \Delta \psi &= -\tau \sin \frac{2\pi y}{L_y}, \\
q &= \Delta \psi - \lambda R^{-2} \psi + \beta y.
\end{align}  \tag{101}

Here $q$ and $\psi$ are the potential vorticity and the streamfunction, respectively, and the Jacobian $J$ gives the advection of potential vorticity by the flow, $J(\psi, q) = \psi_x q_y - \psi_y q_x = \mathbf{u} \cdot \nabla q$, where $\mathbf{u} = (-\psi_y, \psi_x)$, $x$ points east and $y$ points north.

The physical parameters are the strength of the planetary vorticity gradient $\beta$, the Rossby radius of deformation $\lambda R^{-2}$, the eddy-viscosity coefficient $\nu$, the bottom friction coefficient $\mu$, and the wind-stress intensity $\tau$. We use here free-slip boundary conditions $\psi = \Delta^2 \psi = 0$; the qualitative results described below do not depend on the particular choice of homogeneous boundary conditions [39, 78].

We consider the nonlinear PDE system (101) as an infinite-dimensional dynamical system and study its bifurcations as the parameters change. Two key parameters are the wind stress intensity $\tau$ and the eddy viscosity $\nu$. An important property of (101) is its mirror symmetry in the $y = L_y/2$ axis. This symmetry can be expressed as invariance with respect to the discrete $\mathbb{Z}_2$ group $\mathcal{S}$, given by

$$\mathcal{S}[\psi(x, y)] = -\psi(x, L_y - y);$$

any solution of (101) is thus accompanied by its mirror-conjugated solution. Hence the prevailing bifurcations are of either the symmetry-breaking or the Hopf type.

### 11.3 Bifurcations in the double-gyre problem

The development of a comprehensive nonlinear theory of the double-gyre circulation over the last two decades has gone through four main steps. These four steps can be followed through the bifurcation tree in Fig. 16.

#### 11.3.1 Symmetry-breaking bifurcation

The “trunk” of the bifurcation tree is plotted as the solid blue line in the lower part of the figure. When the forcing $\tau$ is weak or the dissipation $\nu$ is large, there is only one steady solution, which is antisymmetric with respect to the mid-axis of the basin. This solution exhibits two large gyres, along with their $\beta$-induced western boundary currents. Away from the western boundary, such a near-linear solution (not shown) is dominated by so-called Sverdrup balance between wind stress curl and the meridional mass transport [69,155].

The first generic bifurcation of this QG model was found to be a genuine pitchfork bifurcation that breaks the system’s symmetry as the nonlinearity becomes large enough with increasing wind stress intensity $\tau$ [77,78,17]. As the wind stress increases, the near-linear Sverdrup solution that lies along the solid blue line in the figure develops an eastward jet along the mid-axis, which penetrates farther into the
domain and also forms two intense recirculation vortices, on either side of the jet and near the western boundary of the domain.

The resulting more intense, and hence more nonlinear solution is still antisymmetric about the mid-axis, but loses its stability for some critical value of the wind-stress intensity, $\tau = \tau_p$. This value is indicated by the filled square on the symmetry axis of Fig. 16 and is labeled “Pitchfork” in the figure.

A pair of mirror-symmetric solutions emerges and it is plotted as the two red solid lines in the figure’s lower part. The streamfunction fields associated with the two stable steady-state branches have a rather different vorticity distribution and they are plotted in the two small panels to the upper-left and upper-right of Fig. 16. In particular, the jet in such a solution exhibits a large meander, reminiscent of the one seen in Fig. 15 just downstream of Cape Hatteras; note that the colors in Fig. 16 were chosen to facilitate the comparison with Fig. 15. These asymmetric flows
are characterized by one recirculation vortex being stronger in intensity than the
other, which deflects the jet accordingly either to the southeast, as is the case in the
observations for the North Atlantic, or to the northeast.

11.3.2 Gyre modes

The next step in the theoretical treatment of the problem was taken in part concur-
rently with the first one above [77, 78] and in part shortly thereafter [150, 40, 138].
It involved the study of time-periodic instabilities through Hopf bifurcation from
either an antisymmetric or an asymmetric steady flow. Some of these studies treated
wind-driven circulation models limited to a stand-alone, single gyre [138, 121]; such
a model concentrates on the larger, subtropical gyre while neglecting the smaller,
subpolar one.

The overall idea was to develop a full, generic picture of the time-dependent be-
havior of the solutions in more turbulent regimes, by classifying the various insta-
bilities in a comprehensive way. However, it quickly appeared that a particular kind
of instability leads to so-called gyre modes [78, 150], and was prevalent across the
full hierarchy of models of the double-gyre circulation; furthermore, this instability
triggers the lowest nonzero frequency present in all such models [37, 39].

These gyre modes always appear after the first pitchfork bifurcation, and it took
several years to really understand their genesis: gyre modes arise as two eigenvalues
merge — one is associated with a symmetric eigenfunction and responsible for the
pitchfork bifurcation, the other is associated with an antisymmetric eigenfunction
[139]; this merging is marked by a filled circle on the left branch of antisymmetric
stationary solutions and labeled $\text{M}$ in Fig. 16.

Such a phenomenon is not a bifurcation stricto sensu: one has topological $C^0$-
equivalence before and after the eigenvalue merging, but not from the $C^1$
point of view. Still, this phenomenon is quite common in small-dimensional dynamical
systems with symmetry, as exemplified by the unfolding of codimension-2 bifurca-
tions of Bogdanov-Takens type [70]. In particular, the fact that gyre modes trigger
the lowest-frequency of the model is due to the frequency of these modes growing
quadratically from zero until nonlinear saturation. Of course these modes, in turn,
become unstable shortly after the merging, through a Hopf bifurcation indicated in
Fig. 16 by a heavy dot marked “$\text{Hopf}$,” from which a stylized limit cycle emerges.
A mirror-symmetric $\text{M}$ and $\text{Hopf}$ bifurcation also occur on the right branch of sta-
tionary solutions, but have been omitted in the figure for visual clarity.

More generally, Hopf bifurcations of various types give rise to features that re-
cur more-or-less periodically in fully turbulent planetary-scale flows, atmospheric,
oceanic and coupled [60, 37, 39, 56]. In the climate sciences, one commonly refers
to this type of near-periodic recurrence as low-frequency variability (LFV).
11.3.3 Global bifurcations

The importance of the gyre modes was further confirmed through an even more puzzling discovery. Several authors realized, independently of each other, that the low-frequency dynamics of their respective double-gyre models was driven by intense relaxation oscillations of the jet \([144, 108, 19, 113, 141, 142, 143]\). These relaxation oscillations, already described in \([28, 150]\), were now attributed to a homoclinic bifurcation, with a global character in phase space \([70, 60]\). In effect, the QG model reviewed here undergoes a genuine homoclinic bifurcation that is generic across the full hierarchy of double-gyre models.

This bifurcation is due to the growth and eventual merging of the limit cycles that arise from the two mutually symmetric Hopf bifurcations. It is marked in the figure by a filled circle and a stylized lemniscate, and labeled “Homoclinic.” This global bifurcation is associated with chaotic behavior of the flow due to the Shilnikov phenomenon \([113, 143]\), which induces horseshoes in phase space.

The connection between such homoclinic bifurcations and gyre modes was not immediately obvious, but Simonnet et al. \([143]\) emphasized that the two were part of a single, global dynamical phenomenon. The homoclinic bifurcation indeed results from the unfolding of the gyre modes’ limit cycles. This familiar dynamical scenario is again well illustrated by the unfolding of a codimension-2 Bogdanov-Takens bifurcation, where the homoclinic orbits emerge naturally.

Since homoclinic orbits have an infinite period, it was natural to hypothesize that the gyre-mode mechanism, in this broader, global-bifurcation context, gave rise to the observed 7-yr and 14-yr North Atlantic oscillations. Although this hypothesis may appear a little farfetched — given the simplicity of the double-gyre models analyzed so far — it is reinforced by results with much more detailed model in the hierarchy, cf. \([37, 39]\) and Section 13 herein.

The successive-bifurcation theory appears therewith to be fairly complete for barotropic, single-layer models of the double-gyre circulation. This theory also provides a self-consistent, plausible explanation for the climatically important 7-year and 14-year oscillations of the oceanic circulation and the related atmospheric phenomena in and around the North Atlantic basin \([39, 37, 112, 171, 30, 51, 52, 49, 50, 88, 142, 143]\). The dominant 7- and 14-year modes of this theory survive, moreover, perturbation by seasonal-cycle changes in the intensity and meridional position of the westerly winds \([154]\).

In baroclinic models, with two or more active layers of different density, baroclinic instabilities \([39, 52, 69, 60, 120, 153, 121, 142, 8, 90]\) surely play a fundamental role, as they do in the observed dynamics of the oceans. However, it is not known to what extent baroclinic instabilities can destroy gyre-mode dynamics. The difficulty lies in a deeper understanding of the so-called rectification process \([83]\), which arises from the nonzero mean effect of the baroclinic component of the flow.

Roughly speaking, rectification drives the dynamics far away from any stationary solutions. In this situation, dynamical systems theory by itself cannot be used as a full explanation of complex, observed behavior resulting from successive bifurcations that are rooted in simple stationary or periodic solutions. Other tools
from statistical mechanics and nonequilibrium thermodynamics should, therefore, be considered [13, 48, 102, 100, 104, 131, 160]. Combining these tools with those of the successive-bifurcation approach may eventually lead to a more general and complete physical characterization of gyre modes in realistic models.

12 Non-autonomous and random dynamical systems

An additional way of improving upon the simple results so far is to include time-dependent forcing and stochasticity. As discussed in Section 10, the appropriate general framework for doing this is the one provided by NDS and RDS theory [59, 27, 10, 43]. In the present section, we summarize this mathematical framework in as straightforward a way as possible, and start with some simple ideas about deterministic vs. stochastic modeling.

12.1 Background and motivation

We are interested in behavior that is robust across a full hierarchy of climate models in general, and of oceanic double-gyre models in particular. Moreover, we would like this robust behavior of the models to be recognizable in the very large and ever-increasing mass of observational data. Finally, the features in whose robustness we are most interested are those that obtain over long time intervals.

Classical DDS theory, going back to H. Poincaré [127], is essentially a geometric approach to studying the asymptotic, long-term properties of solutions to autonomous ODE systems in phase space. Extensions to nonlinear PDE systems are covered, for instance, in [158]. To apply the theory in a reliable manner to a set of complex physical phenomena, one needs a criterion for the robustness of a given model within a class of dynamical systems. Such a criterion should help one deal with the inescapable uncertainties in model formulation, whether due to incomplete knowledge of the governing laws or inaccuracies in determining model parameters.

