Barotropic Waves

ABSTRACT

The aim of this chapter is to describe an assortment of waves that can be supported by an inviscid, homogeneous fluid in rotation and to analyze numerical grid arrangements that facilitate the simulation of wave propagation, in particular for the prediction of tides and storm surges.

9.1 LINEAR WAVE DYNAMICS

Chiefly because linear equations are most amenable to methods of solution, it is wise to gain insight into geophysical fluid dynamics by elucidating the possible linear processes and investigating their properties before exploring more intricate, nonlinear dynamics. The governing equations of the previous section are essentially nonlinear; consequently, their linearization can proceed only by imposing restrictions on the flows under consideration.

The Coriolis acceleration terms present in the momentum equations [(4.21a) and (4.21b)] are, by nature, linear and need not be subjected to any approximation. This situation is extremely fortunate because these are the central terms of geophysical fluid dynamics. In contrast, the so-called advective terms (or convective terms) are quadratic and undesirable at this moment. Hence, our considerations will be restricted to low-Rossby-number situations:

$$Ro = \frac{U}{\Omega L} \ll 1. \qquad (9.1)$$

This is usually accomplished by restricting the attention to relatively weak flows (small U), large scales (large L), or, in the laboratory, fast rotation (large Ω). The terms expressing the local time rate of change of the velocity ($\partial u/\partial t$ and $\partial v/\partial t$) are linear and are retained here in order to permit the investigation of unsteady flows. Thus, the temporal Rossby number is assumed to be on the order of unity:

$$Ro_T = \frac{1}{\Omega T} \sim 1. \qquad (9.2)$$

Contrasting conditions (9.1) and (9.2), we conclude that we are about to consider slow flow fields that evolve relatively fast. Aren't we asking for the impossible? Not at all, for rapidly moving disturbances do not necessarily require large

velocities. In other words, information may travel faster than material particles, and when this is the case, the flow takes the aspect of a wave field. A typical example is the spreading of concentric ripples on the surface of a pond after throwing a stone; energy radiates but there is no appreciable water movement across the pond. In keeping with the foregoing quantities, a scale C for the wave speed (or celerity) can be defined as the velocity of a signal covering the distance L of the flow during the nominal evolution time T, and, by virtue of restrictions (9.1) and (9.2), it can be compared with the flow velocity:

$$C = \frac{L}{T} \sim \Omega L \gg U. \tag{9.3}$$

Thus, our present objective is to consider wave phenomena.

To shed the best possible light on the mechanisms of the basic wave processes typical in geophysical flows, we further restrict our attention to homogeneous and inviscid flows, for which the shallow-water model (Section 7.3) is adequate. With all the preceding restrictions, the horizontal momentum equations (7.9a) and (7.9b) reduce to

$$\frac{\partial u}{\partial t} - fv = - g\frac{\partial \eta}{\partial x} \tag{9.4a}$$

$$\frac{\partial v}{\partial t} + fu = - g\frac{\partial \eta}{\partial y}, \tag{9.4b}$$

where f is the Coriolis parameter, g is the gravitational acceleration, u and v are the velocity components in the x- and y-directions, respectively, and η is the surface displacement (equal to $\eta = h - H$, the total fluid depth h minus the mean fluid thickness H). The independent variables are x, y, and t; the vertical coordinate is absent, for the flow is vertically homogeneous (Section 7.3).

In terms of surface height, η, the continuity equation (7.14) can be expanded in several groups of terms:

$$\frac{\partial \eta}{\partial t} + \left(u\frac{\partial \eta}{\partial x} + v\frac{\partial \eta}{\partial y} \right) + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + \eta\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0$$

if the mean depth H is constant (flat bottom). Introducing the scale ΔH for the vertical displacement η of the surface, we note that the four groups of terms in the preceding equation are, sequentially, on the order of

$$\frac{\Delta H}{T}, \quad U\frac{\Delta H}{L}, \quad H\frac{U}{L}, \quad \Delta H\frac{U}{L}.$$

According to Eq. (9.3) L/T is much larger than U, and the second and fourth groups of terms may be neglected compared with the first term, leaving us with the linearized equation

$$\frac{\partial \eta}{\partial t} + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0, \tag{9.5}$$

the balance of which requires $\Delta H/T$ to be on the order of UH/L or, again by virtue of (9.3),

$$\Delta H \ll H. \tag{9.6}$$

We are thus restricted to waves of small amplitudes.

The system of Eqs. (9.4a) through (9.5) governs the linear wave dynamics of inviscid, homogeneous fluids under rotation. For the sake of simple notation, we will perform the mathematical derivations only for positive values of the Coriolis parameter f and then state the conclusions for both positive and negative values of f. The derivations with negative values of f are left as exercises. Before proceeding with the separate studies of geophysical fluid waves, the reader not familiar with the concepts of phase speed, wavenumber vector, dispersion relation, and group velocity is directed to Appendix B. A comprehensive account of geophysical waves can be found in the book by LeBlond and Mysak (1978), with additional considerations on nonlinearities in Pedlosky (2003).

9.2 THE KELVIN WAVE

The Kelvin wave is a traveling disturbance that requires the support of a lateral boundary. Therefore, it most often occurs in the ocean where it can travel along coastlines. For convenience, we use oceanic terminology such as coast and offshore.

As a simple model, consider a layer of fluid bounded below by a horizontal bottom, above by a free surface, and on one side (say, the y-axis) by a vertical wall (Fig. 9.1). Along this wall ($x=0$, the coast), the normal velocity
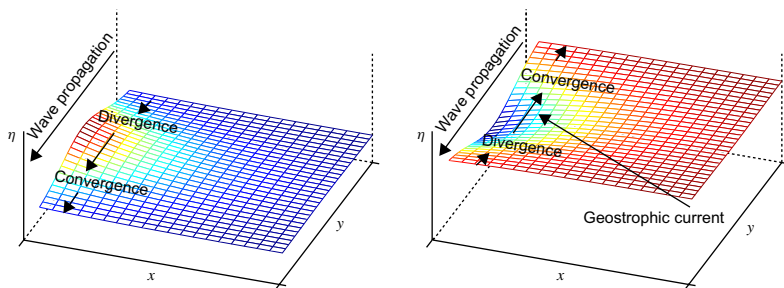

FIGURE 9.1 Upwelling and downwelling Kelvin waves. In the northern hemisphere, both waves travel with the coast on their right, but the accompanying currents differ. Geostrophic equilibrium in the $x-$momentum equation leads to a velocity v that is maximum at the bulge and directed as the geostrophic equilibrium requires. Because of the different geostrophic velocities at the bulge and further away, convergence and divergence patterns create a lifting or lowering of the surface. The lifting and lowering are such that the wave propagates toward negative y in either case (positive or negative bulge).

must vanish ($u=0$), but the absence of viscosity allows a nonzero tangential velocity.

As he recounted in his presentation to the Royal Society of Edinburgh in 1879, Sir William Thomson (later to become Lord Kelvin) thought that the vanishing of the velocity component normal to the wall suggested the possibility that it be zero everywhere. So, let us state, in anticipation,

$$u=0 \tag{9.7}$$

throughout the domain and investigate the consequences. Although Eq. (9.4a) contains a remaining derivative with respect to x, Eqs. (9.4b) and (9.5) contain only derivatives with respect to y and time. Elimination of the surface elevation leads to a single equation for the alongshore velocity:

$$\frac{\partial^2 v}{\partial t^2} = c^2 \frac{\partial^2 v}{\partial y^2}, \tag{9.8}$$

where

$$c = \sqrt{gH} \tag{9.9}$$

is identified as the speed of surface gravity waves in nonrotating shallow waters.

The preceding equation governs the propagation of one-dimensional nondispersive waves and possesses the general solution

$$v = V_1(x, \ y+ct) + V_2(x, \ y-ct), \tag{9.10}$$

which consists of two waves, one traveling toward decreasing y and the other in the opposite direction. Returning to either Eq. (9.4b) or (9.5) where u is set to zero, we easily determine the surface displacement:

$$\eta = -\sqrt{\frac{H}{g}} \, V_1(x, \ y+ct) + \sqrt{\frac{H}{g}} \, V_2(x, \ y-ct). \tag{9.11}$$

(Any additive constant can be eliminated by a proper redefinition of the mean depth H.) The structure of the functions V_1 and V_2 is then determined by the use of the remaining equation, i.e., (9.4a):

$$\frac{\partial V_1}{\partial x} = -\frac{f}{\sqrt{gH}} V_1, \quad \frac{\partial V_2}{\partial x} = +\frac{f}{\sqrt{gH}} V_2$$

or

$$V_1 = V_{10}(y+ct) \, e^{-x/R}, \quad V_2 = V_{20}(y-ct) \, e^{+x/R},$$

where the length R, defined as

$$R = \frac{\sqrt{gH}}{f} = \frac{c}{f}, \tag{9.12}$$

combines all three constants of the problem. Within a numerical factor, it is the distance covered by a wave, such as the present one, traveling at the speed c during one inertial period $(2\pi/f)$. For reasons that will become apparent later, this quantity is called the *Rossby radius of deformation* or, more simply, the radius of deformation.

Of the two independent solutions, the second increases exponentially with distance from shore and is physically unfit. This leaves the other as the most general solution:

$$u = 0 \tag{9.13a}$$

$$v = \sqrt{gH}F(y+ct)\,e^{-x/R} \tag{9.13b}$$

$$\eta = -HF(y+ct)\,e^{-x/R}, \tag{9.13c}$$

where F is an arbitrary function of its variable.

Because of the exponential decay away from the boundary, the Kelvin wave is said to be trapped. Without the boundary, it is unbounded at large distances and thus cannot exist; the length R is a measure of the trapping distance. In the longshore direction, the wave travels without distortion at the speed of surface gravity waves. In the northern hemisphere ($f > 0$, as in the preceding analysis), the wave travels with the coast on its right; in the southern hemisphere, the wave travels with the coast on its left. Note that, although the direction of wave propagation is unique, the sign of the longshore velocity is arbitrary: An upwelling wave (i.e., a surface bulge with $\eta > 0$) has a current flowing in the direction of the wave, whereas a downwelling wave (i.e., a surface trough with $\eta < 0$) is accompanied by a current flowing in the direction opposite to that of the wave (Fig. 9.1).

In the limit of no rotation ($f \to 0$), the trapping distance increases without bound, and the wave reduces to a simple gravity wave with crests and troughs oriented perpendicularly to the coast.

Surface Kelvin waves (as described previously, and to be distinguished from internal Kelvin waves, which require a stratification, see the end of Chapter 13) are generated by the ocean tides and by local wind effects in coastal areas. For example, a storm off the northeast coast of Great Britain can send a Kelvin wave that follows the shores of the North Sea in a counterclockwise direction and eventually reaches the west coast of Norway. Traveling in approximately 40 m of water and over a distance of 2200 km, it accomplishes its journey in about 31 h.

The decay of the Kelvin wave amplitude away from the coast is clearly manifested in the English Channel. The North Atlantic tide enters the Channel from the west and travels eastward toward the North Sea (Fig. 9.2). Being essentially a surface wave in a rotating fluid bounded by a coast, the tide assumes the character of a Kelvin wave and propagates while leaning against a coast on its right, namely, France. This partly explains why tides are noticeably higher along the French coast than along the British coast a few tens of kilometers across (Fig. 9.2).
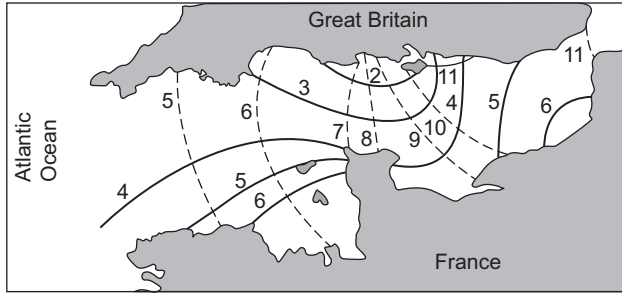
FIGURE 9.2 Cotidal lines (dashed) with time in lunar hours for the M2 tide in the English Channel showing the eastward progression of the tide from the North Atlantic Ocean. Lines of equal tidal range (solid, with value in meters) reveal larger amplitudes along the French coast, namely to the right of the wave progression in accordance with Kelvin waves. (*From Proudman, 1953, as adapted by Gill, 1982*)

9.3   INERTIA-GRAVITY WAVES (POINCARÉ WAVES)

Let us now do away with the lateral boundary and relax the stipulation $u = 0$. The system of Eqs. (9.4a) through (9.5) is kept in its entirety. With f constant and in the presence of a flat bottom, all coefficients are constant, and a Fourier-mode solution can be sought. With u, v, and η taken as constant factors times a periodic function

$$\begin{pmatrix} \eta \\ u \\ v \end{pmatrix} = \Re \begin{pmatrix} A \\ U \\ V \end{pmatrix} e^{i(k_x x + k_y y - \omega t)} \tag{9.14}$$

where the symbol $\Re$ indicates the real part of what follows, k_x and k_y are the wavenumbers in the x-and y-directions, respectively, and ω is a frequency, the system of equations becomes algebraic:

$$-i\omega U - fV = -i g k_x A \tag{9.15a}$$

$$-i\omega V + fU = -i g k_y A \tag{9.15b}$$

$$-i\omega A + H(i k_x U + i k_y V) = 0. \tag{9.15c}$$

This system admits the trivial solution $U = V = A = 0$ unless its determinant vanishes. Thus, waves occur only when the following condition is met:

$$\omega[\omega^2 - f^2 - gH\,(k_x^2 + k_y^2)] = 0. \tag{9.16}$$

This condition, called the *dispersion relation*, provides the wave frequency in terms of the wavenumber magnitude $k = (k_x^2 + k_y^2)^{1/2}$ and the constants of the problem. The first root, $\omega = 0$, corresponds to a steady geostrophic state. Returning to Eqs. (9.4a) through (9.5) with the time derivatives set to zero, we recognize the equations governing the geostrophic flow described in Section 7.1.

In other words, geostrophic flows can be interpreted as arrested waves of any wavelength. If the bottom were not flat, these "waves" would cease to exist and be replaced by Taylor columns.

