Chapter 15

Dynamics of Stratified Rotating Flows

ABSTRACT

Geostrophic motions can arise during the adjustment to density inhomogeneities and maintain a stratified fluid away from gravitational equilibrium. The key is a relationship between the horizontal density gradient and the vertical velocity shear, called the thermal-wind relation. Oceanic coastal upwelling is considered as it is a good example of rotating dynamics in a stratified fluid. Because large gradients and discontinuities (fronts) can form during geostrophic adjustment, the numerical section shows how to treat large gradients in computer models.

15.1 THERMAL WIND

Consider a situation where a cold air mass is wedged between the ground and a warm air mass (Fig. 15.1). The stratification has then both vertical and horizontal components. Mathematically, the density is a function of both height z and distance x (say, from cold to warm). Now, assume that the flow is steady, geostrophic, and hydrostatic:

$$-fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} \tag{15.1}$$

$$\frac{\partial p}{\partial z} = -\rho g. \tag{15.2}$$

Here, v is the velocity component in the horizontal direction y, and p is the pressure field. Taking the z-derivative of Eq. (15.1) and eliminating $\partial p/\partial z$ with Eq. (15.2), we obtain

$$\frac{\partial v}{\partial z} = -\frac{g}{\rho_0 f}\frac{\partial \rho}{\partial x}. \tag{15.3}$$

Therefore, a horizontal density gradient can persist in steady state if it is accompanied by a vertical shear of horizontal velocity. Where density varies in both

Introduction to Geophysical Fluid Dynamics, Volume 101, ISSN 0074-6142, DOI: 10.1016/B978-0-12-088759-0.00015-8

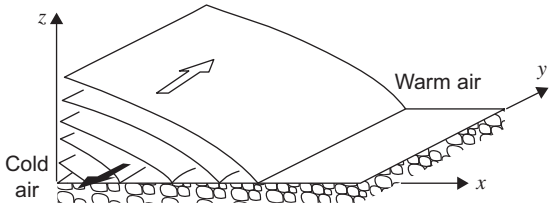


FIGURE 15.1 Vertical shear of a flow in the presence of a horizontal density gradient. The change of velocity with height is called thermal wind.

horizontal directions, the following also holds:

$$\frac{\partial u}{\partial z} = +\frac{g}{\rho_0 f}\,\frac{\partial \rho}{\partial y}. \tag{15.4}$$

These innocent-looking relations have profound meaning. They state that due to the Coriolis force, the system can be maintained in equilibrium, without tendency toward leveling of the density surfaces. In other words, the rotation of the earth can keep the system away from its state of rest without any continuous supply of energy.

Notice that the velocity field (u, v) is not specified, only its vertical shear, $\partial u/\partial z$ and $\partial v/\partial z$. This implies that the velocity must change with height. (In the case of Fig. 15.1, $\partial \rho/\partial x$ is negative and $\partial v/\partial z$ is positive.) For example, the wind speed and direction at some height above the ground may be totally different from those at ground level. The presence of a vertical gradient of velocity also implies that the velocity cannot vanish, except perhaps at some discrete levels. Meteorologists have named such a flow the *thermal wind.*[1]

In the case of pronounced density contrasts, such as across cold and warm fronts, a layered system may be applicable. In this case (Fig. 15.2), the system can be represented by two densities (ρ_1 and ρ_2, $\rho_1 < \rho_2$) and two velocities (v_1 and v_2). Relation (15.3) can be discretized into

$$\frac{\Delta v}{\Delta z} = -\frac{g}{\rho_0 f}\,\frac{\Delta \rho}{\Delta x},$$

where we take $\Delta v = v_1 - v_2$ and $\Delta \rho = \rho_2 - \rho_1$ to obtain

$$v_1 - v_2 = -\frac{g}{\rho_0 f}(\rho_2 - \rho_1)\frac{\Delta z}{\Delta x}. \tag{15.5}$$

The ratio $\Delta z/\Delta x$ is the slope of the interface. The equation is called the *Margules relation* (Margules, 1906), although a more general form of the relation for zonal flows was obtained earlier by Helmholtz (1888).

[1] Although thermal wind is a meteorological expression, oceanographers use it, too, to indicate a sheared current in geostrophic equilibrium with a horizontal density gradient.

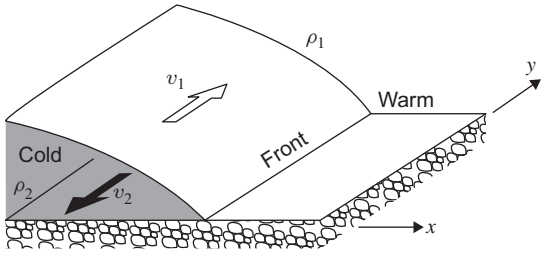


FIGURE 15.2 The layered version of Fig. 15.1, which leads to the Margules relation.

The thermal-wind concept has been enormously useful in analyzing both atmospheric and oceanic data because observations of the temperature and other variables that influence the density (such as pressure and specific humidity in the air, or salinity in seawater) are typically much more abundant than velocity data. For example, knowledge of temperature and moisture distributions with height and of the surface wind (to start the integration) permits the calculation of wind speed and direction above ground. In the ocean, especially in studies of large-scale oceanic circulation, for which sparse current-meter data may not be representative of the large flow due to local eddy effects, the basinwide distribution may be considered unknown. For this reason, oceanographers typically assume that the currents vanish at some great depth (e.g., 2000 m) and integrate the "thermal-wind" relations from there upward to estimate the surface currents. Although the method is convenient (the equations are linear and do not require integration in time), we should keep in mind that the thermal-wind relation of Eqs. (15.3) and (15.4) is rooted in an assumption of strict geostrophic balance. Obviously, this will not be true everywhere and at all times.

15.2 GEOSTROPHIC ADJUSTMENT

We may now ask how situations like the ones depicted in Fig. 15.1 and 15.2 can arise. In the atmosphere, the temperature gradient from the warm tropics to the cold polar regions creates a permanent feature of the global atmosphere, although storms do alter the magnitude of this gradient in time and space. Ocean currents can bring in near contact water masses of vastly different origins and thus densities. Finally, coastal processes such as freshwater runoff can create density differences between saltier waters offshore and fresher waters closer to shore. Thus, a variety of mechanisms exists by which different fluid masses can be brought in contact.

Oftentimes, the contact between different fluid masses is recent, and the flow has not yet had the time to achieve thermal-wind balance. An example is coastal upwelling: Alongshore winds create in the ocean an offshore Ekman drift, and the depletion of surface water near the coast brings denser water from below

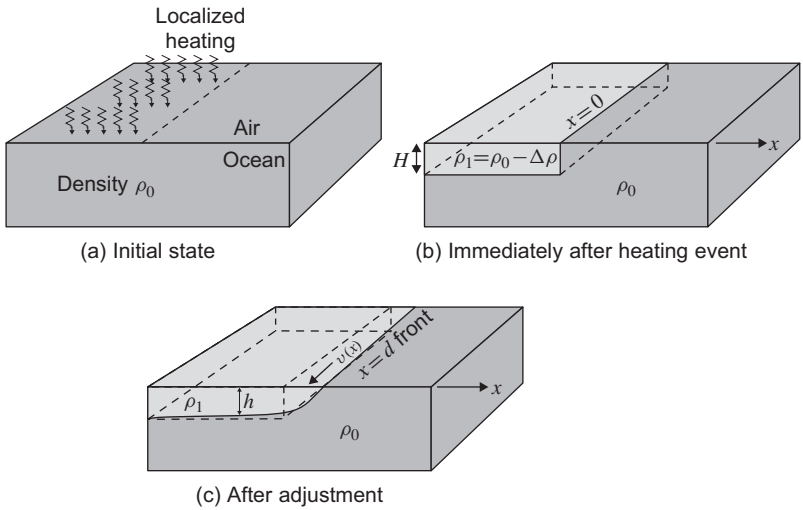


FIGURE 15.3 A simple case of geostrophic adjustment.

(see later section in this chapter). Such a situation is initially out of equilibrium and gradually seeks adjustment.

Let us explore in a very simple way the dynamical adjustment between two fluid masses recently brought into contact. Let us imagine an infinitely deep ocean that is suddenly heated over half of its extent (Fig. 15.3a). A warm upper layer develops on that side, whereas the rest of the ocean, on the other side and below, remains relatively cold (Fig. 15.3b). (We could also imagine a vertical gate preventing buoyant water from spilling from one side to the other.) After the upper layer has been created—or, equivalently, when the gate is removed—the ocean is not in a state of equilibrium, the lighter surface water spills over to the cold side, and an adjustment takes place. In the absence of rotation, spilling proceeds, as we can easily imagine, until the light water has spread evenly over the entire domain and the system has come to rest. But this scenario, as we are about to note, is not what happens when rotational effects are important.

Under the influence of the Coriolis force, the forward acceleration induced by the initial spilling creates a current that veers (to the right in the northern hemisphere) and can come into geostrophic equilibrium with the pressure difference associated with the density heterogeneity. The result is a limited spill accompanied by a lateral flow (Fig. 15.3c).

To model the process mathematically, we use the reduced-gravity model (12.19) on an f-plane and with reduced-gravity constant $g' = g(\rho_0 - \rho_1)/\rho_0$ according to the notation of Fig. 15.3b. We neglect all variations in the y-direction, although we allow for a velocity, v, in that direction, and write

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}-fv=-g'\frac{\partial h}{\partial x} \tag{15.6a}$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+fu=0 \tag{15.6b}$$

$$\frac{\partial h}{\partial t}+\frac{\partial}{\partial x}(hu)=0. \tag{15.6c}$$

The initial conditions (i.e., immediately after the warming event) are $u=v=0$, $h=H$ for $x<0$, and $h=0$ for $x>0$. The boundary conditions are $u, v\to 0$ and $h\to H$ as $x\to-\infty$, whereas the velocity component u at the front is given by the material derivative $u=dx/dt$ where $h=0$ at $x=d(t)$, the moving point where the interface outcrops. This nonlinear problem cannot be solved analytically, but one property can be stated. Fluid parcels governed by the preceding equations conserve the following form of the potential vorticity:

$$q=\frac{f+\partial v/\partial x}{h}. \tag{15.7}$$

Initially, all particles have $v=0$, $h=H$ and share the same potential vorticity $q=f/H$. Therefore, throughout the layer of light fluid and at all times, the potential vorticity keeps the uniform value f/H:

$$\frac{f+\partial v/\partial x}{h}=\frac{f}{H}. \tag{15.8}$$

This property, it turns out, allows us to relate the initial state to the final state without having to solve for the complex, intermediate evolution.

Once the adjustment is completed, time derivatives vanish. Equation (15.6c) then requires that hu be a constant; since $h=0$ at one point, this constant must be zero, implying that u must be zero everywhere. Equation (15.6b) reduces to zero equals zero and tells nothing. Finally, Eq. (15.6a) implies a geostrophic balance,

$$-fv=-g'\frac{\mathrm{d}h}{\mathrm{d}x}, \tag{15.9}$$

between the velocity and the pressure gradient set by the sloping interface. Alone, Eq. (15.9) presents one relation between two unknowns, the velocity and the depth profile. The potential-vorticity conservation principle (15.8), which still holds at the final state, provides the second equation, thereby conveying the information about the initial disturbance into the final state.

Despite the nonlinearities of the original governing Eqs. (15.6a)–(15.6c), the problem at hand, Eqs. (15.8) and (15.9), is perfectly linear, and the solution is relatively easy to obtain. Elimination of either $v(x)$ or $h(x)$ between the two

equations yields a second-order differential equation for the remaining variable, which admits two exponential solutions. Discarding the exponential that grows for $x \to -\infty$ and imposing the boundary condition $h = 0$ at $x = d$ lead to

$$h = H\left[1 - \exp\left(\frac{x-d}{R}\right)\right] \tag{15.10}$$

$$v = -\sqrt{g'H}\exp\left(\frac{x-d}{R}\right), \tag{15.11}$$

where R is the deformation radius, defined by

$$R = \frac{\sqrt{g'H}}{f}, \tag{15.12}$$

and d is the unknown position of the outcrop (where h vanishes). To determine this distance, we must again tie the initial and final states, this time by imposing volume conservation.[2] Ruling out a finite displacement at infinity where there is no activity, we require that the depletion of light water on the left of $x = 0$ be exactly compensated by the presence of light water on the right, that is,

$$\int_{-\infty}^{0} (H - h)\,\mathrm{d}x = \int_{0}^{d} h\,\mathrm{d}x, \tag{15.13}$$

which yields a transcendental equation for d, the solution of which is surprisingly simple:

$$d = R = \frac{\sqrt{g'H}}{f}. \tag{15.14}$$

Thus, the maximum distance over which the light water has spilled in the adjusted state is none other than the radius of deformation, hence the name of the latter.

Notice that R has the Coriolis parameter f in its denominator. Therefore, the spreading distance, R, is less than infinity because f differs from zero. In other words, the spreading is confined because of the earth's rotation via the Coriolis effect. In a nonrotating framework, the spreading would, of course, be unlimited.

Lateral heterogeneities are constantly imposed onto the atmosphere and oceans, which then adjust and establish patterns whereby these lateral heterogeneities are somewhat distorted but maintained. Such patterns are at or near geostrophic equilibrium and can thus persist for quite a long time. This explains why discontinuities such as fronts are common occurrences in both the

[2]The reason why we use a volume conservation to determine the frontal position but let some energy be lost from the system, and not the reverse, is rooted in the very different nature of mass and energy propagation. The latter can be transported far away (to infinity) by waves without net displacement of fluid, while mass propagation demands advection by the flow.

atmosphere and the oceans. As the preceding example suggests, fronts and the accompanying winds or currents take place over distances on the order of the deformation radius. To qualify the activity observed at that length scale, meteorologists refer to the *synoptic scale*, whereas oceanographers prefer to use the adjective *mesoscale*.