In this context, A. A. Andronov and L. S. Pontryagin [2] introduced the concept of structural stability for classifying dynamical systems. Structural stability means that a small, continuous perturbation of a given system preserves its dynamics up to a homeomorphism, i.e., up to a one-to-one continuous change of variables that transforms the phase portrait of our system into that of a nearby system; thus fixed points go into fixed points, limit cycles into limit cycles, etc. Closely related is the notion of hyperbolicity formulated by S. Smale [147]. A system is hyperbolic if, (very) loosely speaking, its limit set can be continuously decomposed into invariant sets that are either contracting or expanding; see [82] for more rigorous definitions.

A very simple example is the phase portrait in the neighborhood of a fixed point of saddle type. The Hartman-Grobman theorem states that the dynamics in this neighborhood is structurally stable. The converse statement, i.e. whether structural
stability implies hyperbolicity, is still an open question: the equivalence between structural stability and hyperbolicity has only been shown in the $C^1$ case, under certain technical conditions \[130, 132, 106, 118\]. Bifurcation theory is quite complete and satisfying in the setting of hyperbolic dynamics. Problems with hyperbolicity and bifurcations arise, however, when one deals with more complicated limit sets.

Hyperbolicity was introduced initially to help pursue the “dynamicist’s dream” of finding — in an abstract space $\mathcal{X}$ of all possible dynamical systems with given dimension and regularity — an open and dense set $\mathcal{S}$ consisting of structurally stable ones. For this set to be open and dense means, roughly speaking, that any dynamical system in $\mathcal{X}$ can be approximated by structurally stable ones in $\mathcal{S}$, while systems in the complement of $\mathcal{S}$ are negligible in a suitable sense.

Smale conjectured that hyperbolic systems form an open and dense set in the space of all $C^1$ dynamical systems. If this conjecture were true then hyperbolicity would be typical of all dynamics. Unfortunately, though, this conjecture is only true for one-dimensional dynamics and flows on disks and surfaces \[122\]. Smale \[146\] himself found several counterexamples to his conjecture. Newhouse \[115\] was able to generate open sets of nonhyperbolic diffeomorphisms using homoclinic tangencies. For the physicist, it is even more striking that the famous Lorenz attractor \[98\] is structurally unstable. Families of Lorenz attractors, classified by topological type, are not even countable \[71, 170\]. In each of these examples, we observe chaotic behavior in a nonhyperbolic situation, i.e., nonhyperbolic chaos.

Nonhyperbolic chaos appears, therefore, to be a severe obstacle to any easy classification of dynamic behavior. As mentioned by J. Palis \[118\], A. N. Kolmogorov already suggested at the end of the sixties that “the global study of dynamical systems could not go very far without the use of new additional mathematical tools, like probabilistic ones.”

Once more, Kolmogorov showed prophetic insight, and nowadays the concept of stochastic stability is an important tool in the study of genericity and robustness for dynamical systems. A system is stochastically stable if its Sinai-Ruelle-Bowen (SRB) measure \[145\] is stable with respect to stochastic perturbations, and the SRB measure is given by $\lim_{n \to \infty} \frac{1}{n} \sum \delta_{z_i}$, with $z_i$ being the successive iterates of the dynamics. This measure is obtained intuitively by allowing the entire phase space to flow onto the attractor \[45\].

To replace the failed program of classifying dynamical systems based on structural stability and hyperbolicity, J. Palis \[113\] formulated his so-called global conjecture: The set $\mathcal{S}$ of systems satisfying the following conditions is dense in the $C^r$-topology for $r \geq 1$. The three conditions are:

(i) the system has only finitely many attractors, i.e. periodic or chaotic sinks;
(ii) the union of the corresponding basins of attraction has full Lebesgue measure in $\mathcal{S}$; and
(iii) each attractor is stochastically stable in its basin of attraction.

Stochastic stability is thus based on ergodic theory. We would like to consider a more geometric approach, which can provide a coarser, more robust classification...
of climate models across a full hierarchy of such models. In this section, we propose such an approach, based on RDS theory [4, 25, 34, 93 and references therein].

RDS theory describes the behavior of dynamical systems subject to external stochastic forcing; its tools have been developed to help study the geometric properties of stochastic differential equations (SDEs). RDS theory is thus the stochastic, SDE counterpart of the geometric theory of ODEs, as presented for instance in [5]. This approach provides a rigorous mathematical framework for a stochastic form of robustness, while the more traditional, topological concepts do not seem to be appropriate for the climate sciences, given the prevalence of nonhyperbolic chaos.

12.2 Pullback and random attractors

Stochastic parametrizations for GCMs aim at compensating for our lack of detailed knowledge on small spatial scales in the best way possible [94, 95, 96, 119, 137, 151]. The underlying assumption is that the associated time scales are also much shorter than the scales of interest and, therefore, the lag correlation of the phenomena being parametrized is negligibly small. Such assumptions clearly hold, for instance, for the spatial and temporal scales of cloud processes, whose parametrization poses a notorious obstacle to realistic climate simulation and prediction [73, 74, 75, 76]. Stochastic parametrizations thus essentially transform a deterministic autonomous system into a nonautonomous one, subject to random forcing.

Explicit time dependence, whether deterministic or stochastic, in a dynamical system immediately raises a technical difficulty. Indeed, the classical notion of attractor is not always relevant, since any object in phase space is moving with time and the concept of forward asymptotics, as considered in autonomous DDS theory, is meaningless. One needs therefore another notion of attractor. In the deterministic nonautonomous framework, the appropriate notion is that of a pullback attractor [129, 85, 15], which we present below. The closely related notion of random attractor in the stochastic framework is also explained briefly below, with further details given in Appendix A.

12.2.1 Pullback attraction

Before defining the notion of pullback attractor, let us recall some basic facts about nonautonomous dynamical systems. Consider the ODE

\[ \dot{x} = f(x, t) \]  

on a vector space \( X \); this space could even be infinite-dimensional, if we were dealing with partial or functional differential equations, as is often the case in fluid-flow and climate problems. Rigorously speaking, we cannot associate a dynamical sys-
tem acting on $X$ with a nonautonomous ODE; nevertheless, in the case of unique solvability of the initial-value problem, we can introduce a two-parameter family of operators $\{S(t,s)\}_{t\geq s}$ acting on $X$, with $s$ and $t$ real, such that $S(t,s)x(s) = x(t)$ for $t \geq s$, where $x(t)$ is the solution of the Cauchy problem with initial data $x(s)$. This family of operators satisfies $S(s,s) = \text{Id}_X$ and $S(t,\tau) \circ S(\tau,s) = S(t,s)$ for all $t \geq \tau \geq s$, and all real $s$. Such a family of operators is called a “process” by Sell [137]. It extends the classical notion of the resolvent of a nonautonomous linear ODE to the nonlinear setting.

We can now define the pullback attractor as the family of invariant sets $\mathcal{A}(t) : -\infty < t < +\infty$ that satisfy, for every real $t$ and all $x_0$ in $X$:

$$\lim_{s \to -\infty} \text{dist}(S(t,s)x_0, \mathcal{A}(t)) = 0.$$  \hspace{1cm} (104)

“Pullback” attraction does not involve running time backwards; it corresponds instead to the idea of measurements being performed at present time $t$ in an experiment that was started at some time $s < t$ in the past: the experiment has been running for long enough, and we are thus looking at the present moment for an “attracting state.” Note that there exist several ways of defining a pullback attractor — the one retained here is a local one, cf. [91, and references therein]; see [7] for further information on nonautonomous dynamical systems in general.

A simple example will be helpful for the general reader. Consider the scalar linear ODE

$$\dot{x} = -\alpha x + \sigma t,$$  \hspace{1cm} (105)

with both $\alpha$ and $\sigma$ positive. Here $\alpha > 0$ implies the system is dissipative, a crucial property both mathematically and physically. In the climate context, its importance has been emphasized by [98, 60], among others.

The situation is depicted in Fig. [17]. The PBA is easily computed in this case, and it is given by

$$\mathcal{A}(t) \equiv a(t) = \frac{\sigma}{\alpha}(t - \frac{1}{\alpha}).$$  \hspace{1cm} (106)

The figure clearly shows that the approximation of the PBA at a given time $t = t_1$ or $t = t_2$ improves as we “pull back” further, from $s = s_1$ to $s = s_2$ (red vs. blue orbits in the figure); it also shows that the accuracy of the approximation depends essentially on the time interval $|t - s|$, rather than on $t$ and $s$ separately. In practice, the characteristic time for obtaining an approximation with given accuracy will depend largely on the rate of dissipation $\alpha$.

In the stochastic context, noise forcing is modeled by a stationary stochastic process. If the deterministic dynamical system of interest is coupled to this stochastic process in a reasonable way — defined below by the “cocycle property” — then random pullback attractors may appear. These pullback attractors will exist for almost each sample path of the driving stochastic process, so that the same probability distribution governs both sample paths and their corresponding pullback attractors. A more detailed and rigorous explanation is given in Appendix A.
Fig. 17 Pullback attractor (PBA) for a scalar linear equation, given by the solid black line. We wish to observe the PBA at times \( t = t_1, t_2 \), marked by dashed vertical lines, and consider the convergence to the straight line of orbits started at past times \( s = s_1, s_2 \); sample orbits for \( s_1 \) and \( s_2 \) are plotted in red and blue, respectively. Courtesy of M. D. Chekroun.

Roughly speaking, the concept of a random attractor provides a geometric framework for the description of asymptotic regimes in the context of stochastic dynamics. Figure 17 showed how additive and linear deterministic forcing modifies a fixed point. We illustrate next, following [27, 56], the way that multiplicative stochastic forcing modifies a strange attractor.

12.2.2 The stochastically perturbed Lorenz model

We consider in this section the Lorenz convection model [98] and its randomly forced version studied in [27]. The model is governed by the well-known three nonlinear, coupled ODEs that are now forced by linearly multiplicative white noise [4, 25].

\[
\begin{align*}
\text{[SLM]} & \quad \begin{cases}
\frac{dx}{dt} = s(y - x) + \sigma x 
+ \sigma dW_t, \\
\frac{dy}{dt} = (rx - y - xz) + \sigma y 
+ \sigma dW_t, \\
\frac{dz}{dt} = (-bz + xy) + \sigma z 
+ \sigma dW_t.
\end{cases}
\end{align*}
\] (107)

In the deterministic context, purely geometric map models were proposed in the 1970s [70] to interpret the dynamics observed numerically by E. N. Lorenz in [98].
These geometric models attracted considerable attention and it was shown that they possess a unique SRB measure \([84, 28]\), i.e., a time-independent measure that is invariant under the flow and has conditional measures on unstable manifolds that are absolutely continuous with respect to Lebesgue measure \([45]\). This result has been extended recently to the Lorenz flow \([98]\) itself, in which the SRB measure is supported by a strange attractor of vanishing volume \([161, 8]\).

