10
Long Waves and Ocean Tides

Myrl C. Hendershott

10.1 Introduction

The main purpose of this chapter is to summarize what was generally known to oceanographers about long waves and ocean tides around 1940, and then to indicate how the subject has developed since then, with particular emphasis upon those aspects that have had significance for oceanography beyond their importance in understanding tides themselves. I have begun with a description of astronomical and radiational tide-generating potentials (section 10.2), but say no more than is necessary to make this chapter self-contained. Cartwright (1977) summarizes and documents recent developments, and I have followed his discussion closely.

The fundamental dynamic equations governing tides and long waves, Laplace's tidal equations (LTE), remained unchanged and unchallenged from Laplace's formulation of them in 1776 up to the early twentieth century. By 1940 they had been extended to allow for density stratification (in the absence of bottom relief) and criticized for their exclusion of half of the Coriolis forces. Without bottom relief this exclusion has recently been shown to be a good approximation; the demonstration unexpectedly requires the strong stratification of the ocean. Bottom relief appears able to make long waves in stratified oceans very different from their flat-bottom counterparts (section 10.4); a definitive discussion has not yet been provided. Finally, LTE have had to be extended to allow for the gravitational self-attraction of the oceans and for effects due to the tidal yielding of the solid earth. I review these matters in section 10.3.

Laplace's study of the free oscillations of a global ocean governed by LTE was the first study of oceanic long waves. Subsequent nineteenth- and twentieth-century explorations of the many free waves allowed by these equations, extended to include stratification, have evolved into an indispensible part of geophysical fluid dynamics. By 1940, most of the flat-bottom solutions now known had, at least in principle, been constructed. But Rossby's rediscovery and physical interpretation, in 1939, of Hough's oscillations of the second class began the modern period of studying solutions of the long-wave equations by inspired or systematic approximation and of seeking to relate the results to nontidal as well as tidal motions. Since then, flat-bottom barotropic and baroclinic solutions of LTE have been obtained in mid-latitude and in equatorial approximation, and Laplace's original global problem has been completely solved. The effects of bottom relief on barotropic motion are well understood. Significant progress has been made in understanding the effects of bottom relief on baroclinic motions. I have attempted to review all those developments in a self-contained manner in section 10.4. In order to treat this

vast subject coherently, I have had to impose my own view of its development upon the discussion. I have cited observations when they appear to illustrate some property of the less familiar solutions, but the central theme is a description of the properties of theoretically possible waves of long period (greater than the buoyancy period) and, consequently, of length greater than the ocean's mean depth.

Although the study of ocean surface tides was *the* original study of oceanic response to time-dependent forcing, tidal studies have largely proceeded in isolation from modern developments in oceanography on account of the strength of the tide-generating forces, their well-defined discrete frequencies, and the proximity of these to the angular frequency of the earth's rotation. A proper historical discussion of the subject, although of great intellectual interest, is beyond the scope of this chapter. To my mind the elements of such a discussion, probably reasonably complete through the first decade of this century, are given in Darwin's 1911 *Encyclopedia Britannica* article "Tides." Thereafter, with a few notable exceptions, real progress had to await modern computational techniques both for solving LTE and for making more complete use of tide gauge observations. Cartwright (1977) has recently reviewed the entire subject, and therefore I have given a discussion in section 10.5 that, although self-contained, emphasizes primarily changes of motivation and viewpoint in tidal studies rather than recapitulates Cartwright's or other recent reviews.

This discussion of tides as long waves continuously forced by lunar and solar gravitation logically could be followed by a discussion of tsunamis impulsively forced by submarine earthquakes. But lack of both space and time has forced omission of this topic.

Internal tides were first reported at the beginning of this century. By 1940 a theoretical framework for their discussion had been supplied by the extension of LTE to include stratification, and their generation was (probably properly) ascribed to scattering of barotropic tidal energy from bottom relief. The important developments since then are recognition of the intermittent narrow-band nature of internal tides (as opposed to the near-line spectrum of surface tides) plus the beginnings of a statistically reliable characterization of the internal tidal spectrum and its variation in space and time. The subject has recently been reviewed by Wunsch (1975). Motivation for studying internal tides has shifted from the need for an adequate description of them through exploration of their role in global tidal dissipation (now believed to be under 10%) to speculation about their importance as energy sources for oceanic mixing. In section 10.6 I have summarized modern observational studies and their implications for tidal mixing of the oceans.

Many features of the presentday view of ocean circulation have some precedent in tidal and long-wave studies, although often unacknowledged and apparently not always recognized. The question of which parts of the study of tides have in fact influenced the subsequent development of studies of ocean circulation is a question for the history of science. In some cases, developments in the study of ocean circulation subsequently have been applied to ocean tides. In section 10.7 I have pointed out some of the connections of which I am aware.

10.2 Astronomical Tide-Generating Forces

Although correlations between ocean tides and the position and phase of the moon have been recognized and utilized since ancient times, the astronomical tide-generating force (ATGF) was first explained by Newton in the *Principia* in 1687. Viewed in an accelerated coordinate frame that moves with the center of the earth but that does not rotate with respect to the fixed stars, the lunar (solar) ATGF at any point on the earth's surface is the difference between lunar (solar) gravitational attraction at that point and at the earth's center. The daily rotation of the earth about its axis carries a terrestrial observer successively through the longitude of the sublunar or subsolar point [at which the lunar (solar) ATGF is toward the moon (sun)] and then half a day later through the longitude of the antipodal point [at which the ATGF is away from the moon (sun)]. In Newton's words, "It appears that the waters of the sea ought twice to rise and twice to fall every day, as well lunar or solar" [Newton, 1687, proposition 24, theorem 19].

The ATGF is thus predominantly semidiurnal with respect to both the solar and the lunar day. But it is not entirely so. Because the tide-generating bodies are not always in the earth's equatorial plane, the terrestrial observer [who does not change latitude while being carried through the longitude of the sublunar (solar) point or its antipode] sees a difference in amplitude between the successive semidiurnal maxima of the ATGF at his location. This difference or "daily inequality" means that the ATGF must be thought of as having diurnal as well as semidiurnal time variation.

Longer-period variations are associated with periodicities in the orbital motion of earth and moon. The astronomical variables displaying these long-period variations appear nonlinearly in the ATGF. The long-period orbital variations thus interact nonlinearly both with themselves and with the short-period diurnal and semidiurnal variation of the ATGF to make the local ATGF a sum of three narrow-band processes centered about 0, 1, and 2 cycles per day (cpd), each process being a sum of motions harmonic at multiples 0,1,2 of the frequencies corresponding to a lunar or a solar day

plus sums of multiples of the frequencies of long-period orbital variations.

A complete derivation of the ATGF is beyond the scope of this discussion. Cartwright (1977) reviews the subject and supplies references documenting its modern development. For a discussion that concentrates upon ocean dynamics (but not necessarily for a practical tide prediction), the most convenient representation of the ATGF is as a harmonic decomposition of the tide-generating potential whose spatial gradient is the ATGF. Because only the horizontal components of the ATGF are of dynamic importance, it is convenient to represent the tide-generating potential by its horizontal and time variation U over some near-sea-level equipotential (the geoid) of the gravitational potential due to the earth's shape, internal mass distribution, and rotation. To derive the dynamically significant part of the ATGF it suffices to assume this surface spherical. U/g (where g is the local gravitational constant, unchanging over the geoid) has the units of sea-surface elevation and is called the *equilibrium tide* $\bar{\zeta}$. Its principle term is (Cartwright, 1977)

$$\bar{\zeta}(\phi,\theta,t) = U(\phi,\theta,t)/g$$

$$= \sum_{m=0,1,2} (A_2^m \cos m\phi + B_2^m \sin m\phi) P_2^m, \qquad (10.1)$$

in which ϕ,θ are longitude and latitude, the $P_2^m(\bar{\theta})$ are associated Legendre functions

$$P_2^0 = \tfrac{1}{3}(3\cos^2\bar{\theta} - 1),$$
$$P_2^1 = 3\sin\bar{\theta}\cos\bar{\theta}, \qquad (10.2)$$
$$P_2^2 = 3\sin^2\bar{\theta}$$

of colatitude $\bar{\theta} = (\pi/2) - \theta$, and A_2^m, B_2^m are functions of time having the form

$$A_n^m(t) = \sum_i M_i \begin{matrix}\cos\\ \sin\end{matrix}\left[\sum_{j=1}^{6} N_j^{(i)} S_j(t)\right]. \qquad (10.3)$$

The M_i are amplitudes obtained from Fourier analysis of the astronomically derived time series $U(\phi,\theta,t)/g$; the $N_j^{(i)}$ are sets of small integers (effectively the Doodson numbers); and the $S_j(t)$ are secular arguments that increase almost linearly in time with the associated periodicity of a lunar day, a sidereal month, a tropical year, 8.847 yr (period of lunar perigee), 18.61 yr (period of lunar node), 2.1×10^4 yr (period of perihelion), respectively.

The frequencies of the arguments $\Sigma N_j^{(i)} S_j(t)$ fall into the three "species"—long period, diurnal, and semidiurnal—which are centered, respectively, about 0, 1, and 2 cpd ($N_1 = 0,1,2$). Each species is split into "groups" separated by about 1 cycle per month, groups are split into "constituents" separated by one cyle per year, etc. Table 10.1 lists selected constituents. In the following discussion they are referred to by their Darwin symbol (see table 10.1).

An important development in modern tidal theory has been the recognition that the ATGF is not the only important tide-generating force. Relative to the amplitudes and phases of corresponding constituents of the equilibrium tide, solar semidiurnal, diurnal, and annual ocean tides usually have amplitudes and phases quite different from the amplitudes and phases of other nearby constituents. Munk and Cartwright (1966) attributed these anomalies in a general way to solar heating and included them in a generalized equilibrium tide by defining an ad hoc radiational potential (Cartwright, 1977)

$$U_R(\phi,\theta,t) = \begin{cases} S(\xi/\bar{\xi})\cos\alpha, & 0 < \alpha < \pi/2: \\ 0, & \pi/2 < \alpha < \pi: \end{cases} \qquad (10.4)$$

which is zero at night, which varies as the cosine of the sun's zenith angle α during the day, and which is proportional to the solar constant S and the sun's parallax ξ (mean $\bar{\xi}$). Cartwright (1977) suggests that for the oceanic S_2 (principal solar) tide, whose anomalous portion is about 17% of the gravitational tide (Zetler, 1971), the dominant nongravitational driving is by the atmospheric S_2 tide. Without entering further into the discussion, I want to point out that the global form of

Table 10.1 Characteristics of Selected Constituents of the Equilibrium Tide

Darwin symbol	N_1, N_2, N_3, N_4	Period (solar days or hours)	Amplitude M (m)	Spatial variation
S_{sa}	0 0 2 0	182.621 d	0.02936	
M_m	0 1 0 −1	27.55 d	0.03227	$\frac{1}{3}(3\cos^2\bar{\theta} - 1)$
M_f	0 2 0 0	13.661 d	0.0630	
O_1	1 −1 0 0	25.82 h	0.06752	
P_1	1 1 −2 0	24.07 h	0.03142	$3\sin\bar{\theta}\cos\bar{\theta} \times \sin(\omega_1 t + \phi)$
K_1	1 1 0 0	23.93 h	0.09497	
N_2	2 −1 0 1	12.66 h	0.01558	
M_2	2 0 0 0	12.42 h	0.08136	$3\sin^2\bar{\theta} \times \cos(\omega_2 t + 2\phi)$
S_2	2 2 −2 0	12.00 h	0.03785	
K_2	2 2 0 0	11.97 h	0.01030	

the atmospheric S_2 tide is well known (Chapman and Lindzen, 1970), so that the same numerical programs that have been used to solve for global gravitationally driven ocean tides could easily be extended to allow for atmospheric-pressure driving of the oceanic S_2 tide.

Ocean gravitational self-attraction and tidal solid-earth deformation are quantitatively even more important in formulating the total tide-generating force than are thermal and atmospheric effects. They are discussed in the following section, since they bring about a change in the form of the dynamic equations governing ocean tides.

10.3 Laplace's Tidal Equations (LTE) and the Long-Wave Equations

Laplace (1775, 1776; Lamb, 1932, §213-221) cast the dynamic theory of tides essentially in its modern form. His tidal equations (LTE) are usually formally obtained from the continuum equations of momentum and mass conservation (written in rotating coordinates for a fluid shell surrounding a nearly spherical planet and having a gravitationally stabilized free surface) by assuming (Miles, 1974a)

(1) a perfect homogeneous fluid,
(2) small disturbances relative to a state of uniform rotation,
(3) a spherical earth,
(4) a geocentric gravitational field uniform horizontally and in time,
(5) a rigid ocean bottom,
(6) a shallow ocean in which both the Coriolis acceleration associated with the horizontal component of the earth's rotation and the vertical component of the particle acceleration are neglected.

The resulting equations are

$$\frac{\partial u}{\partial t} - 2\Omega \sin\theta v = -\frac{\partial}{\partial \phi}(\zeta - \Gamma/g)/a\cos\theta, \tag{10.5a}$$

$$\frac{\partial v}{\partial t} + 2\Omega \sin\theta u = -\frac{\partial}{\partial \theta}(\zeta - \Gamma/g)/a, \tag{10.5b}$$

$$\frac{\partial \zeta}{\partial t} + \frac{1}{a\cos\theta}\left[\frac{\partial}{\partial \phi}(uD) + \frac{\partial}{\partial \theta}(vD\cos\theta)\right] = 0. \tag{10.5c}$$

In these, (ϕ, θ) are longitude and latitude with corresponding velocity components (u, v), ζ the ocean surface elevation, Γ the tide-generating potential, $D(\phi, \theta)$ the variable depth of the ocean, a the earth's spherical radius, g the constant gravitational attraction at the earth's surface, and Ω the earth's angular rate of rotation.

Two modern developments deserve discussion. They are a quanitative formulation and study of the mathematical limit process implicit in assumptions (1) through (6), and the realization that assumptions (4) and (5) are quantitatively inadequate for a dynamic discussion of ocean tides.

