Chapter 20

Oceanic General Circulation

ABSTRACT

The concepts of geostrophy, hydrostaticity, and potential vorticity are merged to study the large-scale baroclinic circulation in the midlatitude oceans. The results lead to the Sverdrup balance, the beta spiral, and a number of properties of large-scale oceanic motions. The numerical part of the chapter provides an overview of the issues raised in constructing a model of the 3D circulation at the scale of ocean basins or the planet.

20.1 WHAT DRIVES THE OCEANIC CIRCULATION

Ocean motions span a great variety of scales in both time and space. At one extreme, we find microturbulence, not unlike that in hydraulics, and at the other, the large-scale circulation, which spans ocean basins and evolves over climatic timescales. The latter extreme is the objective of this chapter.

There are multiple mechanisms that set oceanic water masses in motion: the gravitational pull exerted by the moon and sun, differences in atmospheric pressure at sea level, wind stress over the sea surface, and convection resulting from atmospheric cooling and evaporation. The moon and sun generate periodic tides with negligible permanent circulation, whereas differences in atmospheric pressure play no significant role. On the other hand, deep convection at high latitudes generates currents responsible for a very slow movement in the abyss called the *conveyor belt* (Fig. 20.1). This leaves the stress exerted by the winds along the sea surface as the main driving force of basin-wide circulations in the upper part of the water column.

Ocean waters respond to the wind stress because of their low resistance to shear (low viscosity, even after viscosity magnification by turbulence) and because of the relative consistency with which winds blow over the ocean. Good examples are the *trade winds* in the tropics; they are so steady that shortly after Christopher Columbus and until the advent of steam, ships chartered their courses across the Atlantic according to those winds; hence their name. Further away from the tropics are winds blowing in the opposite direction. While trade

Introduction to Geophysical Fluid Dynamics, Volume 101, ISSN 0074-6142, DOI: 10.1016/B978-0-12-088759-0.00020-1

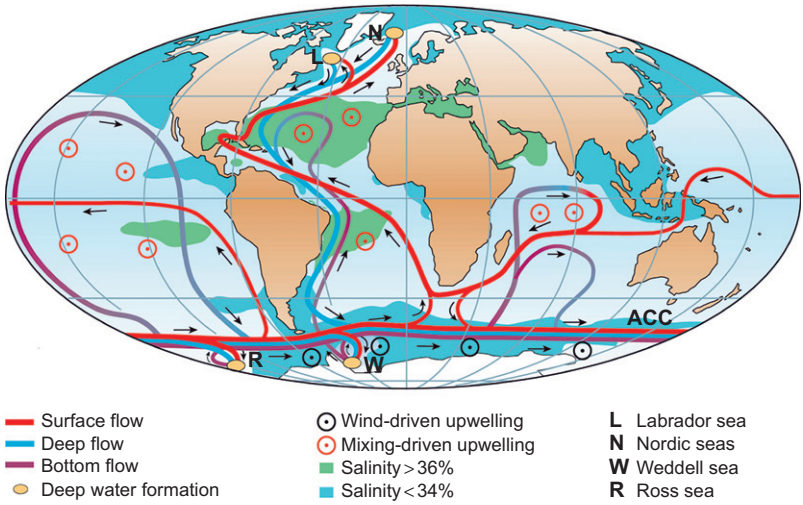


FIGURE 20.1 Cold and salty waters newly formed by deep convection at high latitudes are carried away by the conveyor belt across the ocean basins. These waters eventually resurface in warmer climes and return toward places of deep convection. The time to complete a loop is on the order of several hundred years to millennia. (*Kuhlbrodt et al., 2007*).

winds blow from the east and slightly toward the equator (they are also more aptly called northeasterlies and southeasterlies, depending on the hemisphere), midlatitude winds blow from west to east and are called *westerlies* (Fig. 20.2). Generally, much more variable than the trades, these westerlies nonetheless possess a substantial average component, and the combination of the two wind systems drives significant circulations in all midlatitude basins: North and South Atlantic, North and South Pacific, and Indian Oceans.

In the ocean, the water column can be broadly divided into four segments (Fig. 20.3). At the top lies the *mixed layer* that is stirred by the surface wind stress. With a depth on the order of 10 m, this layer can be assimilated for the purpose of large-scale ocean circulation with the Ekman layer (see Chapter 8) and is characterized by $\partial\rho/\partial z \simeq 0$. Below lies a layer called the *seasonal thermocline*, a layer in which the vertical stratification is erased every winter by convective cooling. Its depth is on the order of 100 m. Below the maximum depth of winter convection is the *main thermocline*, which is fed by water left behind whenever the seasonal thermocline retreats; it is permanently stratified ($\partial\rho/\partial z \neq 0$), and its thickness is on the order of 500–1000 m. The rest of the water column, which comprises most of the ocean water, is the *abyssal layer*. It is very cold, and its movement is very slow.

When considered together, the main thermocline and the abyssal layer form the *oceanic interior*. While mesoscale motions exist in both these layers, under

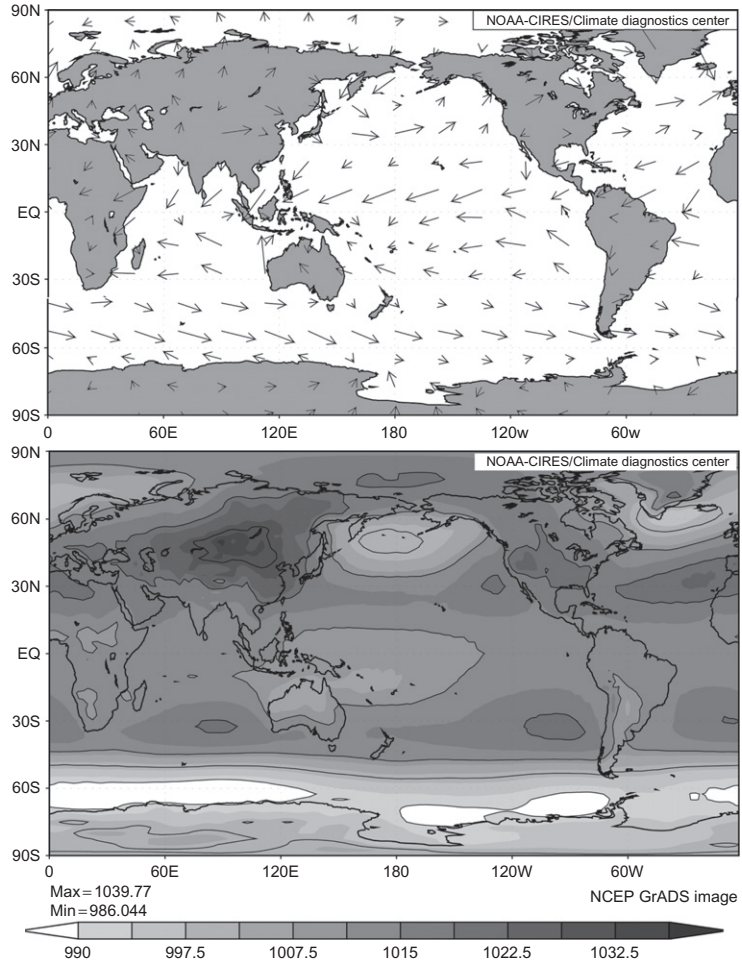


FIGURE 20.2 The major winds over the world ocean for the month of January averaged over the years 1968–1996 and associated sea surface pressure. (*From NCEP*)

the action of pressure fluctuations in the upper layers, it is believed that, in first approximation, the study of the slow background motion in the oceanic interior can proceed independently of the smaller scale, higher-frequency processes.

Although mariners have long been aware of the major ocean currents, such as the Gulf Stream,[1] ocean circulation theory was long in coming, chiefly for lack of systematic data below the surface. The discipline began in earnest

[1]Benjamin Franklin receives credit for publicizing and mapping the Gulf Stream in 1770.

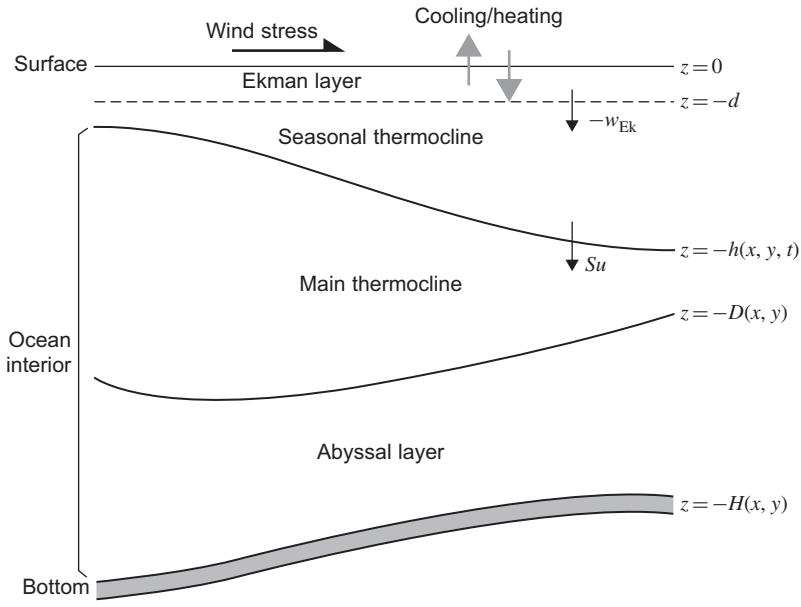


FIGURE 20.3 Vertical structure of the ocean from the point of view of large-scale circulation. Note that the relative thicknesses of the various layers are not to scale, for the abyss is far thicker than all other layers combined.

with the seminal works of Harald Sverdrup,[2] who formulated the equations of large-scale ocean dynamics (Sverdrup, 1947) and Henry Stommel[3], whose major contributions to ocean circulation are many and diverse, beginning with the first correct theory for the Gulf Strean (Stommel, 1948). Today, ocean circulation theory is a significant body of knowledge (Marshall & Plumb, 2008; Pedlosky, 1996; Warren & Wunsch, 1981).

20.2 LARGE-SCALE OCEAN DYNAMICS (SVERDRUP DYNAMICS)

Because oceanic basins have dimensions comparable to the size of the earth, model accuracy demands the use of spherical coordinates, but because the present book only intends to present an introduction to physical oceanography, clarity of exposition trumps accuracy, and we continue to use Cartesian coordinates, with inclusion nonetheless of the beta effect (see Section 9.4). Spherical

[2]Harald Ulrik Sverdrup (1888–1957), Norwegian oceanographer who made his greatest contributions while being director of the Scripps Institution of Oceanography in California. A unit of volumetric ocean transport bears his name: 1 sverdrup = 1 Sv = 10^6 m^3/s.

[3]See biography at end of this chapter.

coordinates (Appendix A) do not change the qualitative nature of the theoretical results exposed here (Pedlosky, 1996, Chapter 1).

Large-scale flows in the main thermocline and abyss are slow and nearly steady. Their long timescales allow us to ignore all time derivatives, whereas their low velocities over long distances make their Rossby number very small and allow us to neglect the nonlinear advection terms in the momentum equations. Furthermore, there is a strong indication that dissipation is not an important feature of large-scale dynamics, at least not at the leading order (Pedlosky, 1996, page 6). Without time derivatives, advection, and dissipation, the horizontal momentum equations reduce to the geostrophic balance:

$$-fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} \tag{20.1a}$$

$$+fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y}, \tag{20.1b}$$

in which the Coriolis parameter f includes the so-called beta effect, which is important over the long length scales under consideration:

$$f = f_0 + \beta_0 y. \tag{20.2}$$

The y-coordinate is thus directed northward, leaving the x-direction to point eastward. The coefficients (see Eq. 9.18) $f_0 = 2\Omega \sin\varphi$ and $\beta_0 = 2(\Omega/a)\cos\varphi$ both depend on the choice of a reference latitude φ, which may be taken as the middle latitude of the basin under consideration.

The geostrophic equations are complemented by the hydrostatic balance

$$\frac{\partial p}{\partial z} = -\rho g, \tag{20.3}$$

the continuity equation (mass conservation for an incompressible fluid)

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{20.4}$$

and the energy equation, which states conservation of heat and salt but is expressed as conservation of density, again with neglect of the time derivative:

$$u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + w\frac{\partial \rho}{\partial z} = 0. \tag{20.5}$$

In the preceding equations, u, v, and w are the velocity components in the eastward, northward, and upward directions, respectively, ρ_0 is the reference density (a constant), ρ is the density anomaly, the difference between the actual density and ρ_0, p is the hydrostatic pressure induced by the density anomaly ρ, and g is the earth's gravitational acceleration (a constant). This set of five equations for five unknowns (u, v, w, p, and ρ) is sometimes referred to as *Sverdrup dynamics*.

Note that the problem is nonlinear owing to product of unknowns in the density equation (20.5). We now proceed with the study of some of its most immediate properties.

20.2.1 Sverdrup Relation

Elimination of pressure between the two momentum equations, by subtracting the y-derivative of Eq. (20.1a) from the x-derivative of Eq. (20.1b) yields

$$\frac{\partial}{\partial x}(fu)+\frac{\partial}{\partial y}(fv)=0, \tag{20.6}$$

or since f is a function of y but not of x,

$$f\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+\beta_0 v=0. \tag{20.7}$$

With the use of continuity equation (20.4), it can be recasted as

$$\beta_0 v=f\frac{\partial w}{\partial z}. \tag{20.8}$$

This most simple equation, called the *Sverdrup relation*, has a clear physical meaning. The factor $\partial w/\partial z$ represents vertical stretching, and any stretching ($\partial w/\partial z>0$) or squeezing ($\partial w/\partial z<0$) demands a change in vorticity for the sake of potential-vorticity conservation, which holds in the absence of dissipation. There is no relative vorticity here,[4] and the only way for a parcel of fluid to change its vorticity is to adjust its planetary vorticity (Fig. 20.4). This requires a meridional displacement, and hence meridional velocity v, to reach the correct f value.