Even though this result was only proven recently, the existence of such an SRB measure was suspected for a long time and has motivated several numerical studies to compute a probability density function (PDF) associated with the Lorenz model, by filtering out the stable manifolds \([42, 111, \text{and references therein}]\). The Lorenz attractor is then approximated by a two-dimensional manifold, called the branched manifold \([70]\), which supports this PDF. Based on such a strategy, Dorfle & Graham \([42]\) showed that the stationary solution of the Fokker-Planck equation for the Lorenz model perturbed by additive white noise possesses a density with two components: the PDF of the deterministic system supported by the branched manifold plus a narrow Gaussian distribution transversal to that manifold.

Fig. 18 Snapshot of the Lorenz \([98]\) model’s random attractor \(A_\omega\) and of the corresponding sample measure \(\mu_\omega\), for a given, fixed realization \(\omega\). The figure corresponds to projection onto the \((y,z)\) plane, i.e. \(\int \mu_\omega(x,y,z)dx\). One billion initial points have been used in both panels and the pullback attractor is computed for \(t = 40\). The parameter values are the classical ones — \(r = 28\), \(s = 10\), and \(b = 8/3\), while the time step is \(\Delta t = 5 \cdot 10^{-3}\). The color bar to the right of each panel is on a log-scale and quantifies the probability to end up in a particular region of phase space. Both panels use the same noise realization \(\omega\) but with noise intensity (a) \(\sigma = 0.3\) and (b) \(\sigma = 0.5\). Notice the interlaced filamentary structures between highly (yellow) and moderately (red) populated regions; these structures are much more complex in panel (b), where the noise is stronger. Weakly populated regions cover an important part of the random attractor and are, in turn, entangled with (almost) zero-probability regions (black). After \([27]\).
It follows that, in the presence of additive noise, the resulting PDF looks very much like that of the unperturbed system, only slightly fuzzier: the noise smoothes the small-scale structures of the attractor. More generally, this smoothing appears in the classical, forward approach that only considers forward asymptotics, with $t \to +\infty$ — as opposed to the pullback approach, introduced here in Section 12.2.1 in which one considers also the pullback asymptotics of $s \to -\infty$ — and it does so for a broad class of additive as well as multiplicative noises, in the sense of [4]. More precisely, this smoothing occurs provided that the diffusion terms due to the stochastic components in the Fokker-Planck equation associated with the SDE system under study are sufficiently non-degenerate; see [27, Appendix C and references therein].

Hörmander’s theorem guarantees that this is indeed the case for hypoelliptic SDEs [6]. The corresponding non-degeneracy conditions allow one to regularize the stationary solutions of the transport equation

$$\partial_t p(x,t) = -\nabla \cdot (p(x,t)F(x)),$$

which is the counterpart of the Fokker-Planck equation in the absence of noise; a measure-theoretic justification for it can be found, for instance in [92, p. 210].

This transport equation is also known as the Liouville equation and it provides the probability density at time $t$ of $S(t)x$ when the initial state $x$ is sampled from a probability measure that is absolutely continuous with respect to Lebesgue measure; here $\{S(t)\}_{t \in \mathbb{R}}$ is the flow of $\dot{x} = f(x)$, for some sufficiently smooth vector field $f$ on $\mathbb{R}^d$. As a matter of fact, when $f$ is dissipative and the dynamics associated with it is chaotic, the stationary solutions of (108) are very often singular with respect to Lebesgue measure; these solutions are therefore expected to be SRB measures. For a broad class of noises — such as those that obey a hypoellipticity condition — the forward approach leads us to suspect that noise effects tend to remove the singular aspects with respect to Lebesgue measure. This smoothing aspect of random perturbations is often useful in the theoretical understanding of any stochastic system, in particular in the analysis of the low- and higher-order moments, which have been thoroughly studied in various contexts.

For chaotic systems subject to noise, however, this noise-induced smoothing observed in the forward approach compresses a lot of crucial information about the dynamics itself; quite to the contrary, the pullback approach brings this information into sharp focus. A quick look at Figs. [18 a,b] and [19] is already enlightening in this respect. All three figures refer to the invariant measure $\mu_\omega$ supported by the random attractor of our stochastic Lorenz model [SLM]. This model obeys the following three SDEs: In system (107), each of the three equations of the classical, deterministic model [98] is perturbed by linearly multiplicative noise in the Itô sense, with $W_t$ a Wiener process and $\sigma > 0$ the noise intensity. The other parameter values are the standard ones for chaotic behavior [60], and are given in the caption of Fig. [18].

Figures [18 a,b] show two snapshots of the sample measure $\mu_\omega$ supported by the random attractor of [SLM] — for the same realization $\omega$ but for two different noise intensities, $\sigma = 0.3$ and 0.5, while Fig. [19] provides four successive snapshots of
µₜ₀, for the same noise intensity σ = 0.5 as in Fig. 18(b), but with t = t₀ + kΔt and k = 0, 1, 2, 3 for some t₀.

The sample measures in these three figures, and in the associated short video given in the SM, exhibit amazing complexity, with fine, very intense filamentation; note logarithmic scale on color bars in the three figures. There is no fuzziness whatsoever in the topological structure of this filamentation, which evokes the Cantor-set foliation of the deterministic attractor [70]. Such a fine structure strongly suggests that these measures are supported by an object of vanishing volume.

Fig. 19 Four snapshots of the random attractor and sample measure supported on it, for the same parameter values as in Fig. 18. The time interval Δt between two successive snapshots — moving from left to right and top to bottom — is Δt = 0.0875. Note that the support of the sample measure may change quite abruptly, from time to time, cf. short video in [27, Supplementary Material] for details. Reproduced from [27], with permission from Elsevier.

Much more can be said, in fact, about these objects. RDS theory offers a rigorous way to define random versions of stable and unstable manifolds, via the Lyapunov spectrum, the Oseledec multiplicative theorem, and a random version of the Hartman-Grobman theorem [4]. These random invariant manifolds can support measures, like in the deterministic context. When the sample measures µₜ₀ of
an RDS have absolutely continuous conditional measures on the random unstable manifolds, then \( \mu_\omega \) is called a random SRB measure.

One can prove rigorously, by relying on Theorem B of [93], that the sample measures of the discretized stochastic system obtained from the [SLM] model share the SRB property. Indeed, it can be shown that a Hörmander hypoellipticity condition is satisfied for our discretized [SLM] model, thus ensuring that the random process generated by this model has a smooth density \( p(t,x) \) [86]; see [27, Appendix C1] for the details. Standard arguments [149] can then be used to prove that the stationary solution \( \rho \) of our model’s Fokker-Planck equation is in fact absolutely continuous with respect to Lebesgue measure.

Since these simulations exhibit exactly one positive Lyapunov exponent, the absolute continuity of \( \rho \) implies that the sample measures seen in Figs. 18 and 19 are, actually, good numerical approximations of a genuine random SRB measure for our discretized [SLM], whenever \( \delta t \) is sufficiently small; see also the next section. In fact, Ledrappier & Young’s [93] Theorem B is a powerful result, which clearly shows that — in noisy systems, and subject to fairly general conditions — chaos can lead to invariant sample measures with the SRB property; see also [27, Appendix C2]. It is striking that the same noise-induced smoothing that was “hiding” the dynamics in the forward approach allows one here to exhibit the existence of an SRB measure from a pullback point of view, and thus to compute explicitly the unstable manifold supporting this invariant measure.

Note that, since the sample measures associated with the discrete [SLM] system are SRB here, they are so-called physical measures [4, 27 and references therein] and can thus be computed at any time \( t \) by simply flowing a large set of initial data from the remote past \( s \ll t \) up till \( t \), for a fixed realization \( \omega \); this is exactly how Figs. 18 and 19 were obtained. Given the SRB property, the nonzero density supported on the model’s unstable manifolds delineates numerically these manifolds; Figs. 18 and 19 provide therefore an approximation of the global random attractor of our stochastic Lorenz system.

Finally, these random measures are Markovian, in the sense that they are measurable with respect to the past \( \sigma \)-algebra of the noise [4]. The latter statement results directly from the fact that these measures are physical, and thus satisfy the required measurability conditions in the pullback limit. The information about the moments that is available in the classical Fokker-Planck approach is complemented here by information about the pathwise moments. These pathwise statistics are naturally associated with the sample measures — when the latter are SRB — by taking appropriate averages.

The evolution of the sample measures \( \mu_{\theta,\omega} \) — as apparent from the short video in [27, Supplementary Material] — is quite complex, and two types of motion are present. First, a pervasive “jiggling” of the overall structure can be traced back to the roughness of the Wiener process \( W_t \) and to the multiplicative way it enters into the [SLM] model. Second, there is a smooth, regular low-frequency motion present in the evolution of the sample measures, which seems to be driven by the deterministic system’s unstable limit cycles and is thus related to the well-known lobe dynamics. The latter motion is clearly illustrated in Fig. 19.
More generally, it is worth noting that this type of low-frequency motion seems to occur quite often in the evolution of the samples measures of chaotic systems perturbed by noise; it appears to be related to the recurrence properties of the unperturbed deterministic flow, especially when energetic oscillatory modes characterize the latter. To the best of our knowledge, there are no rigorous results on this type of phenomenon in RDS theory.

Besides this low-frequency motion, abrupt changes in the global structure occur from time to time, with the support of the sample measure either shrinking or expanding suddenly. These abrupt changes recur frequently in the video associated with Fig. 19 which reproduces a relatively short sequence out of a very long stochastic model integration; see SM.

As the noise intensity $\sigma$ tends to zero, the sample-measure evolution slows down, and one recovers numerically the measure of the deterministic Lorenz system (not shown). This convergence as $\sigma \to 0$ may be related to the concept of stochastic stability [172, 84]. Such a continuity property of the sample measures in the zero-noise limit does not, however, hold in general; it depends on properties of the noise, as well as of the unperturbed attractor [31, 35, 16].

As stated in the theoretical section, the forward approach is recovered by taking the expectation, $E[\mu] := \int_\Omega \mu_\omega P(d\omega)$, of these invariant sample measures. In practice, $E[\mu]$ is closely related to ensemble or time averages that typically yield the previously mentioned PDFs. In addition, when the random invariant measures are Markovian and the Fokker-Planck equation possesses stationary solutions, $E[\mu] = \rho$, where $\rho$ is such a solution. Subject to these conditions, there is even a one-to-one correspondence between Markovian invariant measures and stationary measures of the Markov semi-group [4, 32]. The inverse operation of $\mu \mapsto \rho = E[\mu]$ is then given by $\rho \mapsto \mu_\omega = \lim_{t \to \infty} \Phi(t, \omega)^{-1} \rho$; the latter is in fact the pullback limit of $\rho$ due to the cocycle property [32]. It follows readily from this result that RDS theory “sees” many more invariant measures than those given by the Markov semi-group approach: non-Markovian measures appear to play an important role in stochastic bifurcation theory [4], for instance.

To summarize, one might say that the classical forward approach considers only expectations and PDFs, whereas the RDS approach “slices” the statistics very finely: the former takes a hammer to the problem, while the latter takes a scalpel. Clearly, distinct physical processes may lead to the same observed PDF: the RDS approach and, in particular, the pullback limit are able to discriminate between these processes and thus provide further insight into them.