It has evidently been recognized since the work of Bjerknes, Bjerknes, Solberg, and Bergeron (1933) that assumption (6) (especially the neglect of Coriolis forces due to the horizontal component of the earth's rotation) amounts to more than a minor perturbation of the spectrum of free oscillations that may occur in a thin homogeneous ocean. Thus Stern (1963) and Israeli (1972) found axisymmetric equatorially trapped normal modes of a rotating spherical shell of homogeneous fluid that are extinguished by the hydrostatic approximation. Indeed, Stewartson and Rickard (1969) point out that the limiting case of a vanishingly thin homogeneous ocean is a nonuniform limit: the solutions obtained by solving the equations and then taking the limit may be very different from those obtained by first taking the limit and then solving the resulting approximate (LTE) equations. Quite remarkably, it is the reinstatement of realistically large stratification, i.e., the relaxation of assumption (1), that saves LTE as an approximate set of equations whose solutions are uniformly valid approximations to some of the solutions of the full equations when the ocean is very thin.

The parodoxical importance of stratification for the validity of the ostensibly unstratified LTE appears to have been recognized by Proudman (1948) and by Bretherton (1964). Phillips (1968) pointed out its importance at the conclusion of a correspondence with Veronis (1968b) concerning the effects of the "traditional" approximation (Eckart, 1960; N. A. Phillips, 1966b), i.e., the omission of the Coriolis terms $2\Omega\cos\theta w$ and $-2\Omega\cos\theta u$, in (10.6)-(10.8) below. But it was first explicitly incorporated into the limit process producing LTE by Miles (1974a) who addressed all of assumptions (1) through (6) by defining appropriate small parameters and examining the properties of expansions in them. He found that the simplest set of uniformly valid equations for what I regard as long waves in this review are

$$\frac{\partial u}{\partial t} - 2\Omega\sin\theta v + 2\Omega\cos\theta w = -\frac{\partial p}{\partial \phi}\frac{1}{\bar{\rho}_0 a\cos\theta}, \tag{10.6a}$$

$$\frac{\partial v}{\partial t} + 2\Omega\sin\theta u = -\frac{\partial p}{\partial \theta}\frac{1}{\bar{\rho}_0 a}, \tag{10.6b}$$

$$N^2 w - 2\Omega\cos\theta\frac{\partial u}{\partial t} = -\frac{\partial^2 \rho}{\partial z\,\partial t}\frac{1}{\bar{\rho}_0}, \tag{10.6c}$$

$$\frac{\partial u}{\partial \phi} + \frac{\partial(v\cos\theta)}{\partial \theta} + a\cos\theta\frac{\partial w}{\partial z} = 0, \tag{10.6d}$$

with boundary conditions

$$w = 0 \quad \text{at} \quad z = -D_* \quad \text{(for uniform depth } D_*\text{)}, \tag{10.7}$$

$$p = -\bar{\rho}_0 g \zeta_0 \qquad \text{at} \quad z = 0. \tag{10.8}$$

In these, z is the upward local vertical with associated velocity component w, ρ the deviation of the pressure from a resting hydrostatic state characterized by the stable density distribution $\rho_0(z)$, $\bar{\rho}_0$ a constant characterizing the mean density of the fluid (the Boussinesq approximation has been made in its introduction); the buoyancy frequency

$$N(z) = \left(-\frac{g}{\rho_0}\frac{\partial \rho_0}{\partial z} - \frac{g^2}{c^2} \right)^{1/2} \tag{10.9}$$

is the only term in which allowance for compressibility (c is the local sound speed) is important in the ocean.

Miles (1974a) further found that when $N^2(z) \gg 4\Omega^2$, free solutions of LTE for a uniform-depth (D_*) ocean covering the globe also solve (10.6)–(10.8) with an error which is of order $(\sigma/2\Omega)(4\Omega^2 a/g) \ll 1$, where σ is the frequency of oscillation of the free solutions. LTE surface elevations ζ are consequently in error by order $(D_*/a)(D_* N^2/g)^{-3/4}$, while LTE velocities u,v may be in error by $(D_*/a)(D_* N^2/g)^{-1/4}$. Miles (1974a) obtained this result by taking the (necessarily barotropic) solutions of LTE as the first term of an expansion of the solutions of (10.6)–(10.8) in the parameter $(\sigma/2\Omega)(4\Omega^2 a/g)$, which turns out to characterize the relative importance of terms neglected and terms retained in making assumption (6). The next term in the expansion consists of internal wave modes (see section 10.4.3). Because their free surface displacements are very small relative to internal displacements, the overall free surface displacement remains as in LTE although in the interior of the ocean, internal wave displacements and currents may well dominate the motion.

The analysis is inconclusive at frequencies or depths at which the terms assumed to be correction perturbations are resonant. Finding expressions for all the free oscillations allowed by Miles's simplest uniformly valid system (10.6)–(10.8) involves as yet unresolved mathematical difficulties associated with the fact that these equations are hyperbolic over part of the spatial domain when the motion is harmonic in time (Miles, 1974a).

Application of this analysis to ocean tides is further circumscribed by its necessary restriction to a global ocean of constant depth. I speculate that if oceanic internal modes are sufficiently inefficient as energy transporters that they cannot greatly alter the energetics of the barotropic solution unless their amplitudes are resonantly increased beyond observed levels (section 10.6), and if they are sufficiently dissipative that they effectively never are resonant, then an extension of this analysis to realistic basins and relief would probably confirm LTE as adequate governors of the surface elevation. The ideas, necessary for such an extension, that is, how variable relief and stratification influence barotropic and baroclinic modes, are beginning to be developed (see section 10.4.7).

Miles (1974a) discusses assumptions (3)–(5) with explicit omission of ocean gravitational self-attraction and solid-earth deformation. Self-attraction was included in Hough's (1897, 1898; Lamb, 1932, §222–223) global solutions of LTE. Thomson (1863) evidently first pointed out the necessity of allowing for solid-earth deformation. Both are quantitatively important. The latter manifests itself in a geocentric solid-earth tide δ plus various perturbations of the total tide-generating potential Γ. Horizontal pressure gradients in LTE (10.5) are associated with gradients of the geocentric ocean tide ζ, but it is the observed ocean tide

$$\zeta_0 = \zeta - \delta \tag{10.10}$$

that must appear in the continuity equation of (10.5c).

All these effects are most easily discussed (although not optimally computed) when the astronomical potential U, the observed ocean tide ζ_0, the solid earth tide δ, and the total tide-generating potential Γ are all decomposed into spherical harmonic components U_n, ζ_{0n}, δ_n, and Γ_n. The Love numbers k_n, h_n, k'_n, h'_n, which carry with them information about the radial structure of the solid earth (Munk and McDonald, 1960), and the parameter $\alpha_n = (3/2n + 1)(\bar{\rho}_{\text{ocean}}/\bar{\rho}_{\text{earth}})$ then appear naturally in the development. The total tide-generating potential Γ_n contains an astronomical contribution U_n (primarily of order $n = 2$), an augmentation $k_n U_n$ of this due to solid earth yielding to $-\nabla U_n$, an ocean self-attraction contribution $g\alpha_n\zeta_{0n}$, and a contribution $k'_n g\alpha_n\zeta_{0n}$ due to solid-earth deformation by ocean self-attraction and tidal column weight. Thus

$$\Gamma_n = (1 + k_n)U_n + (1 + k'_n)g\alpha_n\zeta_{0n}. \tag{10.11}$$

There is simultaneously a geocentric solid-earth tide δ made up of the direct yielding $h_n U_n g$ of the solid earth to $-\nabla U_n$ plus the deformation $h'_n g\alpha_n\zeta_{0n}$ of the solid earth by ocean attraction and tidal column weight. Thus

$$\delta_n = h_n U_n/g + h'_n g\alpha_n\zeta_{0n}. \tag{10.12}$$

For computation, Farrell (1972a) has constructed a Green's function such that

$$\sum_n (1 + k'_n - h'_n)\alpha_n\zeta_{0n} = \iint_{\text{ocean}} d\theta'\, d\phi' \cos\theta'\, G(\phi',\theta'|\phi,\theta)\zeta_0(\phi',\theta'). \tag{10.13}$$

With $U_n = U_2$ and with (10.13) abbreviated as $\iint G\zeta_0$, LTE with assumptions (4) and (5) appropriately relaxed become

$$\frac{\partial u}{\partial t} - 2\Omega \sin\theta v = -\frac{g}{a\cos\theta}\frac{\partial}{\partial\phi}[\zeta_0 - (1 + k_2 - h_2)U_2/g - \iint G\zeta_0], \tag{10.14a}$$

$$\frac{\partial v}{\partial t} + 2\Omega \sin\theta u = -\frac{g}{a}\frac{\partial}{\partial\theta}[\zeta_0 - (1 + k_2 - h_2)U_2/g - \iint G\zeta_0], \tag{10.14b}$$

$$\frac{\partial \zeta_0}{\partial t} + \frac{1}{a\cos\theta}\left[\frac{\partial(uD)}{\partial\phi} + \frac{\partial(vD\cos\theta)}{\partial\theta}\right] = 0. \tag{10.14c}$$

The factor $1 + k_2 - h_2 \simeq 0.69$ is clearly necessary for a quantitatively correct solution. The term $\iint G\zeta_0$ was first evaluated by Farrell (1972b). I wrote down (10.14) and attempted to estimate the effect of $\iint G\zeta_0$ on a global numerical solution of (10.14) for the M_2 tide by an iterative procedure which, however, failed to converge (Hendershott, 1972). Subsequent computations by Gordeev, Kagan, and Polyakov (1977) and by Accad and Pekeris (1978) provide improved estimates of the effects (see section 10.5.3).

With appropriate allowance for various dissipative processes (including all mechanisms that put energy into internal tides), I regard (10.14) as an adequate approximation for studying the ocean surface tide. Oceanic long waves should really be discussed using (10.6) but allowing for depth variations by putting

$$w = \mathbf{u}\cdot\nabla D \quad \text{at} \quad z = -D(\phi,\theta) \tag{10.15}$$

in place of (10.7). Miles (1974a) derives an orthogonality relationship that could be specialized to (10.6)–(10.8) in the case of constant depth, but even then the nonseparability of the eigenfunctions into functions of (ϕ,θ) times functions of z has prevented systematic study of the problem. Variable relief compounds the difficulty. Most studies either deal with surface waves over bottom relief, and thus start with LTE (10.5), or else with surface and internal waves over a flat bottom. In the latter case, (10.6)–(10.8) are solved but with the Coriolis terms $2\Omega\cos\theta w$ and $-2\Omega\cos\theta u$ arbitrarily neglected (the "traditional" approximation). Miles' (1974a) results appear to justify this procedure for the former case, but Munk and Phillips (1968) show that the neglected terms are proportional to (mode number)$^{1/3}$ for internal modes so that the traditional approximation may be untenable for high-mode internal waves. The following discussion (section 10.4) of oceanic long waves relies heavily upon the traditional approximation, but it is important to note that its domain of validity has not yet been entirely delineated.

10.4 Long Waves in the Ocean

10.4.1 Introduction

The first theoretical study of oceanic long waves is due to Laplace (1775, 1776; Lamb, 1932, §213–221), who solved LTE for an appropriately shallow ocean covering a rotating rigid spherical earth by expanding the solution in powers of $\sin\theta$. For a global ocean of constant depth D_*, Hough (1897, 1898; Lamb, 1932, §222–223) obtained solutions converging rapidly for small

$$\Lambda = 4\Omega^2 a^2/gD_* \tag{10.16}$$

(sometimes called Lamb's parameter) by expanding the solution in spherical harmonics $P_n^l(\sin\theta)\exp(il\phi)$. He found the natural oscillations to be divided into first- and second-class modes whose frequencies σ are given, respectively, by

$$\sigma = \pm[n(n+1)gD_*/a^2]^{1/2},$$
$$\sigma = -2\Omega l/[n(n+1)] \tag{10.17}$$

as $\Lambda \to 0$. But $\Lambda \simeq 20$ for $D_* = 4000$ m, and is very much larger for internal waves (see section 10.4.3). A correspondingly complete solution of Laplace's problem, valid for large as well as small Λ, was given only recently (Flattery, 1967; Longuet-Higgins, 1968a). Physical understanding of the solutions has historically been developed by studying simplified models of LTE.

10.4.2 Long Waves in Uniformly Rotating Flat-Bottomed Oceans

Lord Kelvin (Thomson, 1879; Lamb, 1932, §207) introduced the idealization of uniform rotation, at Ω, of a sheet of fluid about the vertical (z-axis). LTE become

$$\frac{\partial u}{\partial t} - f_0 v = -g\frac{\partial \zeta}{\partial x}, \tag{10.18a}$$

$$\frac{\partial v}{\partial t} + f_0 u = -g\frac{\partial \zeta}{\partial y}, \tag{10.18b}$$

$$\frac{\partial \zeta}{\partial t} + \frac{\partial(uD)}{\partial x} + \frac{\partial(vD)}{\partial y} = 0. \tag{10.18c}$$

f_0, here equal to 2Ω, is the Coriolis parameter. Lord Kelvin's plane model (10.18) is often called the f-plane.

The solid earth is nearly a spheroid of equilibrium under the combined influence of gravity g and centrifugal acceleration $\Omega^2 a$; the earth's equatorial radius is about 20 km greater than its polar radius. Without water motion, the sea surface would have a congruent spheroidal shape. Taking the depth constant in LTE models this similarity; the remaining error incurred by working in spherical rather than spheroidal coordinates is (Miles, 1974a) of order $\Omega^2 a/g = 10^{-3}$. It is correspondingly appropriate to take the depth constant in Lord Kelvin's plane model (10.18) in order to obtain planar

solutions locally modeling those of LTE with constant depth. But the laboratory configuration corresponding to constant depth in (10.18c) is a container with a paraboloidal bottom $\frac{1}{2}\Omega^2(x^2 + y^2)/g$ rather than a flat bottom [see Miles (1964) for a more detailed analysis].