The remaining two roots,

$$\omega = \sqrt{f^2 + gHk^2} \qquad (9.17)$$

and its opposite, correspond to bona fide traveling waves, called *Poincaré waves*, whose frequency is always superinertial. In the limit of no rotation ($f = 0$), the frequency is $\omega = k\sqrt{gH}$, and the phase speed is $c = \omega/k = \sqrt{gH}$. The waves become classical shallow-water gravity waves. The same limit also occurs at large wavenumbers [$k^2 \gg f^2/gH$, i.e., wavelengths much shorter than the deformation radius defined in Eq. (9.12)]. This is not too surprising, since such waves are too short and too fast to feel the rotation of the earth.

At the opposite extreme of low wavenumbers ($k^2 \ll f^2/gH$, i.e., wavelengths much longer than the deformation radius), the rotation effect dominates, yielding $\omega \simeq f$. At this limit, the flow pattern is virtually laterally uniform, and all fluid particles move in unison, each describing a circular inertial oscillation, as described in Section 2.3. For intermediate wavenumbers, the frequency (Fig. 9.3) is always greater than f, and the waves are said to be *superinertial*. Since Poincaré waves exhibit a mixed behavior between gravity waves and inertial oscillations, they are often called *inertia-gravity waves*.

Because the phase speed $c = \omega/k$ depends on the wavenumber, wave components of different wavelengths travel at different speeds, and the wave is said to be *dispersive*. This is in contrast with the nondispersive Kelvin wave, the signal
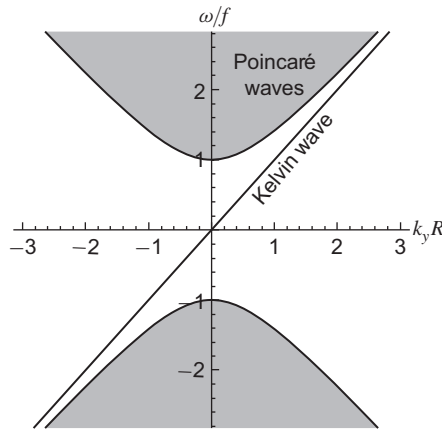

FIGURE 9.3   Recapitulation of the dispersion relation of Kelvin and Poincaré waves on the f-plane and on a flat bottom. Although Poincaré waves (gray shades) can travel in all directions and occupy therefore a continuous spectrum in terms of k_y, the Kelvin wave (diagonal line) propagates only along a boundary.

of which travels without distortion, irrespective of its profile. See Appendix B for additional information on these notions.

Seiches, tides, and tsunamis are examples of barotropic gravity waves. A *seiche* is a standing wave, formed by the superposition of two waves of equal wavelength and propagating in opposite directions due to reflection on lateral boundaries. Seiches occur in confined water bodies, such as lakes, gulfs, and semi-enclosed seas. In the Adriatic Sea, the untimely superposition of a wind-generated seiche with high tide can cause flooding in Venice (Robinson, Tomasin & Artegiani, 1973).

A *tsunami* is a wave triggered by an underwater earthquake. With wavelengths ranging from tens to hundreds of kilometers, tsunamis are barotropic waves, but their relatively high frequency (period of a few minutes) makes them only slightly affected by the Coriolis force. What makes tsunamis disastrous is the gradual amplification of their amplitude as they enter shallower waters, so that what may begin as an innocuous 1m wave in the middle of the ocean, which a ship hardly notices, can turn into a catastrophic multimeter surge on the beach. Disastrous tsunamis occurred in the Pacific Ocean on 22 May 1960, in the Indian Ocean on 26 December 2004, and in the Pacific Ocean off Japan on 11 March 2011. Tsunami propagation is relatively easy to forecast with computer models. The key to an effective warning system is the early detection of the originating earthquake to track the rapid propagation (at speed $\sqrt{gH}$) of the tsunami from point of origin to the coastline on time to issue a warning before the high wave strikes.

Before concluding this section, a note is in order to warn about the possibility of violating the hydrostatic assumption. Indeed, at short wavelengths (on the order of the fluid depth or shorter), the frequency is high (period much shorter than $2\pi/f$), and the vertical acceleration (equal to $\partial^2\eta/\partial t^2$ at the surface) becomes comparable with the gravitational acceleration g. When this is the case, the hydrostatic approximation breaks down, the assumption of vertical rigidity may no longer be invoked, and the problem becomes three dimensional. For a study of nonhydrostatic gravity waves, the reader is referred to Section 10 of LeBlond and Mysak (1978) and Lecture 3 of Pedlosky (2003).

9.4 PLANETARY WAVES (ROSSBY WAVES)

Kelvin and Poincaré waves are relatively fast waves, and we may wonder whether rotating, homogeneous fluids could not support another breed of slower waves. Could it be, for example, that the steady geostrophic flows, those corresponding to the zero frequency solution found in the preceding section, may develop a slow evolution (frequency slightly above zero) when the system is modified somewhat? The answer is yes, and one such class consists of planetary waves, in which the time evolution originates in the weak but important *planetary effect*.

As we may recall from Section 2.5, on a spherical earth (or planet or star, in general), the Coriolis parameter, f, is proportional to the rotation rate, Ω, times

the sine of the latitude, φ:

$$f = 2\Omega \sin\varphi.$$

Large wave formations such as alternating cyclones and anticyclones contribute to our daily weather and, to a lesser extent, Gulf Stream meanders span several degrees of latitude; for them, it is necessary to consider the meridional change in the Coriolis parameter. If the coordinate y is directed northward and is measured from a reference latitude φ_0 (say, a latitude somewhere in the middle of the wave under consideration), then $\varphi = \varphi_0 + y/a$, where a is the earth's radius (6371 km). Considering y/a as a small perturbation, the Coriolis parameter can be expanded in a Taylor series:

$$f = 2\Omega \sin\varphi_0 + 2\Omega \frac{y}{a} \cos\varphi_0 + \cdots \tag{9.18}$$

Retaining only the first two terms, we write

$$f = f_0 + \beta_0 y, \tag{9.19}$$

where $f_0 = 2\Omega \sin\varphi_0$ is the reference Coriolis parameter, and $\beta_0 = 2(\Omega/a)\cos\varphi_0$ is the so-called *beta parameter*. Typical midlatitude values on earth are $f_0 = 8 \times 10^{-5}\,\text{s}^{-1}$ and $\beta_0 = 2 \times 10^{-11}\,\text{m}^{-1}\text{s}^{-1}$. The Cartesian framework where the beta term is not retained is called the *f-plane*, and that where it is retained is called the *beta plane*. The next step in order of accuracy is to retain the full spherical geometry (which we avoid throughout this book). Rigorous justifications of the beta-plane approximation can be found in Veronis (1963, 1981), Pedlosky (1987), and Verkley (1990).

Note that the beta-plane representation is validated at midlatitudes only if the $\beta_0 y$ term is small compared with the leading f_0 term. For the motion's meridional length scale L, this implies

$$\beta = \frac{\beta_0 L}{f_0} \ll 1, \tag{9.20}$$

where the dimensionless ratio β can be called the *planetary number*.

The governing equations, having become

$$\frac{\partial u}{\partial t} - (f_0 + \beta_0 y)v = -g\frac{\partial \eta}{\partial x} \tag{9.21a}$$

$$\frac{\partial v}{\partial t} + (f_0 + \beta_0 y)u = -g\frac{\partial \eta}{\partial y} \tag{9.21b}$$

$$\frac{\partial \eta}{\partial t} + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0, \tag{9.21c}$$

are now mixtures of small and large terms. The larger ones (f_0, g, and H terms) comprise the otherwise steady, f-plane geostrophic dynamics; the smaller ones (time derivatives and β_0 terms) come as perturbations, which, although small,

will govern the wave evolution. In first approximation, the large terms dominate, and thus $u \simeq -(g/f_0)\partial\eta/\partial y$ and $v \simeq +(g/f_0)\partial\eta/\partial x$. Use of this first approximation in the small terms of Eqs. (9.21a) and (9.21b) yields

$$-\frac{g}{f_0}\frac{\partial^2\eta}{\partial y\partial t} -f_0 v -\frac{\beta_0 g}{f_0} y\frac{\partial\eta}{\partial x} = -g\frac{\partial\eta}{\partial x} \qquad (9.22)$$

$$+\frac{g}{f_0}\frac{\partial^2\eta}{\partial x\partial t} +f_0 u -\frac{\beta_0 g}{f_0} y\frac{\partial\eta}{\partial y} = -g\frac{\partial\eta}{\partial y}. \qquad (9.23)$$

These equations are trivial to solve with respect to u and v:

$$u = -\frac{g}{f_0}\frac{\partial\eta}{\partial y} -\frac{g}{f_0^2}\frac{\partial^2\eta}{\partial x\partial t} +\frac{\beta_0 g}{f_0^2} y\frac{\partial\eta}{\partial y} \qquad (9.24)$$

$$v = +\frac{g}{f_0}\frac{\partial\eta}{\partial x} -\frac{g}{f_0^2}\frac{\partial^2\eta}{\partial y\partial t} -\frac{\beta_0 g}{f_0^2} y\frac{\partial\eta}{\partial x}. \qquad (9.25)$$

These last expressions can be interpreted as consisting of the leading and first-correction terms in a regular perturbation series of the velocity field. We identify the first term of each expansion as the geostrophic velocity. By contrast, the next and smaller terms are called *ageostrophic*.

Substitution in continuity equation (9.21c) leads to a single equation for the surface displacement:

$$\frac{\partial\eta}{\partial t} -R^2\frac{\partial}{\partial t}\nabla^2\eta -\beta_0 R^2\frac{\partial\eta}{\partial x} = 0, \qquad (9.26)$$

where ∇^2 is the two-dimensional Laplace operator, and $R = \sqrt{gH}/f_0$ is the deformation radius, defined in Eq. (9.12) but now suitably amended to be a constant. Unlike the original set of equations, this last equation has constant coefficients and a solution of the Fourier type, $\cos(k_x x + k_y y - \omega t)$, can be sought. The dispersion relation follows:

$$\omega = -\beta_0 R^2 \frac{k_x}{1+R^2 (k_x^2 + k_y^2)}, \qquad (9.27)$$

providing the frequency ω as a function of the wavenumber components k_x and k_y. The waves are called *planetary waves* or *Rossby waves*, in honor of Carl-Gustaf Rossby, who first proposed this wave theory to explain the systematic movement of midlatitude weather patterns. We note immediately that if the beta corrections had not been retained ($\beta_0 = 0$), the frequency would have been nil. This is the $\omega = 0$ solution of Section 9.3, which corresponds to a steady geostrophic flow on the f-plane. The absence of the other two roots is explained by our approximation. Indeed, treating the time derivatives as small terms (i.e., having in effect assumed a very small temporal Rossby number, $Ro_T \ll 1$), we have retained only the low frequency, the one much less than f_0. In the parlance of wave dynamics, this is called *filtering*.

That the frequency given by Eq. (9.27) is indeed small can be verified easily. With L $(\sim 1/k_x \sim 1/k_y)$ as a measure of the wavelength, two cases can arise either $L \lesssim R$ or $L \gtrsim R$; the frequency scale is then given by

$$\text{Shorter waves}: \quad L \lesssim R, \quad \omega \sim \beta_0 L \tag{9.28}$$

$$\text{Longer waves}: \quad L \gtrsim R, \quad \omega \sim \frac{\beta_0 R^2}{L} \lesssim \beta_0 L. \tag{9.29}$$

In either case, our premise (9.20) that $\beta_0 L$ is much less than f_0 implies that ω is much smaller than f_0 (subinertial wave), as we anticipated.

Let us now explore other properties of planetary waves. First and foremost, the zonal phase speed

$$c_x = \frac{\omega}{k_x} = \frac{-\beta_0 R^2}{1 + R^2 \, (k_x^2 + k_y^2)} \tag{9.30}$$

is always negative, implying a phase propagation to the west (Fig. 9.4). The sign of the meridional phase speed $c_y = \omega/k_y$ is undetermined, since the wavenumber k_y may have either sign. Thus, planetary waves can propagate only northwestward, westward, or southwestward. Second, very long waves ($1/k_x$ and $1/k_y$ both much larger than R) propagate always westward and at the speed

$$c = -\beta_0 R^2, \tag{9.31}$$
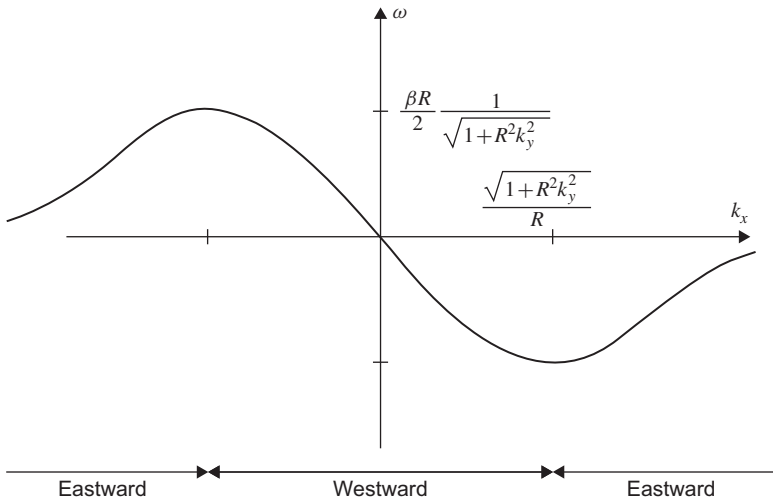
which is the largest wave speed allowed.

FIGURE 9.4   Dispersion relation of planetary (Rossby) waves. The frequency ω is plotted against the zonal wavenumber k_x at constant meridional wavenumber k_y. As the slope of the curve reverses, so does the direction of zonal propagation of energy.

Lines of constant frequency ω in the (k_x, k_y) wavenumber space are circles defined by

$$\left(k_x + \frac{\beta_0}{2\omega} \right)^2 + k_y^2 = \left(\frac{\beta_0^2}{4\omega^2} - \frac{1}{R^2} \right), \tag{9.32}$$

and are illustrated in Fig. 9.5. Such circles exist only if their radius is a real number, that is, if $\beta_0^2 > 4\omega^2/R^2$. This implies the existence of a maximum frequency

$$|\omega|_{\max} = \frac{\beta_0 R}{2}, \tag{9.33}$$

beyond which planetary waves do not exist.