### 13 Time-dependent forcing and ocean–atmosphere coupling

Given the NDS and RDS concepts of pullback attraction and time-dependent invariant measures, as introduced in the previous section, we are ready now to return to the wind-driven ocean circulation of Section 11 and consider two additional steps in its understanding, simulation and prediction, namely time-dependent forcing and
coupling with the atmosphere above. We start by applying a still prescribed but now time-dependent wind-stress to a double-gyre model.

13.1 Time-dependent forcing

The most obvious way in which wind stress varies in mid-latitudes is from summer to winter: the so-called atmospheric jet stream is both stronger and closer to the equator in winter. Such periodic changes in the forcing were already considered in [154], albeit without the theoretical underpinnings discussed herein. The first application of pullback attractors to the double-gyre problem appears in [124], for the case of periodic forcing, and we rely here on [125] for a systematic presentation of this novel application of NDS theory to the wind-driven ocean circulation.

13.1.1 Model formulation

The model, in its autonomous form [123], differs somewhat from that of Eq. (101) by the absence of the Laplacian eddy viscosity — given by the bilaplacian $\Delta^2 \psi$ of the streamfunction $\psi$ — and by a slightly different form of the wind stress. Its discretized form is obtained by projection of Eqs (101) onto a set of cartesian basis functions, and low-order truncation of the expansion thus obtained. The discrete, low-order model is thus governed by four nonlinear coupled ODEs for the variables $\{\psi_i(t), i = 1, \cdots, 4\}$ that are the coefficients of the streamfunction $\psi(x,y,t)$ retained in this expansion. The choice of basis functions follows up on such idealized double-gyre models introduced in [78, 143], in particular on this basis including the exponential form of a current that decays away from the domain’s western boundary.

In the presence of nonautonomous forcing, the model takes the form

$$
\begin{align*}
\dot{\psi}_1 + L_{11} \psi_1 + L_{13} \psi_3 + B_1(\psi, \psi) &= W_1(t), \\
\dot{\psi}_2 + L_{22} \psi_2 + L_{24} \psi_4 + B_2(\psi, \psi) &= W_2(t), \\
\dot{\psi}_3 + L_{33} \psi_3 + L_{31} \psi_1 + B_3(\psi, \psi) &= W_3(t), \\
\dot{\psi}_4 + L_{44} \psi_4 + L_{42} \psi_2 + B_4(\psi, \psi) &= W_4(t).
\end{align*}
$$

Here $\Psi$ denotes the vector of expansion coefficients $\{\psi_1, \psi_2, \psi_3, \psi_4\}$ of the streamfunction, $W$ is the vector of forcing terms $(W_1, W_2, W_3, W_4)$, and the bilinear terms $B_i$ are given by

$$
B_1(\psi, \psi) = 2J_{112} \psi_1 \psi_2 + 2J_{114} \psi_1 \psi_4 + 2J_{123} \psi_2 \psi_3 + 2J_{134} \psi_3 \psi_4, \\
B_2(\psi, \psi) = J_{211} \psi_1^2 + J_{222} \psi_2^2 + J_{233} \psi_3^2 + J_{244} \psi_4^2 + 2J_{213} \psi_1 \psi_3 + 2J_{224} \psi_2 \psi_4,
$$
\begin{align*}
B_3(\Psi, \Psi) &= 2J_{312}\Psi_1\Psi_2 + 2J_{314}\Psi_1\Psi_4 + 2J_{323}\Psi_2\Psi_3 + 2J_{334}\Psi_3\Psi_4, \\
B_4(\Psi, \Psi) &= J_{411}\Psi_1^2 + J_{422}\Psi_2^2 + J_{433}\Psi_3^2 + J_{444}\Psi_4^2 \\
&+ 2J_{413}\Psi_1\Psi_3 + 2J_{424}\Psi_2\Psi_4.
\end{align*}

(110)

In [125], the forcing $W(x, t)$ is defined as

$$W(x, t) = \gamma [1 + \varepsilon f(t)] w(x),$$

(111)

where $x = (x, y)$ and $w(x)$ is the time-constant double-gyre wind stress curl used in [123]; furthermore, $\gamma$ is the dimensionless intensity of this wind stress, while $f(t)$ is an aperiodic time dependence weighed by the dimensionless parameter $\varepsilon$. Clearly, $w(x)$ is projected onto the same four leading-order basis functions as $\Psi$ and $W$.

**Fig. 20** Ensemble behavior of forced solutions of the double-gyre ocean model of [125]. (a) Time dependence of the total forcing $1 + \varepsilon f(t)$, for $\varepsilon = 0.2$. (b,c) Evolution of $N = 644$ initial states emanating from the subset $\Gamma$ in the $(\Psi_1, \Psi_3)$-plane for (b) $\gamma = 0.96$ and (c) $\gamma = 1.1$. (b’,c’) Corresponding time series of $P_{\Psi_3}$. Reproduced from [125], with the permission of the American Meteorological Society.

### 13.1.2 Model behavior

The aperiodic forcing $f(t)$ in Fig. 20 corresponds to an idealized but aperiodic choice with a substantial decadal component. Such decadal climate signals are associated with the so-called Pacific Decadal Oscillation [105, 21], for instance, and climate prediction on such time scales is a matter of great recent interest for the climate sciences [18, 65, 126, 75, 76].
The evolution of the model solutions subject to this forcing is shown in Figs. 20b and 20c for two reference cases, $\gamma = 0.96$ and $\gamma = 1.1$, respectively; the solutions of the corresponding autonomous system, i.e. with $\varepsilon = 0$, are plotted in [109, 125, Figs. 1(b,c)]. It is clear from the two panels that a regime change occurs as the forcing is increased. This regime change occurs in the autonomous system at $\gamma = 1.0$ and corresponds to a global bifurcation associated with a homoclinic orbit; see [143, 123] and Section 11.3.3 here. The persistence of the regime shift in the nonautonomous case, for $\varepsilon$ as large as 0.2, is interesting but none too surprising.

In Figs. 20(b, c), $N$ trajectories $\Psi_k(t)$ emanate from $N$ initial states uniformly distributed at time $t_0 = 0$ on a reference subset $\Gamma$ of the $(\Psi_1, \Psi_3)$-plane, while the $(\Psi_2, \Psi_4)$ coordinates of the initial states (not shown) are chosen the same way as in [124]. The panels (b,c) here provide a first representation of the sets that approximate the corresponding PBAs; the small number of trajectories is in fact chosen for the sake of graphical clarity. The correct identification and characterization of the PBAs requires, however, an analysis of the PDFs that evolve along the trajectories and of the convergence to the appropriate invariant time-dependent sets; see Figs. 20(b',c'). This convergence was shown in [125] to take no longer than about 15 years of model simulation.

13.1.3 Multiplicity of PBAs

The existence of a global PBA is rigorously demonstrated for the weakly dissipative, nonlinear model governed by Eqs. (109)–(111) in [125, Appendix], based on general, NDS-theoretical results [85, 15]. Numerically, though, this unique global attractor seems to possess two separate local PBAs, as apparent from Fig. 21. Panels (a) and (b) in the figure refer again to the two types of regimes apparent in Figs. 20(b,c), for $\gamma = 0.96$ and $\gamma = 1.1$, respectively.

The mean normalized distance plotted in the figure is defined as follows:

$$\sigma(\Psi_1, \Psi_3) = \langle \delta_t \rangle_{T_{\text{tot}}}.$$  

Here $\delta_t$ is the distance, at time $t$, between two trajectories of the model that were a distance $\delta_0$ apart at time $t = t_0$, and the normalized distance $\delta_t = \delta_t/\delta_0$ is averaged over the whole forward time integration $T_{\text{tot}} = t_{\text{fin}} - t_{\text{init}}$ of the available trajectories.

The maps of $\sigma$ in Fig. 21 reveal large chaotic regions where $\sigma \gg 1$ on average (warm colors) but also non-chaotic regions, in which $\sigma \leq 1$ (blue) and thus initially close trajectories do remain close on average. The rectangular regions in the two panels that are labeled by letters A and B and by numbers 1–4 correspond to sub-domains of the initial set $\Gamma$ and are discussed further in [125, Sec. 5]. The numerical evidence in Fig. 21 suggests, furthermore that the boundary between the two types of local attractors has fractal properties.

In the autonomous context, the coexistence of topologically distinct local attractors is well known in the climate sciences [60, 37, 39, 143, 140 and references therein]. The coexistence of local PBAs with chaotic vs. non-chaotic characteris-
Fig. 21 Mean normalized distance $\sigma(\Psi_1, \Psi_2)$ for 15,000 trajectories of the double-gyre ocean model starting in the initial set $I$: (a) $\gamma = 0.96$, and (b) $\gamma = 1.1$. Reproduced from [125], with the permission of the American Meteorological Society.

...tics, within a unique global PBA, as illustrated by Fig. 21 here, seems to be novel, at least in the climate literature.

### 13.2 Coupled ocean–atmosphere variability

Forcing of an ocean model by time-dependent wind stress is but one step in the direction of a fully coupled ocean–atmosphere model. Simple models of this type have been considered, for instance, by S. Vannitsem and associates [165, 166, 167], who also included heat fluxes as an important ingredient in the coupling.

A key conclusion of this group’s work is that, in such a simple coupled model, it is possible to find a fairly broad parameter regime within which slow modes, with interdecadal near-periodicities exist. Moreover, these modes are truly coupled, in the sense that their atmospheric components share the long near-periodicities. We merely refer here to [57] for a summary of these results and to [168] for an observational verification.

### 14 Concluding remarks

Over the last four decades, the climate sciences have become of paramount importance in understanding, predicting and attempting to improve the relations between humanity and the planet it inhabits [116, 73, 74, 75, 76]. This development could not have occurred without a strong contribution from the mathematical sciences.
Several books [37, 38, 60, 99, 104, 164, 107] have documented the mathematics that have been used and that have in turn benefited from this area of applications.

To conclude this contribution, it seems appropriate to formulate a few ideas for possibly fruitful future applications of NDS and RDS theory to the climate sciences. It would appear to the authors that — in spite of increasing efforts in this direction, e.g. [109] and further articles in the same volume — insufficient attention has been paid to understanding the strong interaction between natural or intrinsic climate variability [117, 18], on the one hand, and the anthropogenic forcing upon which a huge fraction of the recent climate literature has focused [75, 76]. We would like, therefore, to suggest an avenue for overcoming this gap, and to summon the necessary mathematics for doing so.

14.1 A generalized definition of climate sensitivity

The usual way of defining climate sensitivity, going back to [116], relies on the conceptual picture of Fig. 22a: climate is in a steady state characterized by a global average temperature $T = T_0$, a parameter that is affected by human activities, say CO$_2$ concentration, is suddenly changed; and $T(t)$ follows — according to the solution of a scalar, linear ODE like Eq. (105), in which $T$ replaces $x$ and a Heaviside function replaces $\sigma t$ on the right-hand side — so that $T(t) \to T_1$, and $T_1 > T_0$ when the CO$_2$ jump is positive.