Without rotation ($f_0 = 0$) and with constant depth ($D = D_*$) Lord Kelvin's plane model reduces to the linearized shallow-water equations (LSWE). For an unbounded fluid sheet, they have plane gravity-wave solutions

$$\zeta = a\exp(-i\sigma t + ilx + iky), \tag{10.19a}$$

$$u = (gl/\sigma)\zeta, \tag{10.19b}$$

$$v = (gk/\sigma)\zeta, \tag{10.19c}$$

$$w = -i\sigma(z/D_* + 1)\zeta, \tag{10.19d}$$

$$\sigma^2 = gD_*(l^2 + k^2), \tag{10.19e}$$

which are dispersionless [all travel at $(gD_*)^{1/2}$] and are longitudinal [(u, v) parallel to (l, k)]. Horizontal particle accelerations are exactly balanced by horizontal pressure-gradient forces while vertical accelerations are negligible. Such waves reflect specularly at a straight coast with no phase shift; thus

$$\zeta = a\exp(-i\sigma t + ilx + iky) + a\exp(-i\sigma t - ilx + iky) \tag{10.20}$$

satisfies $u = 0$ at the coast $x = 0$, the angle $\tan^{-1}(k/l)$ of incidence equals the angle of reflection, and the (complex) reflected amplitude equals the incident amplitude a. The normal modes of a closed basin with perimeter P are the eigensolutions $Z_n(x,y)\exp(-i\sigma_n t)$, σ_n of

$$\nabla_2^2 Z_n + (\sigma_n^2/gD_*)Z_n = 0 \tag{10.21}$$

with

$$\partial Z_n/\partial(\text{normal}) = 0 \qquad \text{on} \quad P. \tag{10.22}$$

If P is a constant surface in one of the coordinate systems in which ∇_2^2 separates, then the normal modes are mathematically separable functions of the two horizontal space coordinates and so are readily discussed in terms of appropriate special functions. Even in general basin shapes, the existence and completeness of the normal modes are assured (Morse and Feshbach, 1953).

With rotation, plane wave solutions of (10.18) with constant depth D_* are

$$\zeta = a\exp(-i\sigma t + ilx + iky), \tag{10.23a}$$

$$u = g[(l\sigma + ikf_0)/(\sigma^2 - f_0^2)]\zeta, \tag{10.23b}$$

$$v = g[(k\sigma - ilf_0)/(\sigma^2 - f_0^2)]\zeta, \tag{10.23c}$$

$$w = -i\sigma(z/D_* + 1)\zeta, \tag{10.23d}$$

$$\sigma^2 = gD_*(l^2 + k^2) + f_0^2. \tag{10.23e}$$

These are often called Sverdrup waves (apparently after Sverdrup, 1926). Rotation has made them dispersive and they propagate only when $\sigma^2 > f_0^2$. The group velocity $\mathbf{c}_g = (\partial\sigma/\partial l, \partial\sigma/\partial k)$ is parallel to the wavenumber and rises from zero at $\sigma = f_0$ toward $(gD_*)^{1/2}$ as $\sigma^2 \gg f_0^2$. Dynamically these waves are LSW waves perturbed by rotation. Particle paths are ellipses with ratio f/σ of minor to major axis and with major axis oriented along (l, k). Particles traverse these paths in the clockwise direction (viewed from above) when $f > 0$ (northern hemisphere). Sverdrup waves are reflected specularly at a straight coast but with a phase shift; the sum

$$\zeta = a\exp(i\sigma t + ilx + iky) + a[(l\sigma + ikf_0)/(l\sigma - ikf_0)] \times \exp(-i\sigma t - ilx + iky) \tag{10.24}$$

of two Sverdrup waves satisfies $u = 0$ at the coast $x = 0$, the angle $\tan^{-1}(k/l)$ of incidence equals the angle of reflection, and the reflected amplitude differs from the incident amplitude by the multiplicative constant $(l\sigma + ikf_0)/(l\sigma - ikf_0)$, which is complex but of modulus unity. The sum (10.24) is often called a *Poincaré wave*. The normal modes of a closed basin with perimeter P are the eigensolutions $Z_n(x,y)\exp(-i\sigma_n t)$, σ_n of

$$\nabla_2^2 Z_n + [(\sigma_n^2 - f_0^2)/gD_*]Z_n = 0 \tag{10.25}$$

with

$$-i\sigma_n \partial Z_n/\partial(\text{normal}) + f_0 \partial Z_n/\partial(\text{tangent}) = 0 \qquad \text{at} \quad P. \tag{10.26}$$

On account of the boundary condition (10.26), they are not usually separable functions of the two horizontal space coordinates. The circular basin (Lamb, 1932, §209-210) is an exception. Rao (1966) discusses the rectangular basin, but the results are not easily summarized. A salient feature, the existence of free oscillations with $\sigma^2 < f_0^2$, is rationalized below.

With rotation, Lord Kelvin (Thomson, 1879; Lamb, 1932, §208) showed that a coast not only reflects Sverdrup waves for which $\sigma^2 > f_0^2$, but makes possible a new kind of coastally trapped motion for which $\sigma^2 \gtrless f_0^2$. This Kelvin wave has the form

$$\zeta = a\exp[-i\sigma t + iky + k(f_0/\sigma)x], \tag{10.27a}$$

$$u = 0, \tag{10.27b}$$

$$v = (gk/\sigma)\zeta, \tag{10.27c}$$

$$w = -i\sigma(z/D_* + 1)\zeta, \tag{10.27d}$$

$$\sigma^2 = gD_*k^2 \tag{10.27e}$$

along the straight coast $x = 0$. The velocity normal to the coast vanishes everywhere in the fluid and not only

at the coast. The wave is dispersionless and propagates parallel to the shore with speed $(gD_*)^{1/2}$ for $\sigma^2 \gtrless f_0^2$ just like a longitudinal gravity wave but with an offshore profile $\exp[k(f_0/\sigma)x]$ that decays or grows exponentially seaward depending upon whether the wave propagates with the coast to its right or to its left (in the northern hemisphere, $f_0 > 0$). For vanishing rotation, the offshore decay or growth scale becomes infinite and the Kelvin wave reduces to an ordinary gravity wave propagating parallel to the coast. The Kelvin wave is dynamically exactly a LSW gravity wave in the longshore direction and is exactly geostrophic in the cross-shore direction.

A pair of Kelvin waves propagating in opposite directions along the two coasts of an infinite canal (at, say, $x = 0$ and $x = W$) gives rise to a pattern of sea-level variation in which the nodal lines that would occur without rotation shrink to amphidromic points, at which the surface neither rises nor falls and about which crests and troughs rotate counterclockwise (in the northern hemisphere) as time progresses. For equal-amplitude oppositely propagating Kelvin waves, the amphidromic points fall on the central axis of the canal and are separated by a half-wavelength (πk^{-1}). When the amplitudes are unequal, the line of amphidromes moves away from the coast along which the highest-amplitude Kelvin wave propagates. For a sufficiently great difference in amplitudes, the amphidromes may occur beyond one of the coasts, i.e., outside of the canal.

Such a pair of Kelvin waves cannot by themselves satisfy the condition of zero normal fluid velocity in a closed canal (say at $y = 0$). Taylor (1921) showed how this condition could be satisfied by adjoining to the pair of Kelvin waves an infinite sum of channel Poincaré modes

$$\zeta = [\cos(m\pi x/W) - (f_0/\sigma)(kW/m\pi) \times \sin(m\pi x/W)] \exp[-i\sigma t \pm iky], \tag{10.28}$$
$$\sigma^2 = (m^2\pi^2/W^2 + k^2)gD_* + f_0^2, \qquad m = 1, 2, \ldots,$$

each of which separately has vanishing normal fluid velocity at the channel walls ($x = 0, W$). These are just the waveguide modes of the canal. Mode m decays exponentially away from the closure (is evanescent) if

$$\sigma^2 < f_0^2 + (m^2\pi^2/W^2)gD_*. \tag{10.29}$$

If σ is so low, W so small, or D so great that all modes $m = 1, 2, \ldots$ are evanescent, then the Kelvin wave incident on the closure has to be perfectly reflected with at most a shift of phase. When (10.29) is violated for the one or more lowest modes, then some of the energy of the incident Kelvin wave is scattered into traveling Poincaré modes.

If (10.29) is satisfied for all m and if the decay scale $(gD_*)^{1/2}/f_0$ of the Kelvin waves is a good deal smaller than the channel width W, then the Poincaré modes sum to an appreciable contribution only near the corners of the closure. The Kelvin wave then proceeds up the channel effectively hugging one coast, turns the corners of the closure with a phase shift [evaluated by Buchwald (1968) for a single corner], and returns back along the channel hugging the opposite coast. Now it becomes apparent that, with allowance for corner phase shifts, closed rotating basins have a class of free oscillation whose natural frequencies are effectively determined by fitting an integral number of Kelvin waves along the basin perimeter. Such free oscillations may have $\sigma^2 \lesseqgtr f_0^2$. They are readily identified in Lamb's (1932, §209–210) normal modes of a uniform-depth circular basin.

10.4.3 The Effect of Density Stratification on Long Waves

All of the foregoing solutions are barotropic surface waves. Stokes (1847; Lamb, 1932, §231) pointed out that surface waves are dynamically very much like waves at the interfaces between fluid layers of differing densities. Allowance for continuous vertical variation of density was made by Rayleigh (1883). Lord Kelvin's plane model (10.18) must be extended to read

$$\frac{\partial u}{\partial t} - f_0 v = -\frac{1}{\bar{\rho}_0}\frac{\partial p}{\partial x}, \tag{10.30a}$$

$$\frac{\partial v}{\partial t} + f_0 u = -\frac{1}{\bar{\rho}_0}\frac{\partial p}{\partial y}, \tag{10.30b}$$

$$N^2 w = -\frac{1}{\bar{\rho}_0}\frac{\partial^2 p}{\partial z\,\partial t}, \tag{10.30c}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{10.30d}$$

with

$$w = 0 \quad \text{at} \quad z = -D_* \tag{10.31}$$

and

$$p = \bar{\rho}_0 g\zeta \quad \text{at} \quad z = 0. \tag{10.32}$$

Notation is as in (10.6).

For the case of constant depth D_*, the principal result is that the dependent variables have the separable form

$$(u,v,w,p) = \{U(x,y,t), V(x,y,t), W(x,y,t), Z(x,y,t)\} \times \{F_u(z), F_u(z), F_w(z), F_p(z)\}^{\mathrm{T}} \tag{10.33}$$

where

$$W = \partial Z/\partial t, \qquad F_u = F_p/\bar{\rho}_0 g = D_n\,\partial F_w/\partial z, \tag{10.34}$$

and

$$\frac{\partial U}{\partial t} - f_0 V = -g\frac{\partial Z}{\partial x}, \tag{10.35a}$$

$$\frac{\partial V}{\partial t} + f_0 U = -g\frac{\partial Z}{\partial y}, \tag{10.35b}$$

$$\frac{\partial Z}{\partial t} + D_n\left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right) = 0, \tag{10.35c}$$

with

$$\frac{\partial^2 F_w}{\partial z^2} + \frac{N^2(z)}{gD_n}F_w = 0, \tag{10.36}$$

$$F_w = 0 \quad \text{at} \quad z = -D_*, \tag{10.37}$$

$$F_w - D_n\,\partial F_w/\partial z = 0 \quad \text{at} \quad z = 0. \tag{10.38}$$

According to (10.35), the horizontal variations of all quantities are exactly as in the homogeneous flat bottom case except that now the apparent depth D_n is obtained by solving the eigenvalue problem (10.36) for the vertical structure. Typically (10.36) yields a barotropic mode $F_{w0} \simeq z + D_*$, $D_0 \simeq D_*$ plus an infinite sequence of baroclinic modes F_{wn} characterized by n zero crossings (excluding the one at $z = -D_*$) and by very small equivalent depths D_n. Baroclinic WKB approximate solutions (10.36)-(10.38) are

$$F_{wn}(z) = N^{-1/2}(z)\sin\left[(1/gD_n)^{1/2}\int_{-D_*}^{z} N(z')\,dz'\right],$$
$$D_n = \left[\int_{-D_*}^{0} N(z')\,dz'\right]^2 \Big/ (gn^2\pi^2). \tag{10.39}$$

These are exact for constant buoyancy frequency N_0. For $D_* = 4000$ m and $N_0 = 10\Omega$, $D_n = (0.1/n^2)$ m. When $N(z) = N_0$, it is easy to show that

[w at the free surface/interior maximum value

$$\text{of } w] = N_0^2 D_*/n\pi g \ll 1. \tag{10.40}$$

Free surface variation is thus qualitatively and quantitatively unimportant for baroclinic modes. They are therefore usually called internal modes.

In a flat-bottomed ocean, stratification is thus seen to make possible an infinite sequence of internal replicas of the barotropic LSW gravity waves, Sverdrup waves, Poincaré waves, Kelvin waves, and basin normal modes discussed above. All these except the Kelvin waves have $\sigma^2 > f_0^2$. They must also have $\sigma^2 < N^2(z)$ over part of the water column, although the present treatment does not make this obvious because $\sigma^2 \ll N^2$ is always assumed. The horizontal variation of these replicas is governed by the equations describing the barotropic mode, except that the depth D_n is a small fraction of the actual (constant) depth D_*. Without rotation, the speed of barotropic long gravity waves is $(gD_*)^{1/2} \simeq 200$ m s^{-1} in the deep sea. Long internal gravity waves move at the much slower $(gD_n)^{1/2} \simeq (1/n)$ m s^{-1}. For comparable frequencies, the internal waves thus have much shorter wavelength than the surface wave.

An important point is that the separation of variables (10.33) works in spherical coordinates as well as in Cartesian coordinates, provided only that the depth is constant. The horizontal variation of flow variables is then governed by LTE with appropriate equivalent depth D_n given by (10.36)-(10.38).

For plane waves of frequency σ, relaxation of $N^2 \gg \sigma^2$ leads to the replacement of (10.36) by

$$\frac{\partial^2 F_w}{\partial z^2} + \frac{[N^2(z) - \sigma^2]}{gD_n}F_w = 0. \tag{10.41}$$

High-frequency waves, for which $[N^2(z) - \sigma^2]$ changes sign over the water column, are discussed in chapter 9.

The simplicity of these flat-bottom results is deceptive, because they are very difficult to generalize to include bottom relief. The reason for this is most easily seen by eliminating (u, v) from (10.30) for harmonic motion [$\exp(-i\sigma t)$] to obtain a single equation in w:

$$\frac{\partial^2 w}{\partial z^2} - \left(\frac{N_0^2}{\sigma^2 - f_0^2}\right)\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right) = 0. \tag{10.42}$$

This equation is hyperbolic in space for internal waves (for which $f_0^2 < \sigma^2 < N_0^2$). Its characteristic surfaces are

$$z = \pm(x^2 + y^2)^{1/2}(\sigma^2 - f_0^2)^{1/2}/N_0. \tag{10.43}$$

Solutions of (10.42) may be discontinuous across characteristic surfaces, and they depend very strongly upon the relative slope of characteristics and bounding surfaces. Without rotation, internal waves of frequency σ are solutions of the hyperbolic equation

$$\frac{\partial^2 p}{\partial z^2} - \frac{(N_0^2 - \sigma^2)}{\sigma^2}\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right) = 0 \tag{10.44}$$

with the simple condition

$$\frac{\partial p}{\partial n} = 0 \quad \text{at solid boundaries.} \tag{10.45}$$

For closed boundaries (i.e., a container filled with stratified fluid) this is an ill-posed problem in the sense that tiny perturbations of the boundaries may greatly alter the structure of the solutions. Horizontal boundaries, although analytically tractable, are a very special case.

