

---

# 5

## Dynamics of Large-Scale Ocean Circulation

*George Veronis*

### 5.1 Introduction and Summary

The past 30 years have witnessed a rapid evolution of circulation theory. Much of the progress can be attributed to the intuition and physical balance that have emerged from the use of simple models that isolate important processes. Major contributions along these lines were made by Stommel, Welander, and others. An excellent presentation of the ideas together with a number of significant advances appears in Stern's (1975a) book. More recently numerical simulations have provided a different attack on the problem. Processes that are difficult to study with analytical models become accessible through the latter approach. Early, climatological-type studies by Bryan have now been supplemented by numerical models oriented toward the isolation of the effects of individual mechanisms. The papers of Rhines and Holland cited below have been especially instructive.

The development of the theory for the dynamics of large-scale oceanic flows is very recent. One has only to look at the chapter on dynamics in Sverdrup, Johnson, and Fleming (1942) to realize how primitive the theory was in the mid-1940s. Sverdrup's (1947) important demonstration of the generation of planetary vorticity by wind stress was the first step in obtaining explicit information about oceanic flow from a simple external observable. Until that time the dynamic method (i.e., geostrophic-hydrostatic balance) was used to obtain flow information, but this hardly constitutes a theory since one internal property must be used to determine another.

Ekman's (1905) theory for what we now call the Ekman layer was a significant early contribution, but its application to large-scale theory was not understood until Charney and Eliassen (1949) showed the coupling to large-scale flows via the spin-up mechanism. Actually, the generation of large-scale flow by Ekman suction in the laboratory was observed and described by Pettersson (1931), who repeated some of Ekman's (1906) early experiments with a stratified fluid to determine the inhibition of vertical momentum transport by stratification. Pettersson found the large-scale circulation to be an annoying interference, however, in his primary objective, determining vertical transfer of momentum by turbulence, and he discarded the approach as unpromising.

Shortly after Sverdrup's paper Stommel (1948) produced the first significant, closed-basin circulation model showing that westward intensification of oceanic flow is due to the variation of the Coriolis parameter with latitude. Hidaka (1949) proposed a closed set of equations for the circulation including the effects of lateral (eddy) dissipation of momentum. Munk (1950) continued the development by obtaining

a solution that resembled Stommel's except for details in the boundary layers near the eastern and western sides of the basin. He applied his solution to an idealized ocean basin with observed wind stresses and related a number of observed oceanic gyres to the driving wind patterns. The first nonlinear correction to these linearized models (Munk, Groves and Carrier, 1950) showed that inertia shifts positive vortices to the south and negative vortices to the north. Nonlinear effects thus introduce the observed north-south asymmetry into a circulation pattern that is predicted by steady linear theory to be symmetric about mid-latitude when the wind driving is symmetric.

Fofonoff (1954) approached the problem from the opposite extreme, treating a completely inertial, non-driven model. His solution exhibits the pure effect of inertia for steady westward flows. The circulation pattern is symmetric in the east-west direction and closes with the center of a cyclonic (anticyclonic) vortex at the south (north) edge of the basin. When linear, frictional effects perturb the nonlinear pattern (Niiler, 1966), the center of the vortex shifts westward. Niiler's model had been proposed independently by Veronis (1966b) after a numerical study of nonlinear effects in a barotropic ocean, and Niiler's solution had been suggested heuristically by Stommel (1965).

The theoretical models leading to these results for wind-driven circulation are discussed below in sections 5.5 and 5.6. More general considerations in section 5.2, based on conservation integrals for the nondissipative equations (Welander, 1971a), prepare the way for the ordered system of quasi-geostrophic equations that are presented in section 5.3. The latter are derived for a fluid with arbitrary stable stratification and for a two-layer approximation to the stratification.<sup>1</sup> A large portion of the remainder of the paper reports results obtained with the simpler two-layer system.<sup>2</sup>

Section 5.7 concludes the discussion of simple models of steady, wind-driven circulation with a suggested simple explanation of why the Gulf Stream and other western boundary currents leave the coast and flow out to sea (Parsons, 1969; Veronis, 1973a). Separation of the Gulf Stream from the coast occurs within an anticyclonic gyre at a latitude where the Ekman drift due to an eastward wind stress in the interior must be returned geostrophically in the western boundary layer. If the mean thermocline depth is sufficiently small, i.e., if the amount of upper-layer water is sufficiently limited, the thermocline surfaces on the onshore side of the Gulf Stream and separation occurs. The surfacing of the thermocline is enhanced by the poleward transport by the Gulf Stream of upper-layer water that eventually reaches polar latitudes and sinks.

A review of models of thermohaline circulation is given in section 5.8. The open models introduced by Welander (1959) and Robinson and Stommel (1959) and

the subsequent developments by them as well as other authors are described. The section concludes with a description of a closed, two-layer model in which the heating and cooling processes are parameterized by an assumed upwelling of lower-layer water across the thermocline (Veronis, 1978). The closure of the model leads to an evaluation of the magnitude of upwelling of  $1.5 \times 10^{-7} \text{ m s}^{-1}$ , in agreement with values obtained from chemical tracers and the estimated age of deep water.

The normal modes for a two-layer system are derived in section 5.9 and the free-wave solutions are obtained for an ocean of constant depth. The derivation is a generalization of the treatment by Veronis and Stommel (1956) but the method is basically the same. The results include barotropic and baroclinic modes of inertio-gravity and quasi-geostrophic Rossby waves. Brief mention is made of observations of these waves and the roles they play in developed flows.

Topography introduces a new class of long-period wave motions. Quasi-geostrophic analysis leads to the three types of waves described by Rhines (1970, 1977) as topographic-barotropic Rossby waves, fast baroclinic (bottom-trapped) waves, and slow baroclinic (surface-trapped) waves. The properties of slow baroclinic waves are independent of topography, yet the creation of these waves may be facilitated by steep topography that inhibits deep motions. For purposes of comparison the analysis is carried out with stratification approximated by two layers and by a vertically uniform density gradient.

Baroclinic instability in a two-layer system is described in section 5.11. The model (Phillips, 1951; Bretherton, 1966a) has convenient symmetries (equal layer depths and equal and opposite mean flows in the two layers) that simplify the analysis and show the nature of the instability more clearly. The stabilizing effect of  $\beta$  is evident after the simpler model has been analyzed. After a discussion of the energetics and of the relative phase of the upper- and lower-layer motions required for instability, the study of linear processes ends with a brief review of the stability study made by Gill, Green, and Simmons (1974) for a variety of mean oceanic conditions.

The last section extends the discussion to include the effects of turbulence and strong nonlinear interactions. Batchelor's (1953a) argument that two-dimensional turbulence leads to a red cascade in wavenumber space is followed by a description of several of Rhines's (1977) numerical experiments exhibiting the red cascade for barotropic quasi-geostrophic flow and the inhibition of the red cascade by lateral boundaries and topography. An initially turbulent flow in a two-layer fluid will evolve toward a barotropic state followed by the red cascade when nonlinear interactions or baro-

clinic instability generate motions on the scale of the internal radius of deformation. The latter scale is the window leading to barotropic behavior. Rough topography can inhibit the tendency toward barotropy by scattering the energy of the flow away from the deformation scale.

The generation of deep motions in wind-driven flows by upper-layer eddies that evolve from barotropic and baroclinic instabilities leads to a mean flow that is very different from the one predicted by the linear theories of the earlier sections. The closed-basin circulation obtained in a two-layer quasi-geostrophic numerical experiment by Holland (1978) and analyzed by Holland and Rhines (1980) shows how many of the processes described earlier come together to generate the mean flow. Simple balances for some of the results are suggested. A significant result of this experiment (and others mentioned) is the enhancement of the mean transport by the circulation resulting from the eddy interactions. A similar enhancement is made possible when topography and baroclinic effects are present (Holland, 1973). A brief discussion of several other numerical studies concludes the review.

Most of the emphasis in this paper is on linear processes and on the remaining features of the dynamics that can be used as building blocks to synthesize the involved, interactive flows observed in the ocean. Only a selected few of the many numerical studies that have emerged in the past few years are discussed, and even for those only some of the generalizable results are mentioned. Some important topics, such as the use of diagnostic models (Sarkisyan, 1977) and the generation of mean circulation by fluctuating winds (Pedlosky, 1964a; Veronis, 1970; Rhines, 1977), are omitted only because time limits forced me to draw the line somewhere. Most of the references are to the literature in the English language because that is the literature with which I am most familiar.

## 5.2 The Equations for Large-Scale Dynamics

The complete equations for conservation of momentum, heat, and salt are never used for studies of large-scale oceanic dynamics because they are much too complicated, not only for analytical studies but even for numerical analyses. Justification for use of an appropriate set of simplified equations requires a much more extensive argument than is feasible here so we shall confine ourselves to a short discussion with references to publications that discuss the different issues. It is appropriate, however, to mention a general result for a fluid with a simple equation of state.

If dissipative processes are ignored, the conservation of momentum for a fluid in a rotating system can be written as

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\frac{1}{\rho} \nabla p - \nabla \Phi, \quad (5.1)$$

or equivalently as

$$\begin{aligned} \frac{\partial \mathbf{v}}{\partial t} + (2\boldsymbol{\Omega} + \nabla \times \mathbf{v}) \times \mathbf{v} \\ = -\frac{1}{\rho} \nabla p - \nabla \left( \frac{1}{2} \mathbf{v} \cdot \mathbf{v} \right) - \nabla \Phi \end{aligned} \quad (5.2)$$

where  $\mathbf{v}$  is the three-dimensional velocity vector,  $\boldsymbol{\Omega}$  is the angular rotation vector of the system,  $\rho$  the density,  $p$  the pressure, and  $\nabla \Phi$  the total gravity term (Newtonian plus rotational acceleration).

Conservation of mass is described by

$$\frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{v} = 0, \quad \frac{d}{dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (5.3)$$

Furthermore, if a state variable  $s(p, \rho)$  is conserved along a trajectory, it satisfies the equation

$$\frac{ds}{dt} = 0. \quad (5.4)$$

These equations can be combined to yield the conservation of potential vorticity (Ertel, 1942):

$$\frac{d}{dt} \left[ \frac{(2\boldsymbol{\Omega} + \nabla \times \mathbf{v}) \cdot \nabla s}{\rho} \right] = 0. \quad (5.5)$$

This general result for a dissipation-free fluid does not apply precisely to sea water where the density is a function not only of temperature and pressure but also of the dissolved salts. The effect of salinity on density is very important in the distribution of water properties. However, for most dynamic studies the effect of the extra state variable is not significant and (5.5) is valid.

Circulation of waters in the world ocean involves trajectories from the surface to the deep sea and from one ocean basin to another. The relative densities of two parcels of water formed at the surface in different locations can be inverted when the parcels sink to great depths. Thus, surface water in the Greenland Sea is denser than surface water in the Weddell Sea, yet when these water masses sink and flow to the same geographic location, the latter (Antarctic Bottom Water) is denser and lies below the former (North Atlantic Deep Water). This inversion is due in large part to the different amounts of thermal expansion of waters of different temperatures and salinities.<sup>3</sup>

Neither compressibility nor individual effects of temperature and salinity on the density are included in the treatment that follows. Use of potential density (not only in the equations but in boundary conditions as well) together with the Boussinesq approximation (Spiegel and Veronis, 1960) makes it possible to treat the dynamic effects of buoyancy forces in a dynami-

cally consistent fashion. Comparison of observed motions (especially long- and short-period waves) with those deduced when potential density is used yields good qualitative, and often quantitative, agreement. But it is clear that some phenomena, such as the relative layering of water masses and small-scale mixing related to double-diffusive processes, cannot be analyzed without the use of a more extended thermodynamic analysis. Therefore, although the present discussion allows a treatment of inertially controlled flows, it does not admit the interesting array of phenomena associated with tracer distributions, except in the crudest sense. By implication, motions related to the largest time and space scales are not accessible either.