The group velocity, at which the energy of a wave packet propagates, defined as the vector $(\partial\omega/\partial k_x, \partial\omega/\partial k_y)$, is the gradient of the function ω in the (k_x, k_y) wavenumber plane (see Appendix B). It is thus perpendicular to the circles of constant ω. A little algebra reveals that the group-velocity vector is directed inward, toward the center of the circle. Therefore, long waves (small k_x and k_y, points near the origin in Fig. 9.5) have westward group velocities, whereas energy is carried eastward by the shorter waves (larger k_x and k_y, points on the
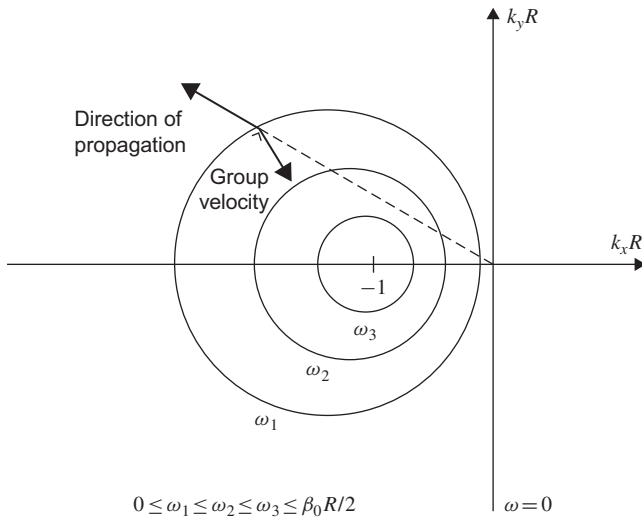

FIGURE 9.5 Geometric representation of the planetary-wave dispersion relation. Each circle corresponds to a fixed frequency, with frequency increasing with decreasing radius. The group velocity of the (k_x, k_y) wave is a vector perpendicular to the circle at point (k_x, k_y) and directed toward its center.

opposite side of the circle). This dichotomy is also apparent in Fig. 9.4, which exhibits reversals in slope ($\partial\omega/\partial k_x$ changing sign).

9.5   TOPOGRAPHIC WAVES

Just as small variations in the Coriolis parameter can turn a steady geostrophic flow into slowly moving planetary waves, so can a weak bottom irregularity. Admittedly, topographic variations can come in a great variety of sizes and shapes, but for the sake of illustrating the wave process in its simplest form, we limit ourselves here to the case of a weak and uniform bottom slope. We also return to the use of a constant Coriolis parameter. This latter choice allows us to choose convenient directions for the reference axes, and, in anticipation of an analogy with planetary waves, we align the y-axis with the direction of the topographic gradient. We thus express the depth of the fluid at rest as:

$$H = H_0 + \alpha_0 y, \tag{9.34}$$

where H_0 is a mean reference depth, and α_0 is the bottom slope, which is required to be gentle so that

$$\alpha = \frac{\alpha_0 L}{H_0} \ll 1, \tag{9.35}$$

where L is the horizontal length scale of the motion. The topographic parameter α plays a role similar to the planetary number, defined in Eq. (9.20).

The bottom slope gives rise to new terms in the continuity equation. Starting with the continuity equation (7.14) for shallow water and expressing the instantaneous fluid layer depth as (Fig. 9.6)

$$h(x, \ y, \ t) = H_0 + \alpha_0 y + \eta(x, \ y, \ t), \tag{9.36}$$

we obtain

$$\frac{\partial \eta}{\partial t} + \left(u\frac{\partial \eta}{\partial x} + v\frac{\partial \eta}{\partial y} \right) + (H_0 + \alpha_0 y)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)$$
$$+ \eta\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + \alpha_0 v = 0.$$

Once again, we strike out the nonlinear terms by invoking a very small Rossby number (much smaller than the temporal Rossby number) for the sake of linear dynamics. The term $\alpha_0 y$ can also be dropped next to H_0 by virtue of Eq. (9.35).
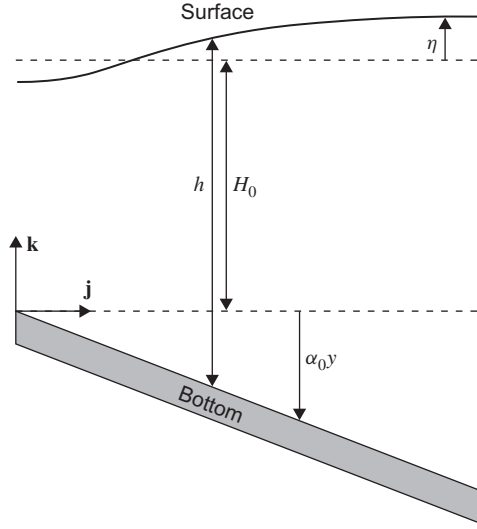
FIGURE 9.6   A layer of homogeneous fluid over a sloping bottom and the attending notation.

With the momentum equations (9.4a) and (9.4b), our present set of equations is

$$\frac{\partial u}{\partial t} - fv = -g\frac{\partial \eta}{\partial x} \tag{9.37a}$$

$$\frac{\partial v}{\partial t} + fu = -g\frac{\partial \eta}{\partial y} \tag{9.37b}$$

$$\frac{\partial \eta}{\partial t} + H_0\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \alpha_0 v = 0. \tag{9.37c}$$

In analogy with the system of equations governing planetary waves, the preceding set contains both small and large terms. The large ones (terms including f, g, and H_0) comprise the otherwise steady geostrophic dynamics, which correspond to a zero frequency. But, in the presence of the small α_0 term in the last equation, the geostrophic flow cannot remain steady, and the time-derivative terms come into play. We naturally expect them to be small and, compared with the large terms, on the order of α. In other words, the temporal Rossby number, $Ro_T = 1/\Omega T$, is expected to be comparable with α, leading to wave frequencies

$$\omega \sim \frac{1}{T} \sim \alpha\Omega \sim \alpha f \ll f$$

that are very subinertial, just as in the case of planetary waves, for which $\omega \sim \beta f_0$.

Capitalizing on the smallness of the time-derivative terms, we take in first approximation the large geostrophic terms: $u \simeq -(g/f)\partial \eta/\partial y$, $v \simeq +(g/f)\partial \eta/\partial x$.

Substitution of these expressions in the small time derivatives yields to the next degree of approximation:

$$u = -\frac{g}{f}\frac{\partial \eta}{\partial y} - \frac{g}{f^2}\frac{\partial^2 \eta}{\partial x \partial t} \tag{9.38a}$$

$$v = +\frac{g}{f}\frac{\partial \eta}{\partial x} - \frac{g}{f^2}\frac{\partial^2 \eta}{\partial y \partial t}. \tag{9.38b}$$

The relative error is only on the order of α^2. Replacement of the velocity components, u and v, by their last expressions (9.38a) and (9.38b) in the continuity equation (9.37c) provides a single equation for the surface displacement η, which to the leading order is

$$\frac{\partial \eta}{\partial t} - R^2 \frac{\partial}{\partial t}\nabla^2 \eta + \frac{\alpha_0 g}{f}\frac{\partial \eta}{\partial x} = 0. \tag{9.39}$$

(The ageostrophic component of v is dropped from the $\alpha_0 v$ term for being on the order of α^2, whereas all other terms are on the order of α.) Note the analogy with Eq. (9.26) that governs the planetary waves: It is identical, except for the substitution of $\alpha_0 g/f$ for $-\beta_0 R^2$. Here, the deformation radius is defined as

$$R = \frac{\sqrt{gH_0}}{f}, \tag{9.40}$$

that is, the closest constant to the original definition (9.12). A wave solution of the type $\cos(k_x x + k_y y - \omega t)$ immediately provides the dispersion relation:

$$\omega = \frac{\alpha_0 g}{f}\frac{k_x}{1 + R^2(k_x^2 + k_y^2)}, \tag{9.41}$$

the topographic analog of Eq. (9.27). Again, we note that if the additional ingredient, here the bottom slope α_0, had not been present, the frequency would have been nil, and the flow would have been steady and geostrophic. Because they owe their existence to the bottom slope, these waves are called *topographic waves*.

The discussion of their direction of propagation, phase speed, and maximum possible frequency follows that of planetary waves. The phase speed in the x-direction—that is, along the isobaths—is given by

$$c_x = \frac{\omega}{k_x} = \frac{\alpha_0 g}{f}\frac{1}{1 + R^2(k_x^2 + k_y^2)} \tag{9.42}$$

and has the sign of $\alpha_0 f$. Thus, topographic waves propagate in the northern hemisphere with the shallower side on their right. Because planetary waves

propagate westward, i.e., with the north to their right, the analogy between the two kinds of waves is "shallow–north" and "deep–south". (In the southern hemisphere, topographic waves propagate with the shallower side on their left, and the analogy is "shallow–south," "deep–north.")

The phase speed of topographic waves varies with the wavenumber; they are thus dispersive. The maximum possible wave speed along the isobaths is

$$c = \frac{\alpha_0 g}{f},$$ (9.43)

which is the speed of the very long waves $\left(k_x^2 + k_y^2 \to 0\right)$. With (9.41) cast in the form

$$\left(k_x - \frac{\alpha_0 g}{2f\omega R^2}\right)^2 + k_y^2 = \left(\frac{\alpha_0^2 g^2}{4f^2 R^4 \omega^2} - \frac{1}{R^2}\right),$$ (9.44)

we note that there exists a maximum frequency:

$$|\omega|_{\max} = \frac{|\alpha_0| g}{2|f| R}.$$ (9.45)

The implication is that a forcing at a frequency higher than the preceding threshold cannot generate topographic waves. The forcing then generates either a disturbance that is unable to propagate or higher frequency waves, such as inertia-gravity waves. However, such a situation is rare because, unless the bottom slope is very weak, the maximum frequency given by (9.45) approaches or exceeds the inertial frequency f, and the theory fails before (9.45) can be applied.

As an example, let us take the West Florida Shelf, which is in the eastern Gulf of Mexico. There the ocean depth increases gradually offshore to 200 m over 200 km ($\alpha_0 = 10^{-3}$) and the latitude (27°N) yields $f = 6.6 \times 10^{-5}$ s^{-1}. Using an average depth $H_0 = 100$ m, we obtain $R = 475$ km and $\omega_{\max} = 1.6 \times 10^{-4}$ s^{-1}. This maximum frequency, corresponding to a minimum period of 11 min, is larger than f, violates the condition of subinertial motions and is thus meaningless. The wave theory, however, applies to waves whose frequencies are much less than the maximum value. For example, a wavelength of 150 km along the isobaths ($k_x = 4.2 \times 10^{-5}$ m^{-1}, $k_y = 0$) yields $\omega = 1.6 \times 10^{-5}$ s^{-1} (period of 4.6 days) and a wave speed of $c_x = 0.38$ m/s.

Where the topographic slope is confined between a coastal wall and a flat-bottom abyss, such as for a continental shelf, topographic waves can be trapped, not unlike the Kelvin wave. Mathematically, the solution is not periodic in the offshore, cross-isobath direction but assumes one of several possible profiles (eigenmodes). Each mode has a corresponding frequency (eigenvalue). Such waves are called *continental shelf waves*. The interested reader can find an

exposition of these waves in LeBlond and Mysak (1978) and Gill (1982, pages 408–415).

9.6   ANALOGY BETWEEN PLANETARY AND TOPOGRAPHIC WAVES

We have already discussed some of the mathematical similarities between the two kinds of low-frequency waves. The objective of this section is to go to the root of the analogy and to compare the physical processes at work in both kinds of waves.

Let us turn to the quantity called potential vorticity and defined in Eq. (7.25). On the beta plane and over a sloping bottom (oriented meridionally for convenience), the expression of the potential vorticity becomes

$$q = \frac{f_0 + \beta_0 y + \partial v/\partial x - \partial u/\partial y}{H_0 + \alpha_0 y + \eta} \ . \tag{9.46}$$

Our assumptions of a small beta effect and a small Rossby number imply that the numerator is dominated by f_0, all other terms being comparatively very small. Likewise, H_0 is the leading term in the denominator because both bottom slope and surface displacements are weak. A Taylor expansion of the fraction yields

$$q = \frac{1}{H_0}\left(f_0 + \beta_0 y - \frac{\alpha_0 f_0}{H_0}\, y + \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} - \frac{f_0}{H_0}\, \eta\right). \tag{9.47}$$

In this form, it is immediately apparent that the planetary and topographic terms (β_0 and α_0 terms, respectively) play identical roles. The analogy between the coefficients β_0 and $-\alpha_0 f_0/H_0$ is identical to the one noted earlier between $-\beta_0 R^2$ of Eq. (9.20) and $\alpha_0 g/f$ of Eq. (9.35), since now $R = (gH_0)^{1/2}/f_0$. The physical significance is the following: Just as the planetary effect imposes a potential-vorticity gradient, with higher values toward the north, so the topographic effect, too, imposes a potential-vorticity gradient, with higher values toward the shallower side.

The presence of an ambient gradient of potential vorticity is what provides the *bouncing* effect necessary to the existence of the waves. Indeed, consider Fig. 9.7, where the first panel represents a north-hemispheric fluid (seen from the top) at rest in a potential-vorticity gradient and think of the fluid as consisting of bands tagged by various potential-vorticity values. The next two panels show the same fluid bands after a wavy disturbance has been applied in the presence of either the planetary or the topographic effect.