This cartoon has been, of course, considerably enriched in successive assessment reports of the Intergovernmental Panel on Climate Change [73, 74, 75, 76] from a scalar, linear ODE to the coupled, nonlinear PDE systems that govern GCMs. But no mathematically consistent picture has emerged for reconciling the intrinsic variability that those PDE systems include, as illustrated in Sections 11 and 13 therein, with the complex nature of the forcing, both natural and anthropogenic. Thus climate sensitivity is still essentially thought of as $\Delta T / \Delta CO_2$, where $\Delta T = T_1 - T_0$, the difference between a final and an initial “equilibrium” temperature, and $\Delta CO_2$ is likewise the difference between a final and an initial forcing level. Consideration has been given to changes in other fields than temperature, such as precipitation [97, 133], to variances and energetics [102, 100], or to regional differences [14, 101].

To include these considerations and follow up on ideas formulated in [56], we would like to suggest a definition of climate sensitivity $\gamma$ that takes full advantage of the NDS and RDS framework and that encompasses changes in the natural variability, in its most general meaning, as the forcing level or some other parameter changes. This definition replaces the more-or-less standard definition of $\gamma = \partial T / \partial \mu$ [60, p. 320], with $\mu$ an arbitrary parameter generalizing CO$_2$ — a definition that corresponds to the equilibrium situation in Fig. 22a — by a more appropriate and self-consistent one that corresponds to the chaotic and random situation in Fig. 22c.

---

3 This $\gamma$ should not be confused with the parameter used in Section 13.1 for the forcing intensity.
14.2 PBAs, invariant measures and Wasserstein distance

The mathematical framework of PBAs, in the presence of time-dependent forcing or coefficients, suggests the definition

$$\gamma = \frac{\partial d_W}{\partial \mu},$$

(113)

where $d_W$ is the Wasserstein distance [41, 56] between the time-dependent invariant measures supported on the system’s PBA at two distinct parameter values. An idea of the large differences that may exist between two such PBAs is given by the three snapshots of the invariant measure on the PBA of the infinite-dimensional, but still relatively simple ENSO model of [53] in Fig. 23 below.

The model’s two dependent variables are sea surface temperature $T$ in the eastern Tropical Pacific and thermocline depth $h$ there, as a function of time $t$:

$$T = f(T(t), h(t)), \quad h(t) = g(T, h, F(t, \tau_1, \tau_2), \quad F(t) = 1 + \varepsilon \cos(\omega t + \phi).$$

(114)

In Eqs. (114), $F$ stands for the seasonal forcing, with period $2\pi/\omega = 12$ months, and all three variables — $T, h$ and $F$ — depend on the time $t$ and the delays $\tau_1$ and $\tau_2$; these delays characterize the traveling times along the Equator of eastward Kelvin and westward Rossby waves. Several authors have studied delay-differential models of ENSO; see [37] for a review and [67, 23] and references therein] for further mathematical details on such models.

The solutions of Eqs. (114) exhibit periodic, quasi-periodic and chaotic behavior, as well as frequency locking to the time-dependent, seasonal cycle. Thus, in principle, an infinite number of scalars are required to define the dependence of these solutions on the parameters $\tau_1$ and $\tau_2$; these scalars need to include not just the means of temperature $T$ and depth $h$, but also their variance and higher-order moments. In [56, Fig. 7] this dependence was shown for the zeroth, second and fourth moments of $h(t)$; more precisely the plotted quantities were the mean, standard deviation and fourth root of the kurtosis of the PDF.
Fig. 23 Time-dependent invariant measures of the $\mathcal{K}_3$ model; snapshots shown at three times $t$: (a) $t_1 = 19.23$ yr, (b) $t_2 = 20$ yr and (c) $t_3 = 20.833$ yr. Courtesy of M. D. Chekroun.
While the figure illustrates successive snapshots of the invariant measure supported on the PBA for the same value of the model parameters, there are similarly large differences between such measures for different parameter values [56, 67]. Closer attention to the definition of Eq. (113) can provide better insights into the changes in time, as well as with respect to a parameter, of higher moments of a model’s PDF but also into the delicate and important matter of changes in the distribution of extreme events [22, 66].

Steps in the application of the NDS and RDS framework to climate change and climate sensitivity are being taken by a number of research groups [10, 11, 38, 43, 101]; see especially [62] and references therein. This area of research is only opening up and should provide many opportunities for contributions by mathematicians and climate scientists, as well as for fruitful interactions among them.

Acknowledgements: It is a pleasure to thank M. D. Chekroun, D. Kondrashov, H. Liu, J. C. McWilliams, J. D. Neelin and I. Zaliapin for many useful discussions and their continuing interest in the questions studied here. The research reported in Section 13 was carried out jointly with M. D. Chekroun, L. De Cruz, J. Demayer, S. Pierini and S. Vannitsem [125, 167]. Figure 23 is due to M. D. Chekroun, a steadfast companion on the road to understanding NDS and RDS theory, along with their applications to climate dynamics in general. H. Liu helped with finalizing Fig. 22. It is a pleasure to thank the organizers of the INdAM Workshop on Mathematical Approaches to Climate Change Impacts — P. Cannarsa, D. Mansutti, and A. Provenzale — for the invitation to deliver a lecture and to contribute to this volume. The writing of this research-and-review paper was supported by grants N00014-12-1-0911 and N00014-16-1-2073 from the Multidisciplinary University Research Initiative (MURI) of the Office of Naval Research and by the US National Science Foundation grant OCE 1243175. This paper is TiPES contribution #4; this project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 820970.

Appendix A. RDS theory and random attractors

We present here briefly the mathematical concepts and tools of random dynamical systems, random attractors and stochastic equivalence. We shall use the concept of pullback attractor introduced in Section 3.2.1 to define the closely related notion of a random attractor, but need first to define an RDS. We denote by $\mathbb{T}$ the set $\mathbb{Z}$, for maps, or $\mathbb{R}$, for flows. Let $(X, \mathcal{B})$ be a measurable phase space, and $(\Omega, \mathcal{F}, \mathbb{P}, (\theta(t))_{t \in \mathbb{T}})$ be a metric dynamical system i.e. a flow in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that $(t, \omega) \mapsto \theta(t)\omega$ is measurable and $\theta(t): \Omega \to \Omega$ is measure preserving, i.e., $\theta(t)\mathbb{P} = \mathbb{P}$.

Let $\varphi: \mathbb{T} \times \Omega \times X \to X$, $(t, \omega, x) \mapsto \varphi(t, \omega)x$, be a mapping with the two following properties:

$(R_1)$: $\varphi(0, \omega) = \text{Id}_X$, and
70 Michael Ghil and Eric Simonnet

\((R_2)\) (the cocycle property): For all \(s,t \in \mathbb{T}\) and all \(\omega \in \Omega\),

\[ \varphi(t + s, \omega) = \varphi(t, \theta(s) \omega) \circ \varphi(s, \omega). \]

If \(\varphi\) is measurable, it is called a measurable RDS over \(\theta\). If, in addition, \(X\) is a topological space (respectively a Banach space), and \(\varphi\) satisfies \((t, \omega) \mapsto \varphi(t, \omega)x\) continuous (resp. \(C^k, 1 \leq k \leq \infty\)) for all \((t, \omega) \in \mathbb{T} \times \Omega\), then \(\varphi\) is called a continuous (resp. \(C^k\)) RDS over the flow \(\theta\). If so, then

\[ (\omega, x) \mapsto \Theta(t)(x, \omega) := (\theta(t) \omega, \varphi(t, \omega)x), \] (115)

is a (measurable) flow on \(\Omega \times X\), and is called the skew-product of \(\theta\) and \(\varphi\). In the sequel, we shall use the terms “RDS” or “cocycle” synonymously.

The choice of the so-called driving system \(\theta\) is a crucial step in this set-up; it is mostly dictated by the fact that the coupling between the stationary driving and the deterministic dynamics should respect the time invariance of the former, as illustrated in Fig. 24. The driving system \(\theta\) also plays an important role in establishing stochastic conjugacy \cite{29} and hence the kind of classification we aim at.

The concept of random attractor is a natural and straightforward extension of the definition of pullback attractor \cite{104}, in which Sell’s \cite{137} process is replaced by a cocycle, cf. Fig. 24 and the attractor \(\mathcal{A}\) now depends on the realization \(\omega\) of the noise, so that we have a family of random attractors \(\mathcal{A}(\omega)\), cf. Fig. 25. Roughly speaking, for a fixed realization of the noise, one “rewinds” the noise back to \(t \rightarrow -\infty\) and lets the experiment evolve (forward in time) towards a possibly attracting set \(\mathcal{A}(\omega)\); the driving system \(\theta\) enables one to do this rewinding without changing the statistics, cf. Figs. 24 and 25.

Other notions of attractor can be defined in the stochastic context, in particular based on the original SDE; see \cite{34, 33} for a discussion on this topic. The present definition, though, will serve us well.

Having defined RDSs and random attractors, we now introduce the notion of stochastic equivalence or conjugacy, in order to rigourously compare two RDSs; it is defined as follows: two cocycles \(\varphi_1(\omega, t)\) and \(\varphi_2(\omega, t)\) are conjugated if and only if there exists a random homeomorphism \(h \in \text{Homeo}(X)\) and an invariant set such that \(h(\omega)(0) = 0\) and

\[ \varphi_1(\omega, t) = h(\theta(t) \omega)^{-1} \circ \varphi_2(\omega, t) \circ h(\omega). \] (116)

Stochastic equivalence extends classic topological conjugacy to the bundle space \(X \times \Omega\), stating that there exists a one-to-one, stochastic change of variables that continuously transforms the phase portrait of one sample system in \(X\) into that of any other such system.
Fig. 24 Random dynamical systems (RDS) viewed as a flow on the bundle \( X \times \Omega = \text{"dynamical space" } \times \text{"probability space."} \) For a given state \( x \) and realization \( \omega \), the RDS \( \varphi \) is such that \( \Theta(t) (x, \omega) = (\theta(t) \omega, \varphi(t, \omega) x) \) is a flow on the bundle. Reproduced from [59], with permission from Elsevier.