However, this is not the end of the story. We can go further with a vertical integration:

$$\beta_0\int_{-H}^{-d} v\,\mathrm{d}z=f[w(z=-d)-w(z=-H)], \tag{20.9}$$

where $z=-H(x,y)$ represents the ocean bottom and $z=-d$ the base of the Ekman layer (Fig. 20.3). Note that in performing this integration, we have taken the liberty of including the seasonal thermocline under the assumption that although it is marked by seasonal variations, that time scale is long compared to the inertial period and its flow is nearly geostrophic.

Abyssal flow is extremely slow, and we may very well take $w(z=-H)=0$ at the bottom regardless of whether the bottom is sloping or is accompanied by

[4]Relative vorticity scales as U/L while planetary vorticity changes are on the order of βL. The former is much smaller than the latter because motion is slow and over a long scale, with $L^2\gg U/\beta$.

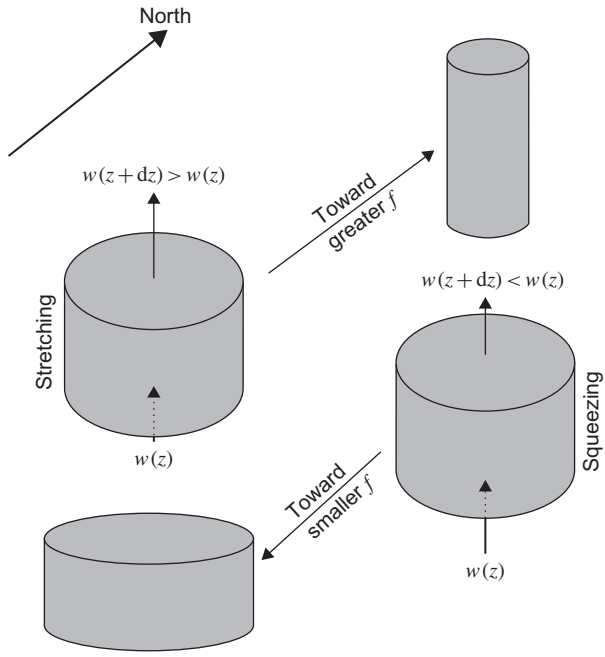


FIGURE 20.4 Meridional migration of fluid parcels induced by vertical stretching or squeezing in the large-scale oceanic circulation.

a bottom Ekman layer. At the base of the Ekman layer, the vertical velocity is the Ekman pumping

$$w_{\mathrm{Ek}} = \frac{1}{\rho_0}\left[\frac{\partial}{\partial x}\left(\frac{\tau^y}{f}\right) - \frac{\partial}{\partial y}\left(\frac{\tau^x}{f}\right)\right], \tag{20.10}$$

as in Eq. (8.36).

20.2.2 Sverdrup Transport

If we define the meridional transport as the vertically integrated north-south velocity, $\mathrm{V} = \int_{-H}^{-d} v\,\mathrm{d}z$, from the ocean's bottom to the base of the Ekman layer, Eqs. (20.9) and (20.10) provide

$$\begin{aligned}\mathrm{V} &= \frac{f}{\beta_0} w(z=-d) \\ &= \frac{f}{\beta_0} w_{\mathrm{Ek}} \\ &= \frac{f}{\rho_0 \beta_0}\left[\frac{\partial}{\partial x}\left(\frac{\tau^y}{f}\right) - \frac{\partial}{\partial y}\left(\frac{\tau^x}{f}\right)\right].\end{aligned} \tag{20.11a}$$

To this we can add the Ekman-layer contribution given by Eq. (8.34b):

$$V_{Ek} = -\frac{\tau^x}{\rho_0 f}, \tag{20.11b}$$

for a total of

$$\begin{aligned} V_{total} &= V + V_{Ek} \\ &= \frac{1}{\rho_0 \beta_0}\left(\frac{\partial \tau^y}{\partial x} - \frac{\partial \tau^x}{\partial y}\right). \end{aligned} \tag{20.11c}$$

We note this surprising result that the vertically cumulated flow component in the north-south direction is not dependent on the basin shape, size or overall wind-stress distribution but is solely dependent on the local curl of the wind stress. This equation is called the *Sverdrup transport*.

However, the same cannot be said of the zonal transport, which we define as the vertical integration of the east-west velocity u. Lumping all layers of the ocean together, we can obtain the total transport $U_{total} = \int_{-H}^{0} u\,dz$ directly from the vertically integrated continuity equation (20.4):

$$\frac{\partial U_{total}}{\partial x} + \frac{\partial V_{total}}{\partial y} = 0, \tag{20.12}$$

which yields

$$U_{total} = -\frac{1}{\rho_0 \beta_0} \int_{x_0}^{x} \frac{\partial}{\partial y}\left(\frac{\partial \tau^y}{\partial x} - \frac{\partial \tau^x}{\partial y}\right) dx, \tag{20.13}$$

where the starting point of integration ($x = x_0$) is to be selected wisely.

Ideally, we wish to impose a boundary condition on the flow at both eastern and western ends of the basin. For example, if we consider a basin limited on both eastern and western sides by north-south coastlines (a fair approximation of the major oceanic basins), the zonal flow and its vertical integral (U_{total}) ought to vanish at those ends. However, this is impossible to require simultaneously because there is only one constant (x_0) to adjust. If we set $x_0 = x_E$, the value of x at the eastern shore of the basin, then we enforce the impermeable-wall condition on the eastern side but make no provision for meeting any boundary condition on the western side, and vice versa if we take $x_0 = x_W$, the value of x at the western shore of the basin. The consequence is that the theory fails at one end of the domain, and as we will see in Section 20.3, a boundary layer must exist at one of the sides, which turns out to be western boundary.

20.2.3 Thermal Wind and Beta Spiral

In the ocean interior, the flow is approximately geostrophic, as we assumed when we wrote Eqs. (20.1). If we now take the vertical derivative of these equa-

tions and then eliminate $\partial p/\partial z$ by use of the hydrostatic balance (20.3), we obtain the thermal-wind relations:

$$\frac{\partial u}{\partial z} = +\frac{g}{\rho_0 f}\frac{\partial \rho}{\partial y} \tag{20.14a}$$

$$\frac{\partial v}{\partial z} = -\frac{g}{\rho_0 f}\frac{\partial \rho}{\partial x}. \tag{20.14b}$$

These are powerful relations in analyzing large-scale oceanic data. While oceanic velocities are difficult to measure,[5] density data are comparatively easy to obtain by dropping a Conductivity-Temperature-Depth (CTD) probe at repeated intervals from a ship cruising across the ocean. After some smoothing over mesoscale wiggles, the data provide the large-scale trends of density across the ocean basin, and it is relatively straightforward to determine the zonal and meridional gradients of density. Thus, we can consider $\partial\rho/\partial x$ and $\partial\rho/\partial y$ as known quantities and, from them, infer the velocity shear (20.14).

To obtain the actual velocity components u and v requires an additional assumption. The traditional approach is to assume a level of no motion, a deep horizon along which the pressure field is assumed to be uniform. Vertical integration of Eq. (20.14) from this level upward then provides the horizontal velocity up to the surface. This approach works well if the deep horizon is chosen within the relatively quiet abyssal layer, and the upward integration is performed to obtain the flow field in the main thermocline (Talley, Pickard, Emery & Swift, 2007).

Let us now return to the thermal-wind relations and extract from them an interesting property. For this, we decompose the horizontal velocity (u, v) in its magnitude U and azimuth θ:

$$u = U\cos\theta, \quad v = U\sin\theta. \tag{20.15}$$

The azimuth angle θ is measured counterclockwise from east and is related to the velocity components by $\theta = \arctan(v/u)$. Its vertical variation is given

$$\begin{aligned}\frac{\partial\theta}{\partial z} &= \frac{1}{u^2+v^2}\left(u\frac{\partial v}{\partial z} - v\frac{\partial u}{\partial z}\right)\\ &= \frac{-g}{\rho_0 f U^2}\left(u\frac{\partial\rho}{\partial x} + v\frac{\partial\rho}{\partial y}\right)\\ &= \frac{gw}{\rho_0 f U^2}\frac{\partial\rho}{\partial z},\end{aligned} \tag{20.16}$$

where the last step makes use of Eq. (20.5).

[5] There are several reasons why horizontal velocities are notoriously difficult to obtain directly from the deep ocean. First, they are almost always fluctuating at the mesoscale, making the time average easily fall within the noise level. Second, deep-water currentmeter moorings are expensive and wobble, thus adding an instrumental drift component to the velocity.

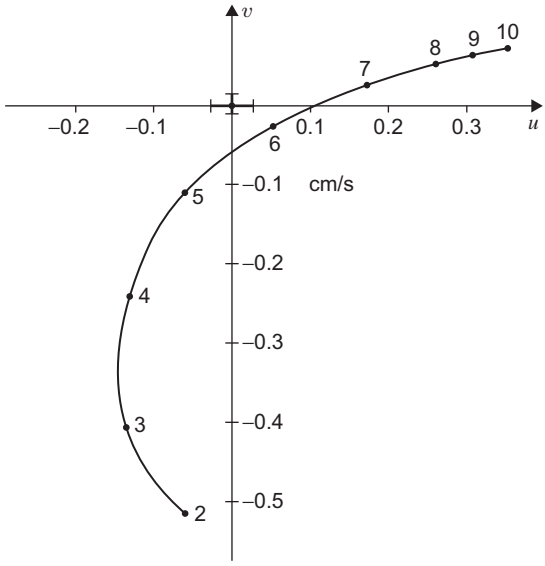


FIGURE 20.5 Beta spiral constructed from hydrographic data in the vicinity of 28°N, 36°W in the North Atlantic ocean. Numbers along the curve indicate depth in units of 100 m. Error bars are shown at the origin. (*Redrawn from Stommel & Schott, 1977*)

As we can see, there is a direct relation between the vertical velocity and the veering (twisting) of the horizontal velocity in the vertical. In the northern hemisphere ($f > 0$) and in the presence of a gravitationally stable water column ($\partial\rho/\partial z < 0$), $\partial\theta/\partial z$ has the opposite sign of w. Thus, in the midlatitudes, where the wind-stress curl is clockwise and Ekman pumping downward, the vertical velocity w is generally negative, and the vector of horizontal velocity turns clockwise with depth. This property has been dubbed the *beta spiral* (Schott & Stommel, 1978; Stommel & Schott, 1977). Figure 20.5 shows an example from the North Atlantic Ocean.

The veering implies that the waters at different levels in the vertical come from different directions and thus possess different origins. However, all levels of motion are under the tight constraint of the Sverdrup transport (20.11c). The local wind-stress curl appears therefore as a constraint on, rather than the forcing of, the flow.

20.2.4 A Bernoulli Function

In Sverdrup dynamics, the Montgomery potential defined as $P = p + \rho g z$ [see Eq. (12.4)] happens to play the role of a Bernoulli function. To show this, we need to prove that P is conserved along streamlines. Thus, we calculate its material derivative, which for steady flow contains only the spatial derivatives:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = u\frac{\partial P}{\partial x} + v\frac{\partial P}{\partial y} + w\frac{\partial P}{\partial z}$$

$$= u\left(\frac{\partial p}{\partial x} + gz\frac{\partial \rho}{\partial x}\right) + v\left(\frac{\partial p}{\partial y} + gz\frac{\partial \rho}{\partial y}\right)$$
$$+ w\left(\frac{\partial p}{\partial z} + \rho g + gz\frac{\partial \rho}{\partial z}\right). \tag{20.17}$$

Use of the geostrophic relations (20.1) and hydrostatic balance (20.3) to eliminate all three derivatives of p leads to the cancellation of several terms, leaving

$$\frac{\mathrm{d}P}{\mathrm{d}t} = gz\left(u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + w\frac{\partial \rho}{\partial z}\right), \tag{20.18}$$

which is identically zero by virtue of density conservation (20.5). Thus, the Montgomery potential P is conserved along the flow.[6] The same result could have been obtained more directly if we had taken the preliminary trouble of expressing the equations in density coordinate (Chapter 12).

While the preceding result holds some appeal, it is rarely useful because it is nearly impossible in the deep ocean to extract the dynamic signal of the pressure field from pressure measurements dominated by the hydrostatic component (that due to ρ_0!) and depth uncertainties caused by the variable sea surface. In contrast, potential vorticity, which depends on the density field, is of greater use with data.

20.2.5 Potential Vorticity

In the low Rossby number regime, momentum advection is negligible compared to the Coriolis acceleration, and consequently, the formulation of potential vorticity does not include relative vorticity next to planetary vorticity. We thus expect that in large-scale ocean dynamics, the expression of potential vorticity reduces to planetary vorticity over layer thickness. Specifically, the form is

$$q = \frac{f}{\Delta z/\Delta \rho}$$
$$= -\frac{f}{\partial z/\partial \rho} \tag{20.19}$$

in density coordinates, to become in depth coordinates

$$q = -f\frac{\partial \rho}{\partial z}. \tag{20.20}$$

[6] The more general expression of the Bernoulli function is $B = \rho_0(u^2 + v^2 + w^2)/2 + p + \rho gz$, but the kinetic-energy term is absent here as a consequence of the neglect of advection in the momentum equations, leaving only the last two terms, which together form the Montgomery potential.