In those cases where a homogeneous fluid model is invoked the effects of stratification are implicitly present since the basic equations would be different for a truly homogeneous fluid (where the direction of the rotation axis could be more important than the local vertical). The fluid is sometimes assumed to be homogeneous only because the feature that is being emphasized is independent of stratification or because the simplified analytical treatment is a helpful preliminary for the more complicated stratified system.

The effects of rotation and Newtonian gravitation lead to an equilibrium shape for the earth that is nearly a planetary ellipsoid. For earth parameters the ellipticity is small (1/298) and an expansion in the ellipticity yields a spherical system with a mean (rather than variable) radius to lowest order (Veronis, 1973b). An additional simplification is to neglect the horizontal component of the earth's rotation. This assumption is not entirely separate from the use of a mean radius (N. A. Phillips, 1966a). It is normally valid for the types of motion treated here, though the effect of the neglected term is discussed for certain physical situations by Needler and LeBlond (1973) and by Stern (1975a). Grimshaw (1975) has reexamined the  $\beta$ -plane approximation and gives a procedure in which the horizontal rotation is retained.

With all these simplifications the foregoing equations simplify to

$$\frac{d\mathbf{v}}{dt} + \mathbf{f} \times \mathbf{v} = -\frac{1}{\rho_m} \nabla p - g \frac{\rho}{\rho_m} \hat{\mathbf{k}}, \quad (5.6)$$

$$\frac{d\rho}{dt} = 0, \quad (5.7)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (5.8)$$

$$\frac{d}{dt} [(\mathbf{f} + \nabla \times \mathbf{v}) \cdot \nabla \rho] \equiv \frac{dq}{dt} = 0, \quad (5.9)$$

where  $\mathbf{f} = 2\Omega \sin \phi \hat{\mathbf{k}}$  is twice the locally vertical (direction  $\hat{\mathbf{k}}$ ) component of the earth's rotation,  $\phi$  is the latitude,  $g$  is gravity,  $\rho_m$  is a mean (constant) density,

and  $\rho$  is the deviation of density from the mean. The hydrostatic pressure associated with the mean density has been subtracted from the system. Equations (5.7) and (5.8) describe the incompressible nature of this Boussinesq fluid. The quantity  $s$  in (5.4) can then be replaced by  $\rho$ , and the potential vorticity  $q$  in (5.9) is simplified accordingly [note the change of dimensions of potential vorticity as defined in (5.5) and (5.9)].

For steady or statistically steady flows we can multiply (5.6) by  $\mathbf{v}$  to obtain a kinetic energy equation which can be written as

$$\mathbf{v} \cdot \nabla \left( \frac{\rho_m}{2} \mathbf{v} \cdot \mathbf{v} + p + g\rho \right) \equiv \mathbf{v} \cdot \nabla B = 0, \quad (5.10)$$

where  $B$  is the Bernoulli function. In this case, since  $q$ ,  $\rho$ , and  $B$  are each conserved along flow paths, any one of them can be expressed in terms of the other two and we obtain

$$\rho = \rho(B, q), \quad B = B(\rho, q), \quad q = q(B, \rho). \quad (5.11)$$

Even though the distributions of the surfaces cannot be determined without knowledge of the flow field, the relationship between  $\rho$ ,  $B$  and  $q$  is conceptually useful.

The quantities  $B$ ,  $q$ , and  $\rho$  are specified by their values in certain source regions where dissipation, mixing, and other physical processes are important. (Obvious source regions are Ekman layers, areas of convective overturning, and boundary layers near coasts.) Having acquired values of  $B$ ,  $q$ , and  $\rho$  at the sources, fluid particles will retain these values along their flow paths. If particles from different sources and with different values of  $B$ ,  $q$ , and  $\rho$  converge to the same geographical location, regions of discontinuity will develop, and mixing, dissipation or some other non-ideal fluid process will be required. The locations of these discontinuous regions can be determined only from a solution to the general problem, and, in general, we may anticipate new sources of  $B$ ,  $q$ , and  $\rho$  to develop there. Hence, the system becomes a strongly implicit one and the closure of the problem is very complicated.

Even though a solution to the general problem may be impossible, these general considerations are important. We should be prepared for the likelihood that the solution at a particular location will not be simply determined by values at solid boundaries that are easily specified. The ocean is more likely a collection of dynamically self-contained pools (some subsurface) that interact along open-ocean boundaries where they join. Perhaps only the most persistent of these are statistically steady features. It is possible that locally the flow is relatively laminar. In that case the solution would be accessible once the source regions were identified and the values of  $B$ ,  $q$ , and  $\rho$  in these regions could be specified.

### 5.3 The Quasi-Geostrophic Equations and the $\beta$ -Plane

Even with the simplifications made in the previous section the equations are more general than required for a study of large-scale dynamics. We shall therefore simplify them further by invoking geostrophic and hydrostatic balances at lowest order and by restricting attention to spatial scales of interest. In so doing we shall derive an appropriate  $\beta$ -plane approximation for the study of oceanic waves and mesoscale motions. A similar procedure is followed by N. A. Phillips (1963).<sup>4</sup>

#### 5.3.1 Continuous Stratification

The spherical components of (5.6) take the form

$$\frac{du}{dt} - \frac{uv \tan \phi}{a} + \frac{uw}{a} - 2\Omega \sin \phi v = -\frac{1}{a \cos \phi} \frac{dP}{d\lambda}, \quad (5.12)$$

$$\frac{dv}{dt} + \frac{u^2 \tan \phi}{a} + \frac{vw}{a} + 2\Omega \sin \phi u = -\frac{1}{a} \frac{\partial P}{\partial \phi}, \quad (5.13)$$

$$\frac{dw}{dt} - \frac{u^2 + v^2}{a} = -\frac{\partial P}{\partial z} - g \frac{\rho}{\rho_m}, \quad (5.14)$$

$$\frac{1}{a \cos \phi} \left[ \frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \phi} (v \cos \phi) \right] + \frac{\partial w}{\partial z} + \frac{2w}{a} = 0, \quad (5.15)$$

$$\frac{d\rho}{dt} = 0, \quad (5.16)$$

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \frac{u}{a \cos \phi} \frac{\partial}{\partial \lambda} + \frac{v}{a} \frac{\partial}{\partial \phi} + w \frac{\partial}{\partial z}, \quad (5.17)$$

where  $(\lambda, \phi, z)$  are longitude, latitude, and upward and have respective velocities  $(u, v, w)$ ;  $P$  is  $p/\rho_m$ ,  $a$  is the mean radius of the earth, and  $\rho$  is the total density minus  $\rho_m$ .

Center attention on a latitude  $\phi_0$ , write  $\phi = \phi_0 + \phi'$ , and consider flows with north-south scale  $L$  substantially smaller than  $a$ . Then with  $a\phi' = y$ , we can expand the trigonometric functions in  $y$ , keeping only terms of  $O(L/a)$ , to obtain

$$\begin{aligned} \sin \phi &\approx \sin \phi_0 (1 + \cot \phi_0 y/a), \\ \cos \phi &\approx \cos \phi_0 (1 - \tan \phi_0 y/a), \end{aligned} \quad (5.18)$$

$$\begin{aligned} f_0 &= 2\Omega \sin \phi_0, \\ \frac{\partial}{\partial x} &\equiv \frac{1}{a \cos \phi_0} \frac{\partial}{\partial \lambda}, \quad \frac{\partial}{\partial y} \equiv \frac{1}{a} \frac{\partial}{\partial \phi}. \end{aligned} \quad (5.19)$$

To first order in  $y/a$  the equations become

$$\frac{du}{dt} + \frac{y}{a} \tan \phi_0 u \frac{\partial u}{\partial x} + \frac{uw}{a}$$

$$\begin{aligned} &- \frac{uv}{a} \tan \phi_0 \left( 1 + \frac{2}{\sin 2\phi_0} \frac{y}{a} \right) - f_0 v \left( 1 + \frac{y}{a} \cot \phi_0 \right) \\ &= -\frac{\partial P}{\partial x} \left( 1 + \frac{y}{a} \tan \phi_0 \right), \end{aligned} \quad (5.20)$$

$$\begin{aligned} &\frac{dv}{dt} + \frac{y}{a} \tan \phi_0 u \frac{\partial v}{\partial x} + \frac{vw}{a} \\ &+ \frac{u^2}{a} \tan \phi_0 \left( 1 + \frac{2}{\sin 2\phi_0} \frac{y}{a} \right) + f_0 u \left( 1 + \frac{y}{a} \cot \phi_0 \right) \\ &= -\frac{\partial P}{\partial y}, \end{aligned} \quad (5.21)$$

$$\frac{dw}{dt} + \frac{y}{a} \tan \phi_0 u \frac{\partial w}{\partial x} - \frac{u^2 + v^2}{a} = -\frac{\partial P}{\partial z} - \frac{g\rho}{\rho_m}, \quad (5.22)$$

$$\begin{aligned} &\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} - \frac{\partial}{\partial y} \left( \frac{y}{a} v \tan \phi_0 \right) \\ &- \frac{y}{a} \tan \phi_0 \frac{\partial w}{\partial z} + \frac{2w}{a} = 0, \end{aligned} \quad (5.23)$$

$$\frac{d\rho}{dt} + \frac{y}{a} \tan \phi_0 u \frac{\partial \rho}{\partial x} = 0, \quad (5.24)$$

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z}. \quad (5.25)$$

Flows with a primary geostrophic balance will satisfy

$$f_0 v \sim \frac{\partial P}{\partial x}, \quad -f_0 u \sim \frac{\partial P}{\partial y}. \quad (5.26)$$

Hydrostatic balance yields

$$\frac{\partial P}{\partial z} \sim -g \frac{\rho}{\rho_m}. \quad (5.27)$$

Variations over the depth  $H$  of the ocean are described by

$$\frac{\partial}{\partial z} \sim \frac{1}{H}, \quad (5.28)$$

so the "pressure" scale derived from (5.27) is

$$P \sim \frac{gH \Delta \rho}{\rho_m}. \quad (5.29)$$

Geostrophic balance then suggests the velocity scale

$$V \sim \frac{gH \Delta \rho}{f_0 L \rho_m}. \quad (5.30)$$

If these scales are used as orders of magnitudes for the respective variables and if we also take

$$\frac{\partial}{\partial x} \sim \frac{1}{L}, \quad \frac{\partial}{\partial y} \sim \frac{1}{L}, \quad \frac{\partial}{\partial t} \sim f_0,$$

$$\frac{y}{a} \sim \frac{L}{a}, \quad \delta = \frac{H}{L}, \quad w \sim V\delta,$$

we note the following.

Relative to the lowest order (in  $y/a$ ) Coriolis terms, the nonlinear terms in  $d/dt$  in (5.19) and (5.20) are  $O(Ro)$  where  $Ro = V/f_0L$ . The remaining nonlinear terms are

$$O\left(\frac{L}{a} Ro\right) \quad \text{or} \quad O\left(\delta \frac{L}{a} Ro\right).$$

In the vertical equation of motion the acceleration terms are

$$O(\delta^2), \quad O(\delta^2 Ro), \quad \text{or} \quad O\left(\delta \frac{L}{a} Ro\right)$$

when compared to the terms on the right. Observations of the flows of interest support the inequalities

$$Ro \ll 1, \quad L/a \ll 1, \quad \delta \ll 1. \quad (5.31)$$

Rather than expand the equations in powers of the small parameters we shall simply make use of (5.31) and drop all terms which involve products of  $Ro$ ,  $\delta$  and  $L/a$ . Also, rather than give a relative ordering of these three parameters we keep all terms up to first order in  $Ro$ ,  $\delta$ , and  $L/a$ , a procedure that yields the following general system of equations

$$\begin{aligned} \frac{du}{dt} - f_0 v \left(1 + \frac{y}{a} \cot \phi_0\right) \\ = -\frac{\partial P}{\partial x} \left(1 + \frac{y}{a} \tan \phi_0\right), \end{aligned} \quad (5.32)$$

$$\frac{dv}{dt} + f_0 u \left(1 + \frac{y}{a} \cot \phi_0\right) = -\frac{\partial P}{\partial y}, \quad (5.33)$$

$$\frac{\partial P}{\partial z} = -g\rho/\rho_m, \quad (5.34)$$

$$\begin{aligned} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} - \frac{\partial}{\partial y} \left(\frac{y}{a} v \tan \phi_0\right) - \frac{y}{a} \tan \phi_0 \frac{\partial w}{\partial z} \\ = 0, \end{aligned} \quad (5.35)$$

$$\frac{d\rho}{dt} + \frac{y}{a} \tan \phi_0 u \frac{\partial \rho}{\partial x} = 0, \quad (5.36)$$

keeping in mind that the nonlinear terms in (5.32) and (5.33) are  $O(Ro)$  compared to the lowest-order Coriolis terms.