We can vary the initial, hypothetical disturbance and generate a variety of geostrophic fronts, all being steady states. A series of examples, taken from published studies, is provided in Fig. 15.4. They are, in order, as follows: surface-to-bottom front on a flat bottom, which can result from sudden and localized heating (or cooling); surface-to-bottom front at the shelf break resulting from the existence of distinct shelf and deep water masses; double, surface-to-surface front; and three-layer front as a result of localized mixing of an otherwise two-layer stratified fluid. The interested reader is referred to

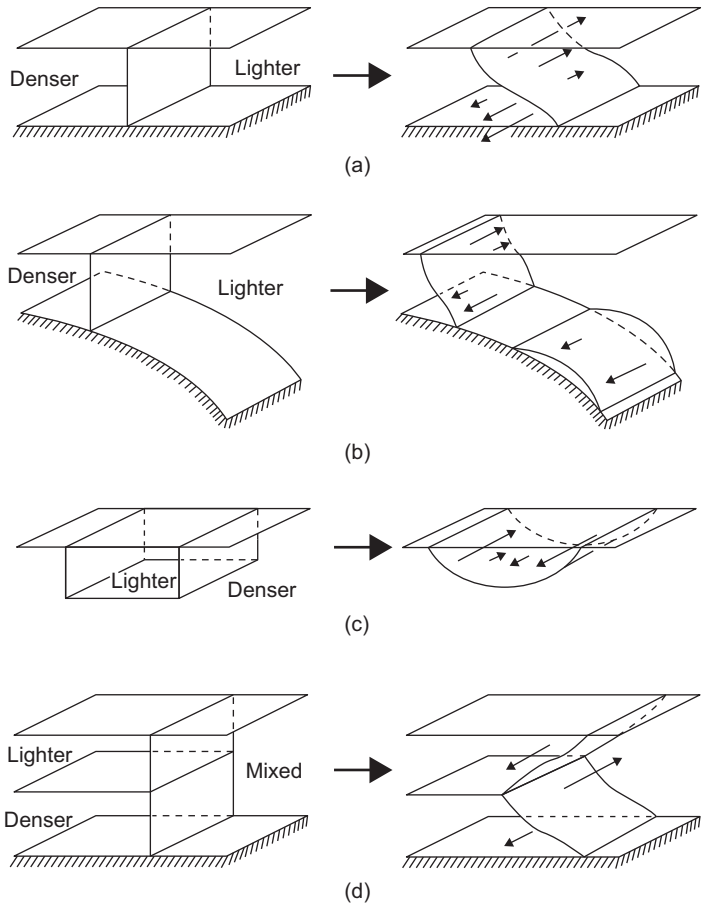


FIGURE 15.4 Various examples of geostrophic adjustment.

the original articles by Rossby (1937, 1938), the article by Veronis (1956), the review by Blumen (1972), and other articles on specific situations by Stommel and Veronis (1980), Hsueh and Cushman-Roisin (1983), and van Heijst (1985). Ou (1984) considered the geostrophic adjustment of a continuously stratified fluid and showed that if the initial condition is sufficiently away from equilibrium, density discontinuities can arise during the adjustment process. In other words, *fronts* can spontaneously emerge from earlier continuous conditions.

The preceding applications dealt with situations in which there is no variation in one of the two horizontal directions. The general case (see Hermann, Rhines & Johnson, 1989) may yield a time-dependent flow that is nearly geostrophic.

15.3 ENERGETICS OF GEOSTROPHIC ADJUSTMENT

The preceding theory of geostrophic adjustment relied on potential vorticity and volume conservation principles, but nothing was said of energy, which must also be conserved in a nondissipative system. All we can do, now that the solution has been obtained, is to check on the budget, where a surprise is awaiting us!

Initially, the system is at rest, and there is no kinetic energy ($KE_i = 0$), whereas the initial potential energy (per unit length in the transverse direction) is[3]

$$PE_i = \frac{1}{2}\,\rho_0 \int_{-\infty}^{0} g' H^2 \,\mathrm{d}x. \tag{15.15}$$

Although this expression is infinite, only the difference with the final potential energy will interest us. So, there is no problem. At the final state, the velocity u is zero, leaving the kinetic energy to be

$$KE_f = \frac{1}{2}\,\rho_0 \int_{-\infty}^{d} hv^2 \,\mathrm{d}x, \tag{15.16}$$

and the potential energy is

$$PE_f = \frac{1}{2}\,\rho_0 \int_{-\infty}^{d} g' h^2 \,\mathrm{d}x. \tag{15.17}$$

During the spreading phase, some of the lighter water has been raised and some heavier water has been lowered to take its place. Hence, the center of

[3]To verify that Eq. (15.15) is the correct form for the potential energy, solve Analytical Problem 12.2 and use as a template for demonstration the approach used in establishing energy conservation for the two-layer system in Section 12.6.

gravity of the system has been lowered, and we expect a drop in potential energy. Calculations yield

$$\Delta PE = PE_i - PE_f = \frac{1}{4}\rho_0 g' H^2 R. \tag{15.18}$$

Some kinetic energy has been created by setting a transverse current. The amount is

$$\Delta KE = KE_f - KE_i = \frac{1}{12}\ \rho_0\ g' H^2 R. \tag{15.19}$$

Therefore, as we can see, only one-third of the potential-energy drop has been consumed by the production of kinetic energy, and we should ask: What has happened to the other two-thirds of the released potential energy? The answer lies in the presence of transients, which occur during the adjustment: some of the time-dependent motions are gravity waves (here, internal waves on the interface), which travel to infinity, radiating energy away from the region of adjustment. In reality, such waves dissipate along the way, and there is a net decrease of energy in the system. The ratio of kinetic-energy production to potential-energy release varies from case to case (Ou, 1986) but tends to remain between 1/4 and 1/2.

An interesting property of the geostrophically adjusted state is that it corresponds to the greatest energy loss and thus to a level of minimum energy. Let us demonstrate this proposition in the particular case at hand. The energy of the system is at all times

$$E = PE + KE = \frac{\rho_0}{2}\int_{-\infty}^{d}\left[g' h^2 + h(u^2 + v^2)\right]\mathrm{d}x, \tag{15.20}$$

and we know that the evolution is constrained by conservation of potential vorticity:

$$f + \frac{\partial v}{\partial x} = \frac{f}{H}h. \tag{15.21}$$

Let us now search for the state that corresponds to the lowest possible level of energy, (15.20), under constraint (15.21) by forming the variational principle:

$$\mathcal{F}(h, u, v, \lambda) = \frac{\rho_0}{2}\int_{-\infty}^{+\infty}\left[g' h^2 + h\left(u^2 + v^2\right) - 2\lambda\left(f + \frac{\partial v}{\partial x} - \frac{fh}{H}\right)\right]\mathrm{d}x \tag{15.22}$$

$$\delta\mathcal{F} = 0 \quad \text{for any } \delta h, \delta u, \delta v \text{ and } \delta\lambda. \tag{15.23}$$

Because expression (15.20) is positive definite, the extremum will be a minimum. The variations with respect to the three state variables h, u, and v and the

Lagrange multiplier λ yield, respectively,

$$\delta h : g'h + \frac{1}{2}\left(u^2 + v^2\right) + \frac{f}{H}\,\lambda = 0 \tag{15.24a}$$

$$\delta u : hu = 0 \tag{15.24b}$$

$$\delta v : hv + \frac{\partial \lambda}{\partial x} = 0 \tag{15.24c}$$

$$\delta \lambda : f + \frac{\partial v}{\partial x} - \frac{f}{H}\,h = 0. \tag{15.24d}$$

Equation (15.24b) provides $u = 0$, whereas the elimination of λ between Eqs. (15.24a) and (15.24c) leads to

$$\frac{\partial}{\partial x}\left(g'h + \frac{1}{2}\,v^2\right) + \frac{f}{H}\,(-hv) = 0,$$

or

$$g'\,\frac{\partial h}{\partial x} + v\left(\frac{\partial v}{\partial x} - \frac{f}{H}\,h\right) = 0.$$

Finally, use of Eq. (15.24d) reduces this last equation to

$$g'\,\frac{\partial h}{\partial x} - fv = 0.$$

In conclusion, the state of minimum energy is the state in which u vanishes, and the cross-isobaric velocity is geostrophic—that is, the steady, geostrophic state.

It can be shown that the preceding conclusion remains valid in the general case of arbitrary, multilayer potential-vorticity distributions, as long as the system is uniform in one horizontal direction. Therefore, it is a general rule that geostrophically adjusted states correspond to levels of minimum energy. This may explain why geophysical flows commonly adopt a nearly geostrophic balance.

15.4 COASTAL UPWELLING

15.4.1 The Upwelling Process

Winds blowing over the ocean generate Ekman layers and currents. The depth-averaged currents, called the Ekman drift, forms an angle with the wind, which was found to be 90° (to the right in the northern hemisphere) according to a simple theory (Section 8.6). So, when a wind blows along a coast, it generates an Ekman drift directed either onshore or offshore, to which the coast stands as an obstacle. The drift is offshore if the wind blows with the coast on its left (right) in the northern (southern) hemisphere (Fig. 15.5). If this is the case, water depletion occurs in the upper layers, and a low pressure sets in, forcing waters from below to move upward and replenish at least partly the space vacated by

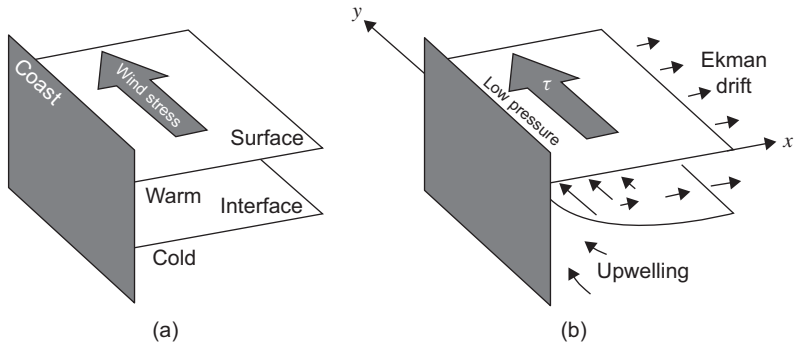


FIGURE 15.5 Schematic development of coastal upwelling.

the offshore drift. This phenomenon is called *coastal upwelling*. The upward movement calls for a replenishment at the lower levels, which is accomplished by an onshore flow at depth. To recapitulate, a wind blowing along the coast (with the coast on the left or the right in, respectively, the northern or the southern hemisphere) sets an offshore current in the upper levels, an upwelling at the coast, and an onshore current at lower levels.

This circulation in the cross-shore vertical plane is not the whole story, however. The low pressure created along the coast also sustains, via geostrophy, an alongshore current, while vertical stretching in the lower layer generates relative vorticity and a shear flow. Or, from a different perspective, the vertical displacement creates lateral density gradients, which in turn call for a thermal wind, the shear flow. The flow pattern is thus rather complex.

At the root of coastal upwelling is a divergent Ekman drift. And, we can easily conceive of other causes besides a coastal boundary for such divergence. Two other upwelling situations are noteworthy: one along the equator and the other at high latitudes. Along the equator, the trade winds blow quite steadily from east to west. On the northern side of the equator, the Ekman drift is to the right, or away from the equator, and on the southern side, it is to the left, again away from the equator (Fig. 15.6). Consequently, horizontal divergence occurs along the equator, and mass conservation requires upwelling (Gill, 1982, Chapter 11; Yoshida, 1959).

At high latitudes, upwelling frequently occurs along the ice edge, in the so-called marginal ice zone. A uniform wind exerts different stresses on ice and open water; in its turn, the moving ice exerts a stress on the ocean beneath. The net effect is a complex distribution of stresses and velocities at various angles, with the likely result that the ocean currents at the ice edge do not match (Fig. 15.6). For certain angles between wind and ice edge, these currents diverge, and upwelling again takes place to compensate for the divergence of the horizontal flow (Häkkinen, 1990).

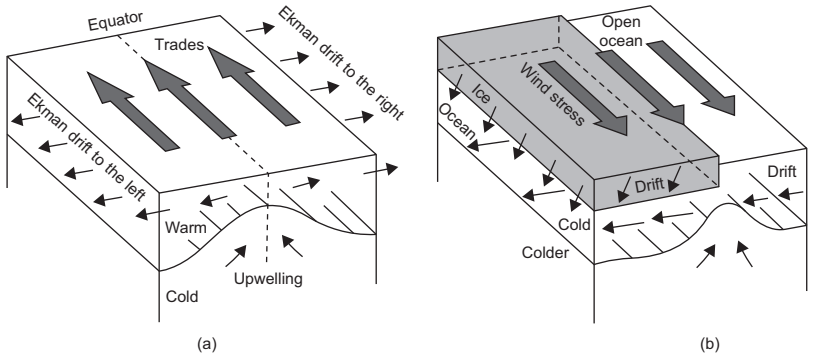


FIGURE 15.6 Other types of upwelling: (a) equatorial upwelling, (b) upwelling along the ice edge.

The upwelling phenomenon, especially the coastal type, has been the subject of considerable attention, chiefly because of its relation to biological oceanography and, from there, to fisheries. In brief, small organisms in the ocean (phytoplankton) proliferate when two conditions are met: sunlight and a supply of nutrients. In general, nutrients lie in the deeper waters, below the reach of sunlight, and so the waters tend to lack either nutrients or sunlight. The major exceptions are the upwelling regions, where deep, nutrient-rich waters rise to the surface, receive sunlight, and stimulate biological activity. Upwelling-favorable winds most generally occur along the west coasts of continents where the prevailing winds blow toward the equator. For a review of observations and a discussion of the biological implications of coastal upwelling, the interested reader is referred to the volume edited by Richards (1981).

15.4.2 A Simple Model of Coastal Upwelling

Consider a reduced-gravity ocean on an f-plane ($f > 0$), bounded by a vertical wall and subjected to a surface stress acting with the wall on its left (Fig. 15.5a). The upper moving layer, defined to include the entire vertical extent of the Ekman layer, supports an offshore drift current. The lower layer is, by virtue of the choice of a reduced-gravity model, infinitely deep and motionless. In the absence of alongshore variations, the equations of motion are

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - fv = -g'\frac{\partial h}{\partial x} \tag{15.25a}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + fu = \frac{\tau}{\rho_0 h} \tag{15.25b}$$

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0, \tag{15.25c}$$

where x is the offshore coordinate, τ is the alongshore wind stress, and all other symbols are conventional (Fig. 15.5b).

Despite its apparent simplicity, the preceding set of equations is nonlinear, and no analytical solution is known. We therefore linearize these equations by assuming that the wind stress τ and, in turn, the ocean's reaction are weak. Noting $h = H - a$, where H is the depth of the undisturbed upper layer and a the small upward displacement of the interface, we write

$$\frac{\partial u}{\partial t} - fv = g' \frac{\partial a}{\partial x} \tag{15.26}$$

$$\frac{\partial v}{\partial t} + fu = \frac{\tau}{\rho_0 H} \tag{15.27}$$

$$-\frac{\partial a}{\partial t} + H\frac{\partial u}{\partial x} = 0. \tag{15.28}$$

This set of equations contains two independent x-derivatives and thus calls for two boundary conditions. Naturally, u vanishes at the coast ($x = 0$) and a vanishes far offshore ($x \to +\infty$).