To show that this expression is indeed conserved for Sverdrup dynamics, we begin by taking the vertical derivative of the density equation (20.5):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\rho}{\partial z}\right)+\frac{\partial u}{\partial z}\frac{\partial\rho}{\partial x}+\frac{\partial v}{\partial z}\frac{\partial\rho}{\partial y}+\frac{\partial w}{\partial z}\frac{\partial\rho}{\partial z}=0. \tag{20.21}$$

If we now eliminate the z-derivatives of u and v by using the thermal-wind relations (20.14) and the Sverdrup relation (20.8) to eliminate $\partial w/\partial z$, the middle terms cancel out, and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\rho}{\partial z}\right)+\frac{\beta_0 v}{f}\frac{\partial\rho}{\partial z}=0. \tag{20.22}$$

Then recognizing that $\beta_0 v = v\,(\mathrm{d}f/\mathrm{d}y) = \mathrm{d}f/\mathrm{d}t$, the equation can be recast as

$$f\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\rho}{\partial z}\right)+\frac{\mathrm{d}f}{\mathrm{d}t}\frac{\partial\rho}{\partial z}=\frac{\mathrm{d}}{\mathrm{d}t}\left(f\frac{\partial\rho}{\partial z}\right)=0, \tag{20.23}$$

which means that the expression $f(\partial\rho/\partial z)$ is conserved along the flow. Thus, the potential vorticity defined above in Eq. (20.20) is conserved by individual water parcels. This conservation law, like conservation of the Bernoulli function shown in the previous subsection, holds wherever the flow is on a large scale and dissipation is weak, that is, in the main thermocline and abyssal layer.

It is interesting to return to the beta spiral and interpret it in the light of potential vorticity. The connection is illustrated in Fig. 20.6, which ties the following ingredients: conservation of potential vorticity, thermal wind, and the effect of

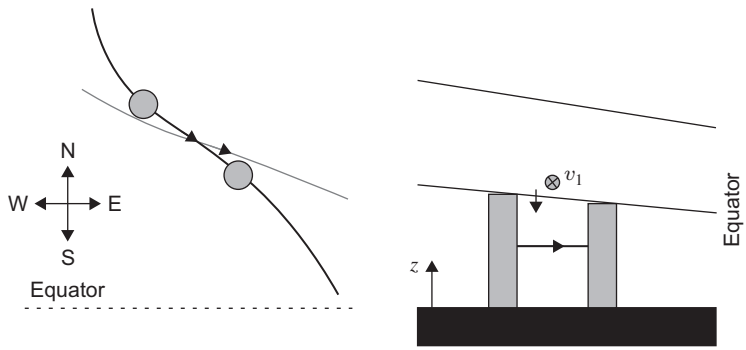


FIGURE 20.6 Along its journey toward the equator, a fluid column experiences a decrease of planetary vorticity f, which is compensated by a decrease in layer height. If we follow the flow in the vertical transect along the trajectory (right panel), there is, by definition of the trajectory, no transverse flow at the level of the water column being followed. Because the layer height decreases approaching the equator, a thermal wind must appear (v_1 on the right panel and gray trajectory in the left panel). With zero transverse flow below, the horizontal velocity vector turns clockwise with depth, in agreement with Eq. (20.16) and downward vertical movement.

vertical velocity. It shows how the necessary squeezing along a southward trajectory must be accompanied by a thermal wind that creates veering with depth and hence the beta spiral.

We have seen that in the ocean's interior, three quantities are conserved simultaneously by traveling water parcels: their density ρ, their Montgomery potential P, and their potential vorticity q. Since it retains its ρ value along the way, a moving parcel is confined to stay on an isopycnal surface. Similarly with P, the same parcel is also confined to stay on a surface of constant P value. Putting both constraints together, we conclude that a trajectory is an intersection line between a ρ surface and a P surface. Since all trajectories are also lines of constant q, it follows that the potential vorticity q must be everywhere a function of density ρ and Montgomery potential P:

$$q = Q(\rho, P). \tag{20.24}$$

In considering data from the North Atlantic Ocean, Williams (1991) found that the density and depth of the seasonal thermocline at its late-winter coldest and deepest stage are such that the waters deposited in the main thermocline are characterized by a potential vorticity that is nearly homogeneous along isopycnal surfaces. This is particularly true for densities in the range 1026.4–1026.75 kg/m^3, corresponding to the main thermocline between 30° and 40°N. Where this occurs, we may reduce the preceding relation to a function of a single variable:

$$q = Q(\rho). \tag{20.25}$$

Since both density and potential vorticity are easily observed variables, mapping these is a standard procedure in the construction of ocean circulation from observations (see also Optimal Interpolation in Section 22.3). Mapping can show isolines of potential vorticity on isopycnal surfaces. Because both density and potential vorticity are nearly conserved, potential vorticity then serves as a tag to trace water movement.

20.3 WESTERN BOUNDARY CURRENTS

In commenting on Eq. (20.13), we concluded that the simplified Sverdrup dynamics do not allow the simultaneous enforcement of impermeability boundary conditions at both eastern and western sides of an ocean basin. Thus, Sverdrup dynamics must break down at one of the sides. The answer is that a boundary layer with scale shorter than the basin scale must exist on the western side of the basin. Let us now verify this.

Over the midlatitude ocean basins, the wind pattern is clockwise (counterclockwise) in the northern (southern) hemisphere because trade winds blow from east to west in the tropics and westerly winds blow from west to east further away from the equator (see Fig. 20.2). The resulting Ekman pumping is

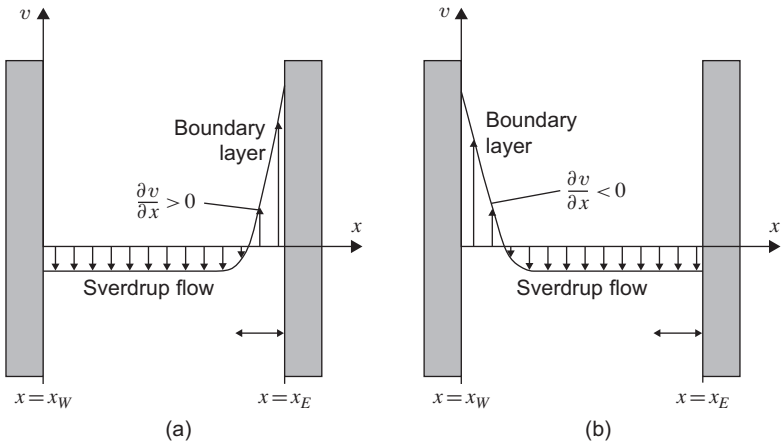


FIGURE 20.7 The two possible configurations for a northward boundary current to compensate the southward Sverdrup flow that exists across most of a midlatitude ocean basin of the northern hemisphere: (a) boundary current on the eastern side, (b) boundary current on the western side. The former is to be rejected on dynamic grounds, leaving the latter as the correct configuration.

downward, and the Sverdrup transport (20.11c) is equatorward (in both hemispheres). Conservation of mass requires that water flow toward the equator must be compensated elsewhere by poleward flow, but poleward flow violates Sverdrup dynamics. Thus, this return flow must exist on a scale shorter than the long length scale invoked in Sverdrup dynamics. In other words, it exists in the form of a narrow boundary current. With two meridional boundaries, one on each side of the basin, there are only two possibilities: either the poleward current follows the eastern boundary (Fig. 20.7a) or it follows the western boundary (Fig. 20.7b). In each case, connection with the equatorward Sverdrup flow in the basin's interior creates a velocity gradient (shear) and thus relative vorticity.

To make the argument easier, let us restrict our attention to the northern hemisphere. (The conclusion continues to hold for the southern hemisphere.) If the boundary layer lies along the eastern wall as shown in Fig. 20.7a, the northward flow has positive $\partial v/\partial x$. This derivative is large because the boundary layer is narrow and also the velocity must be large to accommodate the entire Sverdrup transport. By comparison, $\partial u/\partial y$ is very small. Thus, if flow is returning on the eastern side, it has positive relative vorticity ($\zeta > 0$). On the contrary, if the boundary layer lies along the western wall (Fig. 20.7b), the velocity shear $\partial v/\partial x$ is negative, and because $\partial u/\partial y$ is still negligible, the return flow has negative vorticity ($\zeta < 0$). But which type of vorticity is the return flow allowed to have?

When it makes its return journey back toward higher latitudes, water sees its planetary vorticity f increases, and conservation of potential vorticity

$(f+\zeta)/h$ demands that ζ or h or both change accordingly. Since relative vorticity ζ becomes important in the boundary current, the increase in f must be accompanied by a decrease in ζ, thus the value of ζ must drop from about zero to a negative value as fluid parcels exit the Sverdrup interior and enter the boundary current. The western boundary current does accomplish that, whereas the eastern boundary current does not. So, we reject the existence of a boundary layer along the eastern side of the basin and conclude that the necessary boundary layer must lie along the western side of the basin (Fig. 20.7b).

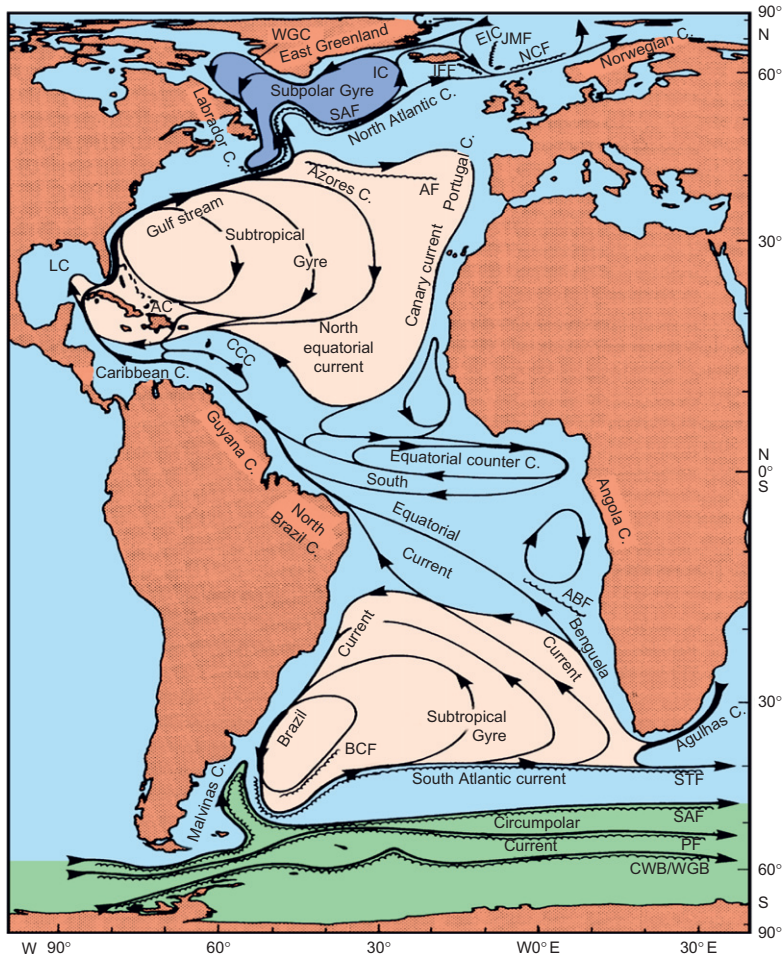


FIGURE 20.8 The average circulation in the Atlantic Ocean, showing equatorward Sverdrup flow in the midlatitudes of each hemisphere and western boundary currents. (*Tomczak & Godfrey, 2003*)

It is evident that the circulation takes the form of an asymmetric gyre, with a slow equatorward flow (the Sverdrup transport) occupying most of the domain and a swift narrow current on the western side that returns the water poleward (Fig. 20.8). The latter current is naturally identified with the Gulf Stream in the North Atlantic Ocean and the respective western-boundary currents in the other ocean basins: Brazil Current in the South Atlantic, Kuroshio in the North Pacific, East Australia Current in the South Pacific, and Agulhas Current in the south Indian Ocean. The circulation closes with zonal currents that connect the entrance and exit of the western boundary current with the interior Sverdrup flow. The funneling of the equatorward Sverdrup transport into a strong return current on the western side of the ocean basin has been termed *westward intensification* by Henry Stommel, who provided the first correct theory for the existence of the Gulf Stream (Stommel, 1948).

With the solution now at hand, let us recapitulate the results. As we are becoming aware, the mechanisms of ocean circulation are intricate and certainly less direct than a simple torque exerted by a surface stress on a viscous fluid. The chief reasons are that viscosity is weak, and the Coriolis effect is strong, including its variation with latitude.

The scenario is as follows. The large-scale atmospheric winds, comprising essentially the trades and westerlies, generate a stress along the ocean surface. Because seawater is only slightly viscous and planetary rotation is strong, the direct effect of the stress is limited to a thin (10 m or so) layer of the ocean. The earth's rotation also generates a component of the upper layer flow transverse to the winds, which converges, resulting in a downward flow into the ocean interior below (Ekman pumping). Although relatively weak on the order of 10^{-6} m/s, which is about 30 m per year, this vertical flow squeezes water parcels vertically. In reaction, fluid parcels flatten and widen, and their planetary vorticity decreases in order to conserve their circulation. They are thus forced to migrate equatorward. As they progress toward the equator, these waters run into a region of slower flow, veering westward and gathering into a zonal flow that intensifies downstream. On arriving at the western boundary, the waters turn and form a swift poleward flow, so swift that their relative vorticity becomes sufficient to compensate for the adverse change in planetary vorticity.