Now write

$$\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1, \quad P = P_0 + P_1, \quad \rho = \rho_0 + \rho_1, \quad (5.37)$$

where  $(\mathbf{v}_1, P_1, \rho_1)$  are  $O(Ro)$  or  $O(L/a)$ . We shall also assume that time variations appear at first order, i.e.,  $\partial/\partial t = O(Ro)$  or  $O(L/a)$ . Then at lowest order we obtain the expected geostrophic hydrostatic system:

$$f_0 v_0 = \frac{\partial P_0}{\partial x}, \quad (5.38)$$

$$f_0 u_0 = -\frac{\partial P_0}{\partial y}, \quad (5.39)$$

$$\frac{\partial P_0}{\partial z} = -g\rho_0/\rho_m, \quad (5.40)$$

$$\frac{\partial u_0}{\partial x} + \frac{\partial v_0}{\partial y} = 0, \quad \frac{\partial w_0}{\partial z} = 0, \quad (5.41)$$

If  $w_0$  vanishes at any level or if it is required to satisfy inconsistent (with  $\partial w_0/\partial z = 0$ ) boundary conditions at top and bottom, it will vanish everywhere. One or the other of these two conditions is satisfied for all of the flows that we shall consider, so we obtain the result

$$w_0 = 0. \quad (5.42)$$

This means that the scaling  $w \sim V\delta$  suggested by the geometry is inappropriate and that a factor  $L/a$  or  $Ro$  should be included on the right-hand side. In other words, quasi-geostrophic flows are quasi-horizontal and the convective derivative in (5.32) reduces to

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u_0 \frac{\partial}{\partial x} + v_0 \frac{\partial}{\partial y}. \quad (5.43)$$

The restriction to flows with less than global scales precludes a treatment leading to the basic stratification. Since vertical density changes  $\Delta\rho$  are generally much larger than the horizontal changes, say  $\Delta\rho'$ , generated by the motion field, we must account for the difference in (5.36). In particular, we write  $\rho = \bar{\rho}(z) + \rho'(x, y, z, t)$  so that

$$\begin{aligned} \frac{\partial \rho'}{\partial t} + u \frac{\partial \rho'}{\partial x} + v \frac{\partial \rho'}{\partial y} + w \frac{\partial \rho'}{\partial z} \\ + w \frac{\partial \bar{\rho}}{\partial z} + \frac{y}{a} \tan \phi_0 u \frac{\partial \rho'}{\partial x} = 0. \end{aligned} \quad (5.44)$$

The considerations leading to (5.43) apply here as well for the terms involving  $\rho'$ . Accordingly, at lowest order we can drop the terms  $w \partial \rho'/\partial z$  and  $(y/a) \tan \phi_0 u (\partial \rho'/\partial x)$  to end up with

$$\frac{d\rho'}{dt} + w \frac{\partial \bar{\rho}}{\partial z} = 0, \quad (5.45)$$

where  $\Delta\rho'$  is assumed to be  $O(Ro)$  or  $O(L/a)$  relative to  $\Delta\rho$ . Since  $w$  is correspondingly smaller than  $u$  or  $v$ , the two terms balance. In terms of our ordering, therefore, we can write

$$\frac{d\rho_0}{dt} + w_1 \frac{\partial \bar{\rho}}{\partial z} = 0, \quad (5.46)$$

where we have used the fact that the density used in the hydrostatic equation is really  $\rho'$  (since the balance  $\partial \bar{P}/\partial z = -g\bar{\rho}$  is valid when there is no motion and hence can be subtracted from the system).

At next order we have

$$\begin{aligned} \frac{du_0}{dt} - f_0 v_1 - \frac{y}{a} f_0 \cot \phi_0 v_0 \\ = -\frac{\partial P_1}{\partial x} - \frac{y}{a} \tan \phi_0 \frac{\partial P_0}{\partial x}, \end{aligned} \quad (5.47)$$

$$\frac{dv_0}{dt} + f_0 u_1 + \frac{y}{a} f_0 \cot \phi_0 u_0 = -\frac{\partial P_1}{\partial y}, \quad (5.48)$$

$$\frac{\partial u_1}{\partial x} + \frac{\partial v_1}{\partial y} + \frac{\partial w_1}{\partial z} - \frac{\partial}{\partial y} \left( \frac{y}{a} \tan \phi_0 v_0 \right) = 0. \quad (5.49)$$

These equations, in addition to providing the balances for first-order quantities, serve the important function of closing the zero-order system when first-order terms are eliminated. Thus, cross differentiating (5.47) and (5.48) and making use of (5.38) to (5.41) and (5.49), we obtain

$$\begin{aligned} \frac{d\zeta_0}{dt} + \beta v_0 = f_0 \frac{\partial w_1}{\partial z}, \quad \beta = \frac{f_0 \cot \phi_0}{a}, \\ \zeta_0 = \frac{\partial v_0}{\partial x} - \frac{\partial u_0}{\partial y}. \end{aligned} \quad (5.50)$$

From (5.46) we observe

$$\frac{\partial w_1}{\partial z} = -\frac{\partial}{\partial z} \left[ \frac{1}{\partial \bar{\rho} / \partial z} \frac{d\rho_0}{dt} \right] = -\frac{d}{dt} \frac{\partial}{\partial z} \left[ \rho_0 / \frac{\partial \bar{\rho}}{\partial z} \right]. \quad (5.51)$$

Also,

$$\beta v_0 = \frac{df}{dt}, \quad (5.52)$$

where  $f = f_0(1 + \cot \phi_0 (y/a))$ . Then using (5.38) to (5.41) to express  $u_0, v_0, \rho_0$  in terms of  $p_0$  we obtain the lowest-order closure

$$\frac{d}{dt} \left[ \nabla^2 p_0 + ff_0 + \frac{\partial}{\partial z} \left( \frac{f_0^2 \partial p_0 / \partial z}{N^2} \right) \right] = 0, \quad (5.53)$$

where  $N^2 = -g(\partial \bar{\rho} / \partial z) / \rho_m$  is the square of the buoyancy frequency. Equation (5.53) describes the conservation of quasi-geostrophic potential vorticity. It is sometimes written in terms of the stream functions  $\psi = P_0 / f_0$ ,

$$\frac{d}{dt} \left[ \nabla^2 \psi + f + \frac{\partial}{\partial z} \left( \frac{f_0^2 \partial \psi / \partial z}{N^2} \right) \right] = 0. \quad (5.54)$$

The derivation given here has been carried out in dimensional form. It is as rigorous, though not as formal, as derivations with nondimensional variables (e.g., Pedlosky, 1964a) and has the advantage of including the intermediate equations in dimensional form. Obviously, the equations are valid only for those motions (smaller than global scale, low frequency, etc.) that satisfy the assumptions.

### 5.3.2 Equations More Commonly Encountered

Instead of the set (5.47)–(5.49) one more often encounters the equations with rectangular cartesian coordinates, no subscripts, and with  $f = f_0 + \beta y$ , i.e.,

$$\frac{du}{dt} - fv = -\frac{\partial P}{\partial x}, \quad (5.55)$$

$$\frac{dv}{dt} + fu = -\frac{\partial P}{\partial y}, \quad (5.56)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (5.57)$$

This system is often used even when the flow is not quasi-geostrophic. For quasi-geostrophic flows, particularly if one makes use principally of the vorticity equation and the fact that  $w$  is really a higher order quantity, one can avoid serious errors.

For flows at low latitudes (small  $\phi_0$ ) the neglected terms ( $\sim \tan \phi_0$ ) are small and (5.55) to (5.57) may be adequate. But errors notwithstanding, a large part of the literature deals with this more approximate system, and we shall have to refer to it frequently.

### 5.3.3 Layered Stratification

Continuous density stratification is frequently approximated by a series of discrete layers each of uniform density. The derivation parallels the one just given but it is easier to make use of what we have done and to note the following.

Number the layers sequentially downward from the top so that the  $i$ th layer has thickness  $h^{(i)}$ , density  $\rho^{(i)}$ , and mean thickness (for linearized cases)  $H^{(i)}$ . Furthermore, write  $h^{(i)} = \eta^{(i)} + H^{(i)} - \eta^{(i+1)}$  so that  $\eta^{(i)}$  and  $\eta^{(i+1)}$  are the deviations of the top and bottom surfaces of the layer from the mean. Integrate the hydrostatic relation downward from the top surface to layer  $i$  to derive the horizontal pressure gradient in terms of gradients of thicknesses

$$\nabla p^{(i)} = g \sum_{n=1}^{i-1} \rho^{(n)} \nabla h^{(n)} + g \rho^{(i)} \nabla \eta^{(i)}. \quad (5.58)$$

Conservation of mass for each homogeneous layer is  $\nabla_3 \cdot \mathbf{v}^{(i)} = 0$ , where  $\nabla_3$  is the three-dimensional operator. The horizontal velocities are independent of  $z$  because the flow is hydrostatic. Therefore, integrating over the depth of the layer yields

$$h^{(i)} \nabla \cdot \mathbf{v}^{(i)} + \frac{dh^{(i)}}{dt} = 0, \quad (5.59)$$

where we have used the free surface conditions

$$\begin{aligned} w^{(i)}(x, y, \eta^{(i)}, t) &= \frac{d\eta^{(i)}}{dt}, \\ w^{(i)}(x, y, \eta^{(i+1)}, t) &= \frac{d\eta^{(i+1)}}{dt}. \end{aligned} \quad (5.60)$$

Also, since  $H^{(i)}$  is constant

$$\frac{d}{dt} (\eta^{(i)} - \eta^{(i+1)}) = \frac{dh^{(i)}}{dt}.$$

We can thus integrate (5.50) over the depth of each layer to obtain the conservation of potential vorticity for the layered system

$$\frac{d}{dt} \left( \frac{\zeta_0 + f}{h_0^{(0)}} \right) = 0. \quad (5.61)$$

The velocity in the convective derivative is  $(u_0^{(0)}, v_0^{(0)})$ .

In subsequent treatments of the two-layer,  $\beta$ -plane, inviscid, momentum equations, we shall use the approximate form (5.55) and (5.56) together with the vertically integrated form of (5.57). The equations are

$$\frac{du_1}{dt} - fv_1 = -g \frac{\partial \eta_1}{\partial x}, \quad (5.62)$$

$$\frac{dv_1}{dt} + fu_1 = -g \frac{\partial \eta_1}{\partial y}, \quad (5.63)$$

$$\frac{dh_1}{dt} + h_1 \left( \frac{\partial u_1}{\partial x} + \frac{\partial v_1}{\partial y} \right) = 0, \quad (5.64)$$

$$\frac{du_2}{dt} - fv_2 = -g \left[ (1 - \epsilon) \frac{\partial \eta_1}{\partial x} + \epsilon \frac{\partial \eta_2}{\partial x} \right], \quad (5.65)$$

$$\frac{dv_2}{dt} + fu_2 = -g \left[ (1 - \epsilon) \frac{\partial \eta_1}{\partial y} + \epsilon \frac{\partial \eta_2}{\partial y} \right], \quad (5.66)$$

$$\frac{dh_2}{dt} + h_2 \left( \frac{\partial u_2}{\partial x} + \frac{\partial v_2}{\partial y} \right) = 0, \quad (5.67)$$

where,  $\epsilon \rho_2 = \rho_2 - \rho_1$ ,  $h_2 = \eta_2 + H_2 - \eta_3$ , and  $\eta_3$  is the height of the bottom above an equilibrium level. The subscripts in (5.62) to (5.67) identify the layer rather than the order of  $L/a$  or  $Ro$ .