The solution to the problem depends on the initial conditions, which may be taken to correspond to the state of the rest ($u = v = a = 0$). Yoshida (1955) is credited with the first derivation of the problem's solution (extended to two moving layers). However, because of the fluctuating nature of winds, upwelling is rarely an isolated event in time, and we prefer to investigate the periodic solutions to the preceding linear set of equations. Taking $\tau = \tau_0 \sin \omega t$, where τ_0 is a constant in both space and time, we note that the solution must be of the type $u = u_0(x) \sin \omega t$, $v = v_0(x) \cos \omega t$ and $a = a_0(x) \cos \omega t$. Substitution and solution of the remaining ordinary differential equations in x yield

$$u = \frac{f\tau_0}{\rho_0 H\left(f^2 - \omega^2\right)}\left[1 - \exp\left(-\frac{x}{R_\omega}\right)\right]\sin \omega t \tag{15.29a}$$

$$v = \frac{\omega\tau_0}{\rho_0 H\left(f^2 - \omega^2\right)}\left[1 - \frac{f^2}{\omega^2}\exp\left(-\frac{x}{R_\omega}\right)\right]\cos \omega t \tag{15.29b}$$

$$a = \frac{-f R_\omega \tau_0}{\rho_0 g' H \omega}\exp\left(-\frac{x}{R_\omega}\right)\cos \omega t, \tag{15.29c}$$

where R_ω is a modified deformation radius defined as

$$R_\omega = \sqrt{\frac{g'H}{f^2 - \omega^2}}. \tag{15.30}$$

From the preceding solution, we conclude that the upwelling or downwelling signal is *trapped* along the coast within a distance on the order of R_ω. Far offshore ($x \to \infty$), the interfacial displacement vanishes, and the flow field includes

the Ekman drift

$$u_{\mathrm{Ek}} = \frac{\tau_0}{\rho_0 f H} \sin \omega t, \quad v_{\mathrm{Ek}} = 0. \tag{15.31}$$

At long periods such as weeks and months ($\omega \ll f$), the distance R_ω becomes the radius of deformation, the vertical interfacial displacements become very large (indeed, the wind blows more steadily in one direction before it reverses), and the far-field oscillations become much smaller than the Ekman drift. Obviously, for very large vertical displacements and low frequency, we must ensure that the linearization hypothesis remains valid, that is, $|a| \ll H$. In terms of the forcing and with $R_\omega \simeq \sqrt{g'H}/f$, this condition translates into

$$\frac{\tau_0}{\rho_0 \omega H} \ll \sqrt{g'H}, \tag{15.32}$$

a condition for which an interpretation will soon be found.

At superinertial frequencies ($\omega > f$), the quantity R_ω becomes imaginary, indicating that the solution does not decay away from the coast but instead oscillates. Physically, the ocean's response is not trapped near the coast and inertia-gravity waves (Section 9.3) are excited. These radiate outward, filling the entire basin. Thus, depending on its frequency, the energy imparted by the wind to the ocean may either remain localized or be radiated away.

15.4.3 Finite-Amplitude Upwelling

If the wind is sufficiently strong or is blowing for a sufficiently long time, the density interface can rise to the surface, forming a front. Continued wind action displaces this front offshore and exposes the colder waters to the surface. This mature state is called *full upwelling* (Csanady, 1977). Obviously, the previous linear theory is no longer applicable.

Because of the added complications arising from the nonlinearities, let us now restrict our investigation to the final state of the ocean after a wind event of finite duration. Equation (15.25b), expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}(v + fx) = \frac{\tau}{\rho_0 h}, \tag{15.33}$$

where $\mathrm{d}/\mathrm{d}t = \partial/\partial t + u\partial/\partial x$ is the time derivative following a fluid particle in the offshore direction, can be integrated over time to yield:

$$(v + fx)_{\text{at end of event}} - (v + fx)_{\text{initially}} = I. \tag{15.34}$$

The *wind impulse* I is the integration of the wind-stress term, $\tau/\rho_0 h$, over time and following a fluid particle. Although the wind impulse received by every parcel cannot be precisely determined, it can be estimated by assuming that the

wind event is relatively brief. The time integral can then be approximated by using the local stress value and replacing h by H:

$$I \simeq \frac{1}{\rho_0 H} \int_{\text{event}} \tau \, \mathrm{d}t. \tag{15.35}$$

If the initial state is one of rest, relation (15.34) implies that a particle initially at distance X from the coast is at distance x immediately after the wind ceases and has an alongshore velocity v such that

$$v + fx - fX = I. \tag{15.36}$$

During the subsequent adjustment and until equilibrium is reached, Eq. (15.33) (with $\tau = 0$) implies that the quantity $v + fx$ remains unchanged, and relation (15.36) continues to hold after the wind has ceased.

If a spatially uniform wind blows over an ocean layer of uniform depth, the drift velocity, too, is uniform, and no vorticity is imparted to fluid parcels. Hence, potential vorticity is conserved during a uniform wind event over a uniform layer (see also Analytical Problem 15.9). After the event, in the absence of further forcing, potential vorticity remains conserved throughout the adjustment phase:

$$\frac{1}{h}\left(f + \frac{\partial v}{\partial x}\right) = \frac{f}{H}. \tag{15.37}$$

Once a steady state has been achieved, there is no longer any offshore velocity ($u = 0$), according to Eq. (15.25c). The remaining equation (15.25a), reduces to a simple geostrophic balance, which together with Eq. (15.37) provides the solution:

$$h = H - A \exp\left(-\frac{x}{R}\right) \tag{15.38}$$

$$v = A\sqrt{\frac{g'}{H}} \exp\left(-\frac{x}{R}\right), \tag{15.39}$$

where R is now the conventional radius of deformation ($\sqrt{g'H}/f$). The constant of integration A represents the amplitude of the upwelled state and is related to the wind impulse via Eq. (15.36). Two possible outcomes must be investigated: either the interface has not risen to the surface (Fig. 15.7, case I) or it has outcropped, forming a front and leaving cold waters exposed to the surface near the coast (Fig. 15.7, case II).

In case I, the particle initially against the coast ($X = 0$) is still there ($x = 0$), and relation (15.36) yields $v(x = 0) = I$. Solution (15.39) meets this condition if $A = I(H/g')^{1/2}$. The depth along the coast, $h(x = 0) = H - A$, must be positive requiring $A \leq H$; that is, $I \leq (g'H)^{1/2}$. In other words, the no-front situation or partial upwelling of case I occurs if the wind is sufficiently weak or sufficiently brief that its resulting impulse is less than the critical value $(g'H)^{1/2}$.

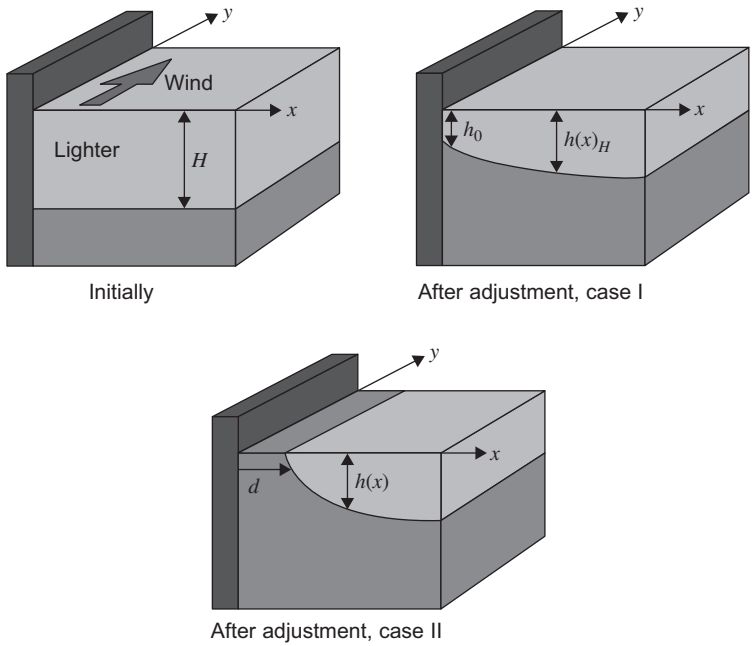


FIGURE 15.7 The two possible outcomes of coastal upwelling following an alongshore wind of finite duration. After a weak or brief wind (case I), the interface has upwelled but not to the point of reaching the surface. A strong or prolonged wind event (case II) causes the interface to reach the surface, where it forms a front; this front is displaced offshore, leaving cold waters from below exposed to the surface. This latter case corresponds to a mature upwelling that favors biological activity.

In the more interesting case II, the front has been formed, and the particle initially against the coast ($X = 0$) is now at some offshore distance ($x = d \geq 0$), marking the position of the front. There the layer depth vanishes, $h(x = d) = 0$, and solution (15.38) yields $A = H \exp(d/R)$. The alongshore velocity at the front is, according to (15.39), $v(x = d) = (g'H)^{1/2}$. Finally, relation (15.36) leads to the determination of the offshore displacement d in terms of the wind impulse:

$$d = \frac{I}{f} - R. \tag{15.40}$$

Since this displacement must be a positive quantity, it is required that $I \geq (g'H)^{1/2}$. Physically, if the wind is sufficiently strong or sufficiently prolonged, so that the net impulse is greater than the critical value $(g'H)^{1/2}$, the density interface rises to the surface and forms a front that migrates away from shore, leaving cold waters from below exposed to the surface. Note how the conditions for the realizations of cases I and II complement each other.

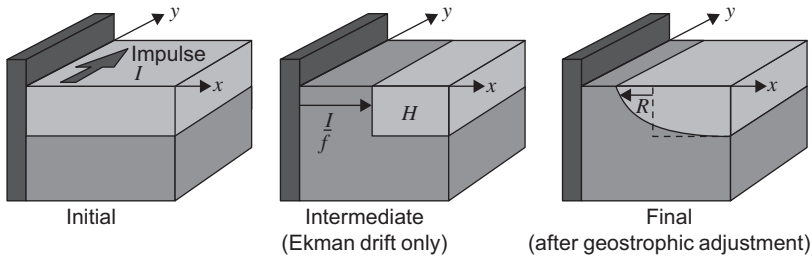


FIGURE 15.8 Decomposition of the formation of a coastal-upwelling front as a two-stage process: first, an offshore Ekman drift in response to the wind, followed by a backward geostrophic adjustment.

Note in passing how condition (15.32) can now be interpreted. Its left-hand side is the wind impulse over a time period $1/\omega$, which must be small compared to the critical value $(g'H)^{1/2}$, in order to remain far away from the outcropping situation, which would invalidate the linearization assumption.

Formula (15.40) has a simple physical interpretation as sketched in Fig. 15.8. The offshore Ekman velocity u_{Ek} is the velocity necessary for the Coriolis force to balance the alongshore wind stress:

$$u_{\mathrm{Ek}} = \frac{\tau}{\rho_0 f h}, \tag{15.41}$$

according to Eq. (8.34a). Integrated over time, this yields a net offshore displacement proportional to the wind impulse

$$x_{\mathrm{Ek}} = \frac{I}{f}. \tag{15.42}$$

If we were now to assume that the wind is responsible for an offshore shift of this magnitude, whereas the surface waters are moving as a solid slab, we would get the intermediate structure of Fig. 15.8. But such a situation cannot persist, and an adjustment must follow, causing an onshore spread similar to that considered in Section 15.2—that is, over a distance equal to the deformation radius. Hence, we have the final structure of Fig. 15.8 and formula (15.40).

15.4.4 Variability of the Upwelling Front

Up to this point, we have considered only processes operating in the offshore direction or, equivalently, an upwelling that occurs uniformly along a straight coast. In reality, the wind is often localized, the coastline not straight, and upwelling not at all uniform. A local upwelling sends a wave signal along the coast, taking the form of an internal Kelvin wave, which in the northern hemisphere propagates with the coast on its right. This redistribution of information not only decreases the rate of upwelling in the forced region but also generates

upwelling in other, unforced areas. As a result, models of upwelling must retain a sizable portion of the coast and both spatial and temporal variations of the wind field (Crépon and Richez, 1982; Brink, 1983).

Because the upwelling front is a region of highly sheared currents, it is a likely region of instabilities. In the two-layer formulation presented in the previous section, this shear is manifested by a discontinuity of the current at the front. The warm layer develops anticyclonic vorticity (i.e., counter to the rotation of the earth) under the influence of vertical squeezing and flows alongshore in the direction of the wind. On the other side of the front, the exposed lower layer is vertically stretched, develops cyclonic vorticity (i.e., in the same direction as the rotation of the earth), and flows upwind. The currents on each side of the front thus flow in opposite directions, causing a large shear, which, as we have seen (Chapter 10), is vulnerable to instabilities. In addition to the kinetic-energy supply in the horizontal shear (barotropic instability), potential energy can also be released from the stratification by a spreading of the warm layer (baroclinic instability; see Chapter 17). Offshore jets of cold, upwelled waters have been observed to form near capes; these jets cut through the front, forge their way through the warm layer, and eventually split to form pairs of counter-rotating vortices (Flament, Armi & Washburn, 1985). This explains why mesoscale turbulence is associated with upwelling fronts (see, e.g., Fig. 15.9 and the article by Strub, Kosro & Huyer, 1991).

The situation is complex and demands careful modeling. Irregularities in the topography and coastline may play influential roles and require adequate spatial resolution, whereas accurate simulation of the instabilities is only possible if numerical dissipation is not excessive in the model.

15.5 ATMOSPHERIC FRONTOGENESIS

Atmospheric fronts are sharp boundaries between cold and warm air masses and have become familiar features of daily weather forecasts. A cold front, depicted in weather charts as a line with spikes (Fig. 15.10), occurs when a colder air mass overtakes a warmer air mass, thus lowering the temperature where it passes. In contrast, a warm front, depicted in weather charts as a line with semicircles, occurs when a warm air mass overtakes a cold air mass, thereby raising the local temperature. The process by which sharp temperature gradients naturally form in the atmosphere is called *frontogenesis* and is readily identified on temperature maps (Fig. 15.11). The word *front* was first coined by Vilhelm Bjerknes[4] who initiated the study of cyclone and front formation during World War I and suggested an analogy between the meeting of two atmospheric air masses and a military line, called a front. The study of frontogenesis has a long history, and the reader may wish to consult the seminal papers of Sawyer (1956),

[4]See biography at the end of Chapter 3.