Obviously, all of our assumptions have eliminated a considerable number of additional processes that can all affect the ocean circulation in one way or another. Inertia (represented by the nonlinear advection terms) is important in the western boundary layer where the flow is swift and narrow (Rossby number becoming of order 1). The consequence is a detachment of this intense current from the coast and its penetration into the ocean interior, where it starts to meander rather freely. Barotropic instabilities (see Chapter 10) are likely. Stratification is another aspect that requires ample consideration. Briefly, the effect of stratification is to decouple the flow in the vertical and thus to make it respond less to bottom friction. On the other hand, a reserve of potential energy

due to the presence of stratification causes baroclinic instabilities (Chapter 17). Barotropic and baroclinic instabilities generate eddies, and these in turn create a net horizontal mixing of momentum. Finally, because poleward western-boundary currents, such as the Gulf Stream, bring warm water masses to higher latitudes, an air–sea heat flux is created, resulting in the cooling of the ocean and a distortion of the circulation pattern. The interested reader will find additional information on ocean-circulation dynamics in the review article by Veronis (1981), the book by Abarbanel and Young (1987), and the article by Cushman-Roisin (1987a).

20.4 THERMOHALINE CIRCULATION

As stated in Section 20.1, the region below the seasonal thermocline is comprised of two subregions, the main thermocline and the abyssal layer, together called the ocean interior. The dynamics expounded in the previous section are applicable to both these regions. We now turn our attention more specifically to the upper of these two layers, the main thermocline.

In contrast to the abyss, which is fed by deep-water convection at high latitudes, the main thermocline is the region of the ocean in which the circulation is primarily caused by the wind-driven Ekman pumping received from the surface layer and is most pronounced at midlatitudes. Figure 20.9, compiled by Talley et al. (2007), summarizes the meridional distribution of density in the North and South Atlantic, North and South Pacific, and Indian Oceans. They reveal similar patterns in all five oceans. The pycnocline is very strong and shallow (100–200 m) at the equator; from there, it spreads vertically downward toward the poles, with a tendency to split into two branches: one surfacing around 25° latitude and the other plunging down to 1000 m around 35° before heading upward again and surfacing around 45° latitude.

20.4.1 Subduction

At the top of the unchanging ocean interior, water is exchanged with the seasonal thermocline. The process called *subduction*, if the water passes from the seasonal thermocline to the steady interior, or *entrainment*, if water is engulfed by the seasonal thermocline, deserves special attention.

There are several processes by which water enters the main thermocline: downward Ekman pumping injection, retreat of the convective mixed layer at the end of winter, and convergence of the flow inside the mixed layer (Cushman-Roisin, 1987b). The sum of these processes creates *subduction*, which can be defined as the deposition of fluid formerly belonging to the seasonal thermocline into the ocean interior. It is expressed mathematically as a volumetric flow rate per unit horizontal area (i.e., with dimensions of velocity, although it is not a vertical velocity).

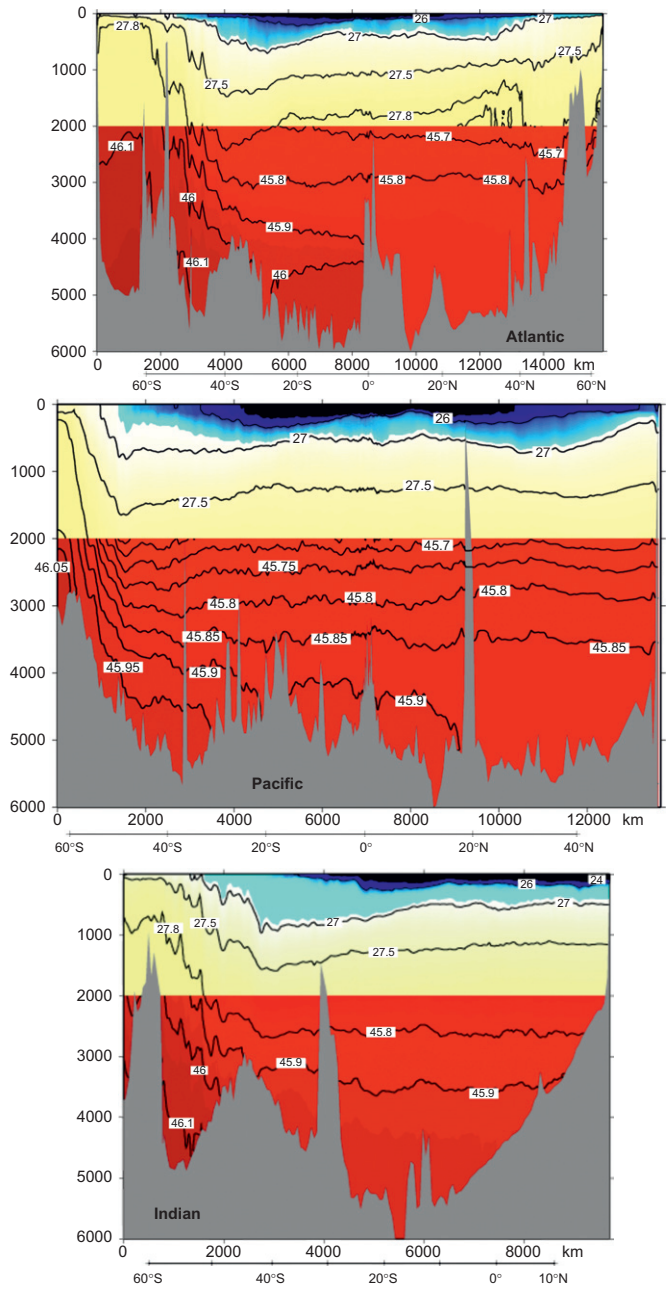


FIGURE 20.9 Meridional structure of potential density (σ_θ top and σ_4 bottom, in kg/m^{-3}) in the five major ocean basins. Top panel: North and South Atlantic. Middle panel: North and South Pacific. Bottom panel: Indian Ocean. Data were gathered during the World Ocean Circulation Experiment (1990–2002). The variables σ_θ and σ_4 are density anomalies corrected for the effect of compressibility under high pressure. *(Talley et al., 2007).*

Consider a fluid column of infinitesimal width and extending from the base of the seasonal thermocline upward to the base fluid of the Ekman layer. The density is this fluid column is vertically uniform because the seasonal thermocline is in a state of mixing (by definition). Denoting its height by $h(x, y, t)$, its volume budget is

$$\frac{\partial h}{\partial t} + \frac{\partial (hu_{\mathrm{ST}})}{\partial x} + \frac{\partial (hv_{\mathrm{ST}})}{\partial y} = -w_{\mathrm{Ek}} - Su, \tag{20.26}$$

where u_{ST} and v_{ST} are the eastward and northward velocity components inside the seasonal thermocline, w_{Ek} the Ekman pumping at its top (base of Ekman layer—see Fig. 20.3), and Su the subduction rate at its bottom. Positive Su represents subduction proper (flow into the main thermocline below), whereas negative Su corresponds to capture of ocean interior water by the seasonal thermocline, that is entrainment. A negative w_{Ek} brings water from the surface mixed layer down and into the seasonal thermocline.

If we now assume that the horizontal velocity (u_{ST}, v_{ST}) is characterized on a seasonal scale by low frequency and long length scale, it must be nearly geostrophic, and we may write

$$-\rho_0 f v_{\mathrm{ST}} = -\frac{\partial p_{\mathrm{ST}}}{\partial x}, \quad +\rho_0 f u_{\mathrm{ST}} = -\frac{\partial p_{\mathrm{ST}}}{\partial y}, \tag{20.27}$$

with $f = f_0 + \beta y$ and p_{ST} the pressure inside the seasonal thermocline. The divergence term of Eq. (20.26) becomes

$$\frac{\partial (hu_{\mathrm{ST}})}{\partial x} + \frac{\partial (hv_{\mathrm{ST}})}{\partial y} = J\left(\frac{p_{\mathrm{ST}}}{\rho_0}, \frac{h}{f}\right), \tag{20.28}$$

and Eq. (20.26) can then be recast as

$$Su = -\frac{\partial h}{\partial t} - w_{\mathrm{Ek}} - J\left(\frac{p_{\mathrm{ST}}}{\rho_0}, \frac{h}{f}\right), \tag{20.29}$$

which shows that subduction is a combination of seasonal thermocline retreat, downward Ekman pumping, and convergence of geostrophic flow in the seasonal thermocline.

Before continuing with motion inside the main thermocline, it is worth making a few remarks concerning the temporal variability of subduction. As Stommel (1979) remarked, much of the water left behind during spring and summer when the thermocline retreats is recaptured the following fall and winter, because it has not had the time to sink deeply enough into the interior before the seasonal thermocline penetrates once again. Much of subduction is in vain, and effective feeding of the ocean interior by subduction occurs only during a relatively brief time interval in late winter. This explains why water properties in the ocean interior systematically reflect surface water properties of late

winter and never those of summer (Stommel, 1979). Cushman-Roisin (1987b) explored in somewhat greater detail the kinematics of seasonal subduction and concluded that the time of effective subduction is not quite as brief as Stommel suggested but that about 30% of the volume subducted escapes recapture and feeds the ocean interior.

20.4.2 Ventilated Thermocline Theory

Early theories of the main thermocline attempted to explain the observed vertical temperature structure as a local advection–diffusion equilibrium between upwelling of cold abyssal water and downward diffusion of heat from the surface or sought analytical solutions based on assumptions dictated by mathematical convenience. Then, the paradigm changed in the 1980s with the publication of a paper by Luyten, Pedlosky and Stommel (1983) titled "The ventilated thermocline." The theory combines subduction from the mixed layer with advective descent into the stratified thermocline.

The scenario is as follows. In the midlatitude ocean, where Ekman pumping is downward, mixed-layer waters are subducted into the main thermocline where they slide along density surfaces, carrying with them their surface properties such as density and potential vorticity, which are set at subduction time. Layers of thermocline water that can be traced back to the base of the mixed layer where Ekman pumping is downward are said to be *ventilated*, and their intersection with the mixed layer where they are supplied is called the *outcrop line*. The vertical structure of the main thermocline in the ventilated area thus reflects the density distribution at the surface in winter (time of effective subduction), a fact long noted by Iselin (1938).

Under these premises, the thermocline problem reduces to solving for the vertical layering of density by tracking individual density layers upstream to their respective outcrop lines, where the (winter) surface density distribution is known. As the idea of subduction and subsequent sliding along density surface suggests a thermocline controlled by advection, a theory using an inviscid and nondiffusive fluid seems appropriate (Luyten et al., 1983). The theory was later extended to continuous stratification by Huang (1989). The situation is complicated by the fact that only a sector of the main thermocline can be traced back to outcrops where surface conditions are known; the remaining portions, on the eastern and western sides, form so-called *shadow zones*. In these zones, the flow is circulating without surface contact, and it has been speculated that it is characterized by a slow but effective homogenization of potential vorticity (Huang, 1989).

Instead of launching in an exposition of the ventilated-thermocline theory and the attending complication of shadow zones (for this, the reader is referred to chapter 4 of Pedlosky, 1996), we shall limit ourselves here to determining a scale for the vertical thickness of the main thermocline.

20.4.3 Scaling of the Main Thermocline

The dynamics of the circulation in the main ocean thermocline are governed by a small number of parameters, namely the constants f_0, β_0, ρ_0, and g that enter the governing equations, and a few external scales that enter through boundary conditions: L_x the width of the basin, W_{Ek} a typical magnitude of the Ekman pumping, and $\Delta\rho$ a typical density variation across the thermocline.

Following Welander (1975), the thermocline depth h_{scale} can be derived by balancing the various terms in the equations of Sverdrup dynamics. First, the scale for pressure is $\Delta P = g h_{\text{scale}} \Delta\rho$ from the hydrostatic balance (20.3), from which follow the north-south velocity scale through (20.1a): $v \sim \Delta P/(\rho_0 f_0 L_x) = g h_{\text{scale}} \Delta\rho/(\rho_0 f_0 L_x)$. The vertical velocity must necessarily scale like the Ekman pumping velocity because it is equal to it at the base of the Ekman layer. With scales for both meridional and vertical velocities known, the Sverdrup relation (20.8) implies

$$\beta_0 \frac{g h_{\text{scale}} \Delta\rho}{\rho_0 f_0 L_x} \sim f_0 \frac{W_{\text{Ek}}}{h_{\text{scale}}} \tag{20.30}$$

from which follows the depth scale h_{scale} of the main thermocline:

$$h_{\text{scale}} = \sqrt{\frac{f_0^2 L_x W_{\text{Ek}}}{\beta_0 g (\Delta\rho/\rho_0)}}. \tag{20.31}$$

To see whether this scale is reasonable, let us use numbers corresponding to the North Atlantic Ocean circa 35°N. At that latitude, the Coriolis parameters are $f_0 = 8.4 \times 10^{-5}$ s^{-1} and $\beta_0 = 1.9 \times 10^{-11}$ m^{-1}s^{-1}, while the width of the basin, stretching from 10° to 80°W, gives $L_x = 6400$ km. With a wind stress curl leading to an Ekman pumping on the order of $W_{\text{Ek}} = 2 \times 10^{-6}$ m/s and a relative density difference $\Delta\rho/\rho_0$ of about 0.002, we obtain $h_{\text{scale}} = 490$ m, about 500 m as observations indicate.

Because ocean basins are much deeper[7] than this scale, there is much water lying below the main thermocline that is not subject to surface Ekman pumping. This is the abyssal layer, which is the subject of the next section. As we shall see, it is driven by deep convection under atmospheric cooling at high latitudes.