For linear steady flows the above system is sometimes used with spherical coordinates.

## 5.4 Ekman Layers

The equations derived above do not contain friction explicitly. However, when the variables are written in terms of a mean (ensemble, time average, etc.) plus a fluctuation and the equations are averaged, Reynolds stresses emerge and these are often parameterized in frictional form through the use of Austausch or eddy coefficients. Though this procedure is often questionable, it may not be a bad approximation near the top surface where wind stresses impart momentum to the ocean and near the bottom where frictional retardation brakes the flow. This was the view taken by Ekman (1905), who introduced the model for what is now called the Ekman layer.<sup>5</sup>

### 5.4.1 Pure Ekman Layers

Ekman first applied the theory to the wind-driven layer near the surface of the ocean. It is preferable to introduce the subject by investigating how a horizontally uniform geostrophic flow given by  $f\mathbf{k} \times \mathbf{v}_g = -\nabla P$  in a fluid occupying the region  $z > 0$  is brought to rest by

frictional processes acting near the bottom.<sup>6</sup> Assuming that horizontal variations of the stress are small compared to vertical variations (easily verified a posteriori), we can write

$$-fv = -fv_g + \nu u_{zz}, \quad (5.68)$$

$$fu = fu_g + \nu v_{zz}, \quad (5.69)$$

where subscript  $z$  corresponds to  $\partial/\partial z$  and the pressure gradient is written in terms of the geostrophic velocity. The velocity vanishes at the (flat) bottom

$$\mathbf{v} = 0 \quad \text{at} \quad z = 0. \quad (5.70)$$

The method of solution is well-known (Lamb, 1932, p. 593). Combining  $u$  and  $v$  as  $u + iv$ ,  $i = \sqrt{-1}$ , the equations (5.68) and (5.69) take the form

$$(u + iv)_{zz} = \frac{if}{\nu} (u + iv) - \frac{if}{\nu} (u_g + v_g) \quad (5.71)$$

and the solution satisfying (5.70) with  $\mathbf{v} \rightarrow 0$  as  $z \rightarrow \infty$  is

$$u + iv = (u_g + iv_g)(1 - e^{-\sqrt{f}\delta z}), \quad (5.72)$$

where  $\delta = \sqrt{f/\nu}$ . Accordingly, the flow vanishes at  $z = 0$ , tends to  $\mathbf{v}_g$  for large  $z$  and is predominantly to the left of  $\mathbf{v}_g$  in between.

The vertically integrated transport of the exponentially decaying part of (5.71) is  $(-1 + i)(u_g + iv_g)(\nu/2f)^{1/2}$ , which suggests  $h_e = (\nu/2f)^{1/2}$  as the scale of the Ekman layer. If we integrate the geostrophic part over the depth,  $h_e$ , we obtain the transport  $(u_g + iv_g)h_e$ . Hence, the net transport is  $i(u_g + iv_g)h_e$ , which is to the left of the geostrophic current, i.e., down the pressure gradient required to support  $\mathbf{v}_g$ , as we would expect. In vector form the net transport is  $(-\mathbf{v}_g, u_g)h_e$ .

Next consider Ekman's problem, with fluid occupying the region  $z < 0$  and with the flow driven by the spatially uniform wind stress (divided by the density) given by  $(\tau^x, \tau^y)$  acting at  $z = 0$ . With  $\mathbf{v}_g = 0$ , the solution is

$$u + iv = \frac{e^{\delta z}}{\sqrt{f\nu}} \{ \tau^x \sin(\delta z + \pi/4) - \tau^y \sin(\delta z - \pi/4) + i[\tau^x \sin(\delta z - \pi/4) + \tau^y \sin(\delta z + \pi/4)] \}. \quad (5.73)$$

In the hodograph  $(u, v)$ -plane the solution has the form of a spiral (called the Ekman spiral). Just as rotation generates a velocity component to the right (for  $f > 0$ ) of the (pressure) force for geostrophically balanced flow, a flow to the right of the tangential-stress force is generated in the Ekman spiral solution. In contrast to geostrophic flow, however, the present system is dissipative, and a velocity component parallel to the force is also present. At the surface the magnitudes of the components are equal so the flow is directed  $45^\circ$  to the right of the wind stress. The velocity component par-

allel to  $\tau$  decreases with depth but near the surface the normal component does not (it cannot since  $\tau$  has the same direction as  $\partial\tau/\partial z$ ). But below that the stress veers to the right as does the velocity vector.

Though Ekman's solution provided a satisfactory explanation of Nansen's observation of surface velocity, the spiral is not normally observed in the field. Ekman failed to observe it in spite of repeated attempts. Hunkins (1966) reported measuring a well-defined Ekman spiral (ironically, in the Arctic Ocean, where Nansen's first observations were made). The spiral structure depends on the form of the stress term, and since the stresses near the surface are turbulent (due to thermal convection, surface waves, and other small-scale processes) and therefore not necessarily of Navier-Stokes form, it is not surprising that the observed current structure differs from the theoretical one. Also, the mixed layer at the surface sits on a stably stratified fluid and the depth  $h$  of the former often does not exceed  $h_e$  when a turbulent eddy viscosity is used. Gonella (1971) showed that when a stress-free condition is applied at the base of the mixed-layer the solution is a function of  $h_e/h$ . For shallow ( $h \ll h_e$ ) mixed layers there is essentially no spiral. Csanady (1972) reported that field measurements in the mixed layer in Lake Huron support Gonella's findings. He also reformulated the problem in terms of external parameters of the system instead of using an eddy viscosity.

In contrast to the detailed velocity structure, the vertically integrated transport of the wind-driven Ekman layer is independent of the form of vertical variation of the stress. If the stress terms in (5.68) and (5.69) are written as  $\partial\tau/\partial z$  and if we integrate the equations vertically, the transports are given by  $(\tau^y, -\tau^x)/f$ . Thus, the total transport is to the right of the wind stress irrespective of the form of  $\tau$  and subject only to these conditions:  $\tau = (\tau^x, \tau^y)$  at the surface and  $\tau = 0$  at the bottom. In vector form, with  $V_e = \int_{-h}^0 \mathbf{v} dz$  (where  $h$  is a depth—finite or infinite—at which  $\tau$  vanishes), the result (called the Ekman drift) is

$$\mathbf{V}_e = (\boldsymbol{\tau} \times \hat{\mathbf{k}})/f. \quad (5.74)$$

where  $\boldsymbol{\tau}$  is now the wind stress vector and  $\hat{\mathbf{k}}$  is the vertical unit vector.

#### 5.4.2 Effect of Ekman Layers on Interior Flows

Although the pure Ekman layer theory given above requires horizontally uniform conditions, the theory is valid with horizontal variations as long as the horizontal scale is substantially larger than  $h_e$ . The neglected horizontal variations of the stress terms are smaller than  $\partial\tau/\partial z$  by the ratio of the squares of vertical to horizontal scales. Furthermore, for the mixed layer near the surface the vertical pressure gradient in the vertical equation of motion vanishes (as long as we consider scales larger than the small-scale turbulence

which generates the mixed layer). Hence, the horizontal pressure gradients associated with Ekman layer processes are negligible at lowest order, and the original equations, and therefore the results given by (5.73), are still applicable.

Accordingly, suppose that  $\boldsymbol{\tau}$  in (5.74) varies horizontally. When the continuity equation  $\nabla \cdot \mathbf{v} = 0$  is integrated in the vertical over the depth of the Ekman layer and the boundary condition ( $w = 0$  at the top) is applied, we find (Charney, 1955a)

$$w_e = \nabla \cdot \mathbf{V}_e, \quad (5.75)$$

where  $w_e$  is the vertical velocity at the base of the Ekman layer. With (5.74) this becomes

$$w_e = \frac{\partial}{\partial x} \left( \frac{\tau^y}{f} \right) - \frac{\partial}{\partial y} \left( \frac{\tau^x}{f} \right) = \hat{\mathbf{k}} \cdot \nabla \times \left( \frac{\boldsymbol{\tau}}{f} \right). \quad (5.76)$$

Thus, horizontal variations in  $\boldsymbol{\tau}$  generate vertical motions which penetrate into the fluid below. Since the Ekman layer is thin relative to the depth of the ocean, this forced vertical velocity (called *Ekman pumping*) can be applied as a boundary condition (approximately at the surface) for the underlying inviscid fluid.

The same analysis can be applied to the bottom (subscript b) Ekman layer, where the vertically integrated transport was found to be  $\mathbf{V}_b = (-v_g, u_g)h_b$ . If the bottom is flat, so that  $w = 0$  there, the vertically integrated continuity equation yields

$$w_b = -\nabla \cdot \mathbf{V}_b,$$

where  $w_b$  is the vertical velocity induced at the top of the Ekman layer. Substituting for  $\mathbf{V}_b$  we obtain

$$w_b = \frac{\partial}{\partial x} (v_g h_b) - \frac{\partial}{\partial y} (u_g h_b)$$

or taking  $h_b$  constant,

$$w_b = h_b \hat{\mathbf{k}} \cdot \nabla \times \mathbf{v}_g. \quad (5.77)$$

This value for  $w$  serves as a boundary condition (approximately at the bottom of the ocean) for the overlying inviscid fluid.

#### 5.4.3 Additional Considerations

Only the simplest results of Ekman layer theory have been given here. A number of important extensions are discussed by Stern (1975a, chapters 7 and 8). Horizontal momentum is imparted to the ocean by the wind stress acting at the surface; yet the momentum vanishes at the base of the Ekman layer. Stern answers the question where the momentum goes by analyzing the angular momentum balance about the axis of rotation for a cylindrical system. The analysis is carried out in an inertial frame of reference where the torque of the wind stress is balanced by the absolute angular momentum of the fluid. The latter is proportional to the absolute

vorticity of the undisturbed (no wind stress) vortex flow (in our case, solid-body rotation). The correspondence between the cylindrical problem and the rectilinear system [with (5.74) as the result] is that for large radii the angular momentum argument is equivalent to saying that the rate of momentum imparted by the wind stress is balanced by the divergence of the radial flux of absolute azimuthal momentum.

Though the Ekman layer depth  $h_e$  is clearly defined for laminar boundary layers, the value for turbulent boundary layers is not. Caldwell, van Atta, and Holland (1972) formed the boundary layer scale  $\tau^{1/2}/f$  from the (only) external parameters  $\tau$  and  $f$ . Assuming that the molecular scale  $(\nu/f)^{1/2}$  is not likely to affect the turbulent scale, they suggest that  $\tau^{1/2}/f$  is the turbulent Ekman boundary layer thickness. Stern (1975a, §8.1) carried out a crude stability analysis to conclude that a layer thicker than  $h_e \sim \tau^{1/2}/f$  will radiate energy to the deep water. He surmised that nonlinear modifications will show that the turbulent energy is thereby reduced as the thickness shrinks to  $\tau^{1/2}/f$ , where the system will stabilize. For typical values of  $\tau$ , the value of  $h_e$  (so defined) is  $O(100 \text{ m})$  at mid-latitudes. These considerations are based on the assumption of a homogeneous fluid. For a stratified fluid like the ocean the stratification may be decisive in determining the boundary layer thickness as Csanady's (1972) report of observed velocities in Lake Huron indicates.

As we saw from the simple analysis presented above, the effect of the top Ekman layer on the underlying water is determined completely by the wind stresses, whereas in the bottom Ekman layer the condition is expressed in terms of the velocity of the overlying water. More generally there will be a nonlinear coupling between the Ekman layer and the interior which can alter the results significantly. Fettis (1955) carried out the analysis for a laboratory model of a nonlinear Ekman layer to show that the results can be approximated by (5.74) but with the absolute vertical vorticity replacing  $f$ . Stern (1966; 1975a, §8.3) and Niiler (1969) have investigated the effect of coupling of Ekman layer flow with geostrophic vorticity (eddies) and have shown that the latter can have a dominant influence since coupling with the interior can occur even for a uniform wind stress.