10.4.4 Rossby and Planetary Waves

In an influential study whose emphasis upon physical processes marks the beginning of the modern period, Rossby and collaborators (1939) rediscovered Hough's second-class oscillations and suggested that they might be of great importance in atmospheric dynamics (see also chapters 11 and 18).

Perhaps of even greater influence than this discovery was Rossby's creation of a new plane model. It amounts to LTE written in the Cartesian coordinates

$$x = (a\cos\theta_0)(\phi - \phi_0), \qquad y = a(\theta - \theta_0) \tag{10.46}$$

tangent to the sphere at (ϕ_0, θ_0). Rossby's creative simplification was to ignore the variation of all metric coefficients ($\cos\theta = \cos\theta_0$) and to retain the latitude variation of

$$\begin{aligned} f &= 2\Omega\sin\theta \simeq 2\Omega\sin\theta_0 + y(2\Omega/a)\cos\theta_0 \\ &= f_0 + \beta y \end{aligned} \tag{10.47}$$

only when f is explicitly differentiated with respect to y. Rossby's notation

$$\beta \equiv \partial f/\partial y \tag{10.48}$$

has since become almost universal. Such Boussinesq-like approximations to the spherical equations are usually called β-plane equations.

Rossby's original approximation, which I shall call Rossby's β-plane, further suppressed horizontal divergence ($\zeta = 0$). It yields rational approximations (Miles, 1974b) for second-class solutions of LTE whose horizontal scale is much smaller than the earth's radius. A quite different approximation yielding rational approximations for both first- and second-class solutions of LTE when they are equatorially trapped is the equatorial β-plane (see section 10.4.5). Both Rossby's β-plane equations and the equatorial β-plane equations differ from those obtained by the often encountered procedure of making Rossby's simplification but retaining divergence. This results in what I shall call simply the β-plane equations, in conformity with common usage. It is a rational approximation to LTE only at low ($\sigma^2 \ll f_0^2$) frequencies and is otherwise best regarded as a model of LTE.

Without divergence, the homogeneous LTE (10.30) may be cross-differentiated to yield a vorticity equation

$$\begin{aligned} &\frac{\partial\nabla_{2s}^2\psi}{\partial t} + 2\Omega\frac{\partial\psi}{\partial\phi} = 0 \\ &\left[\nabla_{2s}^2 \equiv \frac{1}{a^2\cos^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{a^2\cos\theta}\frac{\partial}{\partial\theta}\left(\cos\frac{\partial}{\partial\theta}\right)\right] \end{aligned} \tag{10.49}$$

here written in terms of a streamfunction ψ defined by

$$v\cos\theta = \partial\psi/\partial\phi, \qquad u = -\partial\psi/\partial\theta. \tag{10.50}$$

Rossby's β-plane approximation to this, obtained by locally approximating the spherical Laplacian ∇_{2s}^2 as the plane Laplacian and going to the locally tangent coordinates (10.46), is

$$\frac{\partial\nabla_2^2\psi}{\partial t} + \beta\frac{\partial\psi}{\partial x} = 0, \tag{10.51}$$

where

$$v = \partial\psi/\partial x, \qquad u = -\partial\psi/\partial y. \tag{10.52}$$

The spherical vorticity equation (10.49) has as everywhere bounded solutions

$$\psi = S_n(\phi', \theta'), \qquad S_n(\phi, \theta) \equiv P_n^l(\sin\theta)\exp(il\phi), \tag{10.53}$$

where (ϕ', θ') are spherical coordinates relative to a pole P' displaced an arbitrary angle from the earth's pole P of rotation and rotating about P with angular velocity

$$c = -2\Omega/[n(n+1)] \tag{10.54}$$

(Longuet-Higgins, 1964). When $P' = P$, (10.54) is exactly the second part of (10.17); these *are* Hough's second-class oscillations. Rossby's β-plane equivalents are

$$\psi = Z_n(x - ct, y), \qquad \nabla_2^2 Z_n + \lambda_n^2 Z_n = 0, \tag{10.55}$$

where $c = -\beta/\lambda_n^2$.

Plane Rossby waves are the particular case

$$\psi = a\exp(-i\sigma t + ilx + iky), \tag{10.56}$$

$$\sigma = -\beta l/(l^2 + k^2). \tag{10.57}$$

They are transverse [(u, v) perpendicular to (l, k)] and dispersive, with the propagation of phases always having a westward component ($\sigma/l < 0$). Their frequencies are typically low relative to f_0. Since σ depends both on wavelength and wave direction, these waves do not reflect specularly at a straight coast. Longuet-Higgins (1964) gives an elegant geometrical interpretation (figure 10.1) of their dispersion relation (10.57) and shows that it is the group velocity $(\partial\sigma/\partial l, \partial\sigma/\partial k)$ that reflects specularly in this case (figure 10.2).

Rossby-wave normal modes of a closed basin of perimeter P have the form (Longuet-Higgins, 1964)

$$\psi = Z_n(x,y)\exp[-i\sigma_n t - i(\beta/2\sigma_n)x], \tag{10.58}$$

$$\sigma_n = \beta/2\lambda_n \tag{10.59}$$

where

$$\nabla_2^2 Z_n + \lambda_n^2 Z_n = 0, \qquad Z_n = 0 \quad \text{on} \quad P. \tag{10.60}$$

These Rossby-wave normal modes are in remarkable contrast with the gravity-wave normal modes of the same basin without rotation:

$$\zeta_n = Z_n(x,y)\exp(-i\sigma_n t), \tag{10.61}$$

$$\sigma_n = \lambda_n(gD_*)^{1/2} \qquad [\partial Z_n/\partial(\text{normal}) = 0 \text{ on } P]. \tag{10.62}$$

The gravity-wave modes have a lowest-frequency ($n = 1$) grave mode and the spatial scales of higher-frequency modes are smaller. The Rossby-wave modes have a highest frequency ($n = 1$) mode and the spatial scales of lower frequency modes are smaller.

The physical mechanism that makes Rossby waves possible is most easily seen for nearly zonal waves ($\partial/\partial y \ll \partial/\partial x$). Then Rossby's vorticity equation (10.51) becomes

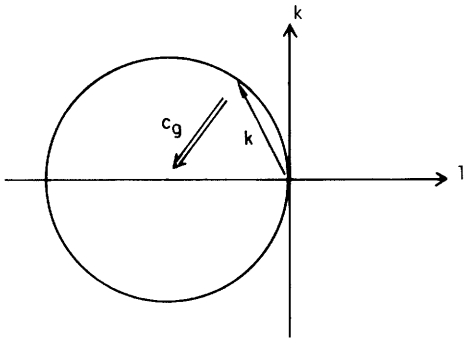


Figure 10.1 The locus of wavenumbers $k = (k,l)$ allowed by the Rossby dispersion relation (10.57)

$$(l + \beta/2\sigma)^2 + k^2 = (\beta/2\sigma)^2$$

is a circle of radius $\beta/2\sigma$ centered at $(-\beta/2\sigma, 0)$. The group velocity vector $\mathbf{c}_g = (\partial\sigma/\partial l, \partial\sigma/\partial k)$ points from the tip of the wavenumber vector toward the center of the circle and has magnitude $|\mathbf{c}_g| = \beta/(l^2 + k^2)$.

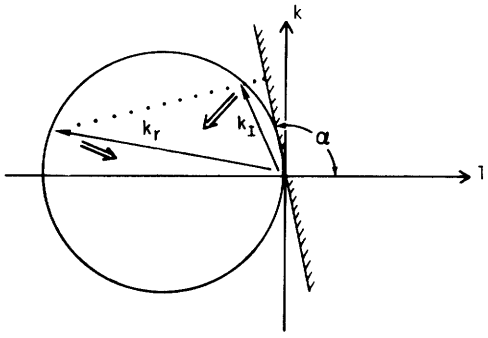


Figure 10.2 A Rossby wave with wavenumber $\mathbf{k}_i$ is incident on a straight coast inclined at an angle α to the east-west direction. The wavenumber $\mathbf{k}_r$ of the reflected wave is fixed by the necessity that $\mathbf{k}_i$ and $\mathbf{k}_r$ have equal projection along the coast. The group velocity reflects specularly in the coast.

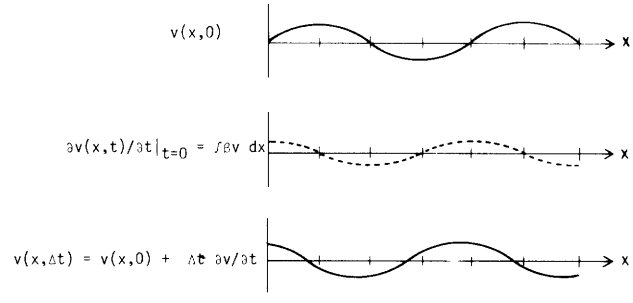


Figure 10.3 The flow $v(x,t)$ evolving from the initial flow $v(x,0) \sim \sin(lx)$ as fluid columns migrate north-south (and so exchange planetary and relative vorticity) is a westward displacement of the initial flow. Notice that although parcels take on clockwise-counterclockwise relative vorticity as they are moved north-south, the westward displacement is not the result of advection of vorticity of one sign by the flow associated with the other as is the case in a vortex street.

$$\frac{\partial^2 v}{\partial x\,\partial t} + \beta v = 0. \tag{10.63}$$

North-south motions v result in changes in the local vorticity $\partial v/\partial x$. When the initial north-south motion is periodic in x, then examination (figure 10.3) of (10.63) shows that the additional north-south motion generated by the vorticity resulting from the initial pattern of north-south motion combines with that pattern to shift it westward, in accordance with (10.57).

Rossby's vorticity equation (10.51) corresponds to the plane equations

$$\frac{\partial u}{\partial t} - fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x}, \tag{10.64a}$$

$$\frac{\partial v}{\partial t} + fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y}, \tag{10.64b}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \tag{10.64c}$$

$$f = f_0, \qquad \partial f/\partial y = \beta, \tag{10.64d}$$

so that Rossby's solutions are almost geostrophic ($\sigma \ll f_0$) and perfectly nondivergent. The absence of divergence and vertical velocity is an extreme of the tendency, in quasigeostrophic flow, for the vertical velocity to be order Rossby number ($\ll 1$) smaller than a scale analysis of the continuity equation would indicate (Burger, 1958). This tendency is absent at planetary length scales, and Rossby's β-plane (10.64) correspondingly requires modification.

Remarkably, Rossby and collaborators (1939) prefaced their analysis with a resumé of a different physical mechanism due to J. Bjerknes (1937), a mechanism that also results in westward-propagating waves but for a different reason, and that supplies the modification of Rossby's β-plane required at planetary scales. The plane equations corresponding to Rossby's summary of Bjerknes' arguments are

$$-fv = -g\frac{\partial \zeta}{\partial x}, \tag{10.65a}$$

$$fu = -g\frac{\partial \zeta}{\partial y}, \tag{10.65b}$$

$$\frac{\partial \zeta}{\partial t} + D_*\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0, \tag{10.65c}$$

$$f = f_0, \qquad \partial f/\partial y = \beta. \tag{10.65d}$$

By physical arguments (figure 10.4), Bjerknes deduced that an initial pressure perturbation would always propagate westward. The corresponding analysis of (10.65) is to form an elevation equation

$$\frac{\partial \zeta}{\partial t} - (gD_*\beta/f_0^2)\frac{\partial \zeta}{\partial x} = 0 \tag{10.66}$$

and to note that it has the dispersionless solutions

$$\zeta = F[x + (gD_*\beta/f_0^2)t,\ y],$$

where $F(x, y)$ is the initial pressure perturbation. In particular,

$$\zeta = a\exp(-i\sigma t + ilx + iky), \tag{10.67}$$

$$\sigma = -(gD_*\beta/f_0^2)l. \tag{10.68}$$

All solutions travel westward at $(gD_*\beta/f_0^2)^{1/2}$. These motions, according to (10.65), are perfectly geostrophic but divergent.

More complete analysis (Longuet-Higgins, 1964) shows that the two dispersion relations (10.57) of Rossby and (10.68) of Bjerknes are limiting cases of the β-plane dispersion relation

$$\sigma = -\beta l/(l^2 + k^2 + f_0^2/gD_*) \tag{10.69}$$

for second-class waves displayed in figure 10.5. It would be appropriate to call the two kinds of second-class waves Rossby and Bjerknes waves, respectively, but in practice both are commonly called Rossby waves. I shall distinguish them as short, nondivergent and long, divergent Rossby waves.

When divergence is allowed, the (constant) depth D_* enters the dispersion relation (10.69) in the length scale

$$a_R = (gD_*/f_0^2)^{1/2}, \tag{10.70}$$

usually called the Rossby radius. There is not one Rossby radius, but rather there are many, since the constant-depth barotropic second-class waves so far discussed have an infinite sequence of baroclinic counterparts with $D_* = D_n$, $n = 1, \ldots$, given by (10.36)-(10.39). Waves longer than the Rossby radius are long, divergent Rossby waves; those shorter than the Rossby radius are short, nondivergent Rossby waves.

The barotropic Rossby radius $a_{Ro} = (gD_0/f_0^2)^{1/2}$ has $D_0 \simeq D_*$ and is thus the order of the earth's radius. Barotropic Rossby waves are consequently relatively high-frequency (typically a few cycles per month) waves and they are able to traverse major ocean basins in days to weeks. Baroclinic Rossby radii $a_{Rn} = (gD_n/f_0^2)^{1/2}$ are the order of 10^2 km or less in mid-latitudes. Baroclinic mid-latitude Rossby waves are consequently relatively low-frequency waves and would take years to traverse major mid-latitude basins. In the tropics, f_0 becomes small and baroclinic Rossby waves speed up to the point where they could traverse major basins in less than a season. But a different discussion is really necessary for the tropics (see chapter 6).