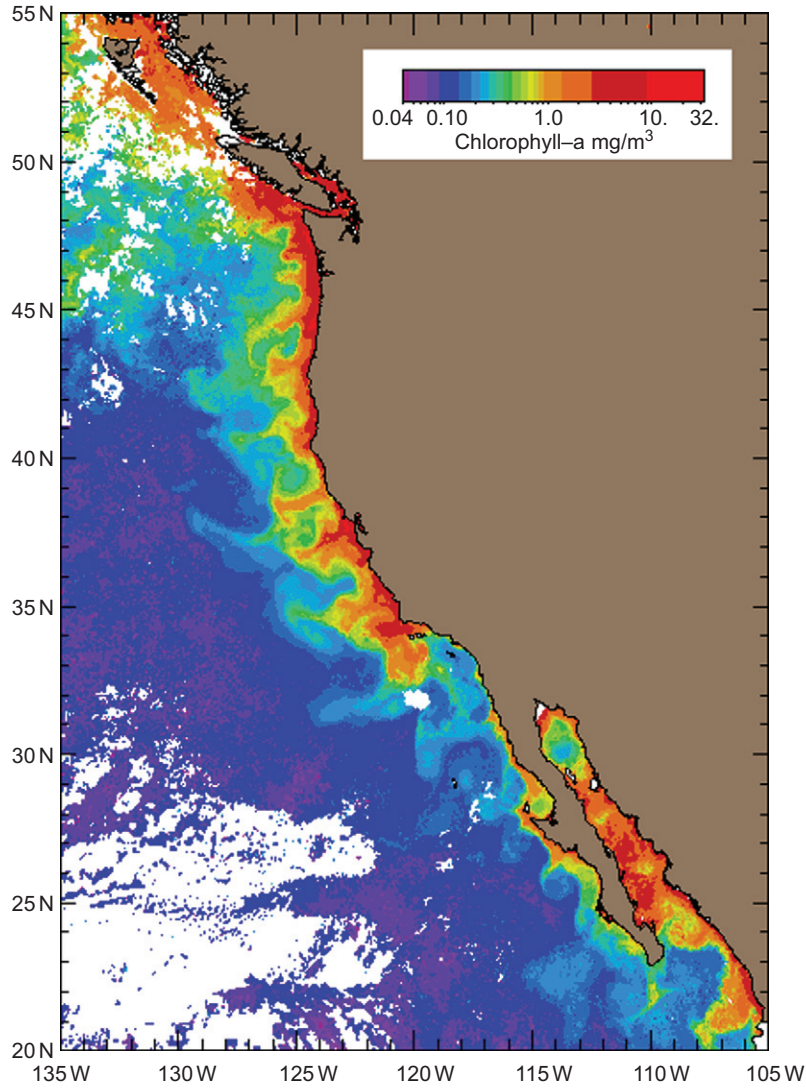


FIGURE 15.9 SeaWiFS satellite image of the North American Pacific coast showing the occurrence of coastal upwelling from Baja California (Mexico) to Vancouver Island (Canada). Shades indicate the amount of chlorophyll concentration in the water, with high values (lighter shades) in regions of high biological activity and low values (darker shades) in biologically inactive waters. Note how instabilities greatly distort the upwelling front. (*Composite image provided courtesy of Dr. Andrew Thomas, School of Marine Sciences, University of Maine, USA*).

FIGURE 15.10 Typical evolution of fronts approaching Belgium (the dark-shaded country). Warm fronts are identified by semicircles, cold fronts by triangles. The side of the front on which the symbols are plotted indicates the direction of the frontal movement. For the cold front in the lower center of the first panel, there is cold air on the western side of the front, which is moving eastward. Indeed, by the next day (middle panel), the front has passed over Belgium from west to east. A day later (right panel), this front has disappeared from the map. Meanwhile, a warm front has appeared from the west followed by a cold front rapidly catching up with it. Once the cold front has overtaken the warm front, the warm air that was in the wedge between fronts has been lifted up and is no longer present at the surface. The new front, called an occluded front, has cold air on both sides. It is depicted by alternating semicircles and triangles (upper-central part of the right panel). (*Royal Meteorological Institute Belgium*)

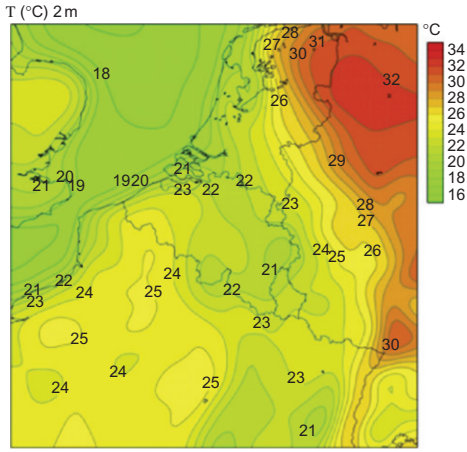


FIGURE 15.11 Temperature field corresponding to the middle panel of Fig. 15.10. Note how the cold front is chasing the warm air eastward. Some of the warm air is also lifted by the advancing cold air, creating condensation. This explains the rainfall accompanying cold fronts. (*Royal Meteorological Institute Belgium*)

Eliassen (1962), and Hoskins and Bretherton (1972). A more detailed mathematical presentation than the one given here can be found in Pedlosky (1987, Section 8.4).

The physical processes involved in frontogenesis are complex, and we will start the analysis with a kinematic study to understand how a given velocity field can lead to a deformation of a thermal distribution that intensifies temperature gradients. Because observations reveal that the generation of a front is a relatively fast process, typically taking no more than a day, we may neglect local heating effects. Also, creating a front by local differential heating would require heat fluxes that exhibit sharp gradients, an unlikely situation. Hence, we focus on temperature changes induced by advection only.

The simplest example (Fig. 15.12) assumes a horizontal velocity field given by

$$u = \omega x, \quad v = -\omega y \tag{15.43}$$

in which ω is a deformation rate. We note that this velocity fields satisfies volume conservation

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \tag{15.44}$$

implying that the vertical velocity is zero over a flat surface, which we assume.

Suppose now that this flow field advects a temperature field with initial gradient in the y-direction. Neglecting turbulent mixing, compressibility, and heating, temperature is conserved by individual fluid parcels, and the temperature field is is governed by the advection equation:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = 0. \tag{15.45}$$

A differentiation of this equation with respect to x gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial x}\right) = -\omega\frac{\partial T}{\partial x}. \tag{15.46}$$

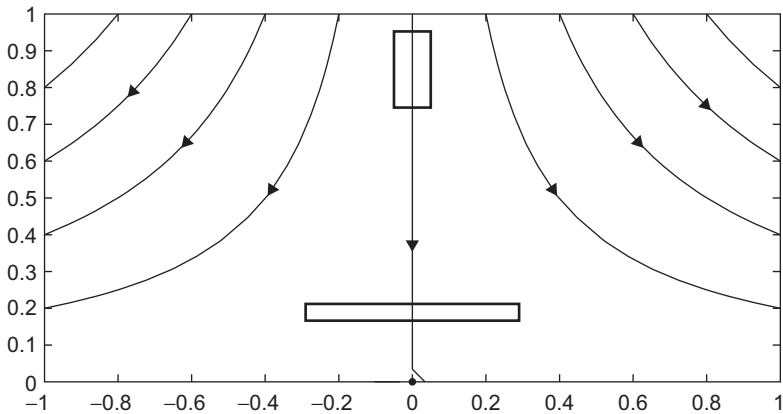


FIGURE 15.12 Frontogenesis induced by the flow field $u = \omega x$, $v = -\omega y$. A collection of fluid parcels forming a rectangle around $x = 0$, $y = 0.9$ stretches in the x-direction as it progresses downstream. By the time parcels have reached $y = 0.2$ (lower rectangle in the figure), convergence in the y-direction has squeezed the fluid parcels, as required by volume conservation to compensate for the divergence in the x-direction. The rear parcels have partly caught up with the front parcels. Without heating or cooling, temperature is conserved by individual fluid parcels, and any pre-existing temperature gradient in the y-direction is intensified. A front occurs when an infinite temperature gradient can be formed in a finite time. (`Frontogenesis.m` may be used to track other groups of fluid parcels.)

Since the initial temperature gradient was exclusively in the y-direction, it follows that $\partial T/\partial x$ was initially zero and remains zero at all subsequent times. Hence, the temperature gradient may change in intensity but not in direction.

More interestingly, differentiation of Eq. (15.45) with respect to y yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial y}\right)=\omega\frac{\partial T}{\partial y}, \tag{15.47}$$

which shows that the magnitude of the temperature gradient increases exponentially with time:

$$\frac{\partial T}{\partial y}=\left.\frac{\partial T}{\partial y}\right|_{t=0}\mathrm{e}^{\omega t}, \tag{15.48}$$

following a fluid parcel. The evolution of the y position of a given air parcel is governed by

$$\frac{\mathrm{d}y}{\mathrm{d}t}=v=-\omega y \quad \Rightarrow \quad y=y_0\mathrm{e}^{-\omega t}. \tag{15.49}$$

Hence, all fluid parcels are converging toward $y=0$. In other words, two parcels with identical x coordinates initially separated by a distance δy_0 see their distance shrink over time. Because each conserves its initial temperature, the temperature gradient increases accordingly.

We now understand how advection can intensify temperature gradients, but by keeping the flow field unchanged, we omitted to consider the fact that the increasing thermal gradient can in turn affect the dynamics. Indeed, a stronger thermal gradient is bound to produce a larger thermal wind, and this is expected to modify the wind velocity that advects the temperature. In other words, there is two-way coupling between velocity and temperature fields. This, it turns out, accelerates the process, and an infinite temperature gradient can be reached in a finite time.

As dynamics accelerate and shorter length scales arise, geostrophy is in jeopardy, and our model needs to retain the effects of nonlinear acceleration (inertia). However, frontal regions are characterized by strong spatial anisotropy, with steep variations across the front and weak variations along the front. Thus, our model may retain geostrophy in one direction. This leads to a semigeostrophic approach (Hoskins & Bretherton, 1972).

With the x-axis aligned with the front, the strong gradients are in the y-direction, and the geostrophic velocity component is u. The weaker v velocity is the one for which geostrophic breaks down. Density (function of temperature) is retained as a dynamically important variable, and the semigeostrophic

equations on the f-plane are

$$\frac{\mathrm{d}u}{\mathrm{d}t} - fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} \tag{15.50a}$$

$$+fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} \tag{15.50b}$$

$$-\alpha g T = -\frac{1}{\rho_0}\frac{\partial p}{\partial z} \tag{15.50c}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{15.50d}$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} = 0, \tag{15.50e}$$

which forms a set of five equations for five variables, namely the three velocity components (u, v, w), pressure p and temperature T. Note that the density was eliminated by using a linear equation of state and that T is measured from the temperature at which density is ρ_0.

The acceleration term $\mathrm{d}u/\mathrm{d}t$ is kept next to the Coriolis term fv in the first equation because u is large and v small, breaking geostrophic balance in the x-momentum budget. Note also that the full material derivative is retained in the first and last equations:

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}. \tag{15.51}$$

The thermal wind balance is obtained by combining the z-derivative of Eq. (15.50b) with the y-derivative of (15.50c):

$$\frac{\partial u}{\partial z} = -\frac{\alpha g}{f}\frac{\partial T}{\partial y}. \tag{15.52}$$

Next, we define the following quantity:

$$q = \left(f - \frac{\partial u}{\partial y}\right)\frac{\partial T}{\partial z} + \frac{\partial u}{\partial z}\frac{\partial T}{\partial y}, \tag{15.53}$$

which is a form of potential vorticity. This q variable is useful because it is conserved by moving fluid parcels. Indeed, some tedious algebra shows that the preceding equations yield the simple conservation equation:

$$\frac{\mathrm{d}q}{\mathrm{d}t} = 0. \tag{15.54}$$

To work with a model that is as simple as possible, we restrict our attention to a flow in which q is initially zero everywhere. With q conserved by fluid

parcels over time, q remains zero everywhere at all subsequent times:

$$\left(f-\frac{\partial u}{\partial y}\right)\frac{\partial T}{\partial z}+\frac{\partial u}{\partial z}\frac{\partial T}{\partial y}=0. \tag{15.55}$$

As we will see shortly, this type of flow has simple attributes that facilitate mathematical developments, but it is not degenerate.

We now become more specific about the flow field, choosing the deformation field used earlier in this section with the addition of terms to reflect the fact that sharpening thermal gradients will affect the thermal wind balance and thus the flow field itself. We assume a solution of the type

$$u=+\omega x+u'(y,z,t) \tag{15.56a}$$

$$v=-\omega y+v'(y,z,t) \tag{15.56b}$$

$$p=-\rho_0 f\omega xy-\frac{1}{2}\rho_0\omega^2x^2+p'(y,z,t) \tag{15.56c}$$

$$w=w(y,z,t) \tag{15.56d}$$

$$T=T(y,z,t). \tag{15.56e}$$

In writing these expressions, care was taken to include in the pressure field terms that are in geostrophic balance with the basic deformation field $(\omega x, -\omega y)$. Further, because of the anisotropy of the front, we anticipate that all components aside from the basic deformation field are independent of the coordinate x. Insertion into Eqs. (15.50) yields

$$\frac{\mathrm{d}u'}{\mathrm{d}t}+\omega u'-fv'=0 \tag{15.57a}$$

$$fu'=-\frac{1}{\rho_0}\frac{\partial p'}{\partial y} \tag{15.57b}$$

$$\alpha\rho_0 gT=\frac{\partial p'}{\partial z} \tag{15.57c}$$

$$\frac{\partial v'}{\partial y}+\frac{\partial w}{\partial z}=0 \tag{15.57d}$$

$$\frac{\mathrm{d}T}{\mathrm{d}t}=0. \tag{15.57e}$$

No linearization was applied, and the material derivative in Eqs. (15.57a) and (15.57e) is the original one except for the x-derivative, which is now nil. We carefully note that the total velocity v appears in this material derivative. The

thermal-wind relation (15.52) and $q=0$ equation (15.55) become

$$\frac{\partial u'}{\partial z} = -\frac{\alpha g}{f}\frac{\partial T}{\partial y} \tag{15.58}$$

$$\left(f - \frac{\partial u'}{\partial y}\right)\frac{\partial T}{\partial z} + \frac{\partial u'}{\partial z}\frac{\partial T}{\partial y} = 0. \tag{15.59}$$

Next, we define the so-called geostrophic coordinate

$$Y = y - \frac{u'}{f}, \tag{15.60}$$

which combines the y coordinate along which gradients occur and the flow in the transverse direction. This quantity has a simple material derivative:

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = -\omega Y \tag{15.61}$$

because $\mathrm{d}y/\mathrm{d}t = v$. With this new variable substituting for u', Eq. (15.59) (expressing $q=0$) becomes

$$\frac{\partial Y}{\partial y}\frac{\partial T}{\partial z} - \frac{\partial Y}{\partial z}\frac{\partial T}{\partial y} = 0, \tag{15.62}$$

which can be recast as

$$-\frac{\partial Y/\partial y}{\partial Y/\partial z} = -\frac{\partial T/\partial y}{\partial T/\partial z} = S. \tag{15.63}$$

This last equation states that the slope S of the Y lines in the vertical plane (y, z) is everywhere equal to the slope of the T lines. This means that the lines of constant Y coincide with the lines of constant T (isotherms), and we can write

$$Y = Y(T, t), \tag{15.64}$$

expressing the fact that in the (y, z) plane, the function Y is constant where T is constant. Here, time t plays the role of a parameter.