Under the planetary effect (middle panel), fluid parcels caught in crests have been displaced northward and have seen their ambient vorticity, $f_0 + \beta_0 y$, increase. To compensate and conserve their initial potential vorticity, they must develop some negative relative vorticity, that is, a clockwise spin. This is
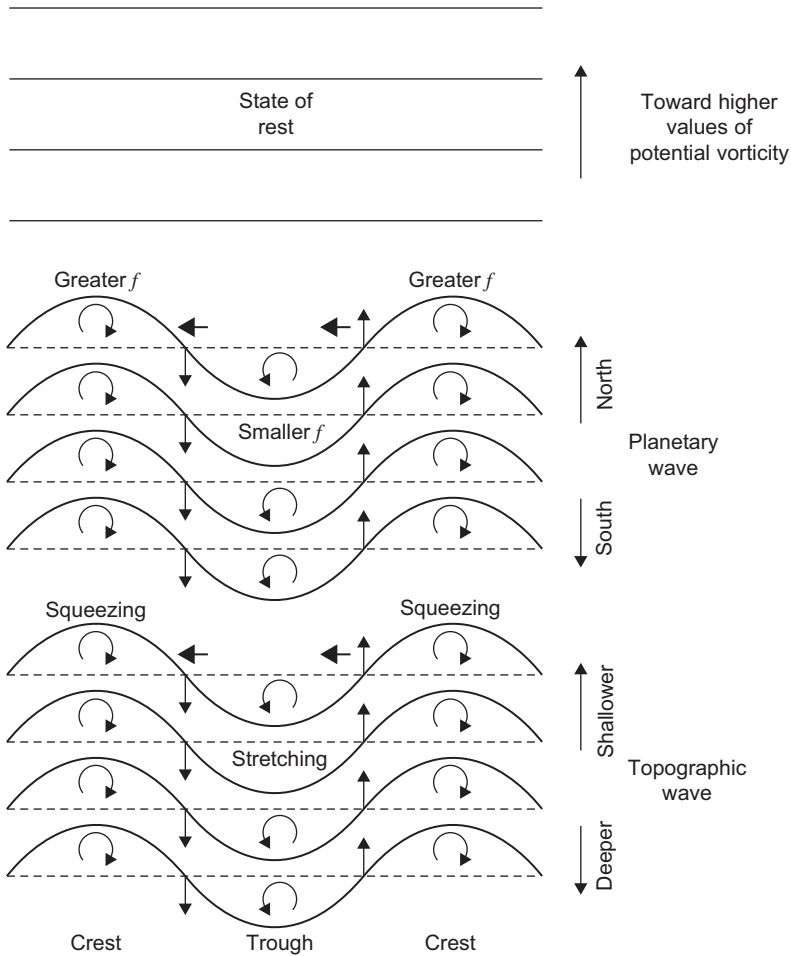
FIGURE 9.7   Comparison of the physical mechanisms that propel planetary and topographic waves. Displaced fluid parcels react to their new location by developing either clockwise or counterclockwise vorticity. Intermediate parcels are entrained by neighboring vortices, and the wave progresses forward.

indicated by curved arrows. Similarly, fluid parcels in troughs have been displaced southward, and the decrease of their ambient vorticity is met with an increase of relative vorticity, that is, a counterclockwise spin. Focus now on those intermediate parcels that have not been displaced so far. They are sandwiched between two counterrotating vortex patches, and, like an unfortunate finger caught between two gears or the newspaper zipping through the rolling press, they are entrained by the swirling motions and begin to move in the meridional direction. From left to right on the figure, the displacements are

southward from crest to trough and northward from trough to crest. Southward displacements set up new troughs, whereas northward displacements generate new crests. The net effect is a westward drift of the pattern. This explains why planetary waves propagate westward.

In the third panel of Fig. 9.7, the preceding exercise is repeated in the case of an ambient potential-vorticity gradient due to a topographic slope. In a crest, a fluid parcel is moved into a shallower environment. The vertical squeezing causes a widening of the parcel's horizontal cross-section (see Section 7.4), which in turn is accompanied by a decrease in relative vorticity. Similarly, parcels in troughs undergo vertical stretching, a lateral narrowing, and an increase in relative vorticity. From there on, the story is identical to that of planetary waves. The net effect is a propagation of the trough-crest pattern with the shallow side on the right.

The analogy between the planetary and topographic effects has been found to be extremely useful in the design of laboratory experiments. A sloping bottom in rotating tanks can substitute for the beta effect, which would otherwise be impossible to model experimentally. Caution must be exercised, however, for the substitution is acceptable so long as the analogy holds. The following three conditions must be met: absence of stratification, gentle slope, and slow motion. If stratification is present, the sloping bottom affects preferentially the fluid motions near the bottom, whereas the true beta effect operates evenly at all levels. And, if the slope is not gentle, and the motions are not weak, the expression of potential vorticity cannot be linearized as in Eq. (9.47), and the analogy is invalidated.

9.7   ARAKAWA'S GRIDS

The preceding developments had for aim to explain the basic physical mechanisms responsible for shallow-water wave propagation, by simplifying the governing equations down to their simplest, yet meaningful ingredients. Numerical models help us do better in cases where such simplifications are questionable, or when it is necessary to calculate wave motions in more realistic geometries. For the sake of clarity, broadly applicable numerical techniques will be illustrated on simplified cases. The simplest situation arises with inertia-gravity waves, for which the core mechanisms are rotation and gravity (see Section 9.3). In a one-dimensional domain of uniform fluid depth H, the linearized governing equations are

$$\frac{\partial \eta}{\partial t} + H \frac{\partial u}{\partial x} = 0 \tag{9.48a}$$

$$\frac{\partial u}{\partial t} - fv = -g\frac{\partial \eta}{\partial x} \tag{9.48b}$$

$$\frac{\partial v}{\partial t} + fu = 0. \tag{9.48c}$$

A straightforward second-order central finite differencing in space yields the following:

$$\frac{\mathrm{d}\tilde{\eta}_i}{\mathrm{d}t} + H \frac{\tilde{u}_{i+1} - \tilde{u}_{i-1}}{2\Delta x} = 0 \tag{9.49a}$$

$$\frac{\mathrm{d}\tilde{u}_i}{\mathrm{d}t} - f\tilde{v}_i = -g \frac{\tilde{\eta}_{i+1} - \tilde{\eta}_{i-1}}{2\Delta x} \tag{9.49b}$$

$$\frac{\mathrm{d}\tilde{v}_i}{\mathrm{d}t} + f\tilde{u}_i = 0. \tag{9.49c}$$

When analyzing this second-order method (upper part of Fig. 9.8), we observe that the effective grid size is $2\Delta x$ in the sense that all derivatives are taken over this distance. This is somehow unsatisfactory because the real grid size, i.e., the distance between adjacent grid nodes is only Δx. To improve the situation, we notice that the spatial derivatives of u are needed to calculate η, while the calculation of u requires the gradients of η. The most natural place to calculate a derivative of η is then at a point midway between two grid points of η, since the gradient approximation there is of second order while using a step of only Δx, and the most natural position to calculate the velocity u is therefore at mid-distance between η-grid nodes. Likewise, the most natural place to calculate the time evolution of η is at mid-distance between u nodes. It appears therefore that locating grid nodes for η and u in an interlaced fashion allows a second-order space differencing of *both* fields over a distance Δx (lower part of Fig. 9.8).
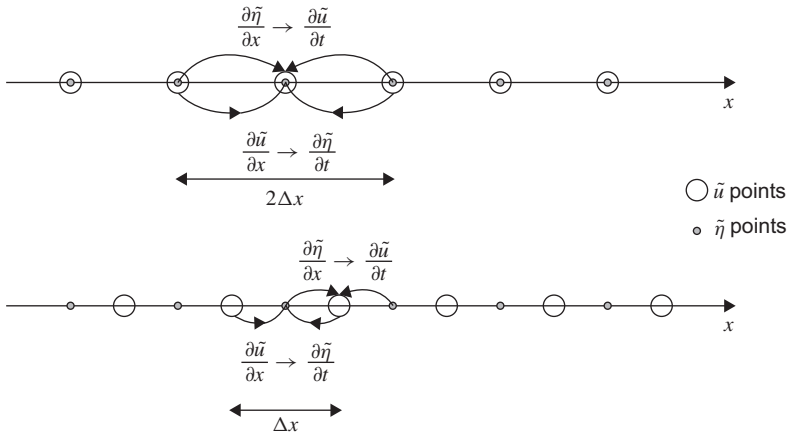

FIGURE 9.8   For variables $\tilde{\eta}$ and $\tilde{u}$ defined at the same grid points (upper panel), the discretization of inertia-gravity waves demands that approximations of spatial derivatives be made over distances $2\Delta x$, even if the underlying grid has a resolution of Δx. However, if variables $\tilde{\eta}$ and $\tilde{u}$ are defined on two different grids (lower panel), shifted one with respect to the other by $\Delta x/2$, the spatial derivatives can conveniently be discretized over the grid spacing Δx. The origin of each arrow indicates which variable influences the time evolution of the node where the arrow ends.

Formally, the discretization on such a *staggered grid* takes the form[1]:

$$\frac{d\tilde{\eta}_i}{dt} + H\frac{\tilde{u}_{i+1/2} - \tilde{u}_{i-1/2}}{\Delta x} = 0 \tag{9.50a}$$

$$\frac{d\tilde{u}_{i+1/2}}{dt} - f\tilde{v}_{i+1/2} = -g\frac{\tilde{\eta}_{i+1} - \tilde{\eta}_i}{\Delta x} \tag{9.50b}$$

$$\frac{d\tilde{v}_{i+1/2}}{dt} + f\tilde{u}_{i+1/2} = 0. \tag{9.50c}$$

Thus, spatial differencing can be performed over a distance Δx instead of $2\Delta x$. Since the scheme is centered in space, it is of second order, and its discretization error is reduced by a factor 4 without any additional calculation.[2] This advantage of a staggered grid over the elementary collocated grid is a prime example of optimization of numerical methods at fixed cost.

But, as we will now show, the performance gain is not the sole advantage of the staggered-grid approach. In the case of a negligible Coriolis force ($f \to 0$), elimination of velocity from Eq. (9.48) leads to a single wave equation for η:

$$\frac{\partial^2 \eta}{\partial t^2} = c^2 \frac{\partial^2 \eta}{\partial x^2}, \tag{9.51}$$

where $c^2 = gH$. This equation is the archetype of a hyperbolic equation, which possesses a general solution of the form

$$\eta = E_1(x+ct) + E_2(x-ct), \tag{9.52}$$

where the functions E_1 and E_2 are set by initial and boundary conditions. The general solution is therefore the combination of two signals, travelling in opposite directions at speeds $\pm c$ (Fig. 9.9). The lines of constant $x+ct$ and $x-ct$ define the characteristics along which the solution is propagated. For $t=0$, it is then readily seen that two initial conditions are needed, one for η and the other on its time derivative, i.e., the velocity field, before one can determine the two functions E_1 and E_2 at the conclusion of the first step. Later on, when characteristics no longer originate from the initial conditions but have their root in the boundaries, the solution becomes influenced first by the most proximate boundary condition and ultimately by both. If the boundary is impermeable, the condition is $u=0$, which can be translated into a zero-gradient condition on η. For discretization (9.50), the necessary numerical boundary conditions are consistent with the analytical conditions, whereas the nonstaggered version (9.53) needs additional conditions to reach the near boundary points. We have already learned (Section 4.7) how to deal with artificial conditions.

---

[1] We arbitrarily choose to place $\tilde{\eta}$ at integer grid indices and $\tilde{u}$ at half indices. We could have chosen the reverse.

[2] Both approaches use the same number of grid points to cover a given domain, and both schemes demand the same number of operations.
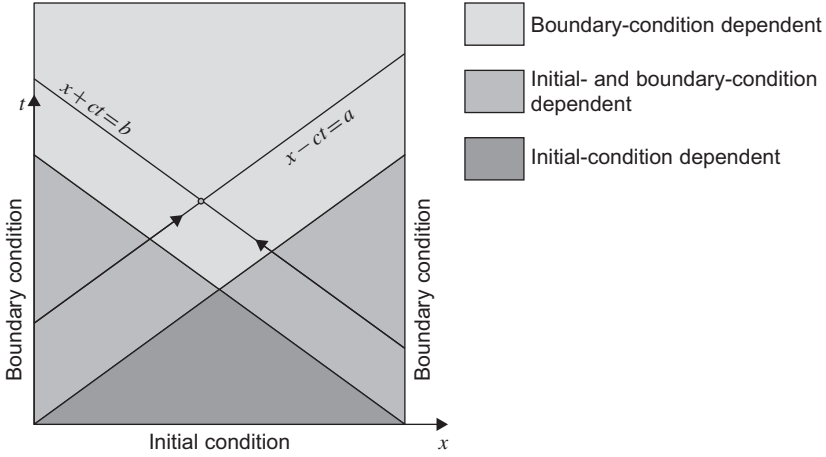
FIGURE 9.9   Characteristics $x + ct$ and $x - ct$ of wave equation (9.51). Information propagates along these lines from two initial conditions and one boundary condition on each side to set in a unique way the value of the solution at any point (x, t) of the domain.

We now turn our attention to the remaining problem of the nonstaggered grid, which is the appearance of spurious, stationary, and decoupled modes within the domain. To illustrate the issue, we use a standard leapfrog time discretization with zero Coriolis force so that the nonstaggered discretization becomes

$$\frac{\tilde{\eta}_i^{n+1} - \tilde{\eta}_i^{n-1}}{2\Delta t} = -H \frac{\tilde{u}_{i+1}^n - \tilde{u}_{i-1}^n}{2\Delta x} \tag{9.53a}$$

$$\frac{\tilde{u}_i^{n+1} - \tilde{u}_i^{n-1}}{2\Delta t} = -g \frac{\tilde{\eta}_{i+1}^n - \tilde{\eta}_{i-1}^n}{2\Delta x}. \tag{9.53b}$$

For the grid-staggered version, we can also introduce a form of time staggering by using a forward-backward approach in time:

$$\frac{\tilde{\eta}_i^{n+1} - \tilde{\eta}_i^n}{\Delta t} = -H \frac{\tilde{u}_{i+1/2}^n - \tilde{u}_{i-1/2}^n}{\Delta x} \tag{9.54a}$$

$$\frac{\tilde{u}_{i+1/2}^{n+1} - \tilde{u}_{i+1/2}^n}{\Delta t} = -g \frac{\tilde{\eta}_{i+1}^{n+1} - \tilde{\eta}_i^{n+1}}{\Delta x}. \tag{9.54b}$$

Although it may first appear that we are dealing with an implicit scheme because of the presence of $\tilde{\eta}^{n+1}$ on the right of the second equation, it is noted that this quantity has just been calculated when marching the preceding equation one step forward in time. The scheme is thus explicit in the sense that we solve the first equation to get $\tilde{\eta}^{n+1}$ everywhere in the domain, and then use it immediately in the second equation to calculate $\tilde{u}^{n+1}$ without having to invert any matrix.