## 5.5 Steady Linear Models of the Wind-Driven Circulation

For steady, linear flow of moderate scale we have  $Ro \ll L/a$  so the term  $d\zeta_0/dt$  in (5.50) can be neglected. The resulting equation is

$$\beta v_0 = f_0 \frac{\partial w_1}{\partial z}. \quad (5.78)$$

When integrated vertically from  $z = -h$  to  $z = 0$  this yields

$$\beta V = f_0 w_1 \Big|_{-h}^0, \quad V = \int_{-h}^0 v_0 dz$$

or

$$\beta V = \hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} - f_0 w_1(x, y, -h), \quad (5.79)$$

where the variation of  $f$  in  $\hat{\mathbf{k}} \cdot \nabla \times (\boldsymbol{\tau}/f)$  is (consistently) neglected at lowest order.

### 5.5.1 Sverdrup Transport

If the stratification is strong enough so that distortion of the density surfaces is negligible at some depth above the bottom, the last term in (5.79) vanishes and we obtain the Sverdrup transport

$$\beta V = \hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau}. \quad (5.80)$$

Thus, the vertically integrated north-south transport is determined by the curl of the wind stress. Sverdrup (1947) introduced this relation to estimate transports in the eastern equatorial Pacific (see chapter 6). Physically, the interpretation of (5.80) is straightforward. With  $\beta V$  written as  $h df/dt$  we see that a column of fluid moves to a new latitude (new value of planetary vorticity  $f$ ) with a speed that compensates for the rate at which the wind stress imparts vorticity to the ocean.

The continuity equation (5.41) can be integrated in the vertical and in  $x$  to yield

$$U = -f \frac{\partial V}{\partial y} dx + F(y)$$

or

$$U = -f \frac{1}{\beta} \frac{\partial}{\partial y} (\hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau}) dx + F(y), \quad (5.81)$$

where  $F(y)$  is arbitrary. The most common procedure for theoretical analyses is to assume that the foregoing is valid eastward to a meridional boundary  $x = L$ , where  $U$  must vanish. Then

$$U = \int_x^L \frac{1}{\beta} \frac{\partial}{\partial y} (\hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau}) dx, \quad (5.82)$$

and the transport is determined in the entire region in which the assumptions are valid. In general, the theory does not determine the flow in a basin bounded on the west as well since it is not possible to satisfy the zero normal flow condition there.

### 5.5.2 Stommel's Frictional Model

If the fluid motion penetrates to the (flat) bottom, the last term in (5.79) is given by (5.75) with  $\mathbf{v}_g = \mathbf{v}_0|_{z=-h}$  and (5.79) becomes

$$\beta V = \hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} - f_0 h_b \left( \frac{\partial v_0}{\partial x} - \frac{\partial u_0}{\partial y} \right)_{z=-h}. \quad (5.83)$$

Thus, one must supplement this equation with additional ones that determine the vertical structure of the velocity field. However, if the fluid is assumed to be homogeneous so that  $v_0$  is independent of  $z$  and if one then writes

$$(U, V) = (u_0, v_0)h,$$

the system closes with

$$\beta V = \hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} - \frac{f_0 h_b}{h} \left( \frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} \right). \quad (5.84)$$

Introducing the transport stream function  $\mathbf{V} = \mathbf{k} \times \nabla \psi$ , yields

$$\beta \psi_x = \hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} - K \nabla^2 \psi, \quad (5.85)$$

where  $K = f_0 h_b / h$ . Stommel (1948) obtained (5.85) by assuming a bottom drag law for the friction term. The derivation using Ekman layer theory makes the assumptions more evident.

The solution with  $\tau^x = -T \cos(\pi y / M)$ ,  $\tau^y = 0$  and with  $\psi = 0$  at  $x = 0, L$  and  $y = 0, M$  is

$$\psi = \frac{MT}{\pi K} \left\{ 1 - \frac{(1 - e^{D_2 L})e^{D_1 x} - (1 - e^{D_1 L})e^{D_2 x}}{e^{D_1 L} - e^{D_2 L}} \right\} \times \sin \frac{\pi y}{M}, \quad (5.86)$$

where

$$D_1 = -\frac{\beta}{2K} + \sqrt{\left(\frac{\beta}{2K}\right)^2 + \left(\frac{\pi}{M}\right)^2},$$

$$D_2 = -\frac{\beta}{2K} - \sqrt{\left(\frac{\beta}{2K}\right)^2 + \left(\frac{\pi}{M}\right)^2}.$$

Values of  $\psi$  versus  $x$  are shown in figure 5.1 for the case with  $L = 6000$  km,  $M = 3000$  km,  $\beta = 2 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$ , and  $K = 2 \times 10^{-6} \text{ s}^{-1}$ . Stommel's model was the first to exhibit the westward intensification of the oceanic response to a symmetric wind-stress curl.

With  $K/\beta L \ll 1$ , (5.85) is a boundary-layer problem, where the highest derivative term (the bottom frictional effect) is important only in a narrow region near the western boundary where the flow is northward. In the remainder of the basin the Sverdrup balance (5.80) is approximately valid (but see below), the flow is slow and southward, and friction is unimportant. The negative vorticity injected into the ocean by the wind is eventually dissipated in the western boundary layer, where the induced northward flow deposits columns of fluid at their original latitudes with the original planetary vorticity restored. Detailed balances and a fairly comprehensive discussion are given by Veronis (1966a).

The westward intensification is normally explained in terms of the vorticity balance, but a qualitative discussion in terms of momentum balance is also possi-

ble. Thus, we note that the Ekman wind drift in the northern half-basin is southward whereas that of the southern half-basin is northward. Water piles up at mid-latitude, raising the free surface level and creating a high pressure ridge at mid-latitude (H in figure 5.2). The induced eastward geostrophic flow in the northern half-basin requires a low pressure along the northern boundary. In the southern half-basin a westward flow of the same magnitude requires less of a north-south pressure difference (because the Coriolis parameter is smaller) so the low pressure (HL in figure 5.2) at the south is higher than the low pressure (LL in figure 5.2) at the north. The solid boundaries at the east and west will divert the flow. A narrow frictional boundary layer at the east would require flow from the low low pressure at the north to the high low pressure at the south, i.e., flow up the (gross) pressure gradient. On the western side, on the other hand, a narrow frictional boundary layer supports flow from high to low pressure. Hence, if a thin frictional boundary layer exists, it must be on the western side. This "explanation" ignores a lot of important details, but the reasoning is consistent with the roles that rotation and friction play in balancing the pressure gradient.

If a system without meridional boundaries (a zonal channel) were subjected to a zonal stress, a zonal flow would be generated (apart from the Ekman drift). Hence, the Sverdrup transport of Stommel's model must depend on the presence of meridional boundaries. Yet it seems likely that if the meridional boundaries are far enough apart, the system should resemble a zonal channel more than an enclosed ocean except in relatively narrow regions near the east and west where meridional flow takes place. Welander (1976) showed that that is the case. With the zonal wind stress given above one can substitute  $\psi = \Phi(x) \sin(\pi y / M)$  to derive

$$K\Phi'' - K\frac{\pi^2}{M^2}\Phi + \beta\Phi' = -\frac{\pi T}{M}. \quad (5.87)$$

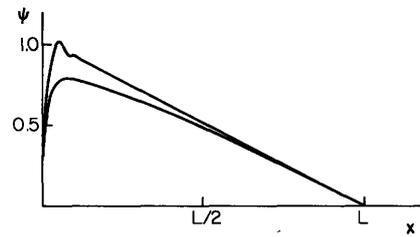


Figure 5.1 The transport streamfunction, normalized with respect to the Sverdrup transport and divided by  $\sin \pi y / M$ , is shown for Munk's solution with lateral diffusion (top curve) and Stommel's solution with bottom friction. The nominal boundary layer thickness is  $L/60$ . Stommel's solution shows the decreased transport because of the effect of friction in a basin with  $\pi L / M \gg 1$ . Munk's solution oscillates near the western boundary, giving rise to a weak countercurrent to the east of the main northward flow.

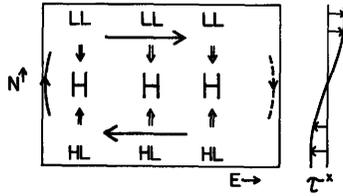


Figure 5.2 A cosine wind stress  $\tau^x$  causes an Ekman drift (double arrows) toward mid-latitude where the free surface is elevated and a high pressure region (H) is created. A geostrophically balanced current flows eastward in the north half-basin and westward in the south. Because of the larger Coriolis parameter a lower low pressure (LL) is required along the north boundary than along the south (HL) to support the same transport geostrophically. If the zonal transport is deflected southward in a frictional boundary layer near the eastern side (dashed curve), the flow must go against the gross pressure difference (from LL to HL). If the flow is in a western boundary layer (solid curve), the gross pressure difference drives the flow against frictional retardation. The latter is a consistent picture.

As we have seen, the second-derivative term is important only in the western boundary layer where the scale of variation is  $K/\beta = 100$  km. North-south diffusion (the undifferentiated  $\Phi$  term) is unimportant when the geometry is square. But when the zonal separation is large ( $\pi L/M \gg 1$ ), the balance is between wind-stress curl and north-south diffusion,  $\Phi \approx MT/\pi K$ , and the flow is zonal. The Sverdrup transport relation holds in an eastern boundary layer with the east-west scale  $\beta M^2 \pi^2 / K$ . Bye and Veronis (1979) pointed out that the northward transport in the western boundary layer is much smaller than the transport calculated by the Sverdrup balance if the aspect ratio  $\pi L/M$  is large, as is the case for nearly all wind-driven oceanic gyres. Of course, these results are contained in the complete solution of the simple model discussed here. But when relatively modest refinements are introduced (e.g., spherical geometry), a complete solution is no longer possible and boundary layer methods must be used. It is then necessary to recognize the correct approximate balance in the different regions of the basin.

### 5.5.3 Topography and Lateral Friction

The principal result of the foregoing analysis, viz., the westward intensification of an oceanic gyre, is verified both by observations and by much more sophisticated analyses. Hence, it is a feature that appears to be insensitive to the drastic simplifications that were made. But it is a simple matter to change the result by relaxing one of the simplifications and then restoring the result with a second, seemingly unrelated, assumption. In other words, the simple model is not as crude as it appears to be.

For example, introduce realistic topography (Holand, 1967; Welander, 1968). Then on vertical integration, we see from (5.60) that, in addition to the effects

of wind stress and bottom friction, the vertical divergence term in (5.78) will also contribute the term  $f dh/dt$  to the right-hand side of (5.84). If the latter is combined with the  $\beta$  term, the result is

$$h^2 \frac{d}{dt} \left( \frac{f}{h} \right) = \hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} - K \left( \frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} \right). \quad (5.88)$$

Hence, the driving and dissipative forces on the right will cause a fluid column to respond by moving to points determined by the value of  $f/h$  rather than  $f$  as before. Since the contours of  $f/h$  are sometimes strongly inclined to latitude circles (Gill and Parker, 1970), the transport pattern is very different from (in fact, less realistic than) Stommel's. Thus, the effect of topography is exaggerated in a homogeneous model.

Stratification can reduce the topographic effect. In fact, if the density surfaces adjust so that the pressure gradient in (5.55) vanishes at and below a given level, there will be no driving force to support a flow. If topography does not project above this level of density compensation, it has no effect on the flow. In an intermediate situation, the density distribution can compensate for part of the pressure gradient so that at the level where it interacts with the bottom the velocity is considerably weaker than the surface velocity. A treatment of the latter case would necessarily incorporate convective processes in some form.