Exploiting the thermal-wind relation (15.58), the slope S of isotherms can be expressed in terms of Y and T as

$$S = -\frac{\partial T/\partial y}{\partial T/\partial z} = -\frac{f^2}{\alpha g}\frac{\partial Y/\partial z}{\partial T/\partial z} = -\frac{f^2}{\alpha g}\frac{\partial Y}{\partial T}, \tag{15.65}$$

which takes advantage of the fact that Y is a function of T. This makes that S, too, is a function of only temperature T (and, parametrically, time t). Logic then imposes that if S is unchanging along an isotherm, this isotherm has uniform slope and is thus a straight line. It follows that all isotherms are straight

lines.[5] Note that the slope may vary from isotherm to isotherm, with some more inclined than others, and that the slope of an individual isotherm may change over time.

Next, we return to Eq. (15.61) governing the temporal evolution of Y. Using the fact that Y is not a function of x but a function of only T and time, and the fact that $\mathrm{d}T/\mathrm{d}t = 0$, we obtain

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \left.\frac{\partial Y}{\partial t}\right|_{T=\text{const}} + \frac{\partial Y}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}t} \tag{15.66}$$

$$\left.\frac{\partial Y}{\partial t}\right|_{T=\text{const}} = -\omega Y. \tag{15.67}$$

Its solution is

$$Y = Y_0(T)\mathrm{e}^{-\omega t}, \tag{15.68}$$

in which Y_0 is the initial distribution of Y, a function of T only, which we do not need to specify.

Passing from Y to S with Eq. (15.65), we have

$$S = -\frac{f^2}{\alpha g}\frac{\mathrm{d}Y_0}{\mathrm{d}T}\mathrm{e}^{-\omega t}. \tag{15.69}$$

The slope of each isotherm is thus reduced over time[6] but more so along certain isotherms than others. It now remains to determine how different isotherms can be compared to one another.

To obtain displacements, we first integrate the v velocity component vertically between horizontal and impermeable ($w = 0$) boundaries, for example, a flat land or sea surface below and the tropopause above (Fig. 15.13) Volume conservation (15.57d) dictates

$$\frac{\partial \bar{v}}{\partial y} = -\omega \tag{15.70}$$

where $\bar{v}$ is the vertical average of the full velocity component v (not just v'!). Assuming that the v velocity is the deformation field $-\omega y$ at large distances from the region of interest, the frontal region, Eq. (15.70) tells that the average velocity $\bar{v}$ is everywhere $-\omega y$. This means that a fluid column, on average, moves toward $y = 0$ and that the y distance between neighboring columns decreases exponentially over time.

---

[5] That isotherms are straight lines can be traced to the choice $q = 0$.

[6] This can also be proven by direct manipulation of the governing equations, see Analytical Problem 15.13.

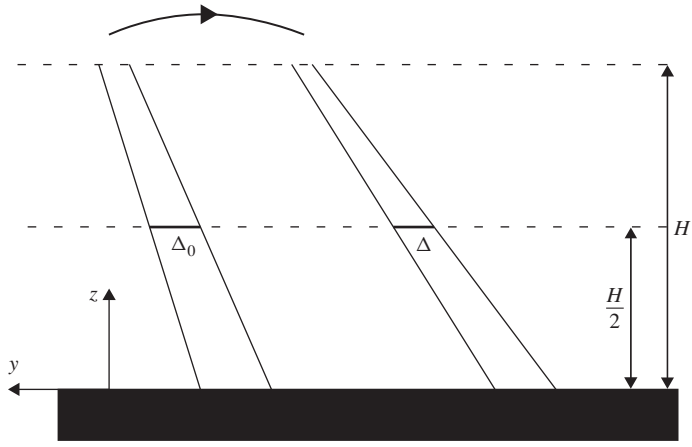


FIGURE 15.13 The distance Δ between two isopycnals at mid-height decreases over time as $\Delta = \Delta_0 e^{-\omega t}$.

Because tilting around $z = H/2$ does not change the volume between two isotherms, which are material surfaces since T is a conserved quantity, the distance Δ between two lines at midlevel $z = H/2$ must decrease exponentially according to $\Delta = \Delta_0 e^{-\omega t}$. This means that the y position of a given isotherm at midlevel $z = H/2$ is none other than the geostrophic coordinate Y. Further, by definition of the slope S, we are entitled to write explicitly

$$Y = y - \frac{z - H/2}{S}. \tag{15.71}$$

While this last equation appears to give Y in terms of y and z, it is best to see it as giving the (y, z) structure of the isotherms in terms of the variables Y and S, which depend only on T and time. Since we know how Y and S vary in time, we can determine the evolution of each isotherm from its initial state.

Given an initial, monotonic temperature distribution at $z = H/2$, say

$$T(y, z = H/2, t = 0) = F(y), \tag{15.72}$$

the inverse function $G = F^{-1}$, which exists because F is monotonic, provides the initial distribution of the geostrophic coordinate in terms of temperature

$$Y_0 = G(T), \tag{15.73}$$

since $Y = y$ at $z = H/2$. Note that, like F, the function G is monotonic, too. The initial slope of an isotherm is known to be

$$S_0(T) = -\frac{f^2}{\alpha g} \frac{dY_0}{dT}. \tag{15.74}$$

We then proceed with the relations $Y = Y_0 e^{-\omega t}$ and $S = S_0 e^{-\omega t}$ to track individual isotherms over time. Figure 15.14 shows a plot made using the code `sgfrontogenesis.m` using an initial temperature distribution F with a slightly enhanced gradient near $y = 0$.

We note that isotherms become gradually less steep, according to Eq. (15.69). This is a form of gravitational relaxation with the denser fluid (colder air in the lower left region) intruding under and lifting the lighter fluid (warmer air in the upper right region). Slacking of isotherms is accentuated in the center where the initial temperature gradient was slightly larger (smaller value of $|dY_0/dT|$). Gradually, some isotherms overtake their neighbors and begin to cross. A pair of discontinuities forms in a finite time. Discontinuities first appear at the top and bottom boundaries and then propagate inward, toward midlevel, where they eventually meet. Physically, a temperature discontinuity is interpreted as a front, a place where temperature varies very rapidly over a very short distance. Note that on the lower (upper) boundary, the discontinuity appears for a positive (negative) value of y, which is on the warm (cold) side of

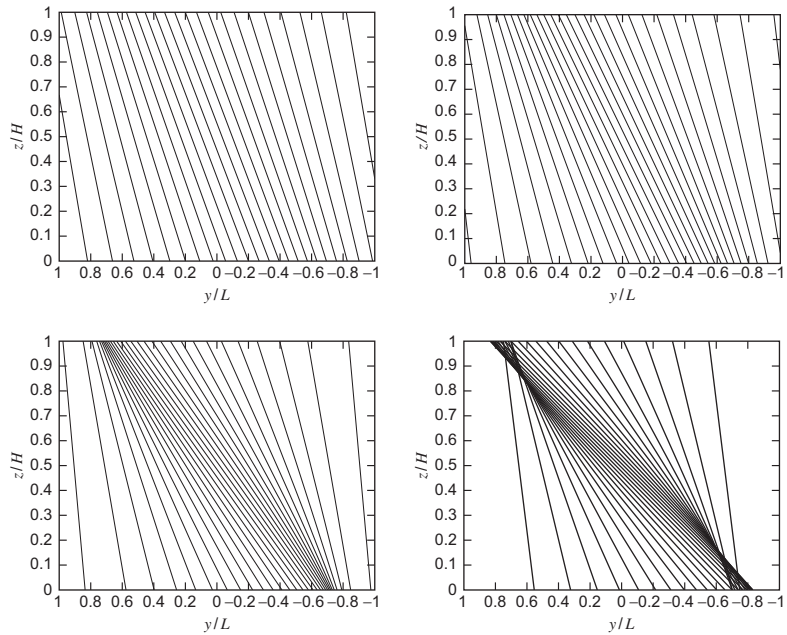


FIGURE 15.14 Evolution of isotherms in a vertical plane during frontogenesis. Note how isotherms become gradually less steep in the center. Eventually (last panel), some isotherms overtake their neighbors and overlap occurs. Physically, a discontinuity, which we call a front, has been formed in a finite time. Note that discontinuities first appear at the top and bottom boundaries. Later (not shown), they propagate inward, toward the midlevel.

the convergence region defined from the basic deformation flow $v = -\omega y$. The shift is attributed to the v' component of the flow field.

According to Eq. (15.71), the intersection of an isotherm with the bottom surface ($z = 0$) occurs at position

$$y_b = Y - \frac{H}{2S} = Y_0 e^{-\omega t} - \frac{H}{2S_0} e^{+\omega t}, \tag{15.75}$$

or, using Eq. (15.74),

$$\begin{aligned} y_b &= Y_0 e^{-\omega t} + \frac{\alpha g H}{2f^2} \frac{1}{dY_0/dT} e^{+\omega t} \\ &= Y_0 e^{-\omega t} + \frac{\alpha g H}{2f^2} \frac{dT}{dY_0} e^{+\omega t}. \end{aligned} \tag{15.76}$$

Two neighboring isotherms begin to intersect when their ground position coincides, that is, when they share the same ground position y_b while retaining their distinct temperatures. Mathematically, this is expressed by a vanishing variation of y_b for a nonzero variation of T, that is

$$\frac{\partial y_b}{\partial T} = 0. \tag{15.77}$$

Switching from the variable T to the variable Y_0, which is in monotonic relation with it, (15.73), we may transform the preceding condition into

$$\frac{\partial y_b}{\partial Y_0} = 0, \tag{15.78}$$

which yields

$$e^{-\omega t} + \frac{\alpha g H}{2f^2} \frac{d^2 T}{dY_0^2} e^{\omega t} = 0. \tag{15.79}$$

In this last equation, the second derivative d^2T/dY_0^2 is known from the initial temperature distribution at midlevel, Eqs. (15.72) and (15.73). Therefore, for every isotherm for which this second derivative is negative, there exits a finite time t given by

$$t = \frac{1}{2\omega} \ln \left[\frac{2f^2}{\alpha g H(-d^2T/dY_0^2)} \right] \tag{15.80}$$

for which condition (15.77) is met. At that time, the isotherm begins to intersect its neighbor, and a temperature discontinuity appears. A front occurs at time t_f that is the shortest of all possible times t given above:

$$t_f = \frac{1}{2\omega} \ln \left[\frac{2f^2}{\alpha g H |d^2T/dY_0^2|_{\max}} \right]. \tag{15.81}$$

To make this result somewhat more concrete, imagine the coordinate y running northward and toward colder air (case of Fig. 15.14). In that case, dY_0/dT is negative and so is dT/dY_0. Finding the maximum of negative d^2T/dY_0^2 is then equivalent to selecting the isotherm marking the place where the initial midlevel temperature decreases fastest with latitude.

Once isotherms begin to intersect, the temperature field becomes multivalued, and the mathematical solution loses physical significance. In reality, dissipative process (dynamic instabilities, friction, and diffusion) become significant, keeping temperature as a unique function of space and thermal gradients large but finite.

15.6 NUMERICAL HANDLING OF LARGE GRADIENTS

A common characteristic of the preceding sections is the appearance and motion of strong gradients, which we call fronts. If we were to apply numerical techniques on a fixed grid to describe such fronts, we would immediately face the problem of needing very high spatial resolution. Indeed, to represent adequately a front in the horizontal, we would need to ensure $\Delta x \ll L$, where Δx is the horizontal grid spacing and L the frontal length scale to be resolved, and this length scale L can become very small across a front. High resolution spanning the model domain is most often very expensive computationally, and it would be preferable to restrict high resolution to the frontal region. This can be achieved by *nesting* approaches, that is, embedding higher resolution models into coarser resolution models where needed (e.g., Barth, Alvera-Azcárate, Rixen & Beckers, 2005; Spall & Holland, 1991). In this case, the abrupt change in grid size at the junction between models may lead to numerical problems and require particular care (Numerical Exercise 15.9). To improve the method, we can allow the grid spacing to vary gradually, rather than suddenly, across the domain.

Such a method based on variable resolution was already suggested when we discussed time discretization (see Fig. 4.10). At that point, we mentioned the problem of a sudden reduction in timescale, which we overcame with shorter time steps during the duration of the event. Our concern now is with space discretization, and we seek a method that uses nonuniform resolution in some optimal way.

We first distribute a series of points x_i according to a known function, say $f(x)$, which may or may not be directly related to one of the variables of the problem, and think that an optimal placement of points is such that, on average, differences in f are similar between adjacent points. In this way, regions of large variations of f will be more densely covered than regions of mild variations. To keep variations constant from point to point, that is

$$|f_{i+1} - f_i| = \text{const}, \tag{15.82}$$

where f_i stands for the known value of f at location x_i, we seek a monotonic coordinate transformation of the type

$$x = x(\xi, t) \tag{15.83}$$

with the variable ξ uniformly distributed while x is not. In terms of the new coordinate ξ, uniform variation in f means

$$\left|\frac{\partial f}{\partial \xi}\right| = \text{const}, \tag{15.84}$$

or, after taking the derivative with respect to ξ,

$$\frac{\partial}{\partial \xi}\left|\frac{\partial f}{\partial \xi}\right| = 0. \tag{15.85}$$

If we express the variations of f in terms of its original and physical variable x, the problem reduces to finding the function $x(\xi)$ that satisfies

$$\frac{\partial}{\partial \xi}\left(\left|\frac{\partial f}{\partial x}\right|\frac{\partial x}{\partial \xi}\right) = 0. \tag{15.86}$$

Since we are interested in discretized problems, we search for the discrete positions x_i that obey

$$\left|\frac{\partial f}{\partial x}\right|_{i+1/2}(x_{i+1} - x_i) - \left|\frac{\partial f}{\partial x}\right|_{i-1/2}(x_i - x_{i-1}) = 0, \tag{15.87}$$

in which we have taken $\Delta\xi = 1$ without loss of generality.