As for all forms of leapfrogging and staggering, we should be concerned by spurious modes. These can be sought here rather simply by eliminating the discrete field $\tilde{u}$ from each set of equations. This elimination can be performed by taking a finite time difference of the first equation and a finite space difference of the second equation.[3] This is the direct analog of the mathematical differentiation used in eliminating the velocity between the two governing equations (9.48) to obtain (9.51). For the nonstaggered and staggered grids, we obtain

$$\tilde{\eta}_i^{n+2} - 2\tilde{\eta}_i^n + \tilde{\eta}_i^{n-2} = \frac{c^2 \Delta t^2}{\Delta x^2} \left(\tilde{\eta}_{i+2}^n - 2\tilde{\eta}_i^n + \tilde{\eta}_{i-2}^n \right) \tag{9.55}$$

$$\tilde{\eta}_i^{n+1} - 2\tilde{\eta}_i^n + \tilde{\eta}_i^{n-1} = \frac{c^2 \Delta t^2}{\Delta x^2} \left(\tilde{\eta}_{i+1}^n - 2\tilde{\eta}_i^n + \tilde{\eta}_{i-1}^n \right). \tag{9.56}$$

These equations are straightforward second-order discretizations of the wave equation (9.51), the first one with spatial and temporal steps twice as large as for the second one. The CFL criterion is $|c|\Delta t/\Delta x \le 1$ in each case, since the propagation speeds of the hyperbolic equation are $\pm c$, and the corresponding characteristics must lie in the numerical domain of dependence.

The discretization (9.55) shows that the nonstaggered grid is prone to decoupled modes. Indeed, for even values of n and i, all grid indexes involved are even, and the evolution is completely independent of that on points with odd values of n or i, which are nonetheless proximate in both time and space. In fact, there are four different solutions evolving independently, with their only link being through the initial and boundary conditions (left panel of Fig. 9.10). Although theoretically acceptable, such decoupled modes typically increase their "distance" from one another in the course of the simulation and induce
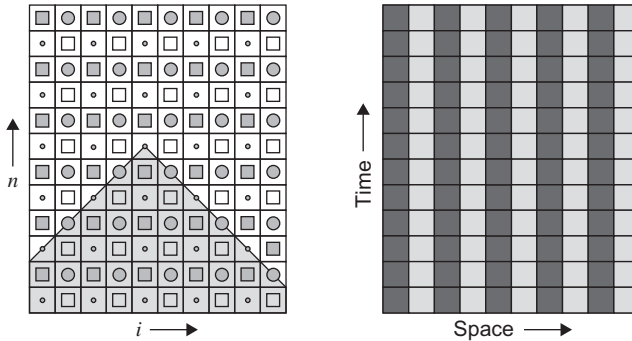

FIGURE 9.10 Four different solutions, each identified by a different symbol, evolve independently on the nonstaggered grid. The numerical domain of dependence is shown as the shaded region (left panel). A spurious stationary mode alternating between two constants (right panel) is incompatible with the original governing equations.

[3]For the nonstaggered version, we make the following formal elimination: Δx [(9.53a)$^{n+1}$ - (9.53a)$^{n-1}$] - $\Delta t H$[(9.53b)$_{i+1}$ - (9.53b)$_{i-1}$]

undesirable space-time oscillations in the solution. Occasionally, this can lead to stationary solutions that are simply unphysical (Fig. 9.10), such as a solution with zero velocity and $\tilde{\eta}$ alternating in space between two different constants. Solutions of this type are clearly spurious. In contrast, the only stationary solution produced by the staggered equation (9.56) is the physical one. This is a desirable property.

More generally, the spurious stationary solutions of a space discretization can be analyzed in terms of the state variable vector **x**, space discretization operator **D**, and the semi-discrete equation

$$\frac{d\mathbf{x}}{dt} + \mathbf{D}\mathbf{x} = 0 \qquad (9.57)$$

so that spurious stationary modes can be found among the nonzero solutions of

$$\mathbf{D}\mathbf{x} = 0. \qquad (9.58)$$

In the jargon of matrix calculation (linear algebra), spurious modes lie within the *null-space* of matrix **D**. In the case of the wave equation, the solution depicted in the right panel of Fig. 9.10 is certainly not a physically valid solution but satisfies Eq. (9.58).

All nonzero stationary solutions (members of the null-space), however, do not need to be spurious, and a physically admissible nonzero stationary solution is possible in the presence of the Coriolis terms, namely the geostrophic equilibrium. In that case, the discretized geostrophic equilibrium solution is also part of the null-space (9.58). It is therefore worthwhile sometimes to analyze the null-space of discretization operators for which the corresponding physical stationary solutions are known.

Having found that staggering has advantages in one dimension, we can now explore the situation in two dimensions but immediately realize that there is no single way to generalize the approach. Indeed, we have three state variables, u, v, and η, which can each be calculated on a different grid. The collocated version, the so-called A-grid model, is readily defined, and discretization[4] of Eqs. (9.4a) through (9.5) with uniform fluid thickness leads to:

$$\frac{d\tilde{\eta}}{dt} = -H\frac{\tilde{u}_{i+1} - \tilde{u}_{i-1}}{2\Delta x} - H\frac{\tilde{v}_{j+1} - \tilde{v}_{j-1}}{2\Delta y} \qquad (9.59a)$$

$$\frac{d\tilde{u}}{dt} = +f\tilde{v} - g\frac{\tilde{\eta}_{i+1} - \tilde{\eta}_{i-1}}{2\Delta x} \qquad (9.59b)$$

$$\frac{d\tilde{v}}{dt} = -f\tilde{u} - g\frac{\tilde{\eta}_{j+1} - \tilde{\eta}_{j-1}}{2\Delta y}. \qquad (9.59c)$$

---

[4]As before, we only write indices that differ from i and j.
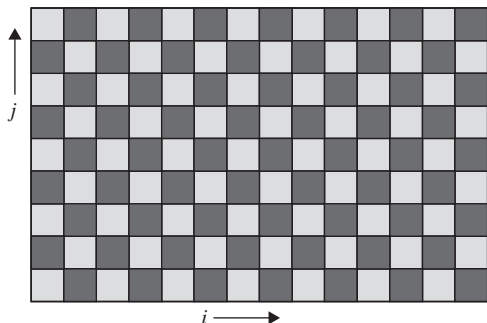
FIGURE 9.11   A spurious stationary $\tilde{\eta}$ mode alternating between two constants (depicted by two different gray levels) on the A-grid. This mode is called for obvious reasons the *checkerboard mode*.

Clearly, a spurious stationary solution exists, again with zero velocity ($\tilde{u} = \tilde{v} = 0$) and $\tilde{\eta}$ alternating between two constants on the spatial grid (Fig. 9.11).

Starting from the collocated grid, we can distribute the variables with respect to one another in different ways and create various staggered grids. These are named Arakawa's grids in honor of Akio Arakawa[5] and bear the letters A, B, C, D, and E depending on where the state variables are located across the mesh (Fig. 9.12; Arakawa & Lamb, 1977). For the linear system of equations considered here, it can be shown (e.g., Mesinger & Arakawa, 1976) that the E-Grid is a rotated B-grid so that we do not need to analyze it further.

A two-dimensional staggered grid we already encountered is the so-called C-grid (bottom left of Fig. 9.12). For advection [recall Eq. (6.58)] and the rigid-lid pressure formulation [recall Eq. (7.38)], we tacitly assumed that the velocity u was being calculated halfway between pressure nodes $(i+1, j)$ and (i, j) and v halfway between nodes (i, j) and $(i, j+1)$. In the present wave problem, this approach yields a straightforward second-order discretization of both divergence and pressure gradient terms

$$\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)_{i,j} \simeq \frac{\tilde{u}_{i+1/2} - \tilde{u}_{i-1/2}}{\Delta x} + \frac{\tilde{v}_{j+1/2} - \tilde{v}_{j-1/2}}{\Delta y} + \mathcal{O}(\Delta x^2, \Delta y^2) \qquad (9.60)$$

$$\left.\frac{\partial \tilde{\eta}}{\partial x}\right|_{i+1/2,j} \simeq \frac{\tilde{\eta}_{i+1} - \tilde{\eta}}{\Delta x} + \mathcal{O}(\Delta x^2) \qquad (9.61a)$$

$$\left.\frac{\partial \tilde{\eta}}{\partial y}\right|_{i,j+1/2} \simeq \frac{\tilde{\eta}_{j+1} - \tilde{\eta}}{\Delta y} + \mathcal{O}(\Delta y^2) \qquad (9.61b)$$

exactly as in the advection and surface pressure problems (Fig. 9.13). But if we proceed with the discretization of the Coriolis term, a problem arises for the C-grid because the velocity components are not defined at the same points. The

---

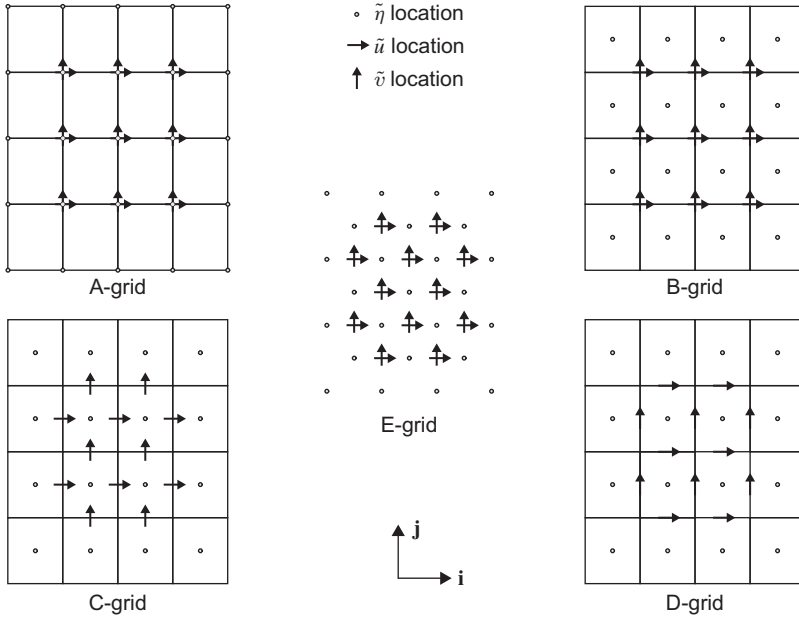[5]See biography at the end of this Chapter.

FIGURE 9.12 The five Arakawa grids. On the A-grid, the variables $\tilde{\eta}$, $\tilde{u}$, and $\tilde{v}$ are collocated, and staggered on other grids, called B-, C-, D-and E-grids. Note that the E-grid (center) has a higher grid-point density than the other grids for the same distance between adjacent nodes.

integration of the du/dt equation at the u grid node $(i+1/2, j)$ requires knowledge of the velocity v, which is only available at node $(i, j \pm 1/2)$. Therefore, an interpolation is necessary. The simplest scheme takes an average of surrounding values:

$$v|_{i+1/2,j} \simeq \frac{\tilde{v}_{j+1/2} + \tilde{v}_{i+1,j+1/2} + \tilde{v}_{j-1/2} + \tilde{v}_{i+1,j-1/2}}{4} \,, \qquad (9.62)$$

where the right-hand side can now be calculated from the available values of $\tilde{v}$. Similar averaging to estimate variables at locations where they are not defined can be used to discretize the equations on the other staggered grids. For example, the B-grid is a grid where η is defined on integer grid indices whereas velocity components are defined at corner points $(i \pm 1/2, j \pm 1/2)$. For this grid, the Coriolis term does not require any averaging, since both velocity components are collocated, but the grid arrangement requires the derivative of η in the x-direction at location $(i+1/2, j+1/2)$. We approximate such a term by the appropriate average

$$\frac{\partial \eta}{\partial x}\bigg|_{i+1/2,j+1/2} \simeq \frac{\frac{\tilde{\eta}_{i+1,j+1} + \tilde{\eta}_{i+1}}{2} - \frac{\tilde{\eta}_{j+1} + \tilde{\eta}}{2}}{\Delta x} \,, \qquad (9.63)$$
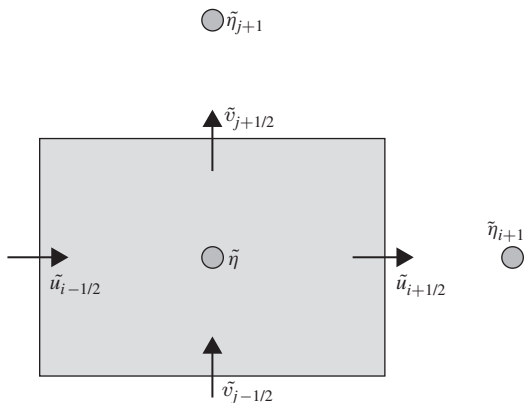
FIGURE 9.13   Discretization on the C-grid. The divergence operator is discretized most naturally by (9.60) while pressure gradients are calculated with (9.61).

and similarly for $\partial u/\partial x$, which is needed at (i,j). The full spatial discretization on each grid can be achieved in similar manner, and the derivation is left as an exercise (Numerical Exercise 9.2).

We can investigate the wave propagation properties on the various grids by a Fourier analysis, for which we take

$$\begin{pmatrix}\tilde{\eta}\\ \tilde{u}\\ \tilde{v}\end{pmatrix}=\Re\begin{pmatrix}A\\ U\\ V\end{pmatrix}e^{i\,(ik_x\Delta x+jk_y\Delta y-\tilde{\omega}t)}. \tag{9.64}$$

Insertion of this type of solution in the various finite-difference schemes and division by the common exponential factor provide the following equations

$$-\,i\tilde{\omega}U-f\alpha V=-\,ig\alpha_x k_x A \tag{9.65a}$$
$$-\,i\tilde{\omega}V+f\alpha U=-\,ig\alpha_y k_y A \tag{9.65b}$$
$$-\,i\tilde{\omega}A+H\,(i\alpha_x k_x U+i\alpha_y k_y V)=0, \tag{9.65c}$$

where the coefficients α, α_x, and α_y vary with the type of grid and are given in Table 9.1.