When complete compensation takes place in a steady model, the topographic influence is eliminated, but our derivation of bottom friction is no longer valid because it is no longer possible to parameterize the frictional processes at the bottom in terms of the mean velocity. The essential results of the model can be preserved, however, by parameterizing frictional effects in terms of an assumed lateral eddy diffusion. The last term in (5.83) is then replaced by a lateral frictional term so that the vorticity equation, in terms of the transport stream function becomes

$$\beta \frac{\partial \psi}{\partial x} = \hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} + A \nabla^2 \psi, \quad (5.89)$$

where  $A$  is the magnitude of eddy viscosity based on the intensity of eddy processes at scales smaller than those being analyzed. Hidaka (1949) introduced this equation together with the vertically integrated continuity equation

$$\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} = 0. \quad (5.90)$$

A convenient set of boundary conditions where the wind stress curl is proportion to  $\sin(\pi y/M)$  is

$$U = 0 = V \quad \text{at } x = 0, L, \quad (5.91)$$

$$V = 0 = \frac{\partial U}{\partial y} \quad \text{at } y = 0, M.$$

The solution is easily obtained (Munk, 1950) and is included in figure 5.1. It contains a Sverdrup transport in the interior; a narrow eastern boundary layer in which  $V$  decreases to zero at the eastern wall; and a western boundary layer with no tangential velocity at the western wall, a northward flow near the wall and a weak, narrow countercurrent just east of the northward flow. Because frictional processes are now associated with higher derivatives, the effect of friction in the interior is considerably weaker than in Stommel's model, and the Sverdrup balance is valid throughout the interior. Accordingly, in this case the aspect ratio of the basin has little influence on the magnitude of the transport. Because the zonal velocity increases linearly with distance from the eastern boundary, for broad ocean basins the flow has a strongly zonal appearance. In Stommel's model the north-south flow essentially vanishes in the western portions of the basin and the flow is truly zonal there.

Although these formal models are steady, the application is to flows that are transient but statistically steady. Transient motions can have a strong barotropic component even when the statistically steady flow is largely baroclinic. With that in mind we may still use a bottom frictional drag for the stratified steady model, though the connection to the mean flow will then be not through the coupling to a steady Ekman layer but through a time averaging of interacting transient motions. Rooth (1972) has made such an estimate for  $K$  and obtains a value considerably smaller than the one normally used.

#### 5.5.4 Laboratory Models

Though these steady, linear models can provide only the crudest approximation to real oceanic flows, they have served an important function in the development of oceanic theory. Stommel (1957b) put together the important components (Ekman suction and  $\beta$ -effect) to construct a comprehensive picture of ocean current theory as determined by these simple processes. The ideas were tested in a laboratory model of ocean circulation (Stommel, Arons, and Faller, 1958) in which the  $\beta$ -effect was simulated by the paraboloidal depth of a homogeneous layer of water in a pie-shaped basin rotating about the apex (see chapter 16). The equivalence of  $\beta$  and variable depth is suggested by the linearized form of potential vorticity,

$$(\zeta + f)/h = (\zeta + f)/H_0 - \frac{f_0 \eta}{H_0^2},$$

where  $\eta$  is the deviation of the free surface from its mean value  $H_0$ , so that a change in  $f_0 \eta/H_0^2$  is equivalent to a change in  $f$ , i.e., to  $\beta$ . When water is being added at the apex, the free surface in the interior rises not by a direct vertical motion but by a radially uniform inward movement of columns of fluid (figure 5.3). The

circulation generated in this way simulates the Sverdrup transport, the inward radial direction corresponding to north (increasing  $f$  or decreasing depth).

In the experiment, boundary layers near the "western" boundary and the rim and apex are required to complete the circulation pattern (figure 5.3). The azimuthal flow and the rising free surface needed to feed the interior radial flow are generated in the rim boundary layer. Near the apex the flow is diverted to the western boundary layer to join the fluid being injected. It is interesting to note that the radially inward flow that causes the free surface to rise is *toward* the source of fluid. Thus, the transport in the western boundary layer is twice that of the source. Half of the former goes to raise the free surface; the other half serves as the vehicle for the indirect circulation. (Also see Figures 16.1 and 16.2 and the accompanying discussion.)

Additional experiments and a rigorous analysis using rotating-fluid theory to treat the various boundary layers were subsequently provided by Kuo and Veronis (1971), who showed that for different parametric ranges the experiment could be used to simulate Stommel's model with a bottom frictional boundary layer or the Hidaka-Munk model with a lateral frictional boundary layer. Veronis and Yang (1972) provided a perturbation treatment of the nonlinear effects and verified the results with a series of experiments. Pedlosky and Greenspan (1967) proposed an alternative laboratory model with the depth variation provided by an inclined boundary at the top and/or bottom of a rotating cylinder. The flow was driven by the differential rotation of the top plate. For this model Beardsley (1969, 1972) carried out a comprehensive set of experiments and

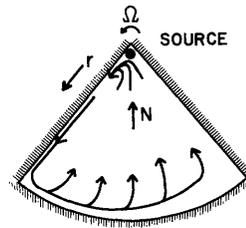


Figure 5.3A A weak source of fluid at the apex of a rotating pie-shaped basin will cause flow toward the rim in a "western" boundary layer. Fluid flows from the rim boundary layer radially inward toward the apex as shown.

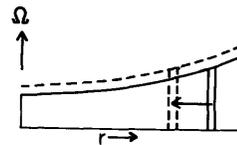


Figure 5.3B A vertical cross section through the apex. The basin is filled by the inward movement of columns of fluid as shown.

extended the theory analytically and numerically to include inertial effects.

The foregoing experiments and theories are more appropriate areas of application than the real ocean is for the ideas introduced by Sverdrup, Stommel, and Munk. At the time that they were introduced, however, these ideas were remarkable advances into unknown territory. They have provided a framework for further development and some of them persist as important elements in more extensive theories.

### 5.6 Preliminary Nonlinear Considerations

The first perturbation analysis of nonlinear effects in a wind-driven gyre was by Munk, Groves, and Carrier (1950), but it is easier to see the qualitative changes by looking at Stommel's model (Veronis, 1966a). From the linear problem we saw that the vorticity and its zonal variation are largest in the western boundary layer, so we expect the largest nonlinearities there. The wind stress is not important in that region, and we start with the vorticity equation, including inertial terms but not the wind-stress curl:

$$\frac{d}{dt}(\zeta + f) = \mathbf{v} \cdot \nabla \zeta + \beta v = -\epsilon \zeta. \quad (5.92)$$

In the southern half of the basin the flow is westward ( $u < 0$ ) into the boundary layer where it is diverted northward. Thus, a fluid particle is carried from the interior, where  $\zeta$  vanishes, into the boundary layer, where  $\zeta$  is large and negative, so  $d\zeta/dt < 0$ . Northward flow implies  $df/dt > 0$ . Hence, the convective term balances part of the  $\beta$ -effect and  $-\epsilon\zeta$  must consequently decrease in size. Since the vorticity is essentially  $\partial v/\partial x$ , it will decrease if  $v$  decreases or if the horizontal scale increases. But from  $v = \partial\psi/\partial x$  we see that a decrease in  $v$  also corresponds to an increase in the horizontal scale. Therefore, we conclude that inertial effects weaken the flow by broadening the scale. This effect will also decrease the dissipation in the inflow region. The same considerations apply to the case with lateral friction.

In the northern half of the basin where the flow emerges ( $u > 0$ ) from the western boundary layer, fluid is carried from a region of negative vorticity to the interior where  $\zeta$  vanishes, so  $d\zeta/dt > 0$ . Since the flow is northward in the boundary layer,  $df/dt$  is also positive. Therefore, the amplitude of the vorticity must be larger since the dissipation  $-\epsilon\zeta$  must be larger than in the linear case. Hence, the horizontal scale of variation must decrease.

The net effect of inertial processes is thus to broaden the boundary layer thickness and to reduce the dissipation in the region of inflow, and to sharpen the boundary layer thickness and increase the dissipation in the region of outflow.

For more nonlinear flows the dissipation takes place largely in the northern half of the boundary layer. Furthermore, the excess inertia of the particles causes them to overshoot their original (interior) latitudes so there must be an additional region where inertial processes and friction restore the particles (southward) to their starting points. The effect is to spread the region of inertial and frictional control first to the north and eventually eastward from the northwest corner of the basin. A discussion of the successively stronger effects of nonlinear processes and a division of the basin into regions where different physical balances obtain is given by Veronis (1966b).

This argument strongly suggests that it may be possible to analyze the region of formation of western boundary currents in terms of a frictionless inertial model. Stommel (1954) proposed such an analysis which he subsequently included in his book (Stommel, 1965).

Fofonoff (1954) focused his attention on nonlinear processes by treating the steady circulation in a frictionless, homogeneous ocean. The starting point is the conservation of potential vorticity in a basin of constant depth, viz.,

$$\frac{d}{dt}(\zeta + f) = 0, \quad (5.93)$$

together with the two-dimensional continuity equation. These equations are satisfied by  $u = -\partial\psi/\partial y =$  constant or

$$\psi = -uy, \quad (5.94)$$

but boundary conditions are not, so it is necessary to add boundary layers at the eastern and western sides of the basin.

A first integral of (5.93) is

$$\nabla^2\psi + f = F(\psi) \quad (5.95)$$

and in the interior where the relative vorticity vanishes

$$F(\psi) = f = f_0 + \beta y. \quad (5.96)$$

But (5.94) yields  $y = -\psi/u$  there, so that  $F(\psi) = f_0 - \beta\psi/u$  and (5.95) becomes

$$\nabla^2\psi + \frac{\beta\psi}{u} = -\beta y. \quad (5.97)$$

This equation is satisfied nearly everywhere by  $\psi = \phi(x)y$  so that

$$\phi'' + \frac{\beta\phi}{u} = -\beta. \quad (5.98)$$

The north-south flow near the meridional boundaries is thus geostrophic. A boundary layer solution with  $\phi = 0$  at  $x = 0, L$  is possible for  $u < 0$  if  $\epsilon = (-\beta/u)^{1/2} \gg L$ . It is

$$\phi = \frac{u}{\sinh \epsilon L} [\sinh \epsilon x - \sinh \epsilon L + \sinh \epsilon(L - x)] \times [y - Me^{-(M-y)\epsilon}]. \quad (5.99)$$

This yields a uniform, westward flow in the interior, and boundary layers of thickness  $\epsilon^{-1}$  with northward flow at the west, southward flow at the east, and a jet across the northern edge (figure 5.4).

It is possible to have the eastward jet at any latitude by adding an appropriate constant to  $\psi$  in (5.94). With  $u > 0$  the system does not have a boundary layer solution but oscillates across the basin (Fofonoff, 1962a).

Although Fofonoff's solution appears to be very artificial, it is one of the survivors of the earlier theories. The strongly nonlinear version of Stommel's model leads to a solution that looks remarkably like Fofonoff's (Veronis, 1966b; Niiler, 1966). The recirculation region just south of the Gulf Stream after the latter has separated from the coast has the appearance of a local inertial circulation. Thus, it is likely that some version of the latter will be part of any successful model of large-scale ocean circulation.

Shortly after Fofonoff's analysis and following Stommel's (1954) suggestions, Charney (1955b) and Morgan (1956) produced models of the Gulf Stream as an inertial boundary layer. By using observed or simulated conditions at the inflow edge of the Gulf Stream to fix the form of  $F(\psi)$ , and working with a two-layer model with potential vorticity  $(f + \zeta)/h$  and geostrophic balance for the northward flow, they were able to calculate the streamfunction pattern and the thermocline depth distribution in the formation region of the Gulf Stream. Charney showed that in a two-layer ocean inertial forces can cause the thermocline to rise to the surface at a latitude corresponding to Cape Hatteras. His solution could not extend beyond that point.

Morgan began his analysis by dividing the ocean into an interior with a Sverdrup balance, a formation region

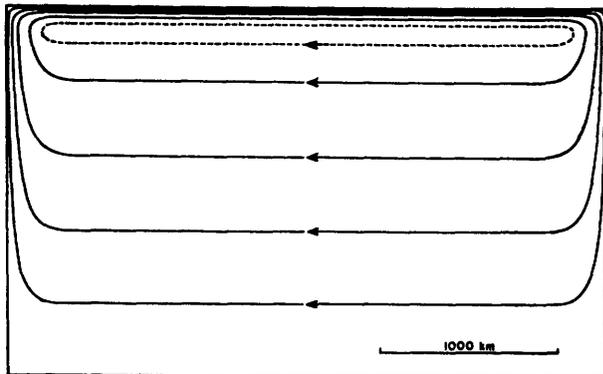


Figure 5.4 Fofonoff's (1954) inertial flow pattern for steady westward flows in the interior. An inertial boundary layer at the west diverts the flow northward and an eastward jet is formed. The latter feeds into an inertial boundary layer on the east that supplies the steady westward flow of the interior.

for the western boundary current, which he analyzed using the same model that Charney did, and a northern region. He speculated that friction and inertial and transient processes would interact in the north, but he did not attempt to analyze that region. He was one of the first to point out that pressure torques at the bottom and sides of the ocean can help to balance the torque exerted by the wind stress about a mid-ocean axis.