20.5 ABYSSAL CIRCULATION

Since the wind stress only affects a relatively small portion of the water column, the bulk of ocean waters forming the abyssal layer must be driven by another

[7] The average depth in the ocean is 3720 m.

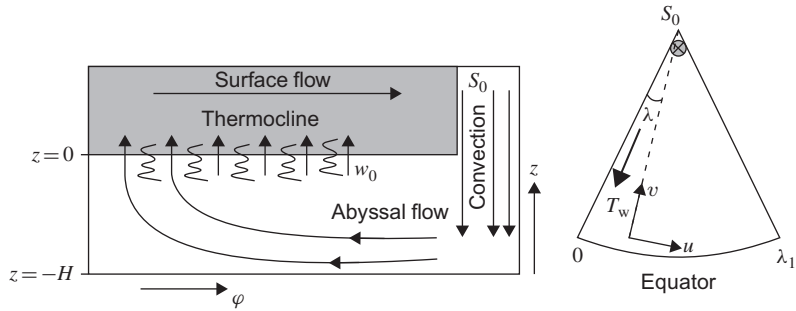


FIGURE 20.10 A highly simplified model of the abyssal circulation, with "longitude" arbitrarily taken as zero at the western boundary (Left panel: side view - Right panel: top view). The volumetric source S_0 is meant to represent dense-water formation at high latitudes.

mechanism. As it turns out, this greatest body of water is set in slow motion by vertical convection taking place over an exceedingly small portion of the ocean surface, narrow zones of extreme cooling at high latitudes.

The scenario is as follows. Exposure to a very cold and dry atmosphere at high latitudes causes both thermal contraction and evaporation. Evaporation takes distilled water away, leaving the salt in the ocean. Brine rejection from water freezing into ice further increases the salinity of the remaining liquid water, and the result is a water that is both very cold and very salty, thus significantly denser than typical seawater. This dense water sinks under the effect of gravity, slowly but effectively filling the abyss of the world ocean. Known areas of dense-water formation by deep convection are the northern reaches of the Atlantic Ocean at the entrance of the Arctic Ocean and several marginal seas along the periphery of Antarctica (Weddell Sea and Ross Sea).

To close the circulation, slow upwelling in lower latitudes must take place to return the waters back to the surface, where they are heated and can once again migrate to high latitudes and complete the circuit (Fig. 20.10). In three dimensions, the totality of this circulation system forms the so-called *conveyor belt* (Fig. 20.1), and a round trip along this path is believed to take thousands of years.

By way of exception, we work here in spherical coordinates because the present analysis is not much more complicated than with Cartesian coordinates, and its results are easier to identify with actual oceanic features. The analysis follows closely that of Stommel and Arons (1960a,b).

The dynamical balance reduces as usual to geostrophy, now in spherical coordinates [see (A.18)] and is complemented by volume conservation:

$$-fv = -\frac{1}{a\cos\varphi}\frac{\partial p}{\partial \lambda} \tag{20.32a}$$

$$+fu = -\frac{1}{a}\frac{\partial p}{\partial \varphi} \tag{20.32b}$$

$$\frac{1}{a}\frac{\partial u}{\partial \lambda} + \frac{1}{a}\frac{\partial}{\partial \varphi}(v\cos\varphi) + \frac{\partial}{\partial z}(w\cos\varphi) = 0, \tag{20.32c}$$

where a is the earth's radius, λ longitude, φ latitude, z the local vertical coordinate, and $f = 2\Omega\sin\varphi$ the latitude-dependent Coriolis parameter. Elimination of pressure between the first two equations yields

$$\frac{\partial}{\partial \lambda}(fu) + \frac{\partial}{\partial \varphi}(fv\cos\varphi) = 0, \tag{20.33}$$

from where, by using volume conservation (20.32c), we recover the Sverdrup relation, now in spherical coordinates

$$\beta v = f\frac{\partial w}{\partial z}, \tag{20.34}$$

with $\beta = 2(\Omega/a)\cos\varphi$.

Vertical integration across the abyssal layer (from $-H$ to 0) from a flat bottom yields the Sverdrup transport

$$\int_{-H}^{0} v\,\mathrm{d}z = \mathrm{V} = aw_0\tan\varphi. \tag{20.35}$$

With positive velocity (upwelling) w_0 almost everywhere (since the sinking of dense water is confined to small corner regions), the abyssal flow must be northward in the northern hemisphere ($\varphi > 0$) and southward in the southern hemisphere ($\varphi < 0$), that is, everywhere poleward. We also note that there is no Sverdrup transport across the equator ($\varphi = 0$). This is evidently problematic since intuition would lead us to believe that convection at polar latitudes should create flow away from, not toward, the poles. The solution to this paradox is that the flow coming from the high-latitude regions is confined to narrow strips along the western margins of the oceanic basins, and the broader abyssal flow consists of the return flow toward higher latitudes.

For the following discussion, we now restrict our attention to the northern hemisphere. The water budget from any northern latitude φ up to the pole, including the deep-water source, demands that the flow carried by the interior (boundary integral of V) plus the deep-water inflow (S_0) be compensated by the western boundary layer flow (T_w), which goes further south, and the cumulated effect of upwelling at the top (surface integral of w_0):

$$S_0 + \int_0^{\lambda_1} \mathrm{V}\,a\cos\varphi\,\mathrm{d}\lambda = T_\mathrm{w}(\varphi) + \int_{\varphi}^{\pi/2}\int_0^{\lambda_1} w_0\,a^2\cos\varphi\,\mathrm{d}\lambda\,\mathrm{d}\varphi. \tag{20.36}$$

In the case of a uniform upwelling speed w_0 and with Eq. (20.35) to eliminate V, we have

$$S_0 + \sin\varphi \lambda_1 a^2 w_0 = T_w(\varphi) + (1 - \sin\varphi)\lambda_1 a^2 w_0. \tag{20.37}$$

The transport in the western boundary layer must therefore be

$$T_w(\varphi) = S_0 + (2\sin\varphi - 1)\lambda_1 a^2 w_0, \tag{20.38}$$

which is maximum at the pole ($\varphi = \pi/2$) and decreases toward the equator ($\varphi = 0$). Three cases can arise.

Case 1: $S_0 = \lambda_1 a^2 w_0$. In this case, upwelling from the equator to the pole exactly matches the source at the pole. The transport in the western boundary current is $T_w = 2S_0 \sin\varphi$, which vanishes at the equator ($\varphi = 0$). Since the Sverdrup transport is also zero at the equator, the two hemispheres are decoupled. We further note that near the pole, the boundary-layer flow is *twice* as strong as the source, while the northward Sverdrup flow is equal to the source implying that the half the flow is pure recirculation. This suprising feature seemed so counterintuitive to Stommel and Arons that they performed laboratory experiments (Stommel & Arons, 1960a,b) to verify their findings.

Case 2: $S_0 > \lambda_1 a^2 w_0$. The dense-water source is stronger than the distributed upwelling, and an excess transport spills across the equator into the other hemisphere. This is the situation encountered in the North Atlantic.

Case 3: $S_0 < \lambda_1 a^2 w_0$. The source is insufficient to sustain the required upwelling, and a northward boundary-layer flow across the equator is necessary to supply the difference. Such a situation is prevailing in the North Pacific.

The preceding solution can be extended to any w_0 distribution in the calculation of the zonal velocity u, requiring zero normal flow at the eastern boundary, consistent with the position of the unresolved boundary layer along the western boundary. From Eq. (20.32a), knowing the meridional velocity in a barotropic abyssal layer (20.35), we can calculate the pressure by imposing a constant value, taken as zero without loss of generality, along the eastern boundary:

$$p = \frac{2\Omega a^2}{H} \sin^2\varphi \int_{\lambda_1}^{\lambda} w_0 \, \mathrm{d}\lambda, \tag{20.39}$$

from which we deduce the zonal velocity according to Eq. (20.32b)

$$u = -\frac{a}{H \sin\varphi} \frac{\partial}{\partial\varphi} \left(\sin^2\varphi \int_{\lambda_1}^{\lambda} w_0 \, \mathrm{d}\lambda \right). \tag{20.40}$$

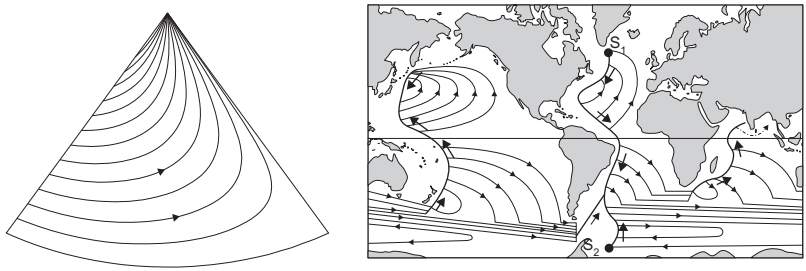


FIGURE 20.11 A few trajectories of the abyssal flow in a basin sector (left panel) and the abyssal circulation inferred by Stommel (1958) (right panel).

For uniform w_0, this reduces to

$$u = 2\frac{a}{H} w_0 (\lambda_1 - \lambda) \cos\varphi, \tag{20.41}$$

which is always directed eastward, implying that the boundary layer feeding this flow must be on the western side of the basin.

The velocity field (20.41) and (20.35) gives rise to the trajectories depicted in the left panel of Fig. 20.11, in the light of which we can understand the abyssal circulation proposed by Stommel (1958), right panel of Fig. 20.11.

The preceding theory serves as an interesting application of geophysical fluid dynamics but barely illustrates the complex dynamics of the abyssal circulation. Chief among our assumptions was that of a flat bottom. The oceanic bathymetry, as we all know, is rather fractured, with a multiplicity of ridges standing as obstacles and passages guiding the flow. Among the special features that ridges and passages inflict on the flow are concentrated zonal currents and recirculation patterns. The interested reader is referred to chapter 7 of Pedlosky (1996).

20.6 OCEANIC CIRCULATION MODELS

A milestone in numerical ocean modeling was the first *Ocean General Circulation Model*, or OGCM in short, developed by a team of scientists at Princeton University's Geophysical Fluid Dynamics Laboratory (Bryan, 1969; Bryan & Cox, 1972). The release to the scientific community of the source code of this model and its successive variants, of the type now called Modular Ocean Model (MOM), contributed to the model's widespread use, especially because the simple numerics lent themselves to various adaptations by many users. The code was based on straightforward second-order centered finite differencing of the governing equations in longitude-latitude coordinates with stepwise representation of coastlines and bottom topography. Time marching was leapfrog.

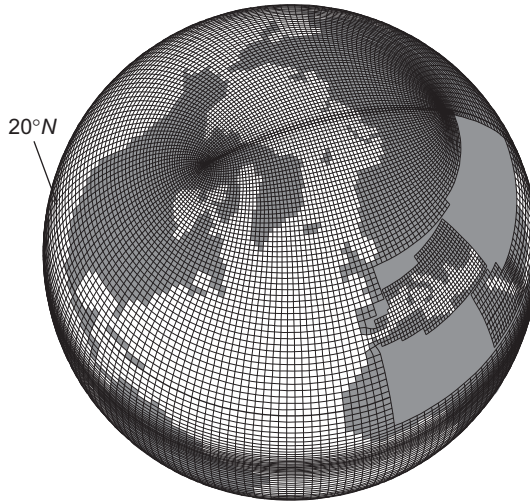


FIGURE 20.12 Ocean model grid with numerical grid poles over continents. (*LODYC, ORCA configuration; Madec, Delecluse, Imbard & Levy, 1998*)

This model enabled the first general circulation studies with the primitive equations (that is, without reliance on the quasi-geostrophic approximation, as it had been the case until then), but it eventually became apparent that a series of improvements were necessary. In particular, the polar convergence of meridians in spherical coordinates led to annoying singularities, creating the so-called *pole problem*. This was later avoided by shifting the spherical coordinate system to relocate its "poles" on continents or by using curvilinear orthogonal grids that maintain a topologically rectangular grid (Fig. 20.12). The stepwise topography (Fig. 20.13) was also ill fitted to simulate weak bottom slopes and the associated potential-vorticity constraint. In response, partially masked cells were introduced (e.g., Adcroft, Hill & Marshall, 1997). A further issue caused by stepwise topography was its poor representation of overflows typical of deep-water formation (Fig. 20.14). To tackle such situations, special algorithms have been developed (e.g., Beckmann & Döscher, 1997). Worthy of special note is DieCAST (Dietrich, 1998), which uses a modified Arakawa "A" grid with fourth-order accuracy in the horizontal directions. The unusually low dissipation of this model provides more accurate simulation of narrower features, such as boundary currents and mesoscale eddies.

In addition to these improvements, new generations of models are constantly being developed. Perhaps the most significant change in terms of numerical implementation is the move from structured to unstructured grids. The presence of continental boundaries remains an obstacle to the use of spectral models (preferred in modeling of the global atmosphere), while unstructured grids offer, by

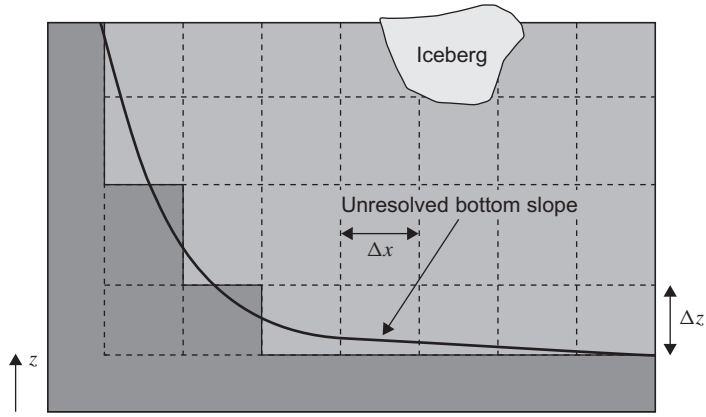


FIGURE 20.13 Masking of a regular grid allows to discretize topography. For rectangular grid boxes, weak slopes can only be resolved if $\Delta z < \Delta x|\partial b/\partial x|$ with bottom given by $z = b(x, y)$. If not, the weak slope is awkwardly approximated by a flat bottom stretching over several grid steps followed by an abrupt step. Masking can also be applied to the top boundary in the presence of icebergs or thick ice (thick compared to the vertical spacing near the surface).