**Pullback attraction to \( A(\omega) \)**

Fig. 25 Schematic diagram of a random attractor \( A(\omega) \), where \( \omega \in \Omega \) is a fixed realization of the noise. To be attracting, for every set \( B \) of \( X \) in a family \( \mathcal{B} \) of such sets, one must have \( \lim_{t \to +\infty} \text{dist}(B(\theta(-t) \omega), A(\omega)) = 0 \) with \( B(\theta(-t) \omega) := \varphi(t, \theta(-t) \omega) B \); to be invariant, one must have \( \varphi(t, \omega) A(\omega) = A(\theta(t) \omega) \). This definition depends strongly on \( \mathcal{B} \); see [33] for more details. Reproduced from [59], with permission from Elsevier.
Appendix B. Mixing in random dynamical systems

In this appendix, we define rigorously the concept of an $\omega$-wise mixing RDS, in the continuous-time context. Recall first the well-known definition of mixing in a deterministic dynamical system. Given a flow $\{\phi_t\}$ on a topological space $X$, which possesses an invariant (Borel) probability measure $\mu$, we say that the dynamical system $(\phi_t, \mu)$ is mixing if for any two measurable sets $A$ and $B$,

$$\mu(A \cap \phi_{-t}(B)) \longrightarrow t \to \infty \mu(A)\mu(B),$$

(117)

or equivalently,

$$\int F \cdot (G \circ \phi_t)d\mu \longrightarrow t \to \infty \int Fd\mu \int Gd\mu,$$

(118)

for any pair of continuous functions $F, G : X \to \mathbb{R}$. Equation (117) states that the set of points in $A$ whose images belong to $B$ by $\{\phi_t\}$ tends towards having the same proportion in $A$ as $B$ has in $X$, with proportions being understood in terms of the measure $\mu$. Hence any measurable set will tend to redistribute itself over the state space according to $\mu$.

Let us now consider a cocycle $\{\Phi(t, \omega)\}_{(t, \omega) \in \mathbb{R} \times \Omega}$ on the base space $(\Omega, \mathcal{F}, P, \{\theta_t\})$, which possesses the sample measures $\{\mu_{\omega}\}$. We say that $\Phi$ is $\omega$-wise mixing or fiber mixing [12] — or even simply mixing, if no confusion is possible — if for any random sets $A(\omega)$ and $B(\omega)$,

$$\mu_{\omega}(A(\omega) \cap \Phi(t, \omega)^{-1}(B(\theta_t \omega))) \longrightarrow t \to \infty \mu_{\omega}(A(\omega))\mu_{\theta_t \omega}(B(\theta_t \omega)),$$

(119)

almost surely with respect to $P$. This mixing concept and its interpretation are natural extensions of their deterministic counterparts just recalled above, except that the mixing property has to be checked across the fibers $\omega$ and $\theta_t \omega$, due to the skew-product nature of the RDS $(\Phi, \theta)$ [12].

Appendix C. Low-frequency variability (LFV) and mixing

Low-frequency variability (LFV) is a widely used, but not clearly defined concept in the atmospheric, oceanic and climate sciences [60, 63, 59]. In general, one just refers to phenomena whose periods are longer than those previously studied. Examples include atmospheric LFV — referring to so-called intraseasonal oscillations whose characteristic time scale of 10–100 days is longer than the 5–10-day life cycle of mid-latitude storms but not longer than a season [63, 64] — or oceanic LFV, referring to interannual or interdecadal variability whose characteristic time scales are longer than the several-months–long ones of mesoscale eddies and the seasonal cycle of a year [55, 63].
In this appendix, we clarify the notion of LFV from a mathematical perspective. Let us reconsider the deterministic Lorenz system [98]. It is known that the power spectral density, or power spectrum, of this system is exponentially decaying [47, 61]. At the same time, one can check numerically that the decay of the autocorrelation function is exponentially decaying, too. Other types of power-spectrum behavior may be encountered for chaotic dynamical systems, though. Aside from pure power-law decay, it may also happen that the power spectrum contains one or several broad peaks that stand out above the continuous background, whether the latter has a power-law [9] or exponential decay. If the central frequencies of these peaks are located in a frequency band that lies close to the lower end of the frequency range being studied, the system is said to exhibit LFV [58, 61].

This climatically motivated, but vague notion of LFV can be formalized mathematically through the mixing concept introduced in Appendix B. Indeed, for a general flow \( \{ \phi_t \} \) on a topological space \( X \), which possesses an invariant (Borel) probability measure \( \mu \), let us define the correlation function by

\[
C_t(F, G) := \left| \int F \cdot (G \circ \phi_t) d\mu - \int F d\mu \int G d\mu \right|
\]

using the same notations as above. If the system \( \{ \phi_t, \mu \} \) is mixing, the rate of approach to zero of \( C_t(F, G) \) is called the rate of decay of correlations for its observables \( F \) and \( G \). A system exhibits a slow decay rate of correlations at short lags if the rate is slower than exponential over some characteristic time interval \([0, T]\).

This emphasis on the nonuniform decay rate of correlations is consistent with the heuristic LFV notion used in the climate sciences, as described above, and it connects the mixing properties of the flow with its power spectral density. In [26], the authors elucidated the relationship between these two approaches, as well as the relationships between the rate of decay of correlations and the occurrence of rough or smooth parameter dependence of the model’s statistics. Moreover, the quantitative knowledge about the system’s LFV — when combined with the pathwise approach discussed in Section 12.2 — can be used, according to [24, 89], to improve ensemble prediction skills of data-driven stochastic models, such as those discussed in [87].

References

1. M.R. Allen, Do-it-yourself climate prediction, Nature 401 (1999), 627.
2. A.A. Andronov and L.S. Pontryagin, Systèmes grossiers, Dokl. Akad. Nauk. SSSR 14 (1937), no. 5, 247–250.
3. V. Araujo, M. Pacifico, R. Pujal, and M. Viana, Singular-hyperbolic attractors are chaotic, Trans. Amer. Math. Soc. 361 (2009), 2431–2485.
4. L. Arnold, Random Dynamical Systems, Springer-Verlag, 1998, 616 pp.
5. V.I. Arnold, Geometrical Methods in the Theory of Differential Equations, Springer-Verlag, 1983, 334 pp.
6. D.R. Bell, Degenerate stochastic differential equations and hypoellipticity, Longman, Harlow, 1995.
7. A. Berger and S. Siegmund, *On the gap between random dynamical systems and continuous skew products*, J. Dyn. Diff. Eq. **15** (2003), 237–279.
8. P. Berloff, A. Hogg, and W. Dewar, *The turbulent oscillator: A mechanism of low-frequency variability of the wind-driven ocean gyres*, J. Phys. Oceanogr. **37** (2007), 2363–2386.
9. K. Bhattacharya, M. Ghil, and I.L. Vulis, *Internal variability of an energy-balance model with delayed albedo effects*, J. Atmos. Sci. **39** (1982), 1747–1773, doi:10.1175/1520-0469.
10. T. Bódai, G. Károlyi, and T. Tél, *A chaotically driven model climate: extreme events and snapshot attractors*, Nonlin. Processes Geophys. **18** (2011), 2363–2386.
11. T. Bódai, V. Lucarini, F. Lunkeit, and R. Boschi, *Global instability in the Ghil-Sellers model*, Clim. Dyn. **44** (2015), 3361–3381.
12. T. Bogenschütz and Z.S. Kowalski, *A condition for mixing of skew products*, Aequationes Math. **59** (2000), 222–234.
13. F. Bouchet and J. Sommeria, *Emergence of intense jets and Jupiter's great red spot as maximum entropy structures*, J. Fluid. Mech. **464** (2002), 165–207.
14. A. Bracco, J.D. Neelin, H. Luo, J.C. McWilliams, and J.E. Meyerson, *High-dimensional decision dilemmas in climate models*, Geosci. Model Dev. **6** (2013), 1673–1687, doi:10.5194/gmdd-6-2731-2013.
15. A. Carvalho, J. A. Langa, and J. C. Robinson, *Attractors for Infinite-Dimensional Non-Autonomous Dynamical Systems*, Springer-Verlag, 2012, 391 pp.
16. A.N. Carvalho, J.A. Langa, and J.C. Robinson, *Lower semicontinuity of attractors for non-autonomous dynamical systems*, Ergod. Th.. Dyn. Syst. **29** (2009), 765–780.
17. P. Cessi and G.R. Ierley, *Symmetry-breaking multiple equilibria in quasigeostrophic wind-driven flows*, J. Phys. Oceanogr. **25** (1995), 1196–1205.
18. C.P. Chang, M. Ghil, M. Latif, and J.M. Wallace, *Climate change: Multidecadal and beyond*, vol. 6, World Scientific Publ. Co./Imperial College Press, 2015, 388 pp.
19. K.I. Chang, K. Ide, M. Ghil, and C.C.A. Lai, *Transition to aperiodic variability in a wind-driven double-gyre circulation model*, J. Phys. Oceanogr. **31** (2001), 1260–1286.
20. P. Chang, L. Ji, H. Li, and M. Flugel, *Chaotic dynamics versus stochastic processes in El Niño-Southern Oscillation in coupled ocean-atmosphere models*, Physica D **98** (1996), 301–320.
21. Y. Chao, M. Ghil, and J.C. McWilliams, *Pacific interdecadal variability in this century’s sea surface temperatures*, Geophys. Res. Lett. **27** (2000), 2261–2264.
22. M. Chavez, M. Ghil, and J. Urrutia Fucugauchi, *Extreme events: Observations, modeling and economics*, vol. 214, American Geophysical Union & Wiley, 2015, 438 pp.
23. M.D. Chekroun, M. Ghil, H. Liu, and S. Wang, *Low-dimensional Galerkin approximations of nonlinear delay differential equations*, Discr. Cont. Dyn. Syst. **36** (8) (2016), 4133–4177.
24. M.D. Chekroun, D. Kondrashov, and M. Ghil, *Predicting stochastic systems by noise sampling, and application to the El Niño-Southern Oscillation*, Proc. Natl. Acad. Sci USA **108** (2011), 11766–11771.
25. M.D. Chekroun, H. Liu, and S. Wang, *Approximation of stochastic invariant manifolds: Stochastic manifolds for nonlinear spdes i*, SpringerBriefs in Mathematics, Springer, 2015.
26. M.D. Chekroun, J.D. Neelin, D. Kondrashov, J.C. McWilliams, and M. Ghil, *Rough parameter dependence in climate models: The role of Ruelle-Pollicott resonances*, Proc. Natl. Acad. Sci USA **111** (2014), 1684–1690, doi:10.1073/pnas.1321816111.
27. M.D. Chekroun, E. Simonnet, and M. Ghil, *Stochastic climate dynamics: Random attractors and time-dependent invariant measures*, Physica D **240** (2011), 1685–1700.
28. P. Collet and C. Tresser, *Ergodic theory and continuity of the Bowen-Ruelle measure for geometrical flows*, Fyzika **20** (1988), 33–48.
29. N.D. Cong, *Topological Dynamics of Random Dynamical Systems*, Oxford Mathematical Monographs, Clarendon Press, Oxford, 1997, 212 pp.
30. E.D. Da Costa and A.C. Colin de Verdière, *The 7.7 year North Atlantic oscillation*, Q. J. R. Meteorol. Soc. **128** (2004), 797–818.
31. H. Crauel, *White noise eliminates instability*, Arch. Math. **75** (2000), 472–480.
32. ____, Random probability measures on Polish spaces, vol. 11, Stochastic Monographs, Taylor & Francis, 2002.
33. ____, A uniformly exponential attractor which is not a pullback attractor, Arch. Math. **78** (2002), 329–336.
34. H. Crauel and F. Flandoli, Attractors for random dynamical systems, Tech. Report 148, cuola Normale Superiore Pisa, 1992.
35. ____, Additive noise destroys a pitchfork bifurcation, J. Dyn. Diff. Eqn. **10** (1998), 259–274.
36. B. Cushman-Roisin and J-M. Beckers, Introduction to geophysical fluid dynamics: Physical and numerical aspects, Academic Press, 2nd Edition, 2011, 875 pp.
37. H.A. Dijkstra, Nonlinear Physical Oceanography: A Dynamical Systems Approach to the Large Scale Ocean Circulation and El Niño, Springer (2nd ed.), 2005, 532 pp.
38. ____, Nonlinear Climate Dynamics, Cambridge Univ. Press, 2013, 367 pp.
39. H.A. Dijkstra and M. Ghil, Low-frequency variability of the large-scale ocean circulation: A dynamical systems approach, Rev. Geophys. **43** (2005), doi:10.1029/2002RG001222.
40. H.A. Dijkstra and C.A. Katsman, Temporal variability of the wind-driven quasi-geostrophic double gyre ocean circulation: basic bifurcation diagrams, Geophys. Astrophys. Fluid Dyn. **85** (1997), 195–232.
41. R.L. Dobrushin, Prescribing a system of random variables by conditional distributions, Theor. Prob. Appl. **15** (1979), 458–486.
42. M. Dorfle and R. Graham, Probability density of the Lorenz model, Phys. Rev. A **27** (1983), 1096–1105.
43. G. Drótos, T. Bódai, and T. Tél, Probabilistic concepts in a changing climate: A snapshot attractor picture, J. Clim. **28** (2015), 3275–3288.
44. M. Dubar, Approche climatique de la période romaine dans l’est du Var : recherche et analyse des composantes périodiques sur un concrétionnement centennal (Ier-IIe siécle apr. J.-C.) de l’aqueduc de Fréjus, Archeoscience **30** (2006), 163–171.
45. J.P. Eckmann and D. Ruelle, Ergodic theory of chaos and strange attractors, Rev. Mod. Phys. **57** (1985), 617–656.
46. D.A. Stainforth et al., Uncertainty in predictions of the climate response to rising levels of greenhouse gases, Nature **433** (2005), 403–406.
47. N. Ohtomo et al., Exponential characteristics of power spectral densities caused by chaotic phenomena, J. Phys. Soc. Japan **64** (1995), 1104–1113.
48. B.F. Farrel and P.J. Ioannou, Structural stability of turbulent jets, J. Atmos. Sci. **60** (2003), 2101–2118.
49. Y. Feliks, M. Ghil, and A. W. Robertson, Oscillatory climate modes in the eastern Mediterranean and their synchronization with the North Atlantic Oscillation, J. Clim. **23** (2010), 4060–4079, doi:10.1175/2010JCLI3181.1.
50. ____, The atmospheric circulation over the North Atlantic as induced by the SST field, J. Clim. **24** (2011), 522–542, doi:10.1175/2010JCLI3839.1.
51. Y. Feliks, M. Ghil, and E. Simonnet, Low-frequency variability in the mid-latitude atmosphere induced by an oceanic thermal front, J. Atmos. Sci. **61** (2004), 961–981.
52. ____, Low-frequency variability in the mid-latitude baroclinic atmosphere induced by an oceanic thermal front, J. Atmos. Sci. **64** (2007), 97–116.
53. E. Galanti and E. Tziperman, ENSO’s phase locking to the seasonal cycle in the fast-SST, fast-wave, and mixed-mode regimes, J. Atmos. Sci. **57** (2000), 2936–2950.
54. M. Ghil, Cryothermodynamics: The chaotic dynamics of paleoclimate, Physica D **77** (1994), 130–159.
55. ____, Hilbert problems for the geosciences in the 21st century, Nonlin. Proc. Geophys. **8** (2001), 211–222.
56. ____, A mathematical theory of climate sensitivity or, How to deal with both anthropogenic forcing and natural variability?, C. P Chang, M. Ghil, M. Latif and J. M. Wallace (Eds.), World Scientific Publ. Co/Imperial College Press, 2015, Climate Change: Multidecadal and Beyond, 31–51.
57. M. Ghil and E. Simonnet, The wind-driven ocean circulation: Applying dynamical systems theory to a climate problem, Discr. Cont. Dyn. Syst. - A 37 (2017), 189–228.