In contrast to the demonstration following (5.92) that inertial effects are consistent with the formation of a western boundary layer by the interior flow, a similar argument for the formation of an eastern boundary layer is not possible. For example, consider an anticyclonic gyre when an eastward interior flow generates an eastern boundary layer with southward flow. The vorticity in the boundary layer is negative, so  $-\epsilon\zeta$  is positive. For southward flow  $df/dt$  is negative and therefore  $d\zeta/dt$  must be positive. But that is not possible since  $\zeta$  must change from a nearly zero value in the interior to a large negative value in the boundary layer. An analysis of the various possibilities for both cyclonic and anticyclonic gyres shows that it is generally not possible to form eastern boundary layers from eastward interior flows (Veronis, 1963). The actual existence of eastern boundary layers means that the necessary physical processes (in my opinion horizontal advection of density must be included) are missing from these simple models.

In an important model of a steady wind-driven gyre in a homogeneous ocean of constant depth, Derek Moore (1963) produced a complete circulation pattern with contributions from frictional and inertial processes in both inflow and outflow regions of the western boundary layer. Moore combined boundary-layer arguments from classical fluid mechanics with most of the features given above. Using a Navier-Stokes form for friction, he proved that frictional and inertial processes cannot be combined consistently to produce a boundary layer confined to the eastern side. In the vorticity equation of his model inertia is included as an east-west convection of the vorticity with a zonal velocity,  $U(y) = U_0 \cos(\pi y/M)$ , consistent with the form of the wind stress. In the southern half-basin (figure 5.5) the incoming (westward) flow forms an inertially controlled western boundary current. In the northern half-basin the emerging flow oscillates eastward and has the appearance of standing, damped Rossby waves imbedded in an eastward current. The center of the gyre is north of mid-latitude, consistent with the effects of inertia mentioned earlier. His results depend on the magnitude of a Reynolds number defined by  $Re = U_0^2/\nu\beta^{1/2}$ , which can be looked upon as the ratio of the inertial boundary layer scale  $(U_0/\beta)^{1/2}$  to the viscous scale  $\nu/U_0$ . The result is shown for  $Re = 5$ . As  $Re$  is decreased, the flow tends toward the Munk pat-

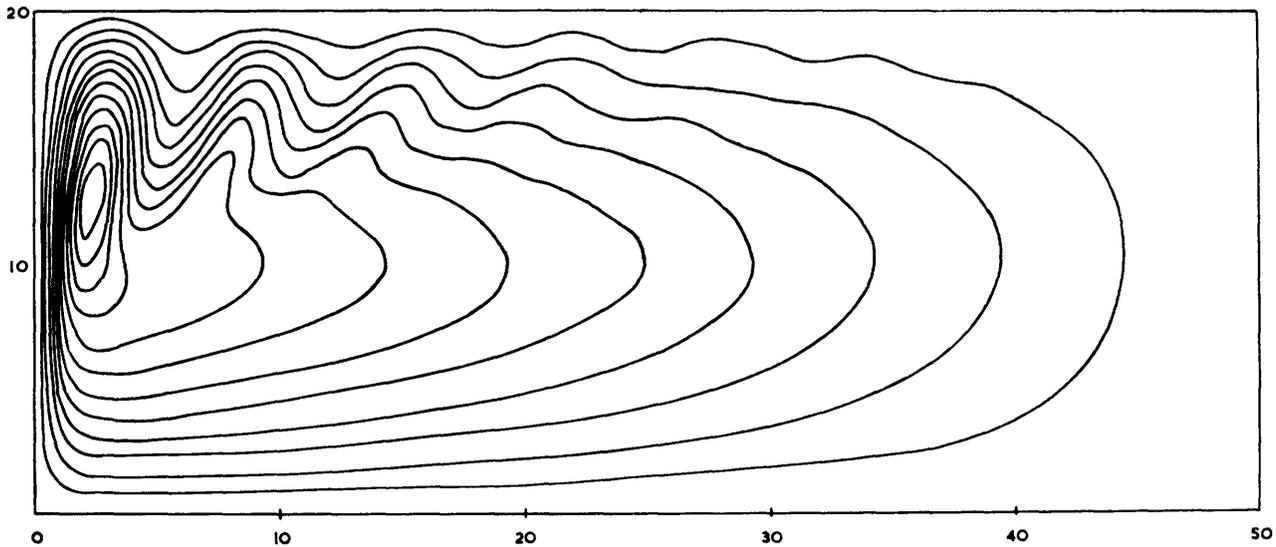


Figure 5.5 Contours of the streamfunction in a homogeneous ocean driven by a wind stress of the form  $-\cos\pi y/M$  as derived by Moore (1963). An Oseen approximation for the non-

linear terms with a mean current  $U(y) \propto \cos\pi y/M$  was used. The wavy contours in the north half-basin are standing Rossby waves imbedded in the mean velocity field.

tern. With larger  $Re$  the oscillations extend farther to the east and eventually fill the northern half of the basin. In the latter case there is a rapid transition across mid-latitude in the interior and the oscillatory flow becomes unstable. Qualitatively this homogeneous model contains a remarkably realistic array of features of oceanic flow, though the observed recirculation in the northwest corner is missing.

We turn to a discussion of stronger nonlinear effects in Stommel's model. As  $Ro$  is increased (Veronis, 1966b), the western boundary layer in the southern half-basin broadens and dissipative effects are more confined to the north. Inertial effects also intensify in the north so that a particle overshoots the northernmost latitude that it had in the interior. Hence, a new boundary layer region must be generated (offshore of the original one) where friction and inertia force the particle southward to its original latitude. In this latter region the relative vorticity is actually positive because the return flow to the south is stronger close to the boundary layer than it is farther to the east. The overshoot can be seen in figure 5.6A.

With even stronger driving the overshoot is larger and eventually the particle is driven close to the northern boundary and then eastward before it starts its southward return to its original latitude (figure 5.6B). Thus, the frictional-inertial region is broadened. In an extreme case (figure 5.6C) fluid particles move eastward in a jet at the north and reach the eastern boundary before turning south. In the latter case, there is essentially no Sverdrup interior, and the flow pattern resembles Fofonoff's free inertial flow with a mild east-west asymmetry as the only evidence that the flow is wind driven. An interesting fact here is that the

northward transport in the western boundary layer does not increase beyond the Sverdrup transport until the eastward moving inertial jet reaches the eastern boundary. In the calculations cited, that happens when the inertial scale  $(U_0/\beta)^{1/2} (=Ro^{1/2}L)$  exceeds the viscous scale  $K/\beta$  by a factor of 2 or so. Here,  $U_0$  is a measure of the Sverdrup velocity. Qualitatively, at least, the observed recirculation to the south and east of the Gulf Stream after it has separated from the coast is simulated by this model. The separation from the coast is not. An analytic model of the highly nonlinear case was suggested by Veronis (1966b) and independently carried out by Niiler (1966). The resulting pattern is consistent with the one shown in figure 5.6C. Stommel (1965) guessed a similar pattern.

Bryan (1963) carried out an extensive set of numerical calculations in a rectangular basin for the nonlinear Hidaka-Munk model with  $\hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} \sim \sin\pi y/M$ , zero velocity boundary conditions at east and west, and zero-shear conditions at north and south. He presented his results in terms of a Reynolds number  $Re$  essentially the same as Moore's, and the Rossby number,  $Ro$ . The results differ greatly from those with bottom friction because for  $Re > 60$  a barotropic (Rayleigh-type) instability can occur near the western boundary where the tangential velocity must vanish. Figure 5.7 illustrates his results for three values of  $Re$ , with  $Ro = 1.28 \times 10^{-3}$  for figures 5.7A and 5.7B and  $Ro = 3.2 \times 10^{-4}$  for figure 5.7C. The first two cases, with  $Re = 20$  and  $Re = 60$ , show the development of the flow with increasing nonlinearity. Only a mild, steady, oscillatory pattern is present with  $Re = 20$ , whereas with  $Re = 60$  the oscillations are more intense and a closed eddy (recirculation) is present near the northwest corner. For  $Re = 100$  the flow is transient with a barotropic

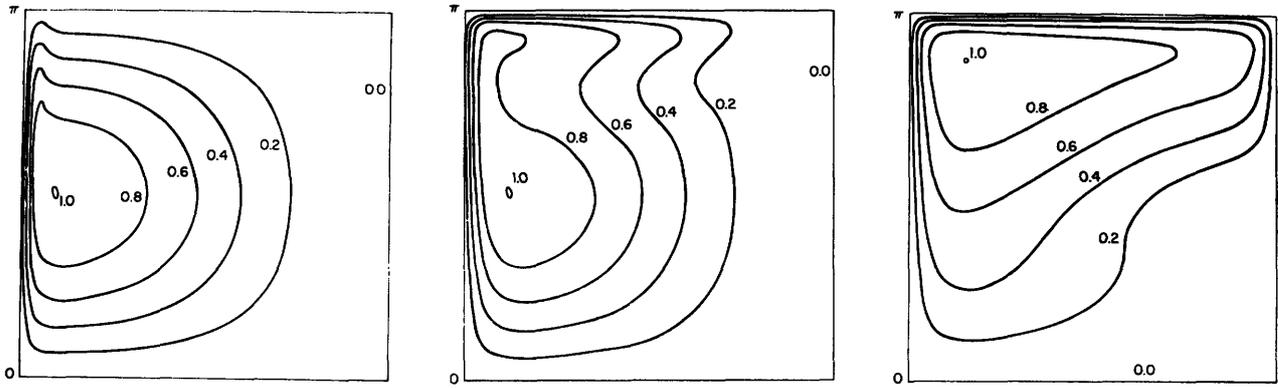


Figure 5.6 Three streamfunction patterns by Veronis (1966b) for an ocean basin with varying degrees of intensity of wind stress. (A) shows the perturbation effect of nonlinearity with fluid particles in the western boundary layer overshooting their equilibrium latitudes. (B) shows a much stronger inertial effect. In (C) inertia dominates the system, creating an eastward jet along the north reminiscent of Fofonoff's solution.