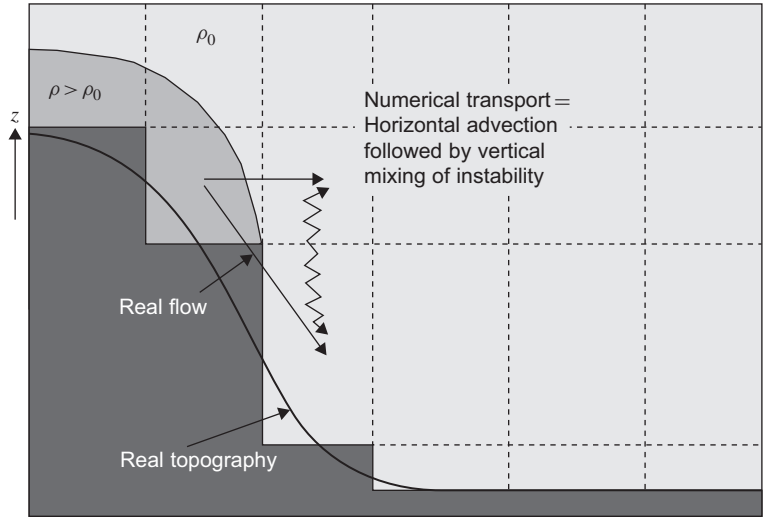


FIGURE 20.14 Vertical section of a schematic overflow where dense bottom water cascades down a slope. With stepwise topography representation, advection carries the dense water horizontally to a model grid point with lower density. It is tempting to remove the resulting unphysical gravitational instability by a mixing algorithm, but this has the unfortunate consequence of diluting the dense-water vein.

FIGURE 20.15 Example of a finite-element model grid for the world ocean. (*Legrand, Legat & Deleersnijder, 2000*)

design, the possibility of following complicated contour lines. Structured grids, natural for models discretized along Cartesian coordinates or longitude-latitude, are topologically similar to a rectangular grid, with every grid point having one and only one neighbor to the "east," "west," "north," "south," above and below. In unstructured grids, on the contrary, every grid cell has a variable number of neighbors (Fig. 20.15), providing great flexibility in terms of geographical coverage. Coastlines can be followed by adding small elements along the side, canyons resolved by adding small elements at the bottom, and open boundaries pushed further away by increasing the size of distant elements. Intense model developments continue unabated (see, e.g., Pietrzak, Deleersnijder & Schroeter, 2005).

An unstructured finite-volume approach is a generalization of the finite-volume approach presented in Section 5.5, in which integration is performed over each finite volume, with the entire group covering the model domain. Physical coupling between the finite volumes arises naturally through the fluxes across the shared interfaces. Finite elements, in contrast, start with a totally different approach based on the Galerkin method (Section 8.8). The solution is expanded as a sum of non-orthogonal basis functions, and each governing equation is multiplied by each basis function before being integrated over the model domain. The basis functions (sometimes also called trial functions) used in finite-element methods are of a special nature, being nonzero only over a given element. Connection between elements then arises because functions are forced to obey some continuity requirement between elements. This necessitates solving a system of linear equations (i.e., inverting a matrix) at every time step, but the nature of the basis functions leads to relatively sparse matrices.

Beside the widespread finite-volume and finite-element methods, additional schemes have been implemented in ocean models such as the spherical cube grid (a three-dimensional generalization of the squaring of a circle, with special connectivity at a few nodes—Adcroft, Campin, Hill & Marshall, 2004) and spectral elements (Haidvogel, Wilkin & Young, 1991) a method in which a domain is covered by large elements, inside of which spectral series are used to approximate the solution.

The advantage of variable resolution exposes the fundamental problem of subgrid-scale parameterization. Scales resolved by the variable grid do change regionally, and therefore the nature of the parameterizations involved should change from place to place within the same model. Among processes requiring parameterization, deep-water formation is perhaps the most crucial example (see Section 11.4). The sinking of dense water is dominated by nonhydrostatic convection, forcing all hydrostatic models to have a parameterization of one form or another. Because the validity of the hydrostatic approximation is related to the aspect ratio of the flow (ratio of vertical scale to horizontal scale, see Section 4.3), nonhydrostatic effects are only relevant at extremely high resolution, and ocean general circulation models are rarely nonhydrostatic. Yet, highly localized dense-water formation does influence broader scale flow and needs to be included in these models. The brutal, and rather common, way to deal with the issue is *convective adjustment* (see Section 11.4). In this scheme, whenever the water column is found to have a density reversal between two vertically aligned grid points, temperature and salinity are merely replaced by their average values, and the process continues downward until no inversion remains. Despite many years of use, this scheme was eventually found to be flawed (Cessi, 1996), for it generates an unstable mode at the smallest resolved horizontal scale. A better way to handle convective overturn is by local mesh refinement with flexible horizontal grids, and today's hydrostatic ocean models often include a nonhydrostatic option, such as used in *MITgcm* (Marshall, Jones & Hill, 1998).

The upper boundary of an ocean circulation model requires special care because this is where atmospheric forcing is applied to the ocean. Rigid-lid models eliminate the fast surface gravity waves in order to permit longer time steps. However, there is a trend toward the use of free-surface models because of their added flexibility and wider range of applicability, including tidal representation. The trick to continue enjoying a reasonably large time step is to treat time stepping of the surface elevation either implicitly, semi-implicitly, or with a mode-splitting scheme (Section 12.7), while the rest of the scheme remains explicit.

The coupling with the atmosphere (see Section 19.4) remains a difficult problem because the ocean model must simulate both surface temperature and mixed-layer depth correctly. The former is essential for air–sea exchanges while the product of temperature and mixed-layer depth is directly related to the heat content of the water column and therefore also the heat budget. Because of the low heat capacity of air, the atmospheric component of the model can be very

sensitive to a small error in the heat content of the upper ocean. Furthermore, because mixed-layer evolution critically depends on turbulence, special care is needed in the specification of the eddy viscosity and diffusivities, particularly if vertical grid spacing is coarse. Turbulence-closure schemes were described in Sections 14.3–14.5.

20.6.1 Coordinate Systems

Vertical resolution near the surface is essential to capture air–sea interactions, and most oceanic models use shorter vertical grid spacing near the surface. For the ocean interior, vertical gridding is again considered crucial, and a significant part of the problem is the separation between vertical and horizontal coordinates, which is necessary because of the small geometric aspect ratio and the vast difference in dynamics between the vertical and horizontal directions. We therefore assume that the horizontal discretization is performed using one among the methods outlined earlier in this text and focus now on vertical gridding.

Vertical gridding can be done quite freely, with variable resolution permitting higher resolution in sensitive segments (near the surface, across a pycnocline, and near the bottom) and lower resolution at less important levels for the sake of computational economy. About the only requirement is that a grid point topologically above (below) another grid point corresponds to a physical point also physically above (below) the other point. The nature of the vertical coordinate can very well change from top to bottom, switching from depth to density to a bottom-following coordinate. The generation of such hybrid-grid models can be achieved in one of two ways: we can subject the governing equations to the finite-volume integration technique over the desired vertical spacings and write a set of governing equations for each cell, or we can first perform a change of coordinates after which we discretize the equations along a uniform grid in that new coordinate system (Fig. 20.16; see also Section 15.6).

Direct integration in physical space exposes the need for parameterization because integrals of nonlinear terms cannot be expressed in terms of cell averages unless some assumptions are made. In contrast, a coordinate transformation followed by finite differencing runs the risk of masking the need for parameterization of subgrid-scale processes. Ultimately, however, the discrete equations that are obtained will be similar, and the choice of approach depends on the path preferred by the modeler.

Here we present the coordinate transformation approach, in close analogy with the isopycnal coordinate change of Chapter 12. In the original Cartesian system of coordinates (x, y, z, t), z is an independent variable, while in the transformed coordinate system (x, y, s, t) the new coordinate s replaces z as an independent variable and $z(x, y, \mathrm{s}, t)$ becomes a dependent variable giving the depth at which the value s is found at location (x, y) and at time t. A surface along which s is constant is called a *coordinate surface*. From a differentiation

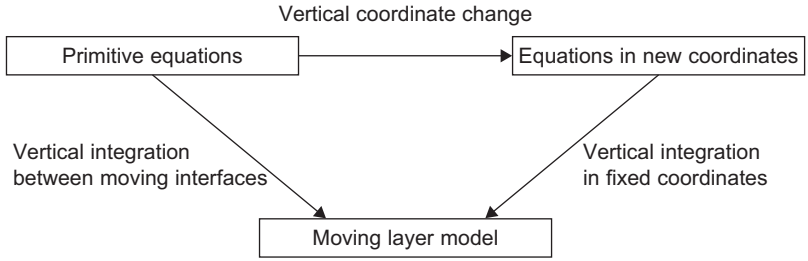


FIGURE 20.16 Equations for discrete layers can be obtained either by integrating the governing equations across each chosen layer or by a transformation to moving coordinates followed by discretization on a fixed grid. The discrete equations to be solved by the computer are the same.

of the expression $a = a(x, y, \mathrm{s}(x, y, z, t), t)$, where a stands for any variable, the rules for the change of variables follow:

$$\begin{aligned}
\frac{\partial}{\partial x} &\longrightarrow \left.\frac{\partial a}{\partial x}\right|_z = \left.\frac{\partial a}{\partial x}\right|_{\mathrm{s}} + \frac{\partial a}{\partial \mathrm{s}} \left.\frac{\partial \mathrm{s}}{\partial x}\right|_z \\
\frac{\partial}{\partial y} &\longrightarrow \left.\frac{\partial a}{\partial y}\right|_z = \left.\frac{\partial a}{\partial y}\right|_{\mathrm{s}} + \frac{\partial a}{\partial \mathrm{s}} \left.\frac{\partial \mathrm{s}}{\partial y}\right|_z \\
\frac{\partial}{\partial z} &\longrightarrow \frac{\partial a}{\partial z} = \frac{\partial a}{\partial \mathrm{s}} \frac{\partial \mathrm{s}}{\partial z} \\
\frac{\partial}{\partial t} &\longrightarrow \left.\frac{\partial a}{\partial t}\right|_z = \left.\frac{\partial a}{\partial t}\right|_{\mathrm{s}} + \frac{\partial a}{\partial \mathrm{s}} \left.\frac{\partial \mathrm{s}}{\partial t}\right|_z
\end{aligned}$$

in analogy with the isopycnal coordinate system presented in Chapter 12. A noteworthy difference with isopycnal coordinates is that s is not necessarily a physical property conserved along the flow. An important expression in the coordinate transformation is the quantity

$$\hbar = \frac{\partial z}{\partial \mathrm{s}}, \tag{20.42}$$

which denotes the change in z for a unit change in s. Hence, it is a measure of the *coordinate layer thickness*, analogous to the thickness of a density layer. It can be positive or negative depending on whether s increases upward or downward. (It is negative if $\mathrm{s} = \rho$, the density coordinate of Chapter 12.)

The material derivative in the new coordinate system takes the form:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{\partial a}{\partial t} + u\frac{\partial a}{\partial x} + v\frac{\partial a}{\partial y} + \omega\frac{\partial a}{\partial \mathrm{s}}, \tag{20.43}$$

where all derivatives are taken in the transformed space ($\partial a/\partial x$ is performed at constant s, *etc.*) and where ω substitutes for the vertical velocity. It is defined as

$$\omega = \left.\frac{\partial \mathrm{s}}{\partial t}\right|_z + u\left.\frac{\partial \mathrm{s}}{\partial x}\right|_z + v\left.\frac{\partial \mathrm{s}}{\partial y}\right|_z + w\frac{\partial \mathrm{s}}{\partial z}. \tag{20.44}$$

The product $\hbar\omega$ is the vertical velocity of the flow relative to the moving s coordinate surface (recall Section 15.6). Clearly, if s is density and if density is conserved with the flow (i.e., in the absence of mixing), $\omega = 0$, and we recover the isopycnal coordinate system, which has the advantage of eliminating the vertical velocity. If the ocean surface and/or ocean bottom are taken as coordinate surfaces, the "vertical" velocity ω vanishes at those boundaries because the flow must follow that material boundary. In general, however, $\hbar\omega$ is not zero.

The transformation of the volume-conservation equation (4.9) using the preceding rules of change of variables and the definition of the vertical velocity leads to

$$\frac{\partial \hbar}{\partial t} + \frac{\partial}{\partial x}(\hbar u) + \frac{\partial}{\partial y}(\hbar v) + \frac{\partial}{\partial \mathrm{s}}(\hbar\omega) = 0 \tag{20.45}$$

with derivatives taken in s space. Interestingly, if we integrate from bottom to top in the case of s constant along both ocean surface and bottom, we obtain

$$\frac{\partial \eta}{\partial t} + \frac{\partial \mathrm{U}}{\partial x} + \frac{\partial \mathrm{V}}{\partial y} = 0 \tag{20.46}$$

because

$$\int_{\text{bottom}}^{\text{surface}} \hbar \, \mathrm{ds} = z_{\text{surface}} - z_{\text{bottom}} \tag{20.47}$$

which, for a time-independent bottom, leads after temporal derivation to the time derivative of the surface elevation η. The other two terms involve U and V, which are the vertically integrated transports

$$\mathrm{U} = \int u \hbar \mathrm{ds} = \int u \, \mathrm{d}z, \qquad \mathrm{V} = \int v \hbar \mathrm{ds} = \int v \, \mathrm{d}z, \tag{20.48}$$

where the integration is performed from bottom to top. We recover a vertically integrated volume-conservation equation.