As for the physical solution, a nonzero solution is only possible when the determinant of the system vanishes, and this provides the dispersion relation of the discretized wave physics:

$$\tilde{\omega}\,[\tilde{\omega}^2-\alpha^2 f^2-gH\,(\alpha_x^2 k_x^2+\alpha_y^2 k_y^2)]=0, \tag{9.66}$$

which is the discrete analog of (9.16).

For small wavenumber values ($k_x\Delta x\ll 1$ and $k_y\Delta y\ll 1$), i.e., long, well-resolved waves, we recover the physical dispersion relation because α, α_x, and α_y all tend towards unity. For shorter waves, the numerical dispersion relation

TABLE 9.1 Definition of the Parameters Involved in the Discrete Dispersion Relation for A-, B-, C-and D-grids with $2\theta_x = k_x \Delta x$ and $2\theta_y = k_y \Delta y$.

Grid	α	$\alpha_x k_x \Delta x$	$\alpha_y k_y \Delta y$
A	1	$\sin 2\theta_x$	$\sin 2\theta_y$
B	1	$2\sin\theta_x \cos\theta_y$	$2\sin\theta_y \cos\theta_x$
C	$\cos\theta_x \cos\theta_y$	$2\sin\theta_x$	$2\sin\theta_y$
D	$\cos\theta_x \cos\theta_y$	$2\cos\theta_x \cos\theta_y \sin\theta_x$	$2\cos\theta_x \cos\theta_y \sin\theta_y$

can be analyzed in detail through the error estimate

$$\frac{\omega^2 - \tilde{\omega}^2}{\omega^2} = \frac{(1-\alpha^2) + R^2\left[(1-\alpha_x^2)k_x^2 + (1-\alpha_y^2)k_y^2\right]}{1 + R^2\left(k_x^2 + k_y^2\right)} \geq 0. \qquad (9.67)$$

Except for the simple statement $\tilde{\omega}^2 \leq \omega^2$, the analysis of the error is rather complex because it involves five length scales[6]: R, $1/k_x$, $1/k_y$, Δx, and Δy. For simplicity, we take $\Delta x \sim \Delta y$ and $k_x \sim k_y$ to reduce the problem. We then define the length scale $L \sim 1/k_x \sim 1/k_y$ of the wave under consideration. In this case $\alpha_x \sim \alpha_y$, and we can distinguish two situations:

$$\text{Shorter waves:} \quad L \lesssim R, \quad \omega^2 \sim \frac{gH}{L^2} \qquad (9.68)$$

$$\text{Longer waves:} \quad L \gtrsim R, \quad \omega^2 \sim f^2. \qquad (9.69)$$

The shorter waves are dominated by gravity, and the relative error on ω^2 behaves as

$$\frac{\omega^2 - \tilde{\omega}^2}{gH/L^2} \sim (1-\alpha_x^2). \qquad (9.70)$$

If the wave is well resolved ($\Delta x \ll L$), the error tends toward zero for all four grids because $\alpha_x \to 1$. For the barely resolved waves ($\Delta x \sim L$), the errors are largest for the discretizations in which α_x and α_y depart most from unity. In this sense, the A-, B-, and D-grids have larger errors than the C-grid (see Table 9.1).

---

[6]Note how the discretization has added two length scales, Δx and Δy, to the discussion.

The longer waves are dominated by rotation, and the relative error on ω^2 behaves as

$$\frac{\omega^2 - \tilde{\omega}^2}{f^2} \sim (1 - \alpha^2). \tag{9.71}$$

Again, α should remain close to unity for all wavenumbers so that the B-grid outperforms both the C- and D-grids. For details on the errors, an exploration with abcdgrid.m in parameter space provides relative error fields as those depicted in Fig. 9.14 for various resolution levels $R/\Delta x$, etc. Errors can be further investigated through the analysis of the group-velocity behavior (Numerical Exercise 9.5) and in the context of generalized dynamics, including planetary waves, with a clear distinction between zonal and meridional wave behaviors (Dukowicz, 1995; Haidvogel & Beckmann, 1999).

Because the A-grid suffers from spurious modes, and the D-grid is always penalized in terms of accuracy, the B-and C-grids are the most interesting ones among the four types. Since wavelengths up to $\Delta x \sim L$ are to be resolved in a
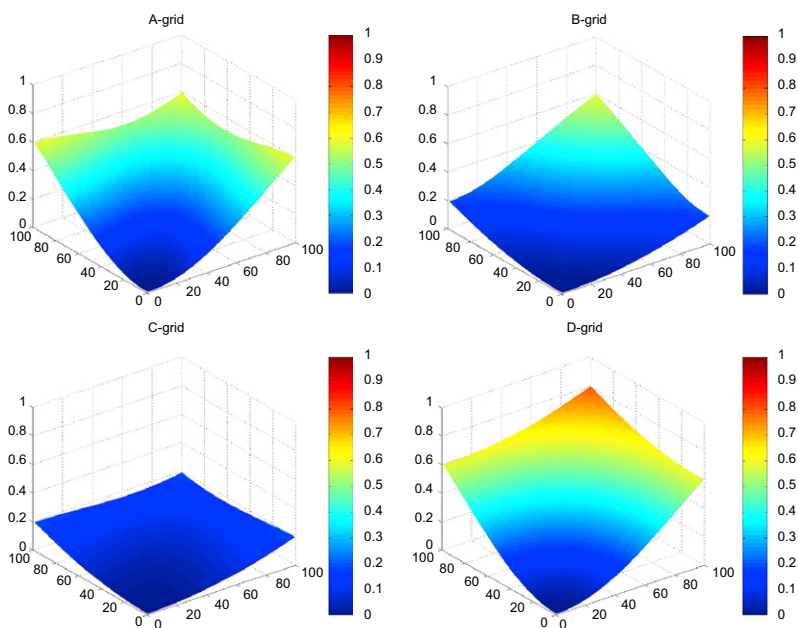

FIGURE 9.14   For medium resolution compared with the deformation radius ($R/\Delta x = R/\Delta y = 2$), the frequency error (9.67) is depicted as a function of $k_x \Delta x$ and $k_y \Delta y$. Waves with wavenumber higher than $k_x \Delta x = \pi/2$ are not shown, and the x- and y-axes are labeled in percents of $\pi/2$. The D-grid clearly exhibits the worst behavior. The B-grid keeps the error low for well-resolved waves, whereas the C-grid creates lower errors for shorter waves. A color version can be found online at http://booksite.academicpress.com/9780120887590/

significant way, the C-grid is the better choice as long as $\Delta x \ll R$, whereas for $R \ll \Delta x$ the B-grid should be preferred on the ground that the error in Eqs. (9.70) and (9.71) is less. This confirms more detailed error analyses of the semidiscrete equations on staggered grids (e.g., Haidvogel & Beckman, 1999; Mesinger & Arakawa, 1976), although additional time discretization or boundary condition implementation can introduce stability problems (e.g., Beckers, 1999; Beckers & Deleersnijder, 1993). Also, time discretization further complicates the error analysis and may sometimes inverse the error behavior (Beckers, 2002). Nevertheless, the choice of the B-grid for larger grid spacing and the C-grid for finer grid spacing is justified by the fact that for large grid spacing we only capture large-scale movements, which are nearly geostrophic. Since the Coriolis force is dominant in this case, its discretization is crucial. Because the B-grid does not require a spatial average of the velocity components, contrary to the C-grid, its use should be advantageous. The pressure gradient, which is the other dominant force, could arguably be better represented on the C-grid. If the grid is very fine, averaging the large-scale geostrophic equilibrium over four closely spaced nodes does not deteriorate the geostrophic solution, whereas smaller-scale processes such as gravity waves and advection are better captured by the C-grid.

From the preceding interpretation, we can establish some general rules for staggering the variables. Starting with the goal of placing variables on the grid so that dominant processes are discretized in the best possible way, we can then afford to represent secondary processes by less accurate discrete operators without affecting overall model accuracy. In practice, however, dominant processes may change in time and space so that no single approach can be guaranteed to work uniformly, but it should at least be tried. For example, if tracer advection is of primary interest, the C-grid can be generalized to three dimensions with vertical velocities defined at the bottom and top of each grid cell. In that case, advection fluxes are readily calculated using one of the advection schemes presented in Section 6.6 without need for velocity interpolations. Similarly, if diffusion in a heterogeneous turbulent environment is the main process at play, the definition of diffusion coefficients between tracer nodes would allow the direct discretization of turbulent fluxes without the need of averaging the diffusion coefficients.

9.8   NUMERICAL SIMULATION OF TIDES AND STORM SURGES

The two-dimensional equations (7.16) describing shallow-water dynamics are the basic equations from which storm-surge models of vertically mixed coastal seas have been developed. The prediction of rising sea level (surge) along a coast depends on remotely generated waves that propagate from the stormy area toward the shore. Because shallow-water equations describe well the propagation of such waves, their prediction is indeed feasible, although a few additional

processes must be taken into account. Among these other processes are the surface wind stress, through which waves are generated, and bottom friction, which causes attenuation during travel. To include these last two stresses, we can start from the observation that the shallow-water equations assume that the flow is independent of the vertical coordinate. If this is not the case, we can at least try to predict the evolution of the depth-averaged velocity,

$$\bar{u} = \frac{1}{h} \int\limits_{b}^{b+h} u \, \mathrm{d}z \quad \bar{v} = \frac{1}{h} \int\limits_{b}^{b+h} v \, \mathrm{d}z \tag{9.72}$$

where $z = b$ is the bottom level and h the water depth. We can derive a governing equation for $\bar{u}$ by integrating vertically the three-dimensional governing equations including the x-momentum equation:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial z}\left(\nu_{\mathrm{E}} \frac{\partial u}{\partial z}\right) + F(u), \tag{9.73}$$

where the term $F(u)$ gathers all terms other than the time derivative and vertical diffusion. We can then integrate vertically to obtain

$$\frac{1}{h} \int\limits_{b}^{b+h} \frac{\partial u}{\partial t} \, \mathrm{d}z = \frac{\tau^{x}}{\rho_{0} h} - \frac{\tau_{b}^{x}}{\rho_{0} h} + \overline{F(u)}, \tag{9.74}$$

where boundary conditions similar to (4.34) have been used for the surface wind stress τ and bottom stress τ_b, respectively. Physically, these stresses appear here as body forces applied to the layer h of fluid moving as a slab with the depth-averaged velocity (Fig. 9.15).

Two difficulties arise, however, during the integration. The first is that the elevation of the surface is time dependent and does not allow a simple permu-
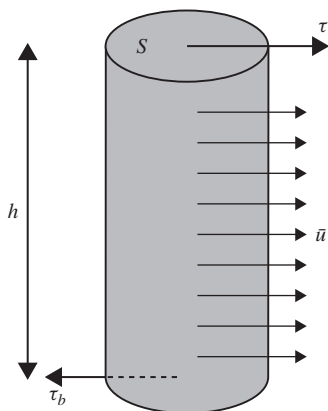

FIGURE 9.15 For a fluid column of volume hS and moving with the average velocity $\bar{u}$, Newton's second law in the absence of lateral friction and pressure force implicates the forces associated with the surface stress τ and bottom friction: $\rho_{0} hS \, \mathrm{d}\bar{u}/\mathrm{d}t = (\tau - \tau_{b})S$.

tation of the integration with the time derivative in the left-hand side of (9.74).[7] The second and more fundamental difficulty is due to the nonlinearities of the equations, which prevent us from equating the average of $F(u)$ with $F(\bar{u})$, i.e.,

$$\overline{F(u)} \neq F(\bar{u}) \tag{9.75}$$

so that we cannot express the right-hand side of (9.74) as a function of only the average velocity. The integrated equation then requires some form of parameterization. Hence, shallow-water models include additional parameterization of the horizontal diffusion type.

    For simplicity, the governing equations are written as if depth averaging had not taken place, and the overbar (¯) operator is ignored. The outcome is that it is sufficient to add the $\tau/(\rho_0 h)$ and $-\tau_b/(\rho_0 h)$ terms to the right-hand side of the two-dimensional momentum equations (7.9a) and (7.9b), and then to include a parameterization of the nonlinear effects. The wind stress τ appears as an externally imposed source term in the equations, whereas the bottom stress is depending on the flow itself, i.e., $\tau_b = \tau_b(u, v)$. A difficulty arises here because the bottom stress depends on the velocity profile near the bottom, whereas the governing equations provide only the vertical average of the velocity. A parameterization is needed here, too. The simplest version is linear bottom friction, in which the frictional term is made linear in, and opposite to, velocity:

$$\tau_b^x = -r\rho_0 u, \quad \tau_b^y = -r\rho_0 v, \tag{9.76}$$

where r is a coefficient with dimensions of velocity (LT^{-1}). The linear formulation is particularly advantageous in analytical studies or with spectral methods. They fail, however, to take into account the turbulent nature of the bottom boundary layer, with stress better expressed as a quadratic function of velocity (see Chapter 14):

$$\tau_b^x = -\rho_0 C_d \sqrt{u^2 + v^2}\, u, \qquad \tau_b^y = -\rho_0 C_d \sqrt{u^2 + v^2}\, v, \tag{9.77}$$

with a dimensionless *drag coefficient* C_d either constant or depending on the flow itself.

    Finally, the direct driving force associated with a moving disturbance of the atmospheric pressure $p_{atm}(x, y, t)$ can be easily taken into account by including it in the pressure boundary condition at the surface (4.32), $p = \rho_0 g\eta + p_{atm}$.