Rossby advanced his arguments to rationalize the motion of mid-latitude atmospheric pressure patterns. In both atmosphere and ocean, the slowness and relatively small scale of most second-class waves must make their occurrence in "pure" form very rare. Oceanic measurements from the MODE experiment

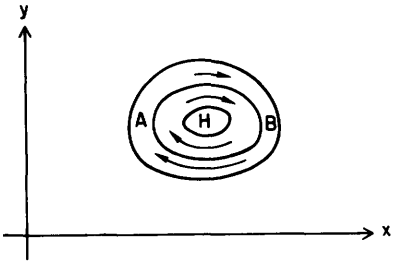


Figure 10.4 If the flow is totally geostrophic but the Coriolis parameter increases with latitude, then the flow at A converges because the geostrophic transport between a pair of isobars south of H is greater than that between the same pair north of H. By (10.65c), pressure thus rises at A. Similarly, the flow at B diverges and pressure there drops. The initial pattern of isobars is then shifted westward.

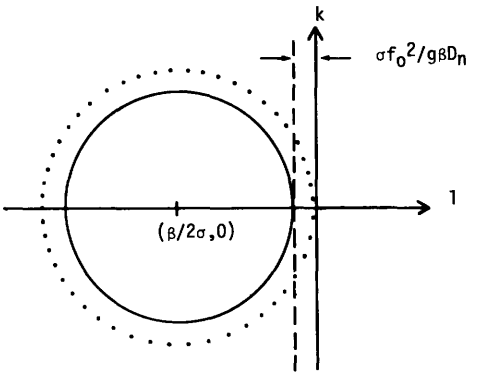


Figure 10.5 The locus of wavenumbers (l, k) allowed by the β-plane dispersion relation (10.69) for second-class waves

$$(l + \beta/2\sigma)^2 + k^2 = (\beta/2\sigma)^2 - (f_0^2/gD_n)$$

is a circle(———) whose radius is $[(\beta/2\sigma)^2 - (f_0^2/gD_n)]^{1/2}$ centered at $(-\beta/2\sigma, 0)$. Dotted circle (· · ·) is the Rossby dispersion relation (10.57) for short waves. Dashed line (---) is the Bjerknes dispersion relation (10.68) appropriate for long waves. The scale a_R dividing short and long waves is

$$a_R = [2(\beta/2\sigma)(\sigma f_0^2/\beta gD_n)]^{-1/2}.$$

do show, however, characteristics of both short baroclinic (figure 10.6) and long barotropic (figure 10.7) Rossby waves.

The oscillations having the two dispersion relations (10.23e) with $D_* = D_n$ for first-class waves and (10.69) for second-class waves are mid-latitude plane-wave approximations of solutions of LTE. Figure 10.8 plots the two dispersion relations together. A noteworthy feature is the frequency interval between f_0 and $(\beta/2f_0)(gD_n)^{1/2}$ within which no plane waves propagate. Taken at face value, this gap suggests that velocity spectra should show a valley between these two frequencies with a steep high-frequency [f_0] wall and a rather more gentle low-frequency [$(\beta/2f_0)(gD_n)^{1/2}$, n = 0,1,2, . . .] wall. Such a gap is indeed commonly observed; but the dynamics of the low frequencies are almost surely more complex than those of the linear β-plane. The latitude dependence implicit in the definition of f_0 and β is consistent with equatorial trapping of low-frequency first-class waves and high-frequency second-class waves. This is more easily seen in approximations, such as the following, which better acknowledge the earth's sphericity.

10.4.5 The Equatorial β-Plane

For constant depth D_*, the homogeneous LTE (10.5) may be equatorially approximated by expanding all variable coefficients in θ and then neglecting $\theta^2, \theta^3, \ldots$. The resulting equatorial β-plane equations are

$$\frac{\partial u}{\partial t} - \beta y v = -g\frac{\partial \zeta}{\partial x}, \tag{10.71a}$$

$$\frac{\partial v}{\partial t} + \beta y u = -g\frac{\partial \zeta}{\partial y}, \tag{10.71b}$$

$$\frac{\partial \zeta}{\partial t} + D_*\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0, \tag{10.71c}$$

where $x = a\phi$, $y = a\theta$, and $\beta = 2\Omega/a$. They govern both barotropic and baroclinic motions provided that D_* is interpreted as the appropriate equivalent depth D_n defined by (10.36)-(10.38). Moore and Philander (1977) and Philander (1978) give modern reviews.

Solutions of these equations can be good approximations to solutions of LTE only when they decay very rapidly away from the equator. But the qualitative nature of their solutions, bounded as $y \to \pm\infty$, closely resembles solutions of LTE bounded at the poles, even

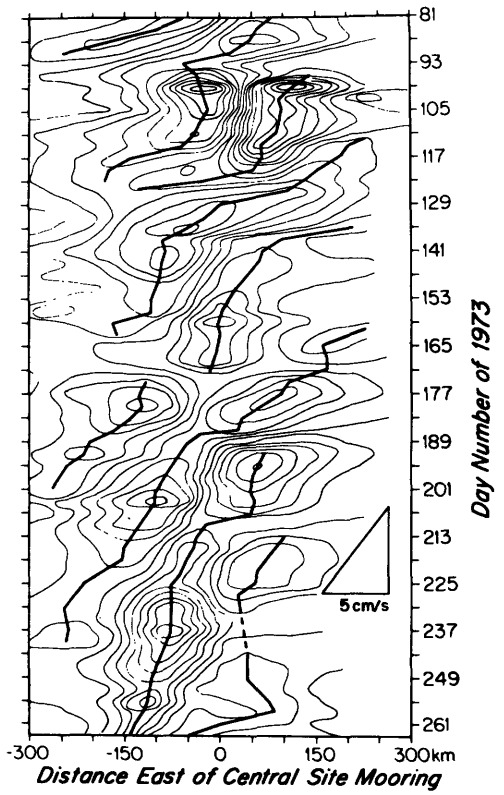


Figure 10.6A Time-longitude plot of streamfunction inferred from objective maps of 1500-m currents along 28°N (centered at 69°49′W) by Freeland, Rhines and Rossby (1975). There is evidence of westward propagation of phases. Currents at this depth are not dominated by "thermocline eddies" (section 10.4.7) but are representative of the deep ocean.

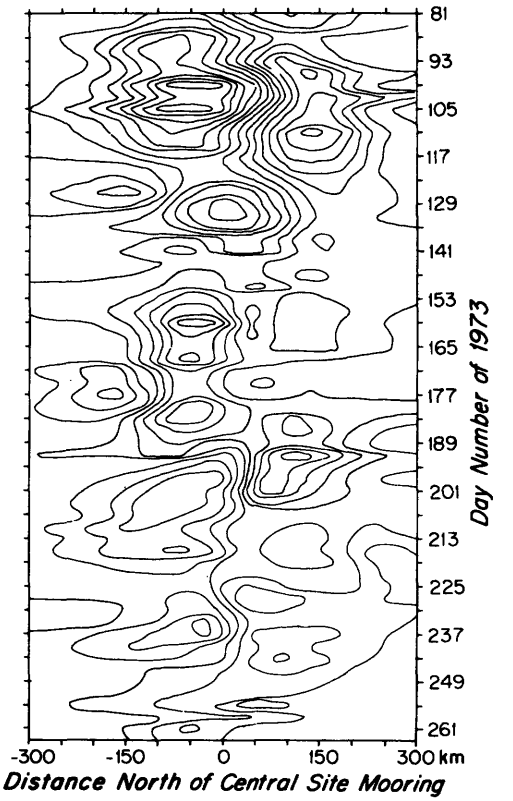


Figure 10.6B As figure 10.6A but in time-latitude plot along 69°40′W. There is no evidence for a preferred direction of latitudinal phase propagation. (Rhines, 1977.)

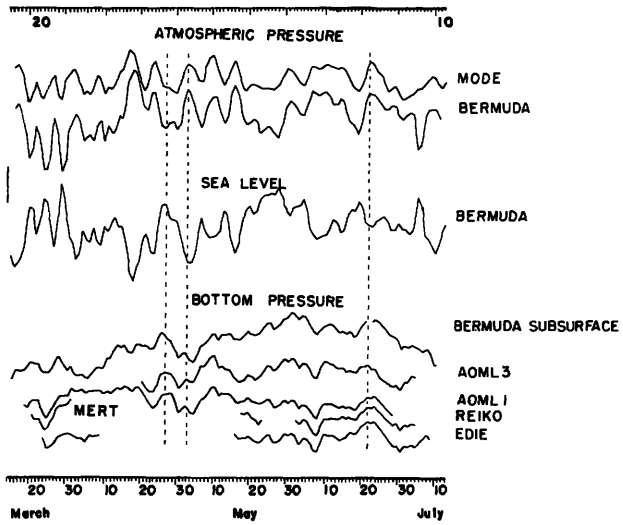


Figure 10.7 Time series of bottom pressure in MODE (Brown et al., 1975). The cluster of named gauges centered at 28°N, 69°40′W show remarkable coherence despite 0 (180-km) separation, and all are coherent with the (atmospheric pressure corrected) sea level at Bermuda (650 km distant, labeled Bermuda bottom). (Brown et al., 1975.)

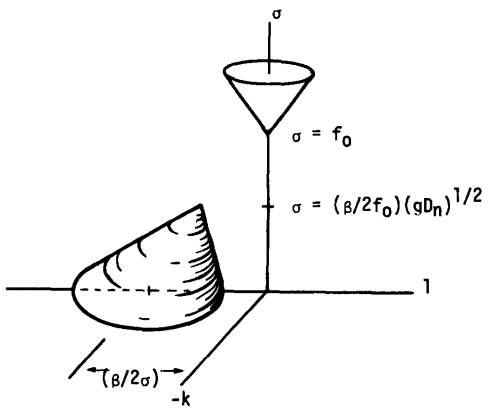


Figure 10.8 The f-plane dispersion relation

$$\sigma^2 = f_0^2 + gD_n(l^2 + k^2)$$

for first-class waves allows no waves with $\sigma^2 < f_0^2$. The β-plane dispersion relation

$$\sigma = -\beta l/[l^2 + k^2 + f_0^2/gD_n]$$

for second-class waves allows no waves with $\sigma > (\beta/2f_0)(gD_n)^{1/2}$.

when the equatorial approximation is transgressed. Historically these approximate solutions provided a great deal of insight into the latitudinal variation of solutions of LTE.

Most of the solutions are obtainable from the single equation that results when u,ζ are eliminated from (10.71). With

$$v = V(y)\exp(-i\sigma t + ilx) \tag{10.72}$$

that equation is

$$\frac{\partial^2 V}{\partial y^2} + \left[\left(\frac{\sigma^2}{gD^*} - l^2 - \frac{l\beta}{\sigma}\right) - \frac{\beta^2}{gD^*}y^2\right]V = 0. \tag{10.73}$$

It also occurs in the quantum-mechanical treatment of the harmonic oscillator. Solutions are bounded as $y \to \pm\infty$ only if

$$\left(\frac{\sigma^2}{gD_*} - l^2 - \frac{l\beta}{\sigma}\right) = (2m+1)\frac{\beta}{(gD_*)^{1/2}}, \qquad m = 0, 1, 2, \ldots, \tag{10.74}$$

and they are then

$$V(y) = H_m[y\beta^{1/2}/(gD_*)^{1/4}]\exp[-y^2\beta/2(gD)^{1/2}], \tag{10.75}$$

wherein the H_m are Hermite polynomials ($H_0(z) = 1$, $H_1(z) = z, \ldots$).

The remaining solution may be taken to be $v = 0$ with $m = -1$ in (10.74). It is obtained by solving (10.71) with $v = 0$. The solution bounded as $y \to \pm\infty$ is

$$\zeta = \exp[-i\sigma t + ilx - (\beta l/\sigma)y^2/2] \tag{10.76}$$

with

$$l = \sigma/(gD_*)^{1/2} \tag{10.77}$$

[(10.77) *is* (10.74) with $m = -1$].

The very important dispersion relation (10.74) with $m = -1, 0, 1, \ldots$ thus governs all the equatorially trapped solutions of (10.71). Introducing the dimensionless variables ω, λ, ζ, η defined by

$$\sigma = \omega(2\Omega\Lambda^{-1/4}), \qquad l = \lambda(a^{-1}\Lambda^{1/4}), \tag{10.78}$$

$$(x, y) = (\xi, \eta)(a\Lambda^{-1/4}), \qquad t = \tau[(2\Omega)^{-1}\Lambda^{1/4}]$$

($\Lambda = 4\Omega^2a^2/gD_*$) allows us to rewrite (10.73) and its solutions (10.72), (10.75), (10.76) as

$$\frac{\partial^2 V}{\partial \eta^2} + [(\omega^2 - \lambda^2 - \lambda/\omega) - \eta^2]V = 0, \tag{10.79}$$

$$v = H_m(\eta)\exp(-i\omega\tau + i\lambda\xi - \eta^2/2), \qquad m = 0, 1, \ldots, \tag{10.80}$$

$$\zeta = \exp[-i\omega\tau + i\lambda\xi - (\lambda/\omega)\eta^2/2], \qquad m = -1, \tag{10.81}$$

while the dispersion relation (10.74) becomes

$$\omega^2 - \lambda^2 - \lambda/\omega = 2m + 1, \qquad m = -1, 0, 1, \ldots. \quad (10.82)$$

These forms allow easy visualization of the solutions and permit a concise graphical presentation of the dispersion relation (figure 10.9).

The dispersion relation is cubic in σ (or ω) for given values of l (or λ) and m. For $m > 1$ the three roots correspond precisely to two oppositely traveling waves of the first class plus a single westward-traveling wave of the second class. The case $m = 0$ (Yanai, or Rossby-gravity, wave) is of first class when traveling eastward but of second class when traveling westward. The case $m = -1$ is an equatorially trapped Kelvin wave, dynamically identical to the coastally trapped Kelvin wave (10.27) in a uniformly rotating ocean.