Irrespective of whether or not the ocean surface and bottom are taken as coordinate surfaces, volume conservation leads to a conservative form of the material derivative

$$\begin{aligned} &\hbar\left(\frac{\partial a}{\partial t} + u\frac{\partial a}{\partial x} + v\frac{\partial a}{\partial y} + \omega\frac{\partial a}{\partial \mathrm{s}}\right) \\ &\quad = \frac{\partial}{\partial t}(\hbar a) + \frac{\partial}{\partial x}(\hbar a u) + \frac{\partial}{\partial y}(\hbar a v) + \frac{\partial}{\partial \mathrm{s}}(\hbar a \omega), \end{aligned} \tag{20.49}$$

which we can interpret as the evolution of the content of a within a layer of s. This form is particularly well suited for integration over a finite volume in the transformed space.

The vertical diffusion term is readily transformed in the new coordinate system:

$$\frac{\partial}{\partial z}\left(\nu_{\mathrm{E}}\frac{\partial a}{\partial z}\right)=\frac{1}{\hbar}\frac{\partial}{\partial \mathrm{s}}\left(\frac{\nu_{\mathrm{E}}}{\hbar}\frac{\partial a}{\partial \mathrm{s}}\right), \tag{20.50}$$

so that the governing equation for a vertically diffusing tracer c becomes

$$\frac{\partial}{\partial t}(\hbar c)+\frac{\partial}{\partial x}(\hbar c u)+\frac{\partial}{\partial y}(\hbar c v)+\frac{\partial}{\partial \mathrm{s}}(\hbar c \omega)=\frac{\partial}{\partial \mathrm{s}}\left(\frac{\kappa_E}{\hbar}\frac{\partial c}{\partial \mathrm{s}}\right). \tag{20.51}$$

We could also transform the horizontal diffusion term, but generally this operation is combined with the parameterization of subgrid-scale processes (see next section).

In view of the isomorphy of (20.51) with a Cartesian-coordinate version, an s-model can thus be implemented in a general way without much additional work once the functional relationship $\mathrm{s}(x, y, z, t)$ is specified. The choice is at the modeler's discretion.

Besides the isopycnal transform with $\mathrm{s}=\rho$ or $\mathrm{s}=-\rho/\Delta\rho$, another coordinate change is the so-called sigma coordinate (σ) system, a particular form of terrain-following coordinates. It is very popular in coastal modeling. This coordinate is defined as

$$\mathrm{s}=\sigma=\frac{z-b}{h}, \quad \hbar=h, \tag{20.52}$$

so that it varies between 0 at the bottom $[z=b(x, y)]$ and 1 at the surface $[z=b(x, y)+h(x, y, t)]$, which are therefore coordinate surfaces (Fig. 20.17). All topographic slopes and free surface movements are naturally followed, avoiding the problem of discretizing equations in a changing domain. Also, calculation points are efficiently used because they all fall into the water column ($0 \leq \sigma \leq 1$), and vertical boundary conditions are straightforward. Moreover, and this is the main advantage in coastal modeling, grid points are more closely spaced in shallow water than in deep water, providing the highest vertical resolution where it is the most needed (Fig. 20.17).

However, the use of this coordinate transform in global-ocean models has raised some concern about the so-called pressure-gradient problem (e.g., Deleersnijder & Beckers, 1992; Haney, 1991). Although the problem was initially identified for the σ-coordinate, it is more general, and we describe it here in a general coordinate framework.

The horizontal pressure gradient, for example, along x, can be evaluated in the new coordinate system by using the following transformation rules:

$$\left.\frac{\partial p}{\partial x}\right|_z=\left.\frac{\partial p}{\partial x}\right|_{\mathrm{s}}+\frac{\partial p}{\partial \mathrm{s}}\left.\frac{\partial \mathrm{s}}{\partial x}\right|_z, \tag{20.53}$$

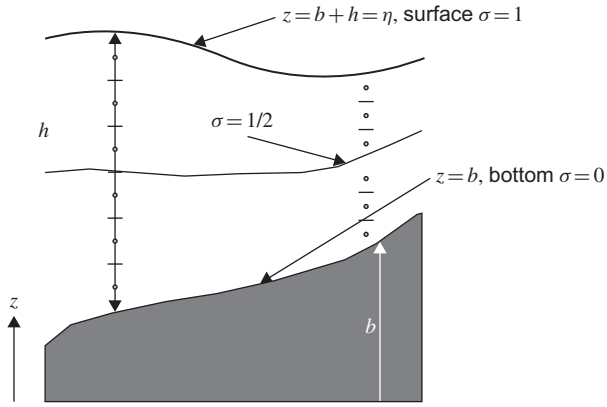


FIGURE 20.17 A sigma coordinate system divides the entire water column into an equal number of vertical grid cells regardless of local depth and surface elevation.

while the hydrostatic balance used to calculate pressure becomes

$$\frac{\partial p}{\partial z} = \frac{1}{\hbar}\frac{\partial p}{\partial \mathrm{s}} = -\rho g. \tag{20.54}$$

This allows us to write several, mathematically equivalent expressions for the horizontal pressure force

$$\frac{1}{\rho_0}\frac{\partial p}{\partial x}\Big|_z = \frac{1}{\rho_0}\frac{\partial p}{\partial x}\Big|_{\mathrm{s}} + \frac{\rho g}{\rho_0}\frac{\partial z}{\partial x}\Big|_{\mathrm{s}} \tag{20.55a}$$

$$= \frac{1}{\rho_0 \hbar} J(p, z) \tag{20.55b}$$

$$= \frac{1}{\rho_0}\frac{\partial p}{\partial x}\Big|_{\text{surface}} + \frac{\rho g}{\rho_0}\frac{\partial \eta}{\partial x} - g\int_{\text{surface}}^{\mathrm{s}} J(\rho, z)\, \mathrm{ds} \tag{20.55c}$$

$$= \frac{1}{\rho_0}\frac{\partial P}{\partial x}\Big|_{\mathrm{s}} - \frac{gz}{\rho_0}\frac{\partial \rho}{\partial x}\Big|_{\mathrm{s}} \tag{20.55d}$$

where the last expression uses the Montgomery potential $P = p + \rho g z$. For the second and third expressions, the Jacobian operator J is defined as $J(a, b) = (\partial a/\partial x)(\partial b/\partial \mathrm{s}) - (\partial a/\partial \mathrm{s})(\partial b/\partial x)$ in the transformed space.

A standard test for terrain-following models is to prescribe density and pressure fields that depend solely on z (e.g., Beckmann & Haidvogel, 1993; Numerical Exercise 20.4). In this way, the horizontal pressure gradient should be identically zero, and no motion should be generated in the absence of external driving forces. The two terms on the right-hand side of Eqs. (20.55a) and (20.55d) cancel each other exactly in the continuous representation and should continue to do so after discretization. Due to the different nature of the two

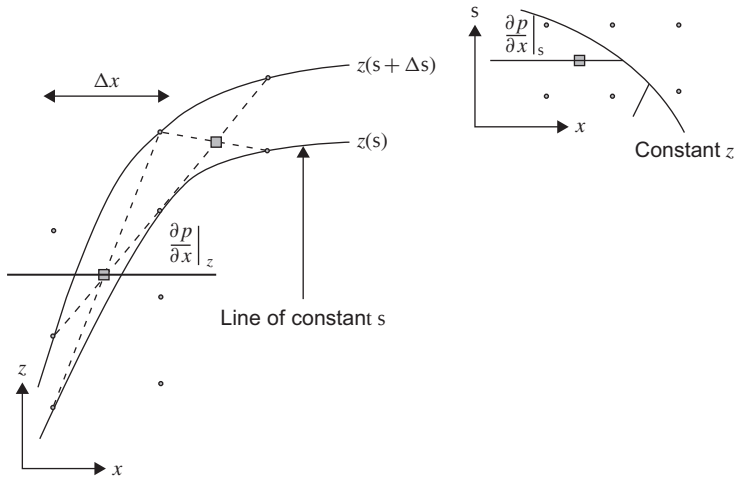


FIGURE 20.18 In the numerical grid (x,s), a standard finite differencing of the pressure gradient at the location of the squares involves the logical neighbors in the discrete mesh, connected with dashed lines in the physical space (left panel). For upper right point, the calculation points are physical neighbors but for the lower left point, the calculation uses far distant points and performs extrapolations.

terms, however, even very careful numerical discretization will most likely leave a residual that acts as a nonzero pressure force and hence generates unphysical horizontal motion. The main problem is that this error is most often not small because the vertical variation of pressure $p(z)$ is very large and taking the pressure gradient along a sloping s direction entails a large vertical component.

Should we argue that the problem would be eliminated with increased resolution, another problem would appear, the problem of so-called hydrostatic consistency: By increasing resolution more rapidly in the vertical than in the horizontal direction, (Fig. 20.18) the numerical stencil used in the calculation of a horizontal pressure gradient may involve grid points that are vertically too far away, and horizontal derivatives may be evaluated by extrapolation instead of interpolation. Such extrapolation leads to large relative errors (see Numerical Exercise 3.5), as well as to an inconsistency. The vertical gradients in Eq. (20.55a)–(20.55d) are then not calculated at the same depth as the horizontal gradient. For simple finite-difference schemes, extrapolations are avoided if the following criterion between slopes is met:

$$\left| \frac{\partial z/\partial x|_{\mathrm{s}}}{\partial z/\partial \mathrm{s}} \right| \leq \frac{\Delta s}{\Delta x}, \tag{20.56}$$

The slope of z lines in the transformed space (left-hand side) cannot be larger than the aspect ratio of the grid in the same space (right-hand side) so that lines of constant z remain within the local stencil. This constraint is not unlike the

constraint on the domain of dependence of advection schemes (Section 6.4). Contrary to the problem in Fig. 20.13, we now have a lower bound for vertical grid spacing, in relation to the slope of the coordinate surfaces. Alternatively, for a fixed vertical grid and given slopes, the requirement imposes a horizontal grid that must be sufficiently fine to resolve the slopes accurately.

For the σ-coordinate,[8] this translates into a constraint that involves the water column height h:

$$\left|\frac{1}{h}\frac{\partial h}{\partial x}\right| \leq \frac{\Delta\sigma}{\Delta x}\frac{1}{1-\sigma}, \tag{20.57}$$

where $\Delta\sigma$ and σ are the values at the numerical grid level under consideration. For a uniform grid in σ space, the constraint is most severe near the bottom layer.

Hence, the worst problems are encountered where the topography is shallow but steep, such as near a shelf break, where the length scale related to topography $|(1/h)(\partial h/\partial x)|^{-1}$ is shortest and must be resolved by the grid spacing. This length scale thus appears as an additional scale to be considered in the design of horizontal grids. Since stratification on the shelf break is typically intersecting topography, problems with the pressure gradient will be exacerbated there: regions of large variations in ρ coincide with regions of large $\partial z/\partial x|_s$, and the two contributions to the horizontal pressure force are large, leading to a significant numerical error. Solutions to this problem include higher-order finite differencing (using more grid points and being less prone to extrapolations—e.g., McCalpin, 1997), subtraction of average density profiles $\rho = \rho(z)$ before any pressure calculation (e.g., Mellor, Oey & Ezer, 1998), specialized finite differencing (e.g., Song, 1998), partial masking of topography (replacing slopes by vertical sections—e.g., Beckers, 1991; Gerdes, 1993), or, simply, smoothing the bottom topography.

With the aforementioned provision for the proper treatment of the pressure gradient, a generalized vertical coordinate can be very attractive, and several models have implemented the approach (e.g., Pietrzak, Jakobson, Burchard, Vested & Peterson, 2002), without actually using them at their full potential but prescribing a priori the position of coordinate surfaces. The general rule for the placement of coordinate surfaces is to match as closely as possible the surface on which physical properties remain smooth (Fig. 20.19; see also adaptive grids in Section 15.6). In ocean circulation models, grids should therefore follow closely density surfaces in the ocean interior, where mixing is weak.

Because a density coordinate is ill-suited to represent mixed layers, the use of the z-coordinate should be preferred near the surface. An implementation that nearly achieves this requirement is *HYCOM* (HYbrid Coordinate Ocean Model—e.g., Bleck, 2002), which is an extension of an earlier density model

[8] We neglect here any surface gradient for the sake of simplicity.