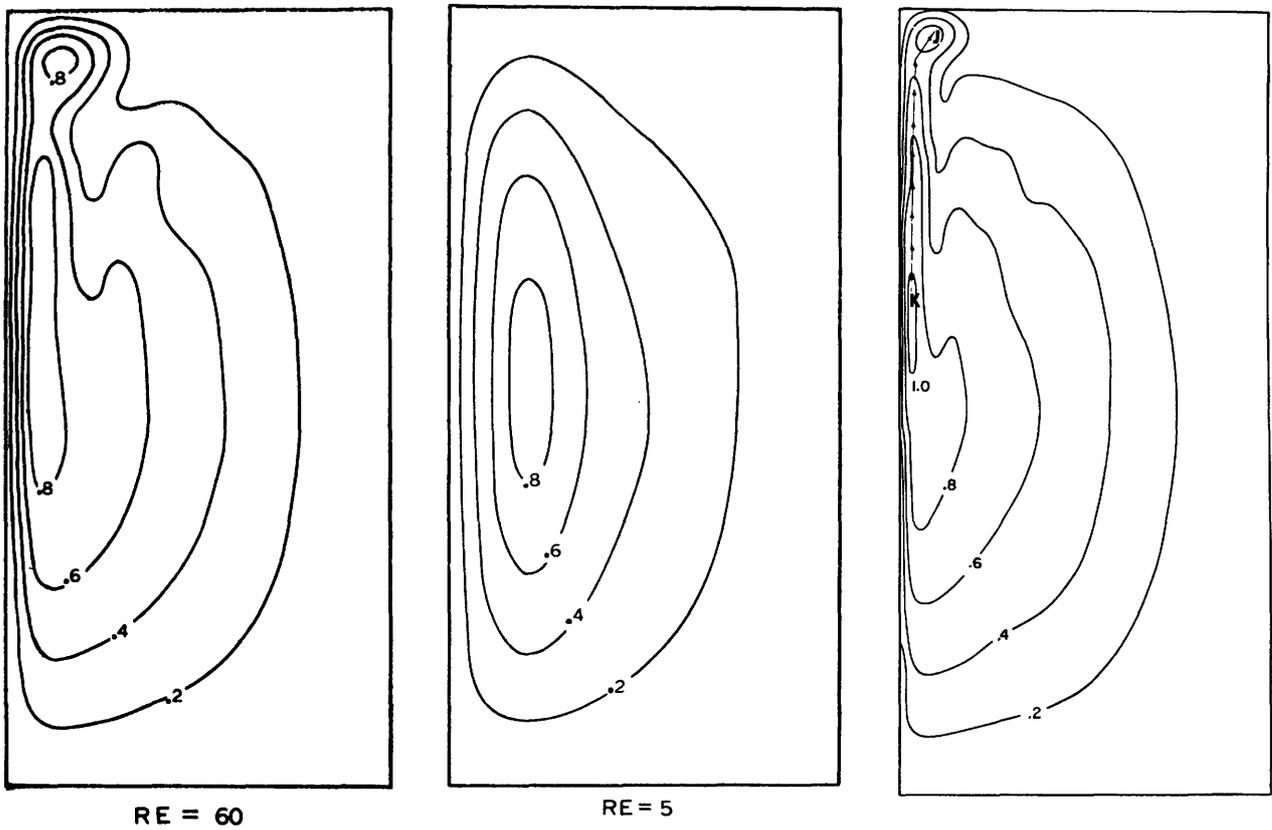


Figure 5.7 Bryan's (1963) streamfunction contours for a homogeneous ocean with lateral friction. The circulation in (A) is nearly linear; that of (B) is near the limit of forcing that still leads to a steady circulation. With even more intense driving a barotropic instability occurs as in (C), where a time-average field is shown after the system approaches a statistically steady state. See also Figure 3.13 and discussion there.  $Re$  is 5 for (A), 60 for (B), and 100 for (C).

instability induced in the northern half of the intense northward jet. Figure 5.7C shows the time-averaged flow after the transients have settled down. In this case also there is an offshore region in the north with positive vorticity where particles return southward to their starting latitudes. It is not possible to obtain an intense recirculation with this model because of the barotropic instability.

Bryan also calculated the flow for a basin with a western boundary directed north, then due east, and then north again. The break was north of mid-latitude. The object was to see whether the break in the boundary would force the western boundary current out to sea. The flow pattern was modified mildly, but the stream turned the corner and hugged the coast.

### 5.7 Why Does the Gulf Stream Leave the Coast?

The Gulf Stream flows along the coast from Florida to Cape Hatteras, where it parts from the coast and flows slightly north of eastward out to sea (see chapter 4). The Kuroshio and all other western boundary currents also separate. The phenomenon is explained here by a very simple argument. Although processes more complicated than the ones discussed below are also present, I believe that the argument given here contains the essential features even though the local dynamical details are not included.

Consider a two-layer system with the lower layer at rest. Then from equations (5.65) and (5.66) it follows that

$$\nabla\eta_2 = -\frac{\rho_1}{\Delta\rho}\nabla\eta_1, \quad \nabla\eta_1 = \frac{\Delta\rho}{\rho_2}\nabla h_1. \quad (5.100)$$

If the motion is geostrophic ( $Ro \ll 1$ ) except for the vertical stress term near the surface, equation (5.62) upon vertical integration over the depth  $h_1$  of the top layer becomes, with  $g' = g\Delta\rho/\rho_2$ ,

$$-fV_1 = -\frac{g'h_1}{a\cos\phi}\frac{\partial h_1}{\partial\lambda} + \tau, \quad (5.101)$$

where spherical coordinates have been retained so there is no geometrical distortion. Here the stress at the interface is assumed negligible and  $\tau$  corresponds to the zonal wind stress. Multiply (5.101) by  $a\cos\phi$  and apply the operator  $\int_{\lambda_e}^{\lambda} d\lambda$ , where  $\lambda_e$  is the meridian of the eastern boundary, to obtain

$$h_1^2 = h_{1e}^2 - \frac{2f}{g'}T_1 - \frac{2f}{g'}T_E, \quad (5.102)$$

where subscript e denotes a value at  $\lambda_e$ ,  $T_1 = \int_{\lambda_e}^{\lambda} a\cos\phi V_1 d\lambda$  is the meridional transport, and  $T_E = \int_{\lambda_e}^{\lambda} a\cos\phi \tau d\lambda/f$  is the Ekman drift.

In all of the calculations reported in the previous section, the downstream velocity in the western boundary layer is geostrophic to a very good approxi-

mation. Hence, (5.102) is valid not only for interior flow but for the entire basin from west to east. Therefore, if we evaluate (5.102) at the western edge  $\lambda_w$ ,  $T_1$  represents the total meridional transport. If the ocean basin is enclosed to the north of the latitude in question,  $T_1$  must vanish in the steady state and (5.102) becomes

$$h_{1w}^2 = h_{1e}^2 - \frac{2f}{g'}T_E. \quad (5.103)$$

Now, for  $\tau > 0$  the Ekman drift,  $T_E$  is toward the south (positive as defined above) and the depth of the upper layer at the western boundary  $h_{1w}$  will be less than  $h_{1e}$ . For sufficiently large  $T_E$ ,  $h_{1w}$  will vanish, i.e., the thermocline (interface) rises to the surface. With observed values for  $\Delta\rho/\rho_2$ ,  $\tau$ , and  $h_{1e}$  for the North Atlantic,  $h_{1w}$  vanishes at about the latitude of Cape Hatteras.

North of that latitude  $\tau$  is even larger and (5.103) cannot be satisfied because  $T_E$  is too large. However, the solution can be extended northward by setting  $h_{1w}$  equal to zero at a new longitude ( $>\lambda_w$ ) which is chosen to reduce  $T_E$  so that the terms on the right of (5.103) balance. This new longitude marks the westernmost edge of the warm-water mass and is the longitude of the Gulf Stream. But  $\lambda > \lambda_w$  means that the Gulf Stream must separate from the coast and extend out to sea. This argument alone does not suffice for higher latitudes where  $\tau$  eventually becomes negative. We shall return to that issue presently.

Before doing so, however, we discuss the simple physical balances given above. The meridional flow in the interior is a combination of geostrophically balanced motion and Ekman drift. If the flow were completely geostrophic, vanishing  $T_1$  would require equal values of  $h_1$  at the eastern and western edges. But the Ekman wind drift, which does not involve a pressure gradient, accounts for part of the southward transport when  $\tau > 0$ . Therefore, since the total transport vanishes, there is a net northward geostrophic transport, of magnitude  $T_E$ , which requires  $h_{1w}^2 < h_{1e}^2$ . Thus, the Ekman drift causes the thermocline to rise to the surface. Separation of the Gulf Stream from the coast simply moves the western edge of the warm-water mass (upper layer) eastward so that the smaller Ekman drift acting on that water mass of more limited east-west width can just balance the geostrophic flow determined by  $h_{1e}^2$  (since  $h_{1w}^2$  vanishes).

It is also interesting to note that the Coriolis parameter does not appear in (5.103). In fact, the result is exactly the one obtained for a nonrotating lake where the wind blows the warm water to the leeward edge and causes the thermocline to rise on the windward side. The principal difference between the two phenomena is that the induced pressure gradient drives a vertical circulation in the lake, whereas it is geostrophically balanced in the rotating ocean, thereby generat-

ing a horizontal cell. But the leeward piling up of water is the same in the two cases.

Returning to the problem at high latitudes, we note first that the analysis given above must be supplemented by the remaining dynamic balances. The reader is referred to Veronis (1973a) for the details for the wind-driven model. The qualitative discussion given here is simpler and clearer than in the original paper.

The first problem is that the Sverdrup transport for the interior vanishes with  $\hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau}$ , and without adding to the simple argument there is no way of supplying warm water to the north of the latitude ( $\approx 40^\circ\text{N}$  in the North Atlantic) where the curl vanishes. Second, even supposing that warm water has somehow been supplied to the north, the Sverdrup transport there is northward ( $\hat{\mathbf{k}} \cdot \nabla \times \boldsymbol{\tau} > 0$ ), so the southward return of the flow by a western boundary current would require that the thermocline be deeper on the western side of the boundary layer. That is not possible with the boundary current in mid-ocean.

Both of these issues can be resolved by considering what happens even farther to the north where warm water flows northward and impinges on the northern boundary. In the real ocean and in a model including thermal driving (Veronis, 1978), this water will sink and give rise to a deep circulation and an overturning cell. In a wind-driven model the water travels counterclockwise as an isolated warm boundary current and rejoins the stream at the point of separation. In the analysis given above, this recirculating current represents an excess transport in the separated boundary current. Because its transport does not depend on local winds, it can transport water past the latitude of vanishing wind-stress curl and supply warm water to the interior at high latitudes. When it is included in the analysis, a revised longitude for the separated boundary current is obtained. The calculation, which can be made consistent and quantitative for both the wind-driven model and the one including thermal driving, is contained in the two papers cited above. The path of the separated Gulf Stream is reproduced in figure 5.8. It is especially interesting to note that the vestigial current in the northeastern corner of the basin corresponds to the Norwegian Current (the Alaskan Current in the Pacific) and that its transport is important for the separation of the Gulf Stream and also for the determination of the longitude of the current after it has separated.

The analysis leading to the separation of the Gulf Stream from the coast is contained in a quasi-geostrophic model by Parsons (1969). It was derived independently by Veronis (1973a) as part of a study of the circulation of the World Ocean. The extension poleward of the latitude where the wind-stress curl vanishes is contained in the latter paper. Kamenkovich and Reznik (1972) included a (bottom friction) analy-

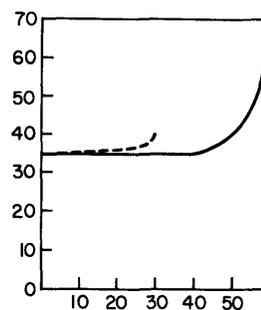


Figure 5.8 The path (solid curve) of the Gulf Stream after it has separated from the coast [from a reduced gravity model by Veronis (1973a)]. The zonal wind stress that drives the system is taken from observations and has zero curl at  $40^\circ\text{N}$ . The Norwegian Current is the narrow jet in the northeast. The dashed curve is the prediction for an isolated anticyclonic wind gyre (Parsons, 1969). The latter solution cannot be extended north of the latitude of zero wind-stress curl. Axes are latitude and longitude.

sis of the deep circulation induced by the separated current.

All of the above make use of a steady, linear, quasi-geostrophic model, and it is certain that the details (e.g., the longitude of the separated current) will be altered when a more complete dynamic model is used. The key elements of the argument, however, are the geostrophic balance of downstream velocity in the western boundary current, the Ekman wind drift, and a limited amount of upper-layer water. As long as a different dynamic model does not drastically change those three features (they are pretty rugged and can withstand a lot of battering) the more complicated dynamics can be incorporated to change the details of the results, leaving the main argument unchanged.

By the same token, the present analysis suggests that an explanation of the separation of western boundary currents from the coast must necessarily include the surfacing of the thermocline (with a possible mixed layer at the surface). Western boundary currents can be forced out to sea between wind-driven gyres of opposite sign, but that occurs at low latitudes as well where the phenomenon is qualitatively different because the thermocline does not surface.

In addition, the argument given here depends on properties of global scale. A more precise dynamic treatment based on local properties can lead to a better understanding of the detailed mechanistic balances of the separated current, but the cause of separation seems to be based on global properties.

## 5.8 Thermohaline Circulation

The physical processes that are involved in the formation of the thermocline have been studied as a separate part of the general circulation. The models incorporate geostrophic dynamics and steady convection