    We can now estimate the wind-induced surge in a shallow sea by considering how the storm piles up water near the coast (Fig. 9.16). This accumulation of water creates a surface elevation (surge) and, consequently, an adverse pressure gradient. Eventually, this adverse pressure gradient can grow strong enough to cancel the wind stress. When this balance is reached, the sea surface slope

---

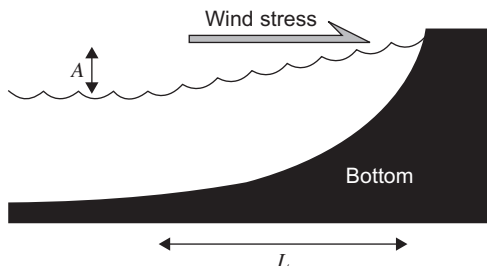[7]This problem can be overcome by using Leibniz rule as done in Section 15.6.

FIGURE 9.16   The piling up of water by a storm near a coast an adverse pressure-gradient force, and an equilibrium can exist if it cancels the wind-stress force.

caused by the wind stress is governed by

$$\frac{\partial \eta}{\partial x} \simeq \frac{\tau}{\rho_0 g h}. \tag{9.78}$$

This relation provides an estimate of the storm-surge amplitude A as a function of the distance L over which the wind blows

$$A \simeq \frac{L\tau}{\rho_0 g h}. \tag{9.79}$$

Note that the shallower the water, the stronger the effect. In other words, storm surges intensify near the coast where the water is shallower.

Storm surges can become dramatic when superimposed to the tide, and it is therefore important to know how to calculate tidal elevations, too. Tides are forced gravity waves caused by the gravitational attraction of the moon and sun. The following development is also valid for the atmosphere, but the velocities associated with atmospheric tides are much smaller than the wind speed due to atmospheric disturbances. Tides, therefore, are generally negligible in the atmosphere, whereas tidal currents in the ocean can be an order of magnitude larger than other currents.

To quantify the net effect of the gravitational acceleration of the moon and sun, we have to realize that the whole system is moving. Therefore, Newton's law cannot simply be written with respect to axes fixed at the earth center as we did in Chapter 2. Instead, using Newton's law in absolute axes $\mathbf{I}$, $\mathbf{J}$, and $\mathbf{K}$ of the solar system, we can calculate the absolute acceleration $\mathbf{A}$ of the fluid parcel and of the earth $\mathbf{A}_e$ under the moon's attraction[8]:

$$\rho \mathbf{A} = \rho \boldsymbol{\gamma} + \rho \mathbf{f} \tag{9.80}$$

$$M_e \mathbf{A}_e = M_e \boldsymbol{\gamma}_e. \tag{9.81}$$

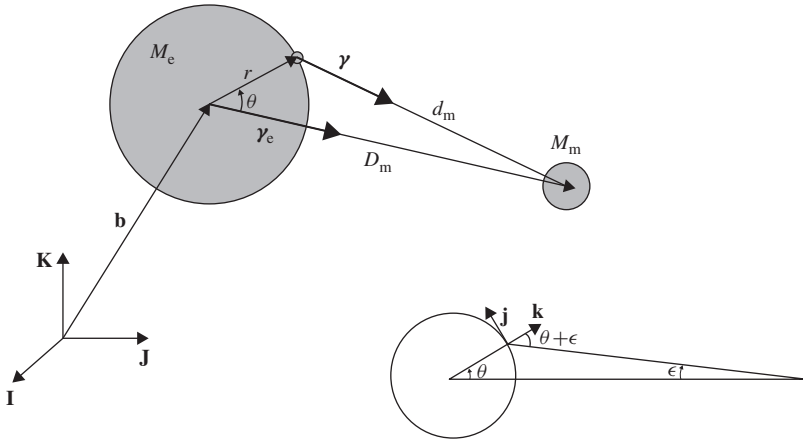[8]The sun's influence can be studied in an analogous way.

FIGURE 9.17   The moon acts simultaneously on a fluid parcel lying on the earth's surface as well as on the entire earth. The tidal force results from the mutual attraction between earth and moon and from the local gravitational attraction on earth.

We regrouped under the force $\rho\mathbf{f}$ all forces acting on the fluid parcel other than the moon's attraction. The gravitational forces[9] $\rho\boldsymbol{\gamma}$ and $M_e\boldsymbol{\gamma}_e$ involve the gravitational constant $G = 6.67 \times 10^{-11}\ \mathrm{N\,m^2/kg^2}$, the earth's mass $M_e = 5.9736 \times 10^{24}$ kg, the moon's mass $M_m = 7.349 \times 10^{22}$ kg, the distance $D_m \sim 385000$ km between earth and moon, and the actual distance d_m of the point under consideration to the moon (Fig. 9.17). The two gravitational accelerations are directed towards the center of the moon and of magnitude

$$\gamma = \frac{GM_m}{d_m^2}, \qquad \gamma_e = \frac{GM_m}{D_m^2}. \tag{9.82}$$

We are not so much interested in the movement of the earth *per se*, but we need its acceleration to subtract it from the fluid parcel's absolute acceleration, which is $\mathbf{A} = \mathbf{A}_e + \mathrm{d}^2\mathbf{r}/\mathrm{d}t^2$. The fluid's parcel's relative acceleration with respect to the earth is thus

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \mathbf{f} + (\boldsymbol{\gamma} - \boldsymbol{\gamma_e}). \tag{9.83}$$

Without the astronomical force, the equation would have been $\mathrm{d}^2\mathbf{r}/\mathrm{d}t^2 = \mathbf{f}$, and so we note that its effect is the addition of a so-called *tidal force* (per unit volume):

$$\rho\mathbf{f}_t = \rho(\boldsymbol{\gamma} - \boldsymbol{\gamma_e}). \tag{9.84}$$

---

[9]It should be clear that for a fluid parcel, forces are per unit volume, whereas for the earth, we speak about full forces.

We notice that this force is the difference between two almost identical forces. Its locally vertical component is as follows:

$$f_\uparrow = \gamma \cos(\theta + \epsilon) - \gamma_e \cos\theta. \tag{9.85}$$

The expression can be simplified because the angle ϵ is extremely small (Fig. 9.17, diagram in lower right). By expanding $\cos(\theta + \epsilon)$ and using

$$\cos\epsilon \simeq 1, \qquad D_m \sin\epsilon \simeq r\sin\theta, \tag{9.86}$$

we obtain

$$f_\uparrow = \frac{GM_m}{D_m^2}\left[\left(\frac{D_m^2}{d_m^2} - 1\right)\cos\theta - \frac{D_m^2}{d_m^2}\frac{r\sin\theta}{D_m}\sin\theta\right]. \tag{9.87}$$

The use of d_m in the formulation of the tidal forcing is not very practical. (Can you tell at any moment the precise distance of your position with respect to the moon?) So, we use the identity $r\cos\theta + d_m\cos\epsilon = D_m$ and the smallness of ϵ to obtain

$$d_m \simeq D_m\left(1 - \frac{r}{D_m}\cos\theta\right). \tag{9.88}$$

For the same reason that ϵ is small, the ratio r/D_m is also considered small,[10] and we drop higher-order terms in r/D_m:

$$\frac{D_m^2}{d_m^2} \simeq 1 + 2\frac{r}{D_m}\cos\theta, \tag{9.89}$$

so that the vertical component of the tidal force can be reduced to

$$f_\uparrow \simeq \frac{GM_m}{D_m^3}\,r\,(3\cos^2\theta - 1). \tag{9.90}$$

To compare its magnitude to $g = GM_e/r^2 = 9.81$ m/s^2, the gravitational acceleration of the earth on its surface, we form the ratio $\delta = f_\uparrow/g$ and find it to be on the order of

$$\delta \sim \frac{r^3 M_m}{D_m^3 M_e} \sim \mathcal{O}(10^{-7}). \tag{9.91}$$

It appears therefore that the tidal force associated with the moon is completely negligible, not only compared with gravity g but also to any of the typical forces acting along the vertical direction. So does it mean tidal forces are not responsible for the observed tides? Of course they are, but not through the local

---

[10]In the case of the earth–moon system, its value is about $6400/385000 \sim 0.017$.
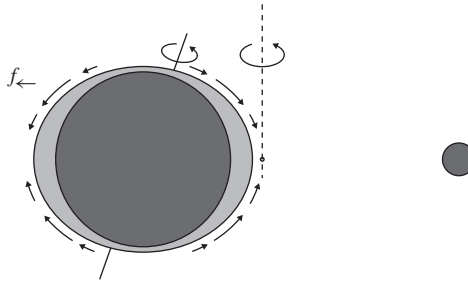
FIGURE 9.18 The movement of earth and moon around their common center of mass creates a centrifugal force that is weaker than the moon's gravitational attraction for earth points facing the moon. The resulting horizontal tidal force $f_\leftarrow$ has a tendency to create a bulge toward the moon. On the earth's face opposite to the moon, the centrifugal force is larger than the moon's gravitational attraction, and the horizontal force creates a second bulge facing away from the moon. Since the earth rotates around its own South–North axis, the two bulges move with respect to the continents.

vertical attraction as sometimes erroneously thought, but through the horizontal component, which we now proceed to calculate.

The component of the tidal force along the local northward axis (**j** in Fig. 2.9) is, after several simplifications similar to those made above,

$$f_\leftarrow \simeq -\frac{GM_m}{D_m^3} 3r \cos\theta \sin\theta. \tag{9.92}$$

The order of magnitude of this force component is the same as that of the vertical one, but since all horizontal forces are much smaller than gravity, the horizontal tidal force is *not* negligible and acts to make the fluid converge or diverge. The spatial distribution of this force along the earth's surface is such that it tends to create a bulge in the region of the earth facing the moon and a second bulge at the diametrically *opposite* place. The explanation is that, for a point closer to the moon than D_m, the gravitational pull of the moon exceeds the centrifugal force associated with the earth–moon corotation, while on the opposite side of the earth the inverse is true; the centrifugal force of the earth–moon corotation exceeds the gravitational pull of the moon. This is the essential mechanism of lunar tides (Fig. 9.18). Solar tides are similar, with the sun taking the place of the moon but being much larger and much further away.

The angle θ involved in our formula is constantly changing in time because of the terrestrial rotation and lunar motion, and it must be determined through astronomical calculations (e.g., Doodson, 1921). These calculations also take into account variations in the earth–moon distance D_m, which induce slow modulations of the tidal force. Trigonometric calculations reveal different periods of motion, the most noticeable one being due to the corotation of the earth and moon, giving rise to an apparent rotation of the moon over a given point on the

earth every 24 h and 50 min (24 h of terrestrial rotation and a delay caused by the moon rotating around the sun in the same direction). But, because there are two bulges half an earth's circumference apart from each other [mathematically because of the product $\cos\theta\sin\theta$ in Eq. (9.92)], the period of the main lunar tide is only half of that, i.e., 12 h 25 min.

For practical purposes, it is worth noting that the tidal force can be derived from the so-called *tidal potential* (see Analytical Problem 9.8). In the local Cartesian coordinate system, the tidal force can be expressed as

$$\mathbf{f}_t = -\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right) \quad \text{with} \quad V = -\frac{GM_m}{D_m^3}\frac{r^2}{2}(3\cos^2\theta - 1). \quad (9.93)$$

All we have to do then is to calculate the local tidal potential, take its local derivatives, and introduce these as tidal forces in the shallow-water equations.

The tidal potential can also be used to estimate tidal amplitudes. Since the tidal force, i.e., the gradient of the potential, has a form similar to the pressure-gradient force associated with the sea surface height, we can ask which distribution of η, denoted η_e, would cancel the tidal force so that no motion would result. Obviously, this is the case when

$$\eta_e = -\frac{V}{g} = \frac{GM_m}{D_m^3}\frac{r^2}{2g}(3\cos^2\theta - 1)$$

$$= \mathcal{O}\left(\frac{GM_m}{D_m^3}\frac{r^2}{g}\right) \sim 0.36 \text{ m.} \quad (9.94)$$

This defines the so-called *equilibrium tide*, first derived by Isaac Newton. It would be the tidal elevation if the fluid could follow the tidal force in order to remain in equilibrium with the pressure gradient generated by the bulges. In reality, however, continents and topographic features in the ocean do not allow sea water to stay at the equilibrium. Not only is the equilibrium tide never reached, but the tidal potential is also in need of further adaptation to take into account the solid earth deformation due to tides and the self-attraction of tides (e.g., Hendershott, 1972).

In the same way as we defined the equilibrium tide, we can determine the sea surface height that would exactly cancel the effect of an atmospheric pressure disturbance p_{atm}:

$$\eta = -\frac{p_{atm}}{\rho_0 g}, \quad (9.95)$$

which can be used as a first approximation to estimate the effect of atmospheric pressure on measurements of η and is called the *inverse barometric response*.

For actual tidal predictions, we resort to numerical methods. For this, we gather all terms previously mentioned in this chapter and add the components of the tidal force. The governing equations used in a shallow-water model to

predict both tides and storm surges are

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}-fv=-\frac{1}{\rho_0}\frac{\partial p}{\partial x}+\frac{\tau^x}{\rho_0 h}-\frac{\tau_b^x}{\rho_0 h}-\frac{\partial V}{\partial x}$$

$$+\frac{1}{h}\frac{\partial}{\partial x}\left(\mathcal{A}\frac{\partial hu}{\partial x}\right)+\frac{1}{h}\frac{\partial}{\partial y}\left(\mathcal{A}\frac{\partial hu}{\partial y}\right) \tag{9.96a}$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+fu=-\frac{1}{\rho_0}\frac{\partial p}{\partial y}+\frac{\tau^y}{\rho_0 h}-\frac{\tau_b^y}{\rho_0 h}-\frac{\partial V}{\partial y}$$

$$+\frac{1}{h}\frac{\partial}{\partial x}\left(\mathcal{A}\frac{\partial hv}{\partial x}\right)+\frac{1}{h}\frac{\partial}{\partial y}\left(\mathcal{A}\frac{\partial hv}{\partial y}\right) \tag{9.96b}$$

$$p=\rho_0 g\eta+p_{atm} \tag{9.96c}$$

together with Eqs. (7.14) and (9.93). Note that the driving forces of wind and tide act very differently. While the wind stress acts as a surface force and therefore appears with a factor $1/h$, the tidal force is a body force acting over the whole water column. Consequently, the tidal force is more important in the deeper parts of the ocean. This might surprise us since we are used to observe the highest tides near the coasts, where h is small! In most cases, tides are generated in the deeper parts of the oceans, where the tidal force acts on a thick layer of water, creates a pattern of convergence/divergence and locally modifies the sea surface height. The sea surface elevation is then propagated as a set of Kelvin and inertia-gravity waves into shelf seas and coastal regions, where the reduced depth increases their amplitudes (Fig. 9.2).