This equation is nonlinear because derivatives of f must be calculated at yet unknown locations. To overcome this quandary, an iterative method is used (see iterative solvers of Section 5.6):

$$x_i^{(k+1)} = x_i^{(k)} + \alpha\Delta t\left[\left|\frac{\partial f}{\partial x}\right|_{i+1/2}\left(x_{i+1}^{(k)} - x_i^{(k)}\right) - \left|\frac{\partial f}{\partial x}\right|_{i-1/2}\left(x_i^{(k)} - x_{i-1}^{(k)}\right)\right], \tag{15.88}$$

in which the superscript (k) is merely an index counting the iterations on the way to the solution, that is, a pseudo-time. If the method converges, $x_i^{(k+1)} = x_i^{(k)}$ eventually, and the vanishing of the bracketed part in Eq. (15.88) tells that the solution has been found.

The preceding iterative method can be interpreted as the numerical solution of a pseudo-evolution equation for the grid nodes:

$$\frac{\partial x}{\partial t} = \alpha\frac{\partial}{\partial \xi}\left(\left|\frac{\partial f}{\partial x}\right|\frac{\partial x}{\partial \xi}\right), \tag{15.89}$$

where the coefficient α is an adjustable numerical parameter that determines how quickly the solution is obtained. If the numerical calculation of the gradients of f proceeds with straightforward centered differences, then the rule of iteration (15.88) reduces to

$$x_i^{(k+1)} = x_i^{(k)} + \alpha \Delta t \left[\left| f_{i+1}^{(k)} - f_i^{(k)} \right| - \left| f_i^{(k)} - f_{i-1}^{(k)} \right| \right], \tag{15.90}$$

in which $f_i^{(k)}$ stands for $f(x_i^{(k)})$. The algorithm is complemented with prescribed values of x on the known boundary positions.

A problem with this formulation appears when the function f is constant over large parts of the domain. By construction, such regions will be void of grid nodes because there are no variations of f there. The remedy is to induce a tendency toward a uniform point distribution where the gradient of f is weak, such as with

$$x_i^{(k+1)} = x_i^{(k)} + \alpha \Delta t \left[w_{i+1/2} \left(x_{i+1}^{(k)} - x_i^{(k)} \right) - w_{i-1/2} \left(x_i^{(k)} - x_{i-1}^{(k)} \right) \right] = 0 \tag{15.91}$$

with the function w replacing $|\partial f / \partial x|$ chosen as

$$w = \left| \frac{\partial f}{\partial x} \right| + w_0. \tag{15.92}$$

In this approach, the parameter w_0 controls the tendency toward a uniform grid distribution. Ideally, its value should fall somewhere between the low and high values of the gradient of f. In this way, wherever the gradient of f is weak, w approaches w_0, and the algorithm leads to solving the equation $\partial^2 x / \partial \xi^2 = 0$, which yields a uniform grid. On the other hand, in places where the gradient of f is steep, w_0 becomes negligible, and we recover Eq. (15.88) that seeds grid points in proportion to the gradient of f. An example is shown in (Fig. 15.15).

The grid positions can thus be obtained by repeated application of a diffusion-type equation (15.91), which is not to be confused with physical diffusion. Here, only positions of grid nodes are calculated. Later, dynamic

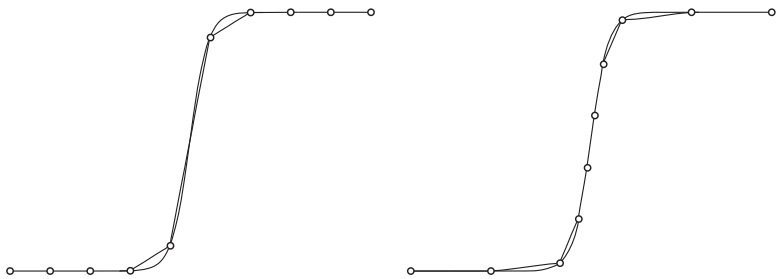


FIGURE 15.15 Grid nodes sampling a strongly varying function, shown by the smooth line. The left panel illustrates uniform sampling and the right panel an adapted grid with higher resolution in the steep region. A linear interpolation between nodes is also shown (line segments).

equations, possibly without any diffusion, may be discretized on this nonuniform grid.

With this adapting technique, we can follow strong gradients as long as the chosen function f effectively mimics the variations that are expected of the solution. Other techniques to distribute grid points nonuniformly exist (e.g., Liseikin, 1999; Thompson, Warsi & Mastin, 1985), always with the objective of reducing some measure of the discretization error.

Spatially nonuniform grids can be used in a frozen or adaptive way. In the frozen version, the grid is generated once at the beginning and then kept unchanged during the remainder of the calculations. This is done when the positions of steep gradients are known in advance, such as those attached to topographic features. Alternatively, it allows the modeler to zoom into a particular region of interest. In the adaptive version, the grid is allowed to move in time, following to the extent possible the dynamically relevant features (e.g., Burchard & Beckers, 2004). The challenge is then to find an effective rule of adaptation, which needs to be reflected in the discretization operators. The modification compared to standard methods on a fixed grid can be illustrated with the one-dimensional tracer equation

$$\frac{\partial c}{\partial t}+\frac{\partial (uc)}{\partial x}=\frac{\partial}{\partial x}\left(\mathcal{A}\frac{\partial c}{\partial x}\right). \tag{15.93}$$

The adaptive grid can be constructed via a coordinate transformation similar to the density coordinate substituting for depth. For the one-dimensional problem, we calculate $c(\xi(x,t),t)$ with ξ as the new coordinate. Since our grid generation provides $x(\xi,t)$, the rules for the transformation follow as in Section 12.1:

$$\frac{\partial \xi}{\partial t}=-\frac{\partial x/\partial t}{\partial x/\partial \xi}, \qquad \frac{\partial \xi}{\partial x}=\frac{1}{\partial x/\partial \xi}. \tag{15.94}$$

The equation for c in the new coordinate system (ξ, t) is then

$$\frac{\partial x}{\partial \xi}\frac{\partial c}{\partial t}-\frac{\partial x}{\partial t}\frac{\partial c}{\partial \xi}+\frac{\partial (uc)}{\partial \xi}=\frac{\partial}{\partial \xi}\left[\mathcal{A}\left(\frac{\partial x}{\partial \xi}\right)^{-1}\frac{\partial c}{\partial \xi}\right]. \tag{15.95}$$

All spatial derivatives with respect to ξ are performed in the new coordinate system, which is uniformly gridded, and standard discretization techniques can be applied. Furthermore, it is advantageous to use a flux form of the equation:

$$\frac{\partial}{\partial t}\left(\frac{\partial x}{\partial \xi}c\right)+\frac{\partial}{\partial \xi}\left[\left(u-\frac{\partial x}{\partial t}\right)c\right]=\frac{\partial}{\partial \xi}\left[\mathcal{A}\left(\frac{\partial x}{\partial \xi}\right)^{-1}\frac{\partial c}{\partial \xi}\right]. \tag{15.96}$$

The factor $\partial x/\partial \xi$ is readily interpreted as the grid spacing in physical space (Δx under the choice of $\Delta \xi = 1$), whereas the term $\partial x/\partial t$ is the velocity at which the grid nodes move. (The partial time derivative in the new coordinate system measures the x displacement per time unit, for fixed ξ.)

In the numerical space of the ξ grid, the advection term involves the velocity difference $u-\partial x/\partial t$, which is the velocity of the flow relative to the moving grid. This is indeed the velocity needed to advect the information relative to the nodes. Should we move the grid with the flow velocity, the relative velocity would be zero, and we would be using a Lagrangian method. The reader may have recognized that a particular case of an adaptive grid is the layered model of Section 12, where the vertical positions of discrete levels are moved in a Lagrangian fashion to follow density interfaces and the vertical velocity disappears from the formalism.

With an adaptive grid, however, the movement of the grid does not necessarily correspond to the flow velocity but must be chosen to follow large gradients. Care must be taken to ensure numerical stability of the scheme because the Courant number now includes the effective velocity, which differs from the actual velocity and may be larger if there are places where the drift speed of the grid is counter to the fluid velocity.

An alternative to performing a change of coordinate is to discretize the equation on a moving grid by directly applying the space integration in physical space between moving grid points. On integrating Eq. (15.93) between the two consecutive moving grid points $x_i(t)$ and $x_{i+1}(t)$, in a way similar to the finite-volume approach (Section 3.9), we can write

$$\int_{x_i(t)}^{x_{i+1}(t)} \frac{\partial c}{\partial t}\mathrm{d}x + q_{i+1} - q_i = 0, \qquad q = uc - \mathcal{A}\frac{\partial c}{\partial x}. \tag{15.97}$$

Our goal is to make explicit the unknown, that is, the grid-averaged concentration, and this requires that we move the time derivative from inside to the outside of the integral. For this, we must be mindful that the integration boundaries vary in time and use Leibniz rule:

$$\frac{\partial}{\partial t}\int_{x_i(t)}^{x_{i+1}(t)} c\,\mathrm{d}x + c(x_i,t)\frac{\partial x_i}{\partial t} - c(x_{i+1},t)\frac{\partial x_{i+1}}{\partial t} + q_{i+1} - q_i = 0. \tag{15.98}$$

Defining the modified flux

$$\hat{q} = \left(u - \frac{\partial x}{\partial t}\right)c - \mathcal{A}\frac{\partial c}{\partial x}, \tag{15.99}$$

the finite-volume equation on the moving grid reads

$$\frac{\partial}{\partial t}\int_{x_i(t)}^{x_{i+1}(t)} c\mathrm{d}x + \hat{q}_{i+1} - \hat{q}_i = 0, \tag{15.100}$$

which is similar to a straightforward finite-volume budget, except for the subtraction of the grid drift velocity $\partial x/\partial t$ from the flow velocity u. Defining $\tilde{c}_i$ as

the cell-averaged concentration and relabeling the grid positions (by using half indices for clarity), the preceding equation can be recast as

$$\frac{\partial}{\partial t}\left[(x_{i+1/2}-x_{i-1/2})\tilde{c}_i\right]+\hat{q}_{i+1/2}-\hat{q}_{i-1/2}=0. \tag{15.101}$$

In this form, the equation is now discrete in space but still continuous in time.

Generalization to three-dimensions proceeds along similar lines, and the outcome is again the subtraction in the advection terms of the grid drift speed from the physical velocity field. Issues in the implementation then concern the placement of nodes with respect to the grid cells (cell boundaries or interval centers in 1D, corners or centers of finite volumes in 2D and 3D; see Numerical Exercises 15.4 and 15.5).

Finally, it is also important to handle correctly the way the changing grid size is discretized in time. As usual, mathematical properties of the original budget equations are not necessarily shared by the numerical operators. In particular, we must ensure that the time discretization of Eq. (15.101) conserves the "volume" $\partial x/\partial \xi$ of the numerical grid in the sense that for a constant c, the time-discretized equation (15.101) is identically satisfied. If it is not, an artificial source of c will appear. This is similar to the advection problem in which the divergence operator of the fluxes has to be consistent with the one used in the physical volume conservation (Section 6.6).

15.7 NONLINEAR ADVECTION SCHEMES

Rather than chasing steep gradients (fronts) with a moving grid, we can also try to capture them with a fixed grid at the cost of appropriate numerical discretization. In this class of methods are the TVD (Total Variation Diminishing) advection schemes mentioned in Section 6.4. It is clear that advection schemes play a crucial role in the context of frontal displacements, and it is no surprise that intense research has been directed toward designing accurate advection schemes.

As we saw in Section 6.4, the basic upwind scheme is monotonic and does not create artificial extrema, but it rapidly smears out strong variations. In contrast, higher-order advection schemes better keep the gradients but at the cost of wiggles (unphysical extrema) in the numerical solution. Several attempts at designing schemes that are monotonic and more accurate than the upwind scheme can be mentioned. Flux-corrected transport (FCT) methods (Boris & Book, 1973; Zalesak, 1979) make two passes on the numerical grid, the first one with an upwind scheme and the second one adding as much anti-diffusion as possible (to restore the steep gradients) without generating wiggles. Flux-limiter methods (Hirsch, 1990; Sweby, 1984), presented hereafter in more detail, degrade the higher-order flux calculations toward upwind fluxes near problem zones. Finally, essentially nonoscillatory (ENO) methods (Harten, Engquist, Osher & Chakravarthy, 1987) adapt different interpolation functions near discontinuities.

The common characteristic of these methods (e.g., Thuburn, 1996) is that they allow the scheme to change its operation depending on the local solution itself. This feedback avoids the annoying consequence of the Godunov theorem (see Section 6.4), which states that the only linear scheme that is monotonic is the first-order upwind scheme, by violating its premise of linearity. Thus, we hope to find a monotonic scheme by introducing some clever nonlinearity in the formulation, even if the underlying physical problem is linear! The general strategy is the following: The nonlinearity is activated whenever over- or under-shooting is likely to occur, in which case the scheme increases numerical diffusion. When the solution is smooth, the scheme is allowed to remain of higher-order to be an improvement over the upwind scheme, of first order.

To design such an adaptive scheme in one dimension, we begin by defining a measure of the variation of the solution called *TV* for *Total Variation*:

$$TV^n = \sum_i |\tilde{c}^n_{i+1} - \tilde{c}^n_i|, \tag{15.102}$$

in which the sum is taken over all grid points of interest. A scheme is said TVD (Total Variation Diminishing) if

$$TV^{n+1} \leq TV^n. \tag{15.103}$$

The *TV* value is meant to be a quantification of the wiggles that appear, such as those arising when the leapfrog or Lax–Wendrofff advection schemes are used.