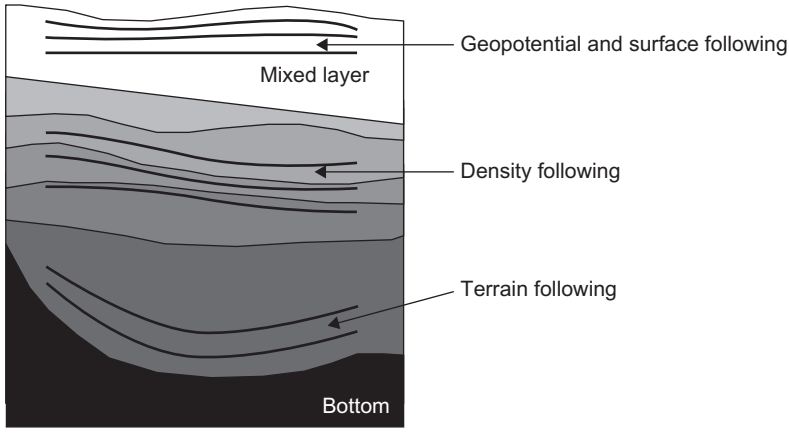


FIGURE 20.19 Most physically meaningful types of coordinate surfaces as a function of water depth. Near the surface and in the mixed layer, lateral coordinates run horizontally, or nearly horizontally, in order to follow the surface. In the ocean interior, where motion proceeds with little or no mixing, density becomes the best vertical coordinate, and coordinate surfaces follow density surfaces. Near the bottom, which the flow is forced to follow, a terrain-following coordinate is best.

allowing for an upper mixed layer spread over several z levels. For further improvement, the vertical coordinate near the bottom may be made to follow the topography and thus behave as a σ coordinate.

20.6.2 Subgrid-Scale Processes

Once the grid is defined, and the shortest resolved scale known, subgrid-scale processes must be considered. Up to this point, subgrid-scale processes other than turbulence were modeled by horizontal diffusion such as

$$\mathcal{D}(c) = \frac{\partial}{\partial x}\left(\mathcal{A}\frac{\partial c}{\partial x}\right) + \frac{\partial}{\partial y}\left(\mathcal{A}\frac{\partial c}{\partial y}\right) \tag{20.58}$$

for any quantity c, be it a tracer concentration, temperature, salinity, or even a velocity component. One question that arises is whether the derivatives in this expression ought to be those in Cartesian coordinates or in any other set of coordinates. The choice of Cartesian coordinates implies that there exists a tendency to mix properties along the horizontal plane, but this may not always be representative of what is actually happening. The case in point is preferential mixing along (possibly sloping) density surfaces because this is the direction in which mixing motions are not inhibited by buoyancy forces. So, let us suppose that, to be faithful to the physics, mixing occurs along surfaces of constant density ρ, and the x- and y-derivatives in the preceding expression are to be taken at constant ρ.

The diffusion operator expressed as classical diffusion in the coordinate of choice can then be translated back into Cartesian coordinates, as done by Redi (1982). The price to pay is the presence of additional terms and non-constant coefficients, and, contrary to the original operator, its discretization may no longer be monotonic (e.g., Beckers, Burchard, Campin, Deleersnijder & Mathieu, 1998).

Diffusion-like parameterizations are, of course, based on the expectation that unresolved eddies act similarly to diffusion. However, depending on the scales under consideration, some subgrid-scale processes may not be considered randomly mixing the ocean, especially at the larger scales. The internal radius of deformation is a locus of energetic motions, in large part due to baroclinic instability (see Chapter 17). Because the deformation radius in the ocean is at most a few tens of kilometers, global ocean models typically do not resolve baroclinic instability and hence the eddies that it sheds. Unless regional models are used, coarse-resolution ocean models must parameterize the effect of mesoscale motions. Baroclinic instability releases potential energy by flattening density surfaces, a process quite different from pure mixing. Isopycnal diffusion cannot account for such a flattening since by construction it diffuses only along isopycnals, hence not forcing them to flatten out.

Instead of a diffusion-type parameterization, the so-called Gent–McWilliams parameterization (Gent & McWilliams, 1990; Gent, Willebrand, McDoughall & McWilliams, 1995) is recommended. This scheme proceeds by adding a so-called bolus velocity to the large-scale currents, the components of which, marked with a star, are

$$u^\star = -\frac{\partial Q_x}{\partial z}, \quad v^\star = -\frac{\partial Q_y}{\partial z}, \quad w^\star = \frac{\partial Q_x}{\partial x} + \frac{\partial Q_y}{\partial y}, \tag{20.59}$$

with the pair (Q_x, Q_y) taken as

$$Q_x = -\frac{\kappa}{\rho_z}\frac{\partial \rho}{\partial x} = \kappa \left.\frac{\partial z}{\partial x}\right|_\rho \tag{20.60a}$$

$$Q_y = -\frac{\kappa}{\rho_z}\frac{\partial \rho}{\partial y} = \kappa \left.\frac{\partial z}{\partial y}\right|_\rho \tag{20.60b}$$

where $\rho_z = \partial\rho/\partial z$.

As it immediately appears, these quantities are proportional to the x- and y-slopes of the isopycnals, with the coefficient κ of proportionality being a tunable model parameter with same dimension as a diffusion coefficient (length squared per time). Since the derivatives are expressed in Cartesian coordinates, it is easily verified that the bolus velocity is divergence-free, a property that should be preserved by the numerical discretization. The bolus velocity effectively advects the density field and, with the chosen signs, leads to a reduction of the frontal slope (Fig. 20.20), substituting therefore for baroclinic instability. It is important to note that this advection is performed without any underlying dynamical

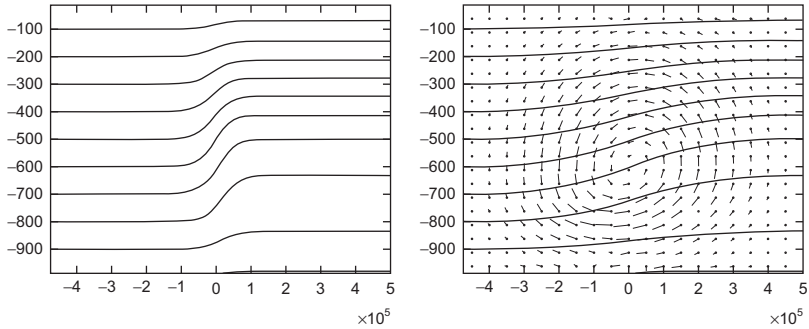


FIGURE 20.20 Vertical section across a density field with frontal structure as shown by sloping isopycnals (left panel) and the corresponding bolus velocity (right panel, with velocities originating from the dotted ends). Note how the bolus velocity acts to relax the density field in time so that some time later the front has indeed been weakened (right panel, solid lines). Since the bolus velocity depends on the existing slope, the flattening of isopycnals slows down over time.

equation, reflecting the fact that this is a parameterization of unresolved dynamics. The strength of the effect is controlled by the parameter κ, and Griffies (1998) shows how one may combine bolus advection with the isopycnal diffusion into a single operator as long as both "diffusion" coefficients are equal. For additional information and a recent review of ocean model developments, the reader is referred to Griffies et al. (2000) and references therein.

ANALYTICAL PROBLEMS

20.1. Derive the expression for the Sverdrup transport on an irregular bottom.

20.2. Derive the veering of the horizontal velocity with respect to depth, working with density ρ as the vertical coordinate. Show that one recovers Eq. (20.16) by changing to z-coordinates.

20.3. Given that the North Pacific Ocean is about twice as wide as the North Atlantic Ocean and that both basins are subjected to equally strong winds, compare their Sverdrup transport cumulated over the width of the basin. Express your answers in Sverdrup units (1 Sv $= 10^6$ m^3/s).

20.4. Demonstrate that western intensification would still occur if the global wind patterns were reversed, that is if Ekman pumping were upward and Sverdrup transport poleward at midlatitudes. In other words, show that the return flow from the Sverdrup transport must be squeezed against the western boundary regardless of the sign of the wind-stress curl over the ocean.

20.5. Imagine that a single ocean were covering the entire globe as the atmosphere does. With no western wall to support a boundary current returning

the equatorward Sverdrup flow, what would the circulation pattern be? Relate your results to the existence of the Antarctic Circumpolar Current. [For a succinct description of this major current, see Section 7.2 of the book by Pickard and Emery (1990) or some other physical oceanography textbook.]

20.6. Consider the Stommel–Arons model of the abyssal circulation with zero total flow across the equator. Show that the travel time to reach the pole is independent of the initial longitude of a water parcel released at latitude φ_0. Calculate this travel time by time integration of the trajectory in the case of uniform upwelling velocity. (*Hint*: Use the definition of velocities in spherical coordinates from the material derivative (A.17) and integrate trajectories in the (λ, φ) domain.)

20.7. Consider the bolus velocity of equation (20.59), with constant diffusion coefficient κ. Investigate the possibility of stationary solutions for the density field solely advected by this bolus velocity in the (x, z) vertical plane for convenience. After finding a general condition on Q_x and ρ, assume a linear relationship between those variables and a uniform vertical stratification to determine a particular, stationary solution for ρ.

NUMERICAL EXERCISES

20.1. Take the density data used in `iwavemed` and calculate the geostrophic velocities. Work in z-coordinates with z levels corresponding to the data. First assume a level of no motion at 500 m and calculate currents at the surface and 2000 m down. Then repeat with level of no motion at 1500 m.

20.2. Experiment with `bolus` to see the flattening of isopycnals in a vertical section with flat bottom. Then implement a sigma-coordinate change in a vertical section with flat surface and topography given by

$$h(x) = h_0 + \alpha x, \tag{20.61}$$

where the domain extends from $x = -L/2$ to $x = L/2$ and where $\alpha L = h_0$. Start from the same physical density distribution as in the flat-bottom case. (*Hint*: Express the bolus advection term as a Jacobian in the vertical plane. Then use the rules of change of variables to express this Jacobian in the new coordinate system. Do not forget to calculate the slopes using the rules of changes of variables.)

20.3. Obtain a density section from somewhere and calculate bolus velocity using `bolus`. Which problems do you expect near boundaries? (*Hint*: You might consider κ as a calibration parameter that varies in space.)

20.4. Use `pgerror` to explore the pressure-gradient error for a fixed density anomaly profile depending only on z according to

$$\rho = \Delta\rho \tanh\left(\frac{z+D}{W}\right), \tag{20.62}$$

in which D and W control the position and thickness of the pycnocline. Bottom topography is given by

$$h(x) = H_0 + \Delta H \tanh\left(\frac{x}{L}\right), \tag{20.63}$$

where L and ΔH control the steepness of the slope. Calculate the error and associated geostrophic velocity for $f = 10^{-4}\ \mathrm{s}^{-1}$. Vary the number of vertical grid points, horizontal grid points, the position of the pycnocline, its depth, and strength. What happens if you increase only the number of vertical grid points? Implement the discretization of another pressure gradient expression of (20.55) in `bcpgr` and compare.

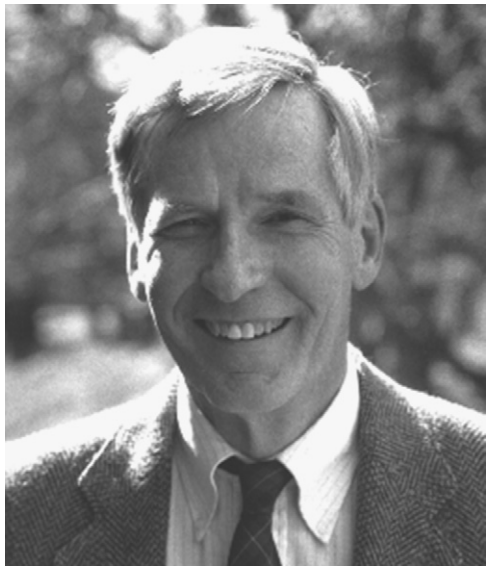
20.5. Bottom topography is generally smoothed before it is used in a model, by the repeated application of a Laplacian-type diffusion. In view of the hydrostatic consistency constraint, which adapted filter technique would you advocate? (*Hint*: Remember that a Laplacian filter applied to a function F decreases the norm of the gradient $\int\int \left[(\partial F/\partial x)^2 + (\partial F/\partial y)^2\right] \mathrm{d}x\,\mathrm{d}y$ over the domain.)

Henry Melson Stommel
1920–1992

At an early age, Henry Stommel considered a career in astronomy but turned to oceanography as a way to make a peaceful living during World War II. Having been denied admission to graduate school at the Scripps Institution of Oceanography by H. U. Sverdrup, then its director, Stommel never obtained a doctorate. This did not deter him; having soon realized that, in those years, oceanography was largely a descriptive science almost devoid of physical principles, he set out to develop dynamic hypotheses and to test them against observations. To him, we owe the first correct theory of the Gulf Stream (1948), theories of the abyssal circulation (early 1960s), and a great number of significant contributions on virtually all aspects of physical oceanography.

Unassuming and avoiding the limelight, Stommel relied on a keen physical insight and plain common sense to develop simple models that clarify the roles and implications of physical processes. He was generally wary of numerical models. Particularly inspiring to young scientists, Stommel continuously radiated enthusiasm for his chosen field, which, as he was the first to acknowledge, is still in its infancy. (*Photo by George Knapp*)

Kirk Bryan
1929–



As soon as computer mainframes became available for scientific research, in the 1960s, Kirk Bryan in collaboration with colleague Michael Cox (1941–1989) and student Bert Semtner began to develop codes for the simulation of oceanic circulation. This was truly pioneering work not only in the face of stringent hardware limitations but also because still little was known at the time about numerical stability, accuracy, spurious modes, etc. The so-called Bryan-Cox code of Princeton University's Geophysical Fluid Dynamics Laboratory quickly became a staple in oceanic modeling, often at the root of others' codes.

Concerns over climate change prompted Bryan later in his career to construct fully coupled atmosphere–ocean models, which are extremely challenging in view of their complexity and vastly different temporal scales. Rather than being daunted by this complexity, Bryan stresses the complementarity between atmospheric and oceanic processes and scales of motion.

Fame and numerous awards have come his way, but Kirk Bryan has retained a gentlemanly demeanor, with a kind word for all with whom he comes in contact. (*Photo courtesy of Princeton University*)