The most useful aspect of these exact solutions is their provision of a readily understandable dispersion relation [(10.74) or (10.82)]. The latitudinal variation of flow variables is more readily discussed in terms of WKB solutions of (10.73). One can easily see *the* salient feature of the solutions, a transition from oscillatory to exponentially decaying latitudinal variation as the turning latitudes y_T of (10.73) (at which the coefficient [ ] of that equation vanishes), are crossed poleward. For waves of the first class, the term $l\beta/\sigma$ is small relative to the other terms in the dispersion relation and in the coefficient [ ]. The corresponding turning latitudes $y_T^{(1)}$ are therefore approximately given by

$$[y_T^{(1)}]^2 = (\sigma/\beta)^2[1 - l^2(gD_*)/\sigma^2] < (\sigma/\beta)^2. \quad (10.83)$$

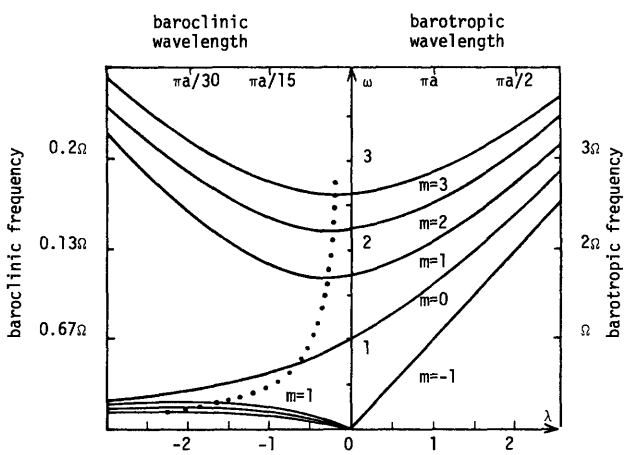


Figure 10.9 The equatorial β-plane dispersion relation (10.82) $\omega^2 - \lambda^2 - \lambda/\omega = 2m + 1$.

Dimensional wavelengths and frequencies are obtained from the scaling (10.78) and are given for the barotropic mode (D_0 = 4000 m, $\Lambda = 20$) and for the first baroclinic mode ($D_0 = 0.1$ m, $\Lambda = 10^6$). For all curves but $m = 0$, intersections with dotted curve are zeros of group velocity.

For waves of the second class, the term σ^2/gD_* is small relative to the other terms in the dispersion relation and in the coefficient [ ]. The corresponding turning latitudes $y_T^{(2)}$ are therefore approximately given by

$$[y_T^{(2)}]^2 = (gD_*/\beta^2)(-l^2 - l\beta/\sigma) < gD_*/4\sigma^2. \quad (10.84)$$

Increasingly low-frequency waves of the first class and increasingly high-frequency waves of the second class are thus trapped increasingly close to the equator.

Only first-class waves having frequency greater than the inertial frequency βy penetrate poleward of latitude y [by (10.83)]. Only second-class waves having frequency below the cutoff frequency $(gD_*/4y^2)^{1/2}$ penetrate poleward of latitude y [by (10.84)]. This frequency-dependent latitudinal trapping corresponds to the mid-latitude frequency gap between first- and second-class waves discussed in the previous section and illustrated in figure 10.8. The correspondence correctly suggests that trapping and associated behavior characterize slowly varying (in the WKB sense) packets of waves propagating over the sphere as well as the globally standing patterns corresponding to the Hermite solutions (10.75). Waves thus need *not* be globally coherent to exhibit trapping and the features associated with it.

Near the trapping latitudes, (10.73) becomes

$$\frac{\partial^2 V}{\partial y^2} + (-2\beta^2 y_T/gD_*)(y - y_T)V = 0. \quad (10.85)$$

The change of variable $\eta = (2\beta^2 y_T/gD_*)^{1/3}(y - y_T)$ reduces this to Airy's equation

$$\frac{\partial^2 V}{\partial \eta^2} - \eta V = 0, \quad (10.86)$$

whose solution $Ai(\eta)$ bounded as $\eta \to \infty$ is plotted in figure 10.10. This solution has two important features: (1) gentle amplification (like $\eta^{-1/4}$) of the solution as the turning latitude ($\eta = 0$) is approached from the equator; and (2) transition from oscillatory to exponentially decaying behavior in a region $\Delta\eta$ of order roughly unit width surrounding the turning latitude. Consequently the interval Δy over which the solution of (10.85) changes from oscillatory to exponential behavior is $\Delta y = [2\beta^2 y_T/gD_*]^{-1/3}\Delta\eta$ or, since $\beta = 2\Omega/a$ and $\Delta\eta \simeq 1$,

$$\Delta y = a(\Lambda\sigma/2\Omega)^{-1/3} \quad (10.87)$$

$$= a(2^{-1/2}\Lambda^{-1/2}\sigma/2\Omega)^{1/3} \quad (10.88)$$

for maximally penetrating first- and second-class waves. For diurnal ($\sigma = \Omega$) first-class barotropic ($\Lambda = 20$) waves, $\Delta y \simeq 0.5a$; for diurnal first-class baroclinic ($\Lambda \simeq 10^6$) waves, $\Delta y \simeq 0.013a$. For 10-day ($\sigma = 0.1\Omega$) second-class barotropic waves, $\Delta y \simeq 0.25a$. For 1-month

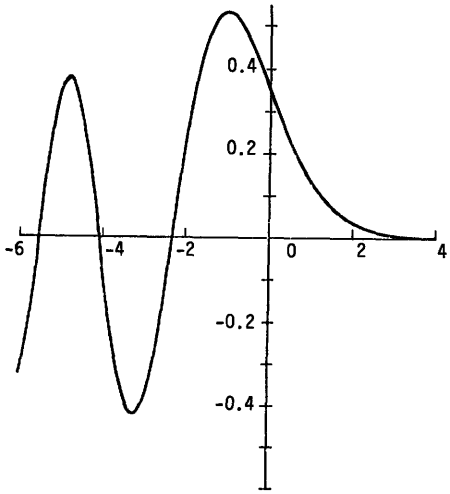


Figure 10.10 The Airy function *Ai*.

($\sigma \simeq .03\Omega$) second-class baroclinic waves, $\Delta y = 0.03a$. We thus obtain the important result that barotropic modes are not noticeably trapped (Δy is a fair fraction of the earth's radius and the Airy solution is only qualitatively correct anyway) but baroclinic modes are abruptly trapped (Δy is a few percentage points of the earth's radius).

The abrupt trapping of baroclinic waves at their inertial latitudes means that the Airy functions may describe quite accurately the latitude variation of near-inertial motions. Munk and Phillips (1968) and Munk (chapter 9) discuss the structure.

The clearest observations of equatorial trapping are by Wunsch and Gill (1976), from whose paper figure 10.11 is taken. Longer-period fluctuations at and near the equator have been observed, but their relation to the trapped solutions is not yet clear.

When an equatorially trapped westward-propagating wave meets a north–south western boundary (at, say, $x = 0$) it is reflected as a superposition of finite numbers of eastward-propagating waves including the Kelvin ($m = -1$) and Yanai ($m = 0$) waves (Moore and Philander, 1977). But when an equatorially trapped eastward-propagating wave meets a north–south eastern boundary, some of the incident energy is scattered into a poleward-propagating coastal Kelvin wave (10.29) and thus escapes the equatorial region (Moore, 1968). In latitudinally bounded basins, the requirement that solutions decay exponentially away from the equator is replaced by the vanishing of normal velocity at the boundaries. Modes closely confined to the equator will not be greatly altered by such boundaries; modes that have appreciable extraequatorial amplitude will behave like the β-plane solutions of sections 10.4.1–10.4.4 near the boundaries. A theory of free oscillations in idealized basins on the equatorial β-plane could be constructed on the basis of such observations, but powerful techniques for dealing with the spherical problem now exist (section 10.4.8).

10.4.6 Barotropic Waves over Bottom Relief

Stokes (1846; Lamb, 1932, §260) had shown that shoaling relief results in the trapping of an edge wave whose amplitude decays exponentially away from the coast, but the motion was not thought to be important.

Eckart (1951) solved the shallow-water equations [(10.18) with $f_0 = 0$] with the relief $D = ax$. Solutions of the form

$$\zeta = h(x)\exp(-i\sigma t + iky) \tag{10.89}$$

are governed by

$$x\frac{\partial^2 h}{\partial x^2} + \frac{\partial h}{\partial x} + [\sigma^2/(ag) - xk^2]h = 0. \tag{10.90}$$

Solutions of this are bounded as $x \to \infty$ only if

$$\sigma^2 = k(2n + 1)ag, \qquad n = 0, 1, \ldots, \tag{10.91}$$

and they are then

$$h(x) = L_n(2kx)\exp(-kx), \tag{10.92}$$

where the L_n are Laguerre polynomials [$L_0(z) = 1$, $L_1(z) = z - 1, \ldots$]. The $n = 0$ mode corresponds to Stokes's (1846) edge wave.

Eckart's solutions are LSW gravity waves refractively trapped near the coast by the offshore increase in shallow water wave speed $(g\alpha x)^{1/2}$. The Laguerre solutions (10.92) are correspondingly trigonometric shoreward of the turning points x_T at which the coefficient [ ] of (10.90) vanishes, and decay exponentially seaward.

Eckart's use of the LSW equations is not entirely self-consistent, since $D = ax$ increases without limit. Ursell (1952) removed the shallow-water approximation by completely solving

$$\frac{\partial^2 \phi}{\partial z^2} + \frac{\partial^2 \phi}{\partial x^2} - k^2\phi = 0 \tag{10.93}$$

subject to

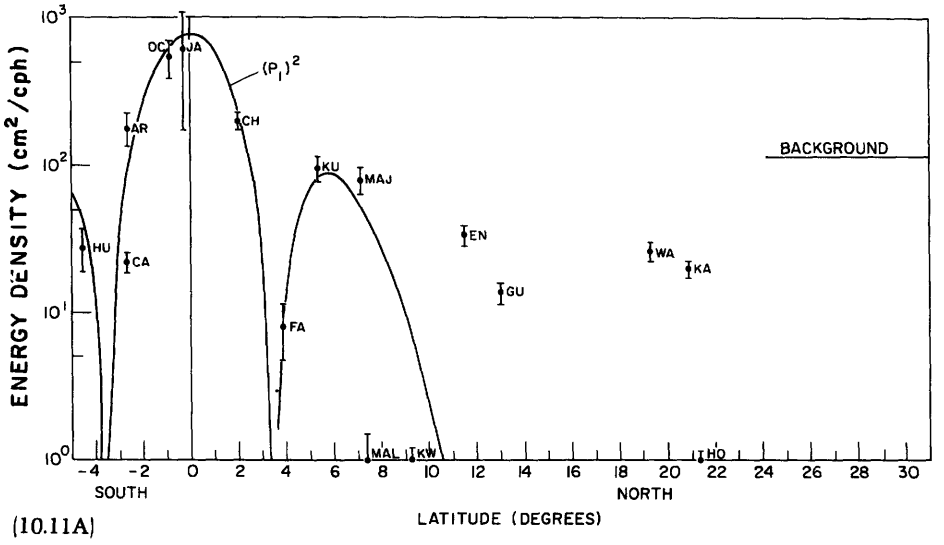
$$\frac{\partial \phi}{\partial z} = (\sigma^2/g)\phi \qquad \text{at} \quad z = 0 \tag{10.94}$$

and

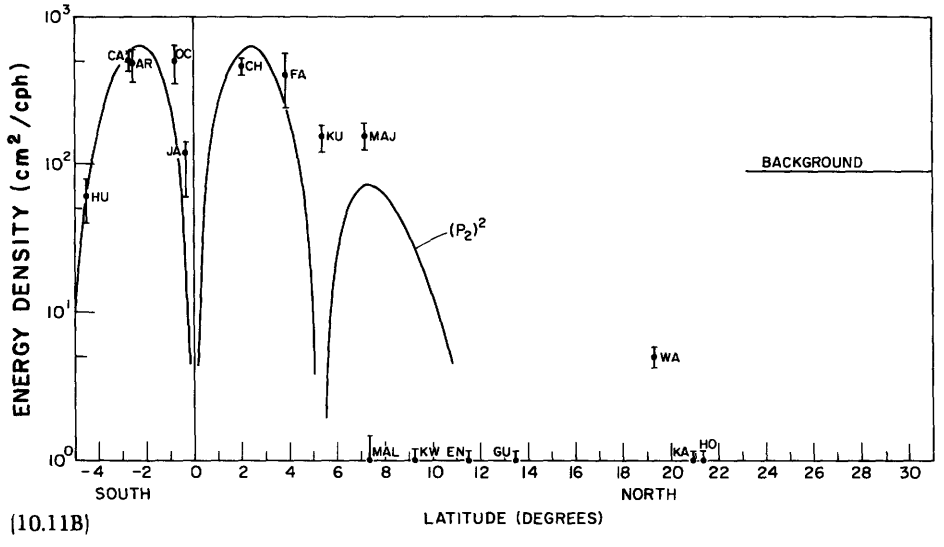
$$\frac{\partial \phi}{\partial \eta} = 0 \qquad \text{at} \quad z = -ax \tag{10.95}$$

plus boundedness of the velocity field ($\partial\phi/\partial x, ik\phi, \partial\phi/\partial z$) as $x \to \infty$. He found (1) a finite number of coastally trapped modes with dispersion relation

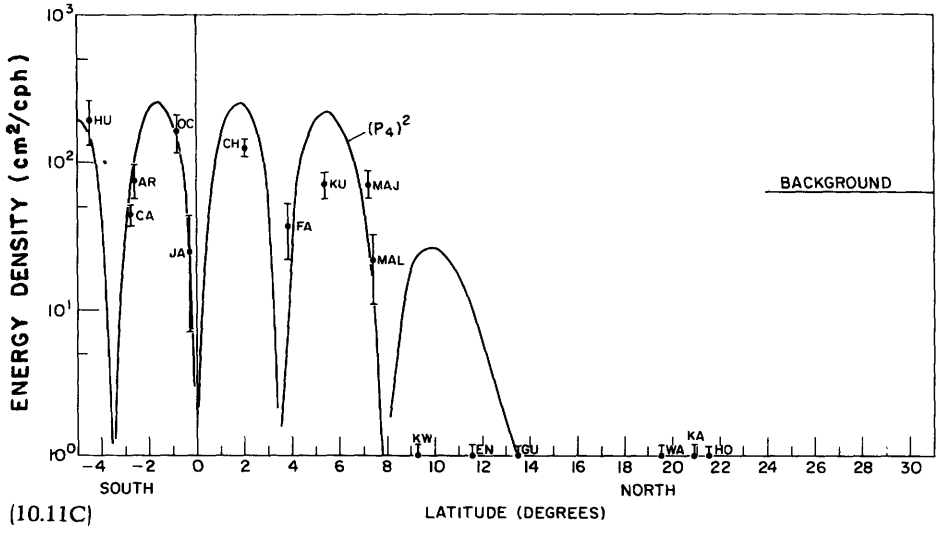
$$\sigma^2 = kg\sin[(2n + 1)\tan^{-1}a], \qquad n = 0, 1, \ldots < [\pi/(4\tan^{-1}a)^{-1/2}] \tag{10.96}$$



(10.11A)



(10.11B)



(10.11C)

Figure 10.11 Energy at periods of 5.6d, $m = 1$ (A); 4.0d, $m = 2$ (B); 3.0d, $m = 4$ (C); in tropical Pacific sea-level records as a function of latitude. A constant (labeled BACKGROUND) representing the background continuum has been subtracted from each value. Error bars are one standard deviation of χ^2. The solid curves are the theoretical latitudinal structure from the equatorial β-plane. (Wunsch and Gill, 1976.)

corresponding, for low n, to Eckart's results, plus (2) a continuum of solutions corresponding to the coastal reflection of deep-water waves incident from $x = \infty$ and correspondingly not coastally trapped. Far from the coast, the continuum solutions have the form $\phi = \cos(lx + \text{phase})\exp[-i\sigma t + iky + (l^2 + k^2)^{1/2}z]$ and their dispersion relation must require $\sigma^2 \geq gk$. They are filtered out by the shallow-water approximation. Figure 10.12 compares Eckart's (1951) and Ursell's (1952) dispersion relations.