Suppose now that the numerical scheme can be cast into the following form:

$$\tilde{c}^{n+1}_i = \tilde{c}^n_i - a_{i-1/2}\left(\tilde{c}^n_i - \tilde{c}^n_{i-1}\right) + b_{i+1/2}\left(\tilde{c}^n_{i+1} - \tilde{c}^n_i\right), \tag{15.104}$$

where the coefficients a and b may depend on $\tilde{c}$. We now prove that the so-defined scheme is TVD when

$$0 \leq a_{i+1/2} \quad \text{and} \quad 0 \leq b_{i+1/2} \quad \text{and} \quad a_{i+1/2} + b_{i+1/2} \leq 1. \tag{15.105}$$

Note that $b_{i+1/2}$ appears in combination with $a_{i+1/2}$ in the TVD condition but with $a_{i-1/2}$ in the numerical scheme. Scheme (15.104) can also be written for point $i+1$:

$$\tilde{c}^{n+1}_{i+1} = \tilde{c}^n_{i+1} - a_{i+1/2}\left(\tilde{c}^n_{i+1} - \tilde{c}^n_i\right) + b_{i+3/2}\left(\tilde{c}^n_{i+2} - \tilde{c}^n_{i+1}\right),$$

from which we can subtract Eq. (15.104) to form a marching equation for the variations

$$\begin{aligned}\tilde{c}^{n+1}_{i+1} - \tilde{c}^{n+1}_i = &\left(1 - a_{i+1/2} - b_{i+1/2}\right)\left(\tilde{c}^n_{i+1} - \tilde{c}^n_i\right) \\ &+ b_{i+3/2}\left(\tilde{c}^n_{i+2} - \tilde{c}^n_{i+1}\right) + a_{i-1/2}\left(\tilde{c}^n_i - \tilde{c}^n_{i-1}\right).\end{aligned}$$

We then take the absolute value of each side (remembering that the absolute value of a sum is smaller than the sum of the absolute values of its individual terms), assume that conditions (15.105) are satisfied, and sum over all grid

points to obtain

$$\sum_i \left|\tilde{c}_{i+1}^{n+1} - \tilde{c}_i^{n+1}\right| \le \sum_i \left(1 - a_{i+1/2} - b_{i+1/2}\right) \left|\tilde{c}_{i+1}^n - \tilde{c}_i^n\right| + \sum_i b_{i+3/2} \left|\tilde{c}_{i+2}^n - \tilde{c}_{i+1}^n\right| + \sum_i a_{i-1/2} \left|\tilde{c}_i^n - \tilde{c}_{i-1}^n\right|.$$

Ignoring boundary effects or assuming cyclic conditions, we may shift the index i in the last two sums in order to gather all sums with $|\tilde{c}_{i+1}^n - \tilde{c}_i^n|$. Then, utilizing the TVD conditions (15.105), we have

$$\begin{aligned}\sum_i \left|\tilde{c}_{i+1}^{n+1} - \tilde{c}_i^{n+1}\right| &\le \sum_i \left(1 - a_{i+1/2} - b_{i+1/2}\right) \left|\tilde{c}_{i+1}^n - \tilde{c}_i^n\right| \\ &\quad + \sum_i b_{i+1/2} \left|\tilde{c}_{i+1}^n - \tilde{c}_i^n\right| + \sum_i a_{i+1/2} \left|\tilde{c}_{i+1}^n - \tilde{c}_i^n\right| \\ &\le \sum_i \left|\tilde{c}_{i+1}^n - \tilde{c}_i^n\right|.\end{aligned}$$

The last inequality is none other than Eq. (15.103), proving that discretization (15.104) is TVD if conditions (15.105) are satisfied.

Advection does not increase variance (see Section 6.1), and discretizations with the TVD property appear interesting in this context. Let us therefore design nonlinear TVD schemes for the one-dimensional advection problem with positive velocity u. We combine explicit Euler schemes

$$\tilde{c}_i^{n+1} = \tilde{c}_i^n - \frac{\Delta t}{\Delta x}\left(\tilde{q}_{i+1/2} - \tilde{q}_{i-1/2}\right), \tag{15.106}$$

$$\tilde{q}_{i-1/2} = \tilde{q}_{i-1/2}^L + \Phi_{i-1/2}\left(\tilde{q}_{i-1/2}^H - \tilde{q}_{i-1/2}^L\right), \tag{15.107}$$

where the lower-order flux

$$\tilde{q}_{i-1/2}^L = u\tilde{c}_{i-1}^n$$

is the upwind flux, which is too diffusive, and the higher-order flux

$$\tilde{q}_{i-1/2}^H = u\tilde{c}_{i-1}^n + u\frac{1-\mathrm{C}}{2}\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right)$$

leads to a second-order nonmonotonic scheme (Section 6.4). The weighting factor Φ is allowed to vary with the solution by tending toward zero when excessive variations are present (exploiting the damping properties of the upwind scheme) but otherwise remaining close to 1 for maintaining accuracy where the solution is smooth. In other words, Φ controls the amount of antidiffusion applied to the scheme and bears the name of *flux limiter*. For the following, we assume that Φ

is positive, and the Courant number C satisfies the CFL condition ($C \leq 1$). Then this scheme with weighted flux can be expanded as

$$\begin{aligned}\tilde{c}_i^{n+1} = \tilde{c}_i^n - C\left[\tilde{c}_i^n + \Phi_{i+1/2}\frac{(1-C)}{2}\left(\tilde{c}_{i+1}^n - \tilde{c}_i^n\right)\right] \\ + C\left[\tilde{c}_{i-1}^n + \Phi_{i-1/2}\frac{(1-C)}{2}\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right)\right].\end{aligned} \tag{15.108}$$

It is readily seen that this is not yet a form that ensures TVD because the coefficient $b_{i+1/2}$ multiplying $\tilde{c}_{i+1}^n - \tilde{c}_i^n$ is always negative. However, we can group this term with the upwind part as follows:

$$\tilde{c}_i^{n+1} = \tilde{c}_i^n - C\left[1 - \Phi_{i-1/2}\frac{(1-C)}{2} \quad + \Phi_{i+1/2}\frac{(1-C)}{2}\frac{\left(\tilde{c}_{i+1}^n - \tilde{c}_i^n\right)}{\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right)}\right]\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right).$$

This last form is of the type

$$\tilde{c}_i^{n+1} = \tilde{c}_i^n - a_{i-1/2}\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right) \tag{15.109}$$

even though the coefficient $a_{i-1/2}$ depends on the solution. After all, we are designing a nonlinear method, and this dependence should be no surprise. The scheme can then be made TVD by imposing conditions (15.105), which reduce to $0 \leq a_{i-1/2} \leq 1$, with

$$a_{i-1/2} = C\left[1 - \Phi_{i-1/2}\frac{(1-C)}{2} + \Phi_{i+1/2}\frac{(1-C)}{2}\frac{\left(\tilde{c}_{i+1}^n - \tilde{c}_i^n\right)}{\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right)}\right]. \tag{15.110}$$

We now try to find a simple strategy for specifying $\Phi_{i-1/2}$ and $\Phi_{i+1/2}$ such that the scheme remains TVD in all cases.

The function $\tilde{c}$ appears in parameter $a_{i-1/2}$ in the form of a ratio of differences:

$$r_{i+1/2} = \frac{\tilde{c}_i^n - \tilde{c}_{i-1}^n}{\tilde{c}_{i+1}^n - \tilde{c}_i^n}, \tag{15.111}$$

which is a measure of the variability of $\tilde{c}$: for $r_{i+1/2} = 1$, $\tilde{c}$ varies linearly over the three points involved, whereas for $r_{i+1/2} \leq 0$ there is a local extremum. The parameter $r_{i+1/2}$ will thus be involved in deciding the value of the local weighting Φ to be applied. If $r_{i+1/2}$ is negative (local extremum exists), we require $\Phi_{i+1/2} = 0$ because the local variation on the grid scale would create new extrema if the higher-order scheme were activated.

The TVD condition requires

$$0 \leq C + \frac{C(1-C)}{2}\left[\frac{\Phi_{i+1/2}}{r_{i+1/2}} - \Phi_{i-1/2}\right] \leq 1. \tag{15.112}$$

Ideally, we would like to choose a value of Φ at $i-1/2$ independently of its value at $i+1/2$, otherwise a simultaneous system of equations would have to be solved. To do this, we plan for the worst case, which happens when $\Phi_{i+1/2}=0$. We then need to ensure that $\Phi_{i-1/2}$ is not too large so that $a_{i-1/2}>0$, which implies

$$\Phi_{i-1/2} \leq \frac{2}{(1-\mathrm{C})}. \tag{15.113}$$

Reversing the roles of $\Phi_{i-1/2}$ and $\Phi_{i+1/2}$, the worst case happens when $\Phi_{i-1/2}=0$, demanding that

$$\frac{\Phi_{i+1/2}}{r_{i+1/2}} \leq \frac{2}{\mathrm{C}} \tag{15.114}$$

to ensure that $a_{i-1/2} \leq 1$. Since both conditions must be satisfied for all i values, the following conditions on Φ are required to make a TVD scheme:

$$\Phi \leq \frac{2}{(1-\mathrm{C})} \quad \text{and} \quad \frac{\Phi}{r} \leq \frac{2}{\mathrm{C}}, \tag{15.115}$$

in which the index has become unimportant. In practice, the parameter C varies and inequalities become cumbersome. To circumvent this, we resort again to sufficient conditions built on the worst cases. Since $0 \leq \mathrm{C} \leq 1$, the sufficient conditions to ensure the TVD property are

$$\Phi \leq 2 \quad \text{and} \quad \frac{\Phi}{r} \leq 2. \tag{15.116}$$

Finally, we look for a function $\Phi(r)$ that meets this pair of conditions, and here various choices are open to us. We use this freedom at our advantage by trying to keep the scheme at the highest possible order. If r falls close to 1, the solution is smooth, behaving nearly as a straight line, and a second-order method should do well. So, we impose $\Phi(1)=1$. Incidentally, this is the case for both Lax–Wendroff and Beam–Warming schemes, but these schemes otherwise fail to meet the TVD conditions (Fig. 15.16).

Examples of acceptable limiters are (Fig. 15.16):

- van Leer: $\Phi = \frac{r+|r|}{1+|r|}$
- minmod: $\Phi = \max(0, \min(1, r))$
- Superbee: $\Phi = \max(0, \min(1, 2r), \min(2, r))$
- MC: $\Phi = \max(0, \min(2r, (1+r)/2, 2))$

Note that flux-limiter calculations depend on the direction of the flow, and the ratio r involved in the calculation of Φ must be adapted if the velocity changes sign (Numerical Exercise 15.8).

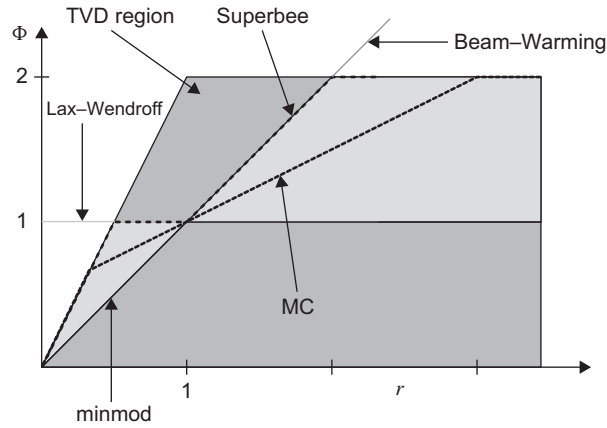


FIGURE 15.16 TVD domain and some standard limiters. The Lax–Wendroff scheme, which corresponds to $\Phi = 1$, has a portion that does not lie within the TVD region (for $r < 0.5$), and likewise the Beam–Warming scheme, with $\Phi = r$, does not lie entirely within the TVD region. The other limiters (minmod, MC, and Superbee) are TVD schemes.

After this lengthy exposition of the design of a TVD scheme, the attentive reader may ask what is the relation between this and a monotonic scheme, which was our original goal. From (15.109), we see that $\tilde{c}_i^{n+1}$ is obtained by linear interpolation of $\tilde{c}_i^n$ and $\tilde{c}_{i-1}^n$ because $0 \le a_{i-1/2} \le 1$. Therefore, $\min(\tilde{c}_i^n, \tilde{c}_{i-1}^n) \le \tilde{c}_i^{n+1} \le \max(\tilde{c}_i^n, \tilde{c}_{i-1}^n)$, and no new local extremum can be created (no over- or under-shooting in the jargon of numerical analysis), and if all values are initially positive, they are guaranteed to remain positive at all times.

This can be verified on the standard test case with the Superbee limiter (Fig. 15.17). The scheme, indeed, keeps the solution within the initial bounds, and results are quite improved compared with previous methods. The method is particularly well suited to GFD applications with large gradients and strong fronts. However, the occasional use of the upwind scheme during parts of the calculations tends to degrade the formal truncation error below second order. Also, absent large gradients when the solution is smooth, fourth-order methods generally outperform second-order TVD schemes. For this reason, fourth-order TVD schemes have been formulated (e.g., Thuburn, 1996). Clearly, the choice of one scheme over another is a question of modeling priorities (conservation, monotonicity, accuracy, ease of implementation, robustness, stability) and expected behavior on the part of the solution (strongly varying, gentle, steady state, etc.).

ANALYTICAL PROBLEMS

15.1. In a certain region, at a certain time, the atmospheric temperature along the ground decreases northward at the rate of 1°C every 35 km, and there are good reasons to assume that this gradient does not change much with

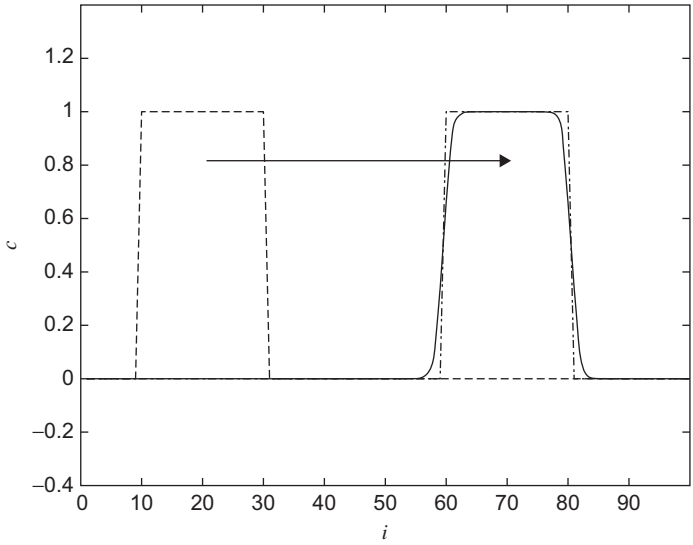


FIGURE 15.17 Advection scheme with TVD Superbee limiter applied to the transport of a "hat" signal with C = 0.5 using 100 time steps. Note that no new extremum is created by the numerical advection and that the diffusion is drastically reduced compared to the upwind scheme.

height. If there is no wind at ground level, what are the wind speed and direction at an altitude of 2 km? To answer, take latitude = 40°N, mean temperature = 290 K, and uniform pressure on the ground.

15.2. A cruise to the Gulf Stream at 38°N provided a cross-section of the current, which was then approximated to a two-layer model (Fig. 15.18) with a warm layer of density $\rho_1 = 1025$ kg/m^3 and depth $h(y) = H - \Delta H \tanh(y/L)$, overlying a colder layer of density $\rho_2 = 1029$ kg/m^3. Taking $H = 500$ m, $\Delta H = 300$ m, and $L = 60$ km and assuming that there is no flow in the lower layer and that the upper layer is in geostrophic balance, determine the flow pattern at the surface. What is the maximum velocity of the Gulf Stream? Where does it occur? Also, compare the jet width (L) to the radius of deformation.

15.3. Derive the discrete Margules relation (15.5) from the governing equations written in the density-coordinate system (Chapter 12).