Some shelf models can provide tidal predictions by imposing tidal elevations at distant open boundaries and propagating the waves into the domain while discarding the local tidal force. This is consistent with the idea that, in shallow seas, the wind stress is the dominant local forcing. Indeed, in a 10,000-m-deep basin, the tidal force is equivalent to the surface friction of a 75 m/s wind, whereas in a shallow sea of 100 m, a wind of 7.5 m/s already matches the local tidal force. An example of a tidal calculation in which the tides are imposed along an open boundary is given in Fig. 9.19. In this figure, we note in passing the presence of nodes where the tidal amplitude is nil and the phase undefined. Each such node, called an *amphidromic point*, is a place where the various wave components cancel each other (destructive interference).

The numerical implementation of the model we have just developed is readily feasible since we have already encountered all its ingredients: time stepping, advection, Coriolis term, pressure gradient, diffusion, which were all treated in detail in previous sections. The only remaining term is that including the bottom stress, and for it we suggest to discretize it with the Patankar technique (to be discussed in Section 14.6) if the quadratic relationship is selected:

$$\tau_b^x=-\rho_0 C_d\sqrt{(u^n)^2+(v^n)^2}\,u^{n+1},\quad \tau_b^y=-\rho_0 C_d\sqrt{(u^n)^2+(v^n)^2}\,v^{n+1}. \tag{9.97}$$
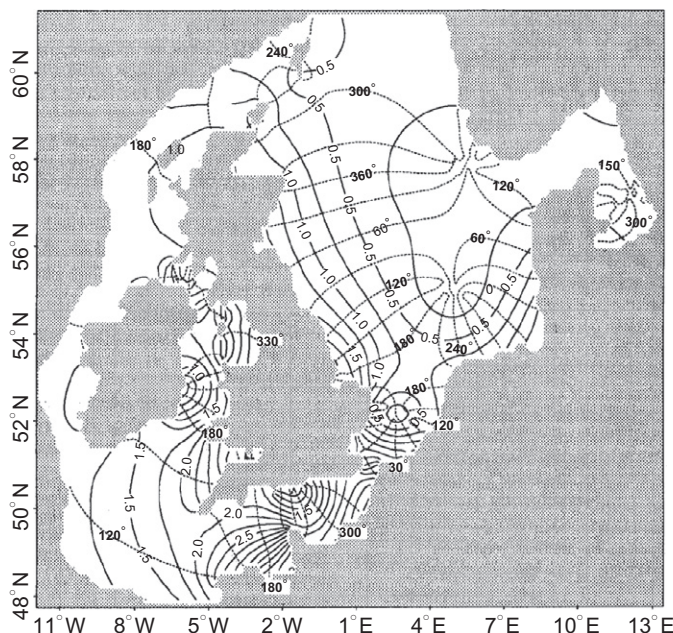
FIGURE 9.19 Tidal amplitudes (full lines) and phases (dotted lines) over the Northwestern European continental shelf, generated by the moon. (*Eric Delhez*)

Since several methods are available for each process, the combination of the various processes leads to a very wide array of possible numerical implementations, all at relatively low cost with two spatial dimensions. This explains the large number of two-dimensional numerical models that were developed relatively early in geophysical fluid modeling (e.g., Backhaus, 1983; Heaps, 1987; Nihoul, 1975).

ANALYTICAL PROBLEMS

9.1. Prove that Kelvin waves propagate with the coast on their left in the southern hemisphere.

9.2. The Yellow Sea between China and Korea (mean latitude: 37°N) has an average depth of 50 m and a coastal perimeter of 2600 km. How long does it take for a Kelvin wave to go around the shores of the Yellow Sea?

9.3. Prove that at extremely large wavelengths, inertia-gravity waves degenerate into a flow field where particles describe circular inertial oscillations.

9.4. An oceanic channel is modeled by a flat-bottom strip of ocean between two vertical walls. Assume that the fluid is homogeneous and inviscid,

and that the Coriolis parameter is constant. Describe all waves that can propagate along such a channel.

9.5. Consider planetary waves forced by the seasonal variations of the annual cycle. For $f_0 = 8 \times 10^{-5}\,\text{s}^{-1}$, $\beta_0 = 2 \times 10^{-11}\,\text{m}^{-1}\text{s}^{-1}$, $R = 1000\,\text{km}$, what is the range of admissible zonal wavelengths?

9.6. Because the Coriolis parameter vanishes along the equator, it is usual in the study of tropical processes to write

$$f = \beta_0 y,$$

where y is the distance measured from the equator (positive northward). The linear wave equations then take the form

$$\frac{\partial u}{\partial t} - \beta_0 y v = -g \frac{\partial \eta}{\partial x} \tag{9.98}$$

$$\frac{\partial v}{\partial t} + \beta_0 y u = -g \frac{\partial \eta}{\partial y} \tag{9.99}$$

$$\frac{\partial \eta}{\partial t} + H \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0, \tag{9.100}$$

where u and v are the zonal and meridional velocity components, η is the surface displacement, g is gravity, and H is the ocean depth at rest. Explore the possibility of a wave traveling zonally with no meridional velocity. At which speed does this wave travel and in which direction? Is it trapped along the equator? If so, what is the trapping distance? Does this wave bear any resemblance to a midlatitude wave (f_0 not zero)?

9.7. Seek wave solutions to the nonhydrostatic system of equations with nonstrictly vertical rotation vector:

$$\frac{\partial u}{\partial t} - f v + f_* w = -\frac{1}{\rho_0} \frac{\partial p}{\partial x} \tag{9.101a}$$

$$\frac{\partial v}{\partial t} + f u = -\frac{1}{\rho_0} \frac{\partial p}{\partial y} \tag{9.101b}$$

$$\frac{\partial w}{\partial t} - f_* u = -\frac{1}{\rho_0} \frac{\partial p}{\partial z} \tag{9.101c}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \tag{9.101d}$$

The fluid is homogeneous ($\rho = 0$), inviscid ($v = 0$) and infinitely deep. Consider in particular the equivalent of the Kelvin wave ($u = 0$ at $x = 0$) and Poincaré waves.

9.8. Prove by using a local polar coordinate system that tidal forces derive from the tidal potential (9.93).

9.9. Estimate the average travel time for a gravity wave to circle the earth along the equator, assuming that there are no continents, and that the average depth of the ocean is 3800 m. Compare with the tidal period.

9.10. Based on the mass of the sun and its distance to the earth, how intense do you expect solar tides to be compared with lunar tides? At what period do the combined forces give rise to the strongest tides?

9.11. Knowing that a hurricane approaching Florida has a diameter of 100 km and wind-speeds U of 150 km/h, which storm surge height do you expect in a 10-m-deep coastal sea? Use the following wind-stress formula: $\tau = 10^{-6} \rho_0 U^2$.

9.12. Assuming the earthquake near Indonesia's Sumatra Island on 26 December 2004 generated a surface wave (tsunami) by an upward motion of the sea floor during 10 min, estimate the wavelength of the wave. For simplicity, assume a uniform depth $h = 4$ km. Estimate also the time available between the detection of the earthquake and the moment the tsunami reaches a coastline 4000 km away. If instead of a uniform depth, you use the depth profile $h(x)$ provided in sumatra.m, how would you estimate the travel time? Investigate under which conditions you can use the local wave speed of gravity waves over uneven topography. (*Hint*: Compare the wavelength with the length scale of topographic variations.)

9.13. In order to avoid the problem in Section 9.5 of an infinitely deep layer at large distances, assume now that the flow takes place in a channel of width L. How are the topographic waves modified by the presence of the lateral boundaries?

9.14. Consider an inertia-gravity wave of wavelength $\lambda = 2\pi/k$ on the f-plane and align the x-axis with the direction of propagation (i.e., $k_x = k$ and $k_y = 0$). Write the partial differential equations and solve them for u and η proportional to $\cos(kx - \omega t)$ and v proportional to $\sin(kx - \omega t)$. Then, calculate the kinetic and potential energies per unit horizontal area, defined as

$$KE = \frac{1}{\lambda} \int_0^\lambda \frac{1}{2} \rho_0 (u^2 + v^2) H \, dx \qquad (9.102a)$$

$$PE = \frac{1}{\lambda} \int_0^\lambda \frac{1}{2} \rho_0 g \eta^2 \, dx, \qquad (9.102b)$$

each in terms of the amplitude of η and show that the kinetic energy is always greater than the potential energy, except in the case $f = 0$ (pure gravity waves), in which case there is equipartition of energy.

9.15. Take a hurricane or typhoon from last summer season and note the pressure anomaly in its eye as it approached the coast. Determine the inverse barometric response at that time.

NUMERICAL EXERCISES

9.1. Establish the numerical stability condition of schemes (9.53) and (9.54). Can you provide an interpretation for the parameter $c\Delta t/\Delta x$? Compare with the CFL criterion.

9.2. Spell out the spatial discretization on the B-, C-, and D-grids of Eqs. (9.4a) through (9.5).

9.3. Implement the C-grid in Matlab for Eqs. (9.4a) through (9.5) with a variable fluid thickness given on a grid at the same location as η. Use a time discretization as in Eq. (9.54) and a fractional-step approach for the Coriolis term. Then, use your code to simulate a pure Kelvin wave for different values of $\Delta x/R$ and $k_x^2 R^2$ by initializing with the exact solution. (*Hint*: Use a periodic domain in the x-direction and a second impermeable boundary in the y-direction, to be justified, at $y=10R$. Start from `shallow.m`.)

9.4. Analyze the way geostrophic equilibrium is represented in discrete Fourier modes on the B- and C-grids.

9.5. Investigate group-velocity errors for the different Arakawa grids using the numerical dispersion relation given in Eq. (9.66). Use $\Delta x = \Delta y$ and distinguish two types of waves: $k_x \neq 0$, $k_y = 0$ and $k_x = k_y$. Vary the resolution by taking $R/\Delta x = 0.2, 1, 5$, where R is the deformation radius.

9.6. Design the ideal staggering strategy for a model in which the eddy viscosity ν_E is chosen proportional to $|\partial u/\partial z| l_m^2$, where l_m is a specified mixing length and velocity u is determined numerically from a governing equation that includes vertical turbulent diffusion.

9.7. Assume you need to calculate the vertical component of relative vorticity from a discrete velocity field provided on the two-dimensional C-grid. Where is the most natural node to calculate the relative vorticity? Can you see an advantage to using a D-grid here?

9.8. Can you think of possibilities to include bottom topographic variations as those inducing tsunamis in shallow-water equations?

9.9. Take the variable depth implementation of Numerical Exercise 9.3 and apply it to the following topography

$$h = H_0 + \Delta H \left[1 + \tanh\left(\frac{x - \frac{L}{2}}{D}\right) \right]$$

with $H_0 = 50\,\mathrm{m}$, $D = L/8$, and $\Delta H = 5H_0$. Use a solid wall at $x = 0$ and $x = L = 100\,\mathrm{km}$ and periodic boundary conditions in the y-direction across a domain of length $5L$. Start with zero velocities and a Gaussian sea surface elevation of width $L/4$ and height of $1\,\mathrm{m}$ in the center of the basin. Use linear bottom friction with friction coefficient $r = 10^{-4}\,\mathrm{m/s}$. Trace the evolution of the sea surface elevation for $f = 10^{-4}\,\mathrm{s}^{-1}$.

9.10. Perform a storm-surge simulation with the implementation of Numerical Exercise 9.3 by using a uniform wind stress over a square basin with a uniform topography and then with the topography given in Numerical Exercise 9.9. Use the quadratic law (9.77) for bottom friction.

William Thomson, Lord Kelvin
1824–1907
(Standing at right, in laboratory of Lord Rayleigh, left)
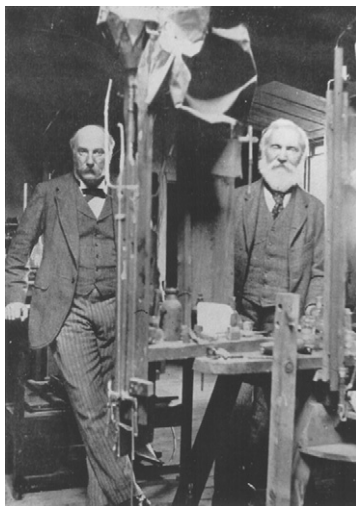
Named professor of natural philosophy at the University of Glasgow, Scotland, at age 22, William Thomson became quickly regarded as the leading inventor and scientist of his time. In 1892, he was named Baron Kelvin of Largs for his technological and theoretical contributions leading to the successful laying of a transatlantic cable. A friend of James P. Joule, he helped establish a firm theory of thermodynamics and first defined the absolute scale of temperature. He also made major contributions to the study of heat engines. With Hermann von Helmholtz, he estimated the ages of the earth and sun and ventured into fluid mechanics. His theory of the so-called Kelvin wave was published in 1879 (under the name William Thomson). His more than 300 original papers left hardly any aspect of science untouched. He is quoted as saying that he could understand nothing of which he could not make a model. (*Photo by A.G. Webster*)

Akio Arakawa
1927–2010

Akio Arakawa entered the Japanese Meteorological Agency in 1950 and received his doctorate at the University of Tokyo in 1961. He then went to the University of California in Los Angeles (UCLA) to pursue research, at a time when the atmospheric circulation computer models could reproduce weather-like motion but not for long. Beyond a two-week simulation, the computed patterns no longer looked like weather, and Arakawa's work demonstrated that the problem lied in the artificial generation of energy by inadequate numerical procedures. He also found the remedy. This remedy consisted of enforcing conservation of energy and of enstrophy (the square of vorticity) at the grid level.

The grids, which he proposed and later came to bear his name, were developed in the context of a study (Arakawa & Lamb, 1977) on the effects of grid topology on the dispersion of inertia-gravity waves. Arakawa's legacy to the science of weather prediction by computer modeling is significant and enduring. (*Photo credit: Akio Arakawa*)