58. M. Ghil, R.M. Allen, M.D. Dettinger, K. Ide, D. Kondrashov, M.E. Mann, A. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou, Advanced spectral methods for climatic time series, Rev. Geophys. 40(1) (2002), 3.1–3.41.

59. M. Ghil, M.D. Chekroun, and E. Simonnet, Climate dynamics and fluid mechanics: Natural variability and related uncertainties, Physica D 237 (2008), 2111–2126.

60. M. Ghil and S. Childress, Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics, Springer-Verlag, Berlin/Heidelberg/New York, 1987, 512 pp.

61. M. Ghil and N. Jiang, Recent forecast skill for the El Niño/Southern Oscillation, Geophys. Res. Lett. 25 (1998), 171–174.

62. M. Ghil and V. Lucarini, The physics of climate variability and climate change, Rev. Mod. Phys. (2019), 1–84, submitted, arXiv:1910.00583.

63. M. Ghil and A.W. Robertson, Solving problems with gcms: General circulation models and their role in the climate modeling hierarchy, in: General Circulation Model Development: Past, Present and Future, D. Randall (Ed.), Academic Press, San Diego, 2000, pp. 285–325.

64. M. Ghil and R. Vautard, Interdecadal oscillations and the warming trend in global temperature time series, Nature 350 (1991), 324–327.

65. M. Ghil and I. Zaliapin, Understanding enso variability and its extrema: A delay differential equation approach, vol. 214, ch. 6, pp. 63–78, M. Chavez, M. Ghil and J. Urrutia-Fucugauchi (Eds.), American Geophysical Union & Wiley, 2015, Extreme Events: Observations, Modeling and Economics.

66. M. Ghil, I. Zaliapin, and S. Thompson, A delay differential model of ENSO variability:Parametric instability and the distribution of extremes, Nonlin. Processes Geophys. 18 (2011), 295–350, 10.5194/npg-18-295-2011.

67. M. Ghil and I. Zaliapin, Understanding enso variability and its extrema: A delay differential equation approach, vol. 214, ch. 6, pp. 63–78, M. Chavez, M. Ghil and J. Urrutia-Fucugauchi (Eds.), American Geophysical Union & Wiley, 2015, Extreme Events: Observations, Modeling and Economics.

68. M. Ghil, I. Zaliapin, and S. Thompson, A delay differential model of ENSO variability: Parametric instability and the distribution of extremes, Nonlin. Processes Geophys. 15 (2008), 417–433.

69. A.E. Gill, Atmosphere-ocean dynamics, Academic Press, 1982, 662 pp.

70. J. Guckenheimer and P. Holmes, Nonlinear oscillations, dynamical systems and bifurcations of vector fields, Springer-Verlag (2nd ed.), 1991, 453 pp.

71. J. Guckenheimer and R.F. Williams, Structural stability of Lorenz attractors, Publ. Math. I.H.E.S. 50 (1979), 59–72.

72. J. Holton and G.J. Hakim, An introduction to dynamic meteorology, Academic Press, 5th Edition, 2012, 552 pp.

73. IPCC, Climate Change, The IPCC Scientific Assessment, J. T. Houghton, G. J. Jenkins and J. J. Ephraums (Eds.), Cambridge Univ. Press, Cambridge, MA, 1991, 365 pp.

74. IPCC, Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell and C. A. Johnson (Eds.), Cambridge University Press, Cambridge, U.K., 2001, 944 pp.

75. IPCC, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller (Eds.), Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007.

76. IPCC, Climate Change 2013. The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker, T. F. and Qin, D. and Plattner, G. K. and Tignor, M. and Allen, S. K. and Boschung, J. and Nauels, A. and Xia, Y. and Bex, B. and Midgley, B. M. (Eds.). Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.
77. S. Jiang, F.-F. Jin, and M. Ghil, *The nonlinear behavior of western boundary currents in a wind-driven, double-gyre, shallow-water model*, pp. 64–67, in: Ninth Conf. Atmos. & Oceanic Waves and Stability (San Antonio, TX), American Meteorological Society, Boston, Mass., 1993.

78. ______*, Multiple equilibria, periodic, and aperiodic solutions in a wind-driven, double-gyre, shallow-water model*, J. Phys. Oceanogr. 25 (1995), 764–786.

79. F.-F. Jin, J. D. Neelin, and M. Ghil, *El Niño on the Devil’s Staircase: Annual subharmonic steps to chaos*, Science 264 (1994), 70–72.

80. ______*, El Niño/Southern Oscillation and the annual cycle: Subharmonic frequency locking and aperiodicity*, Physica D 98 (1996), 442–465.

81. E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge Univ. Press Cambridge/London, UK, 2003, 341 pp.

82. A. Katok and B. Haselblatt, *Introduction to the modern theory of dynamical systems*, vol. 54, Cambridge Univ. Press, Encycl. Math. Appl., 1995, 822 pp.

83. C.A. Katsman, H.A. Dijkstra, and S.S. Drijfhout, *The rectification of the wind-driven ocean circulation due to its instabilities*, J. Mar. Res. 56 (1998), 559–587.

84. Y. Kifer, *Ergodic theory of random perturbations*, Birkhäuser, 1988.

85. P.E. Kloeden and M. Rasmussen, *Nonautonomous dynamical systems*, vol. 176, Mathematical Surveys and Monographs, American Mathematical Society, 2011.

86. J.J. Kohn, *Pseudo-differential operators and hypoellipticity*, Proc. Amer. Math. Soc. Symp. Pure Math. 23 (1973), 61–69.

87. D. Kondrashov, M.D. Chekroun, and M. Ghil, *Data-driven non-Markovian closure models*, Physica D 297 (2015), 33–55.

88. D. Kondrashov, Y. Feliks, and M. Ghil, *Oscillatory modes of extended Nile River records (A.D. 622–1922)*, Geophys. Res. Lett. 32 (2005), L10702, doi:10.1029/2004GL022158.

89. K. Kondrashov, M.D. Chekroun, A.W. Robertson, and M. Ghil, *Low-order stochastic model and “past-noise forecasting” of the Madden-Julian oscillation*, Geophys. Res. Lett. 40 (2013), 5303–5310.

90. S. Kravtsov, P. Berloff, W.K. Dewar, M. Ghil, and J.C. McWilliams, *Dynamical origin of low-frequency variability in a highly nonlinear mid-latitude coupled model*, J. Climate 19 (2007), 6391–6408.

91. J.A. Langa, J.C. Robinson, and A. Suarez, *Stability, instability, and bifurcation phenomena in non-autonomous differential equations*, Nonlinearity 15 (2002), 887–903.