15.4. Through the Strait of Gibraltar, connecting the Mediterranean Sea to the North Atlantic Ocean, there is an inflow of Atlantic waters near the top and an equal outflow of much more saline Mediterranean waters below. At its narrowest point (Tarifa Narrows), the strait is 11 km wide and 650 m deep. The stratification closely resembles a two-layer configuration with a relative density difference of 0.2% and an interface sloping from 175 m along the Spanish coast (north) to 225 m along the

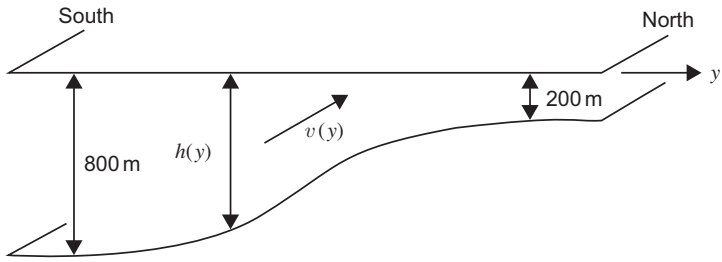


FIGURE 15.18 Schematic cross section of the Gulf Stream, represented as a two-layer geostrophic current (Analytical Problem 15.2).

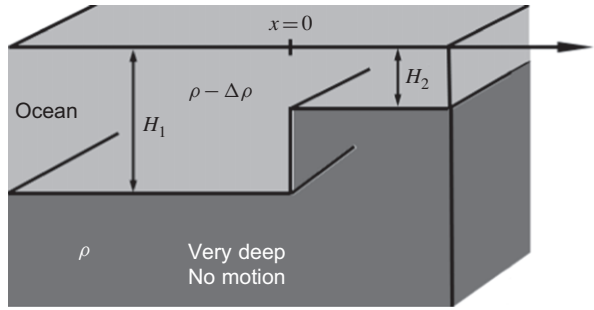


FIGURE 15.19 State prior to geostrophic adjustment (Analytical Problem 15.6).

African coast (south). Taking $f = 8.5 \times 10^{-5}$ s^{-1}, approximating the cross section to a rectangle, and assuming that the volumetric transport in one layer is equal and opposite to that in the other layer, estimate this volumetric transport.

15.5. Determine the geostrophically adjusted state of a band of warm water as depicted in Fig. 15.4c. The variables are $\rho_0 =$ density of water below, $\rho_0 - \Delta\rho =$ density of warm water, $H =$ initial depth of warm water, $2a =$ initial width of warm-water band, and $2b =$ width of warm-water band after adjustment. In particular, determine the value b, and investigate the limits when the initial half-width a is much less and much greater than the deformation radius R.

15.6. Find the solution for the geostrophically adjusted state of the initial configuration shown on Fig. 15.19, and calculate the fraction of potential-energy release that has been converted into kinetic energy of the final steady state.

15.7. In a valley of the French Alps ($\simeq 45°$N), one village (**A**) is situated on a flank 500 m above the valley floor and another (**B**) lies on the opposite side 200 m above the valley floor (Fig. 15.20). The horizontal distance

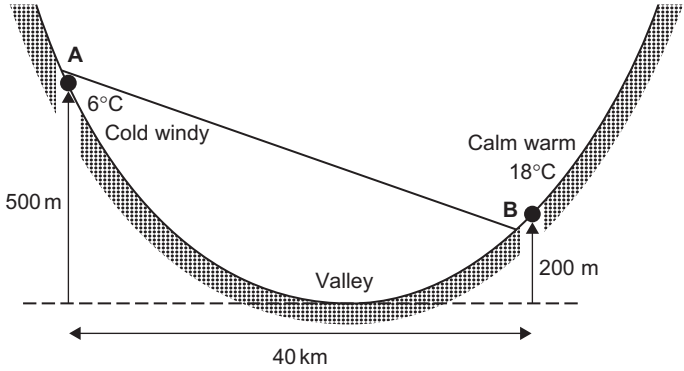


FIGURE 15.20 Air masses in a valley. What is the wind? (Analytical Problem 15.7)

between the two is 40 km. One day, a shepherd in the upper village, who is also a fine meteorologist, notes a cold wind with temperature 6°C. Upon calling her cousin, a blacksmith in the lower village across the valley, she learns that he is enjoying a calm afternoon with a comfortable 18°C! Assuming that the explanation of this perplexing situation resides in a cold wind blowing along one side of the valley (Fig. 15.20), she is able to determine a lower bound for its speed. Can you? Also, in which direction is the wind blowing? (*Hint*: Do not ignore compressibility of air.)

15.8. Using the linearized equations for a two-layer ocean (undisturbed depths H_1 and H_2) over a flat bottom and subject to a spatially uniform wind stress directed along the coast,

$$\frac{\partial u_1}{\partial t} - fv_1 = g'\frac{\partial a}{\partial x} - \frac{1}{\rho_0}\frac{\partial p_2}{\partial x}, \quad \frac{\partial u_2}{\partial t} - fv_2 = -\frac{1}{\rho_0}\frac{\partial p_2}{\partial x}$$
$$\frac{\partial v_1}{\partial t} + fu_1 = \frac{\tau}{\rho_0 H_1}, \quad \frac{\partial v_2}{\partial t} + fu_2 = 0$$
$$-\frac{\partial a}{\partial t} + H_1\frac{\partial u_1}{\partial x} = 0, \quad \frac{\partial a}{\partial t} + H_2\frac{\partial u_2}{\partial x} = 0,$$

study the upwelling response to a wind stress oscillating in time. The boundary conditions are no flow at the coast ($u_1 = u_2 = 0$ at $x = 0$), and no vertical displacements at large distances ($a \to 0$ as $x \to +\infty$). Discuss how the dynamics of the upper layer are affected by the presence of an active lower layer and what happens in the lower layer.

15.9. Investigate under which conditions it is valid to make the assertion made in the text above Eq. (15.37) that the potential vorticity is conserved if the wind stress is spatially uniform.

15.10. A coastal ocean at midlatitude ($f = 10^{-4}$ s^{-1}) has a 50-m-thick warm layer capping a much deeper cold layer. The relative density difference between the two layers is $\Delta\rho/\rho_0 = 0.002$. A uniform wind exerting a surface stress of 0.4 N/m^2 lasts for 3 days. Show that the resulting upwelling includes outcropping of the density interface. What is the offshore distance of the front?

15.11. Generalize to the two-layer ocean, the theory for the steady adjusted state following a wind event of given impulse. For simplicity, consider only the case of equal initial layer depths ($H_1 = H_2$).

15.12. Because of the roughness of the ice, the stress communicated to the water is substantially larger in the presence of sea ice than in the open sea. Assuming that the ice drift is at 20° to the wind, that the water drift is at 90° to the wind (in the open) and to the ice drift (under ice), and that the stress on the water surface is twice as large under ice than in the open sea, determine which wind directions with respect to the ice-edge orientation are favorable to upwelling.

15.13. Show from (15.67) that during frontogenesis with zero potential vorticity, the isopycnal slope

$$S = -\frac{\partial T/\partial y}{\partial T/\partial z} = -\frac{\partial Y/\partial y}{\partial Y/\partial z} = -\frac{f^2}{\alpha g}\frac{\partial Y/\partial y}{\partial T/\partial y} \tag{15.118}$$

evolves according to

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\omega S. \tag{15.119}$$

(*Hint*: Use the fact that $q = 0$.)

NUMERICAL EXERCISES

15.1. By looking into `upwelling.m`, guess which governing equations are being discretized. Then use the program to simulate coastal upwelling and see if the outcropping condition (15.40) is realistic. Examine the algorithm `flooddry.m` used to deal with the outcropping and explore what happens when you deactivate it.

15.2. Add a discretization of momentum advection to `upwelling.m` and redo Numerical Exercise 15.1. Then diagnose

$$I \simeq \frac{1}{\rho_0}\int_{\text{event}} \frac{\tau}{h}\mathrm{d}t \tag{15.120}$$

during the simulation with `upwelling.m` and compare to the estimate (15.35). (*Hint*: Remember that I is calculated for an individual water parcel.)

15.3. Use `adaptive.m` and see how the linearly interpolated function using a uniform or adapted grid approximates the original function `functiontofollow.m`. Quantify the error by sampling the linear interpolations on a very high–resolution grid and calculate the rms error between this interpolation and the original function. Show how this error behaves during the grid-adaptation process.

15.4. Analyze file `adaptiveupwind.m` used to simulate an advection problem with upwind discretization and optional adaptive grid (Fig. 15.21). Explain how the grid size changes and how grid velocities must be discretized in a consistent way. Verify your analysis by using a constant value for c. Modify the parameters involved in the grid adaptation. Try to implement a Lagrangian approach by moving the grid nodes with the physical velocity. What problem appears in a fixed domain?

15.5. Redo Numerical Exercise 15.4 but define and move the grid nodes at the interface and calculate concentration-point positions at the center.

15.6. Prove that the Beam–Warming scheme of Section 6.4 can be recovered using a flux-limiter function $\Phi = r$. Implement the scheme and apply it to the standard problem of the top-hat signal advection. How does the solution compare to the Lax–Wendroff solution of Fig. 6.9?

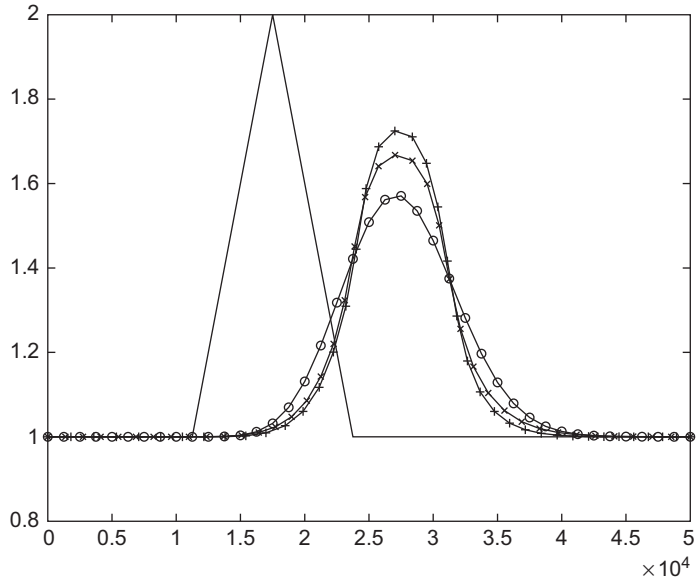


FIGURE 15.21 Upwind advection of a triangular distribution with fixed grid (most diffusive solution) and an adaptive grid with or without added Lagrangian-type advection. The best solution is obtained with the Lagrangian-type advection. See Numerical Exercise 15.4.

15.7. Apply the Superbee TVD scheme to the advection of the top-hat signal and then to the advection of a sinusoidal signal. What do you observe? Can you explain the behavior and verify that your explanation is correct by choosing another limiter? Experiment with `tvdadv1D.m`.

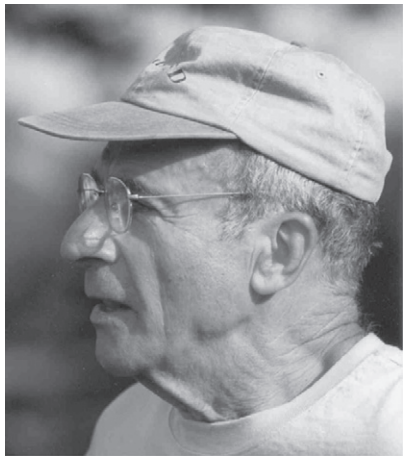
15.8. Write out the flux-limiter scheme in the 1D case for $u \leq 0$ by exploiting symmetries of the problem.

15.9. Implement a leapfrog advection scheme on a nonuniform grid with scalar c defined at the center of the cells of variable width. The domain of interest spans $x = -10L$ to $x = 10L$. For initialization and downstream boundaries, use an Euler upwind scheme. Use a spatial grid spacing of Δx for $x < 0$ and $r\Delta x$ for $x > 0$. Use a constant velocity $u = 1$ m/s. Simulate a time window of $15L/u$ and show the solution at every time step. Use a maximum value of $C = 0.5$. Advect a Gaussian distribution of width L initially located at $x = -5L$:

$$c(x, t = 0) = \exp\left[-(x + 5L)^2/L^2\right]. \tag{15.121}$$

Use all combinations of $\Delta x = L/4, L/8, L/16$ with $r = 1, 1/2, 2, 1/10, 10$. In particular, observe what happens when the patch crosses $x = 0$.

George Veronis
1926–

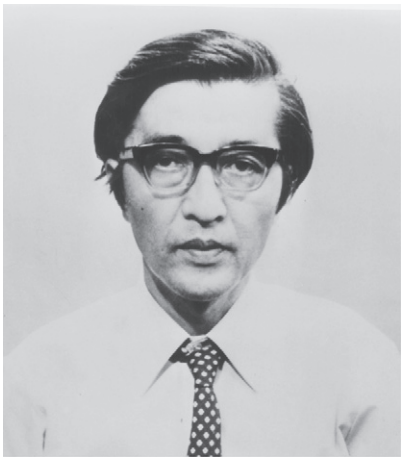


An applied mathematician turned oceanographer, George Veronis has been a driving force in geophysical fluid dynamics since its early days. With Willem V. R. Malkus, he cofounded the GFD Summer Program at the Woods Hole Oceanographic Institution, which continues after more than 50 years to bring together oceanographers, meteorologists, physicists, and mathematicians to debate problems related to geophysical flows.

Veronis is best known for his theoretical studies on oceanic circulation, rotating and stratified fluids, thermal convection with and without rotation, and double-diffusion processes. His model of the circulation of the world ocean was an analytical study based on planetary geostrophic dynamics and the nonlinearity of thermal processes, in which he showed how western boundary currents cross the boundaries of wind gyres and connect all of the world's oceans into a single circulating system.

Veronis has earned a reputation as a superb lecturer, who can explain difficult concepts with amazing ease and clarity. (*Photo credit: G. Veronis*)

Kozo Yoshida
1922–1978



In the early years of his professional career, Kozo Yoshida studied long (tsunami) and short (wind) waves. Later, during a stay at the Scripps Institution of Oceanography, he turned to the investigation of the upwelling phenomenon, which was to become his lifelong interest. His formulations of dynamic theories for both coastal and equatorial upwelling earned him respect and fame. A wind-driven surface eastward current along the equator is called a Yoshida jet. In his later years, he also became interested in the Kuroshio, a major ocean current off the coast of Japan, wrote several books, and promoted oceanography among young scientists.

Known to be very sincere and logical, Yoshida did not shun administrative responsibilities and emphasized the importance of international cooperation in postwar Japan. (*Photo courtesy of Toshio Yamagata*)