With rotation f_0 restored to (10.18), (10.90) becomes (Reid, 1958)

$$x\frac{\partial^2 h}{\partial x^2} + \frac{\partial h}{\partial x} + \left[\left(\frac{\sigma^2 - f_0^2}{ag} - \frac{f_0 k}{\sigma}\right) - xk^2\right]h = 0. \qquad (10.97)$$

Solutions still have the form (10.89), (10.92) but now the dispersion relation is

$$\sigma^2 - f_0^2 - f_0 kag/\sigma = k(2n + 1)ag, \qquad (10.98)$$

which is cubic in σ, whereas with $f_0 = 0$ it was quadratic. Rotation has evidently introduced a new class of motion.

That this should be so is clear from the β-plane vorticity equation [obtained by cross-differentiating (10.30a,b) and with $\partial f_0/\partial y = \beta$]:

$$\frac{\partial}{\partial t}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) - \frac{f_0}{D}\frac{\partial \zeta}{\partial t} - u\frac{f_0}{D}\frac{\partial D}{\partial x} - v\left(-\beta + \frac{f_0}{D}\frac{\partial D}{\partial y}\right)$$

$$= 0. \qquad (10.99)$$

We have already seen (section 10.4.4) that the term βv gives rise to short and long Rossby waves (with the vortex-stretching term $f_0/D\,\partial\zeta/\partial t$ important only for the latter) when the depth is constant. But in (10.99), the topographic vortex-stretching term $\mathbf{u}f_0/D\cdot\nabla D$ plays a role entirely equivalent to that of $\beta v = \mathbf{u}\cdot\nabla f_0$. Hence, we expect it to give rise to second-class waves, both short and long, even if $\beta = 0$. Such waves are called *topographic Rossby waves.* Over the linear beach $D = -ax$ they are all refractively trapped near the coast.

The nondimensionalization

$$s = \sigma/f, \qquad K = kag/f^2 \qquad (10.100)$$

casts the dispersion relation into the form

$$s^3 - s[1 + (2n + 1)K] - K = 0, \qquad (10.101)$$

which is remarkably similar to (10.82) and plotted in figure 10.13.

The linear beach $D = ax$ is most unreal in that there is no deep sea of finite depth in which LSW plane waves can propagate. When the relief is modified to

$$D(x) = \begin{cases} ax, & 0 < x < D_0 a^{-1} \quad \text{(shelf)} \\ D_0, & D_0 a^{-1} < x < \infty \quad \text{(sea)}, \end{cases} \qquad (10.102)$$

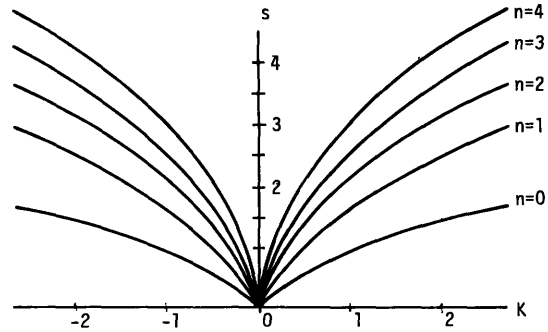


Figure 10.12A Eckart's (1951) dispersion relation (10.91)

$s^2 = K(2n + 1)$

for the shallow-water waves over a semi-infinite uniformly sloping, nonrotating beach. For convenience in plotting, $s = \sigma/f$ and $K = gak/f^2$ even though problem is not rotating.

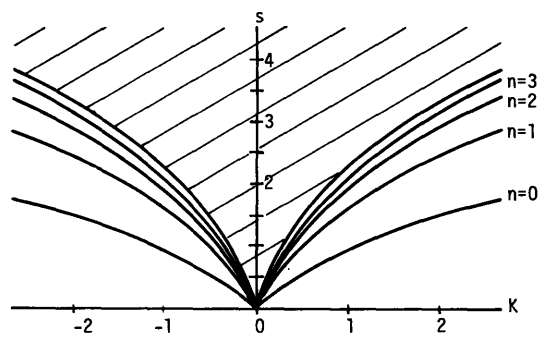


Figure 10.12B Ursell's (1952) dispersion relation (10.96)

$s^2 = Ka^{-1}\sin[(2n + 1)\tan^{-1}a]$

for edge waves, and the continuum

$s^2 > Ka^{-1}$

of deep-water reflected waves. For convenience in plotting, s and K are defined as above. Plot is for $a = 0.2$.

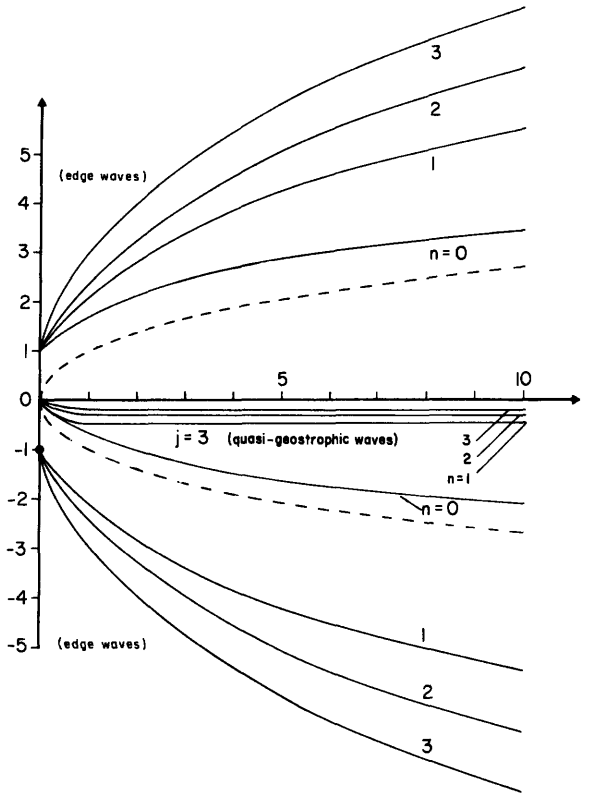


Figure 10.13 The dispersion relation (10.91) for edge and quasi-geostrophic shallow-water waves over a semi-infinite uniformly sloping beach. The dashed curves are the Stokes solution without rotation [(10.91) with $n = 0$]. Axes are as in figure 10.12. (LeBlond and Mysak, 1977.)

the most important alteration of the dispersion relation is an "opening up" of the long-wavelength part of the dispersion relation to include a continuum analogous to that of Ursell (1952) but now consisting of LSW first-class waves incident from the deep sea and reflected back into it by the coast and shelf. These waves are not coastally trapped. They are often called *leaky modes* because they can radiate energy that is initially on the shelf out into the deep sea. There are no second-class counterparts because the deep sea with constant depth and $\beta = 0$ cannot support second-class waves.

The dispersion relation corresponding to (10.102) is plotted in figure 10.14. All of Eckart's modes are modified so that $\sigma^2 > f_0^2$ save one ($n = 0$, traveling with the coast to its right), which persists as $\sigma \to 0$ and is a Kelvin-like mode. The others cease to be refractively trapped at superinertial ($>f_0$) individual cutoff frequencies bordering the continuum of leaky modes. At subinertial frequencies there is an infinite family of refractively trapped topographic Rossby waves, all traveling with the coast to their right (like the Kelvin mode) and all tending toward the constant frequency $s = -1/(2n + 1)$ at small wavelengths. This dispersion relation is qualitatively correct for most other shelf shapes. It differs from its equatorial β-plane counterpart (10.82) only in the absence of a mixed Rossby-gravity (Yanai) mode and in the tendency of short second-class modes to approach constant frequencies.

Topographic vortex stretching plus refraction of both first- and second-class waves are effective over any relief. Thus islands with beaches, submerged plateaus, and seamounts can in principal trap both first- and second-class barotropic waves (although these topographic features may have to be unrealistically large for their circumference to span one or more wavelengths of a trapped first-class wave). A submarine escarpment can trap second-class waves (then called *double Kelvin waves*; Longuet-Higgins, 1968b). Examples of such solutions are summarized by Longuet-Higgins (1969b) and by Rhines (1969b).

First-class waves trapped over the Southern California continental shelf have been clearly observed by Munk, Snodgrass, and Gilbert (1964), who computed the dispersion relation for the actual shelf profile and found (figure 10.15) sea-level variation to be closely confined to the dispersion curves thus predicted for periods of order of an hour or less. Both first- and second-class coastally trapped waves may be variously significant in coastal tides [Munk, Snodgrass, and Wimbush (1970) and section 10.5.2]. At longer periods, a number of observers claim to have detected coastally trapped second-class modes (Leblond and Mysak, 1977). A typical set of observations is shown in figure 10.16, after R. L. Smith (1978).

10.4.7 Long Waves over Relief with Rotation and Stratification

The two mechanisms of refraction and vortex stretching that govern the behavior of long waves propagating in homogeneous rotating fluid over bottom relief are sufficiently well understood that qualitatively correct dispersion relations may be found intuitively for quite complex relief even though their quantitative construction might be very involved. Stratification complicates the picture greatly. In this section, emphasis is upon problems with stratification that may be solved with sufficient completeness that they augment our intuition.

By appealing to the quasi-geostrophic approximation, Rhines (1975; 1977) has given a far-reaching treatment of the interplay between beta, weak bottom slope, and stratification for second-class waves. If equations (10.30a) and (10.30b) are cross-differentiated to eliminate pressure, and continuity (10.30d) is then invoked, the result is

$$\frac{\partial}{\partial t}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + \beta v - f_0 \frac{\partial w}{\partial z} = 0$$

in the β-plane approximation (section 10.4.4) $f = f_0$, $\beta = \partial f/\partial y$. Now this equation is recast as an approxi-

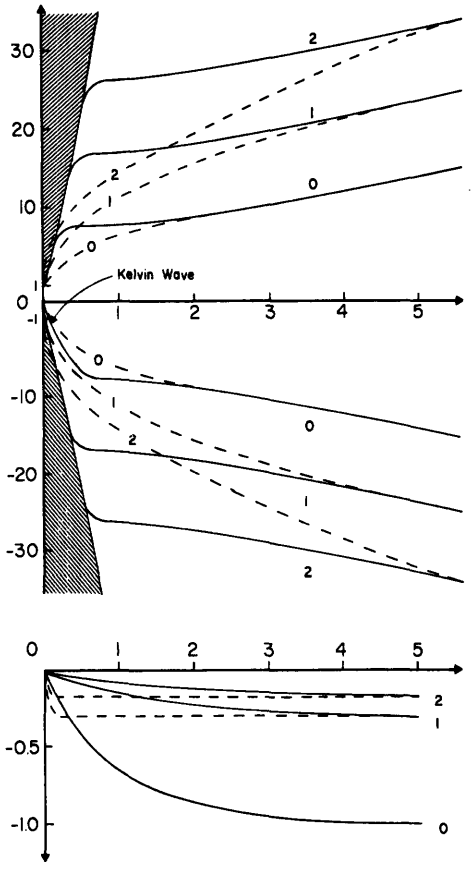


Figure 10.14 The dispersion relation (solid lines) for edge and quasi-geostrophic shallow-water waves over a uniformly sloping beach (slope $a = 2 \times 10^{-3}$) terminating in a flat ocean floor at a depth of $D = 5000$ m. The dotted lines are for the semi-infinite uniformly sloping beach. The shaded region is the continuum of leaky modes. Axes are as in figure 10.12. (Leblond and Mysak, 1977.)

mate equation in p by using (10.30c) plus the geostrophic approximation to obtain

$$\frac{\partial}{\partial t}\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{f_0^2}{N_0^2}\frac{\partial^2 p}{\partial z^2}\right) + \beta\frac{\partial p}{\partial x} = 0. \tag{10.103}$$

This result emerges from a more systematic treatment (Pedlosky, 1964a) as the linearized quasi-geostrophic approximation. It is here specialized to the case of constant buoyancy frequency N_0. The free surface may be idealized as rigid without loss of generality; the corresponding condition on p is

$$\frac{\partial p}{\partial z} = 0 \qquad \text{at} \quad z = 0. \tag{10.104}$$

The inviscid bottom boundary condition $w = av$ at the north-south sloping bottom $z = -D_0 + ay$ becomes, in quasi-geostrophic approximation,

$$-\frac{\partial^2 p}{\partial t\,\partial z} + a\frac{N_0^2}{f_0}\frac{\partial p}{\partial x} = 0 \qquad \text{at} \quad z = -D_0, \tag{10.105}$$

and is linearized about $z = -D_0$ for sufficiently small a.

With no bottom slope, solutions of (10.103)–(10.105) are

$$P = \cos(\lambda z)\exp(-i\sigma t + ilx + iky), \tag{10.106}$$

$$\sigma = -\beta l/(l^2 + k^2 + \lambda^2 f_0^2/N_0^2), \tag{10.107}$$

with λ given by (10.105) as a solution of

$$\sin(\lambda D_0) = 0 \tag{10.108}$$

i.e.,

$$\lambda = n\pi/D_0, \qquad n = 0, 1, 2, \ldots. \tag{10.109}$$

These correspond to the barotropic ($n = 0$) and baroclinic ($n = 1, 2, \ldots$) Rossby waves of section 10.4.4.