92. A. Lasota and M.C. Mackey, *Chaos, fractals, and noise: Stochastic aspects of dynamics*, vol. 97, Applied Mathematical Sciences, Springer-Verlag, 1994.

93. F. Ledrappier and L.-S. Young, *Entropy formula for random transformations*, Prob. Theory Related Fields 80 (1988), 217–240.

94. J.W.B. Lin and J.D. Neelin, *Influence of a stochastic moist convective parameterization on tropical climate variability*, Geophys. Res. Lett. 27 (2000), 3691–3694.

95. ______*, Considerations for stochastic convective parameterization*, J. Atmos. Sci. 59 (2002), 959–975.

96. ______*, Toward stochastic deep convective parameterization in general circulation models*, Geophys. Res. Lett. 30 (2003), 1162, doi:10.1029/2002GL016203.

97. P.C. Loikith and J.D. Neelin, *Short-tailed temperature distributions over North America and implications for future changes in extremes*, Geophys. Res. Lett. 42 (2015), 8577–8585, doi:10.1002/2015GL065502.

98. E.N. Lorenz, *Deterministic nonperiodic flow*, J. Atmos. Sci. 20 (1963), 130–141.

99. ______*, The essence of chaos*, Univ. of Washington Press, 1995, 240 pp.

100. V. Lucarini, R. Blender, C. Herbert, F. Ragone, S. Pascale, and J. Wouters, *Mathematical and physical ideas for climate science*, Rev. Geophys. 52 (2014), 809–859.

101. V. Lucarini, F. Ragone, and F. Lunkeit, *Predicting climate change using response theory: Global averages and spatial patterns*, (2016), submitted.

102. V. Lucarini and S. Sarno, *A statistical mechanical approach for the computation of the climatic response to general forcings*, Nonlin. Processes Geophys. 18 (2011), 7–28.
103. R.A. Madden and P.R. Julian, Observations of the 40–50-day tropical oscillations - a review, Mon. Wea. Rev. 122 (1994), 814–837.
104. A. Majda and X. Wang, Nonlinear dynamics and statistical theories for basic geophysical flows, Cambridge Univ. Press, 2006, 551 pp.
105. N.J. Mantua, S. Hare, Y. Zhang, J.M. Wallace, and R.C. Francis, A Pacific interdecadal climate oscillation with impacts on salmon production, Bull. Am. Meteor. Soc. 78 (1997), 1069–1079.
106. R. Mañe, A proof of the C¹-stability conjecture, Publ. Math. I.H.E.S. 66 (1987), 161–210.
107. J.C. McWilliams, Fundamentals of Geophysical Fluid Dynamics, 2nd ed., Cambridge Univ. Press, 2011, 272 pp.
108. S.P. Meacham, Low-frequency variability in the wind-driven circulation, J. Phys. Oceanogr. 30 (2000), 269–293.
109. G.A. Meehl, Decadal climate variability and the early-2000s hiatus, vol. 13(3), pp. 1–6, D. Menemenlis and J. Sprintall (Eds.), US CLIVAR Variations in Understanding the Earth’s Climate Warming Hiatus: Putting the Pieces Together, 2015.
110. S. Minobe, A. Kuwano-Yoshida, N. Komori, S.-P. Xie, and R.J. Small, Influence of the Gulf Stream on the troposphere, Nature 452 (2008), 206–209.
111. A.K. Mittal, S. Dwivedi, and R.S. Yadav, Probability distribution for the number of cycles between successive regime transitions for the Lorenz model, Physica D 233 (2007), 14–20.
112. V. Moron, R. Vautard, and M. Ghil, interannual oscillations in global sea-surface temperatures, Clim. Dyn. 14 (1998), 545–569.
113. N.T. Nadiga and B.P. Luce, Global bifurcation of Shilnikov type in a double-gyre ocean model, J. Phys. Oceanogr. 31 (2001), 2669–2690.
114. J.D. Neelin, D.S. Battisti, A.C. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S. Zebiak, ENSO Theory, J. Geophys. Res. 104(C7) (1998), 14261–14290.
115. S. Newhouse, The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms, Publ. Math. I.H.E.S. 50 (1979), 101–150.
116. NRC, Carbon Dioxide and Climate: A Scientific Assessment, J.G. Charney et al. (Eds.), National Academies Press, Washington, D.C., 1979.
117. , Natural Climate Variability on Decade-to-Century Time Scales, D. G. Martinson and K. Bryan and M. Ghil et al. (Eds.), National Academies Press, Washington, D.C., 1995, 630 pp.
118. J. Palis, A global perspective for non-conservative dynamics, Ann. I.H. Poincaré 22 (2005), 485–507.
119. T.N. Palmer, The prediction of uncertainty in weather and climate forecasting, Rep. Prog. Phys. 63 (2000), 71–116.
120. J. Pedlosky, Geophysical Fluid Dynamics, Springer-Verlag (2nd ed.), 1987, 710 pp.
121. , Ocean Circulation Theory, Springer, New York, 1996.
122. M. Peixoto, Structural stability on two-dimensional manifolds, Topology 1 (1962), 101–110.
123. S. Pierini, Low-frequency variability, coherence resonance, and phase selection in a low-order model of the wind-driven ocean circulation, J. Phys. Oceanogr. 41 (2011), 1585–1604.
124. Ensemble simulations and pullback attractors of a periodically forced double-gyre system, J. Phys. Oceanogr. 44 (2014), 3245–3254.
125. S. Pierini, M. Ghil, and M.D. Chekroun, Exploring the pullback attractors of a low-order quasigeostrophic ocean model: The deterministic case, J. Clim. (2015), in press, doi:10.1175/JCLI-D-15-0848.1
126. G. Plaut, M.Ghil, and R. Vautard, Interannual and interdecadal variability in 335 Years of Central England temperatures, Science 268 (1995), 710–713.
127. H. Poincaré, Sur les équations de la dynamique et le problème des trois corps, Acta Math. 13 (1890), 1–270.
128. V.D. Pope, M. Gallani, P.R. Rowntree, and R.A. Stratton, The impact of new physical parameterisations in the Hadley Centre climate model HadAM3, Clim. Dyn. 16 (2000), 123–146.
129. M. Rasmussen, Attractivity and bifurcation for nonautonomous dynamical systems, Springer, 2007.
130. J. Robbin, A structural stability theorem, Annals Math. 94 (1971), 447–449.
131. R. Robert and J. Sommeria, Statistical equilibrium states for two-dimensional flows, J. Fluid Mech. 229 (1991), 291–310.
132. C. Robinson, Structural stability of C^1 diffeomorphisms, J. Diff. Equations 22 (1976), 28–73.
133. T.W. Ruff and J.D. Neelin, Long tails in regional surface temperature probability distributions with implications for extremes under global warming, Geophys. Res. Lett. 39 (2012), doi:10.1029/2011GL050610
134. R. Salmon, Lectures on geophysical fluid dynamics, Oxford University Press, 1998, 378 pp.
135. A. Saunders and M. Ghil, A Boolean delay equation model of ENSO variability, Physica D 160 (2001), 54–78, doi:10.1029/2011GL050610
136. S. Speich, H.A. Dijkstra, and M. Ghil, Successive bifurcations in a shallow-water model applied to the wind-driven ocean circulation, Nonlin. Processes Geophys. 2 (1995), 241–268.
137. H. Stommel, Thermohaline convection with two stable regimes of flow, Tellus 13 (1961), 224–230.
138. L. Sushama, M. Ghil, and K. Ide, Spatio-temporal variability in a mid-latitude ocean basin subject to periodic wind forcing, Atmosphere-Ocean 45 (2007), 227–250, doi:10.3137/ao.450404
139. H.U. Sverdrup, Wind-driven currents in a baroclinic ocean; with application to the equatorial currents of the eastern Pacific, Proc. Natl. Acad. Sci. USA 33 (1947), 318–326.
156. H.U. Sverdrup, M.W. Johnson, and R.H. Fleming, *The Oceans: Their Physics, Chemistry and General Biology*, Prentice-Hall, New York, 1942, available at [http://ark.cdlib.org/ark:/13030/kt167nb66y/](http://ark.cdlib.org/ark:/13030/kt167nb66y/).

157. T.N. Palmer, T. Jung and G.J. Shufts, *Influence of a stochastic parameterization on the frequency of occurrence of North Pacific weather regimes in the ECMWF model*, Geophys. Res. Lett. **32** (2005), Art. No. L23811.

158. R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, 2nd Ed., 1997, 648 pp.

159. P.D. Thompson, *Numerical Weather Analysis and Prediction*, Macmillan, New York, 1961, 170 pp.

160. L.N. Trefethen, A. Trefethen, S.C. Reddy, and T.A. Driscoll, *Hydrodynamic stability without eigenvalues*, Science **261** (1993), 578–584.

161. W. Tucker, *Lorenz attractor exists*, C. R. Acad. Sci. Paris **328** (12) (1999), 1197–1202.

162. E. Tziperman, M.A. Cane, and S.E. Zebiak, *Irregularity and locking to the seasonal cycle in an ENSO prediction model as explained by the quasi-periodicity route to chaos*, J. Atmos. Sci. **50** (1995), 293–306.

163. E. Tziperman, L. Stone, M. Cane, and H. Jarosh, *El Niño chaos: Overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator*, Science **264** (1994), 72–74.

164. G. Vallis, *Atmospheric and Oceanic Fluid Dynamics*, Cambridge University Press, UK, 2006, 745 pp.

165. S. Vannitsem, *Dynamics and predictability of a low-order wind-driven ocean–atmosphere coupled model*, Clim. Dyn. **42** (2014), 1981–1998.

166. S. Vannitsem, J. Demaeyer, L. De Cruz, and M. Ghil, *A 24-variable low-order coupled ocean-atmosphere model: OA-QG-WS v2*, Geoscientific Model Development **7** (2014), 649–662.

167. ______, *Low-frequency variability and heat transport in a low-order nonlinear coupled ocean-atmosphere model*, Physica D **309** (2015), 71–85, doi:10.1016/j.physd.2015.07.006.

168. S. Vannitsem and M. Ghil, *Evidence of coupling in ocean–atmosphere dynamics over the North Atlantic*, Geophys. Res. Lett. **44** (2017), 2016–2026, doi:10.1002/2016GL072229.

169. E.R. Weeks, Y. Tian, J.S. Urbach, K. Ide, H.L. Swinney, and M. Ghil, *Transitions between blocked and zonal flows in a rotating annulus with topography*, Science **278** (1997), 1598–1601.

170. R.F. Williams, *The structure of Lorenz attractors*, Publ. Math. I.H.E.S. **50** (1979), 73–99.

171. C. Wunsch, *The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations*, Bull. Am. Meteorol. Soc. **80** (1999), 245–255.

172. L.S. Young, *What are SRB measures, and which dynamical systems have them?*, J. Stat. Phys. **106** (2002), 733–754.