With no beta but with bottom slope, solutions of (10.103)–(10.105) are

$$P = \cosh(\lambda z)\exp(-i\sigma t + ilx + iky), \tag{10.110}$$

$$\lambda = (N_0/f_0)(l^2 + k^2)^{1/2}, \tag{10.111}$$

$$\sigma = \frac{alN_0^2\coth(\lambda D_0)}{\lambda f_0}. \tag{10.112}$$

If $\lambda D_0 \ll 1$, p is virtually depth independent and the dispersion relation (10.112) becomes

$$\sigma = D_0^{-1} a f_0 l/(l^2 + k^2). \tag{10.113}$$

This is a barotropic topographic Rossby wave with vortex stretching over the relief playing the role of beta. If $\lambda D_0 \gg 1$, p decays rapidly away from the bottom and the dispersion relation (10.112) becomes

$$\sigma = N_0 l/(l^2 + k^2)^{1/2}. \tag{10.114}$$

Such *bottom-trapped* motions are of theoretical importance because they allow mid-latitude quasi-geostrophic vertical shear and density perturbations at periods much shorter than the very long ones predicted by the flat-bottom baroclinic solutions (10.106)–(10.109). Rhines (1970) generalizes this bottom-trapped solution to relief of finite slope and points out that it reduces to the usual baroclinic Kelvin wave at a vertical boundary. Figure 10.17 shows what appear to be motions of this type.

With both beta and bottom slope, solutions of (10.103) and (10.104) are

$$P = \begin{matrix}\cos \\ \cosh\end{matrix}(\lambda z)\exp(-i\sigma t + ilx + iky) \tag{10.115a,b}$$

$$\sigma = -\beta l/(l^2 + k^2 \pm \lambda^2 f_0^2/N_0^2), \tag{10.116a,b}$$

with λ given by (10.105) as

$$\lambda\tan(\lambda D_0) = -\frac{aN_0^2}{f_0}\frac{l}{\sigma} \tag{10.117a}$$

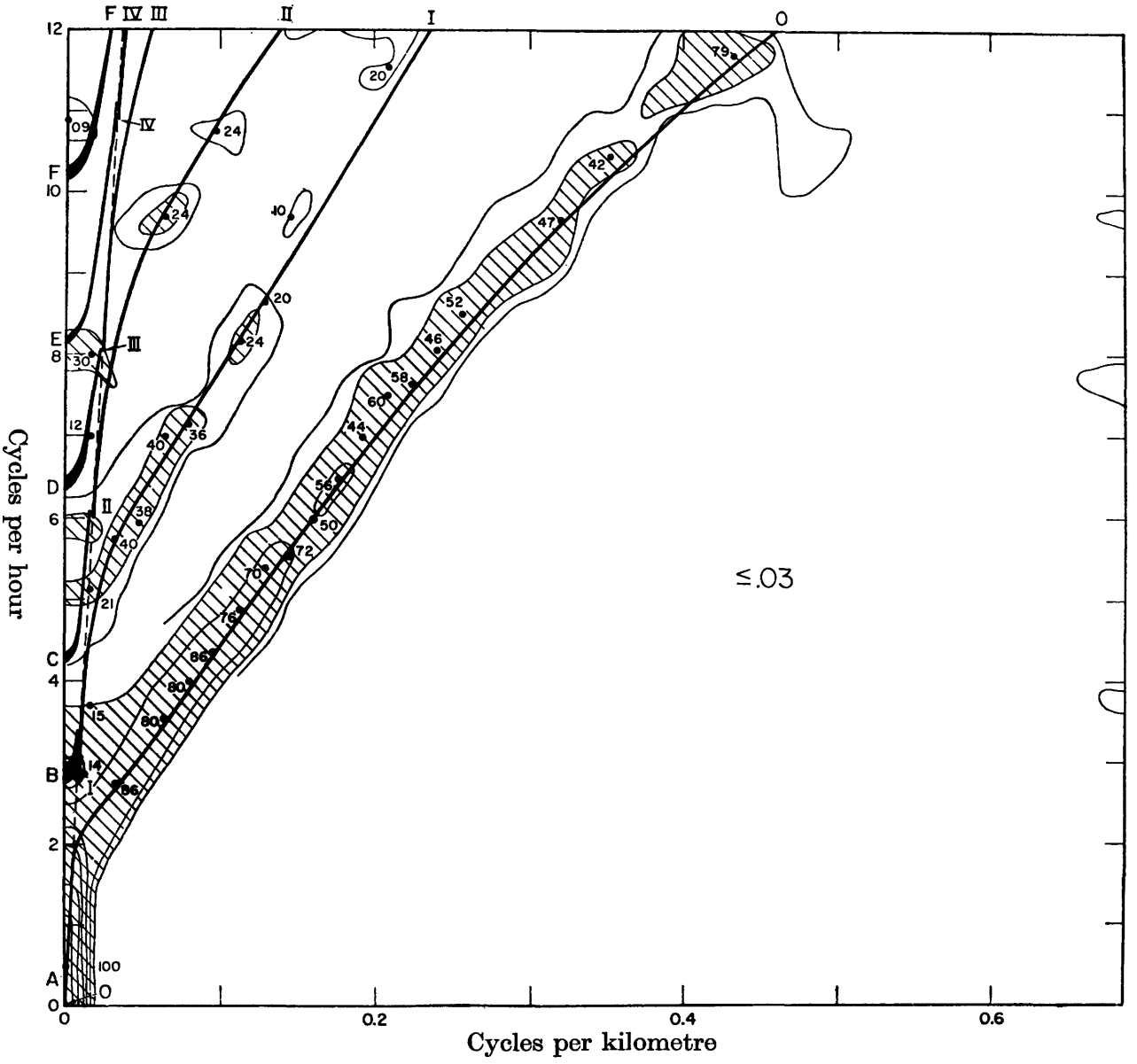


Figure 10.15 Comparison of theoretical and observed dispersion for the California continental shelf. Heavy lines 0–IV correspond to theoretical dispersion relations for the first five trapped first-class modes. The dashed line bounds the continuum of leaky modes. The observed normalized two-dimensional cospectrum of bottom pressure is contoured for values of 0.03, 0.05, 0.10, 0.25, 0.50, 0.75, and 0.90 with the area above 0.05 shaded. (Munk, Snodgrass, and Gilbert, 1964.)

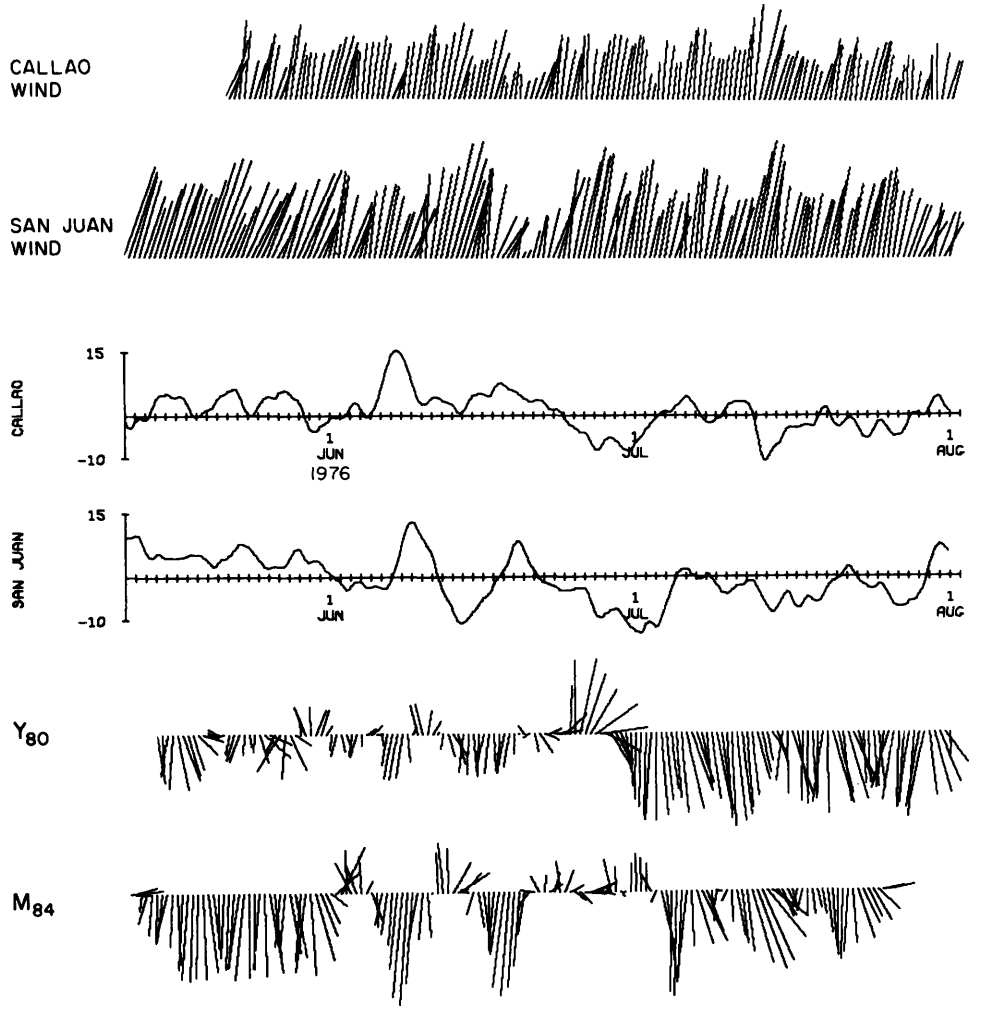


Figure 10.16 Low-passed wind vectors and sea-level records from Callao (12°04′ S) and San Juan (15°20′ S), Peru, and current vectors from Y (80 m below surface off Callao) and from M (84 m below surface off San Juan). Sea level and currents show propagation of events along the coast; wind records do not. (R. L. Smith, 1978.)

for case (a) and

$$\lambda \tanh(\lambda D_0) = \frac{aN_0^2}{f_0}\frac{1}{\sigma} \tag{10.117b}$$

for case (b). Equation (10.117b) has one root λ corresponding to a bottom-trapped wave for large a and to a barotropic β-wave for vanishing a. For vanishing a, (10.117a) reproduces the familiar flat-bottom barotropic and low-frequency baroclinic modes $\lambda = n\pi/D_0$, $n = 0, 1, \ldots$. When a is large the baroclinic roots are shifted toward $\lambda \simeq (\pi/2 + n\pi)/D_0$, so that the pressure (10.115a) and hence the horizontal velocity have a node at the bottom. There is thus a tendency for relief to result in the concentration of low-frequency baroclinic energy away from the bottom. With more realistic stratification this concentration is increasingly in the upper ocean. Rhines (1977) therefore calls such motions "thermocline eddies" and suggests that they are relevant to the interpretation of the observations of figure 10.18. It is straightforward to allow for an arbitrary direction of the bottom slope, but the results are not easy to summarize. Rhines (1970) gives a complete discussion.

A powerful treatment of second-class motion in a rotating stratified fluid over the linear beach $D = ax$ has been provided by Ou (1979), Ou (1980), and Ou and Beardsley (1980). They have generously permitted me to make use of their results in this discussion. Neglecting free surface displacement (so that $w = 0$ at $z = 0$) and eliminating u,v,w from (10.30) in favor of p yields

$$\frac{\partial^2 p}{\partial z^2} + \frac{N^2}{f_0^2 - \sigma^2}\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right) = 0, \tag{10.118}$$

$$\frac{\partial p}{\partial z} = 0 \qquad \text{at} \quad z = 0, \tag{10.119}$$

$$a\left(i\sigma\frac{\partial p}{\partial x} - f_0\frac{\partial p}{\partial y}\right) + i\sigma\frac{f_0^2 - \sigma^2}{N_0^2}\frac{\partial p}{\partial z} = 0 \qquad \text{at} \quad z = -ax \tag{10.120}$$

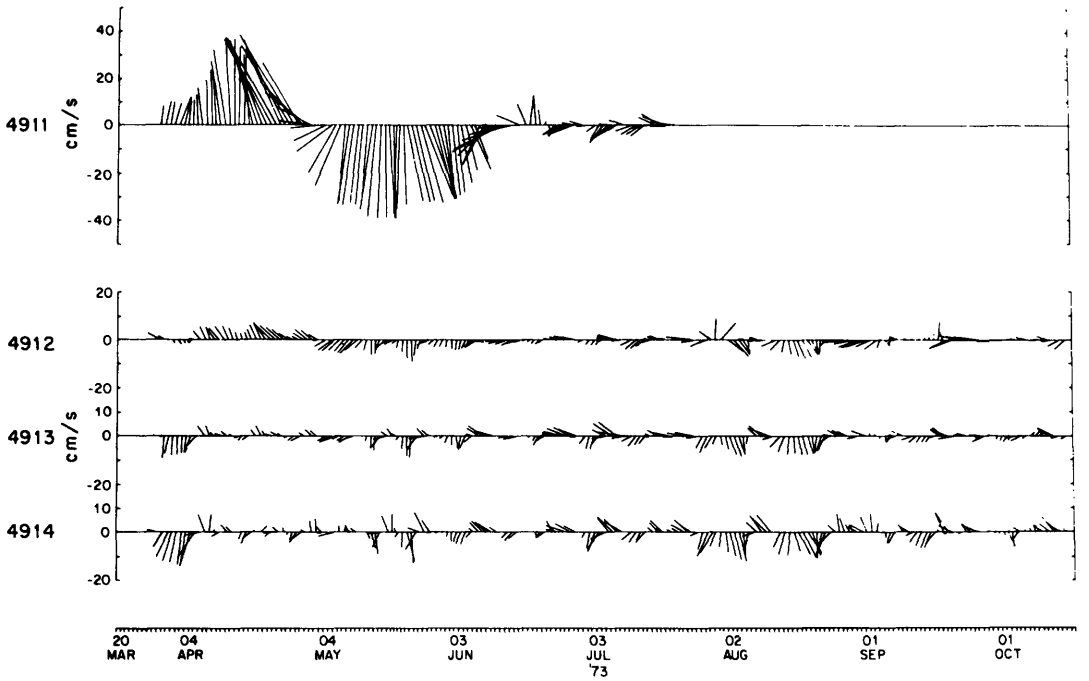


Figure 10.17A Currents at 39°10'N, 70°W (site D) at 205, 1019, 2030, and 2550 m. The total depth is 2650 m. A "thermocline" eddy initially dominates the upper flow.

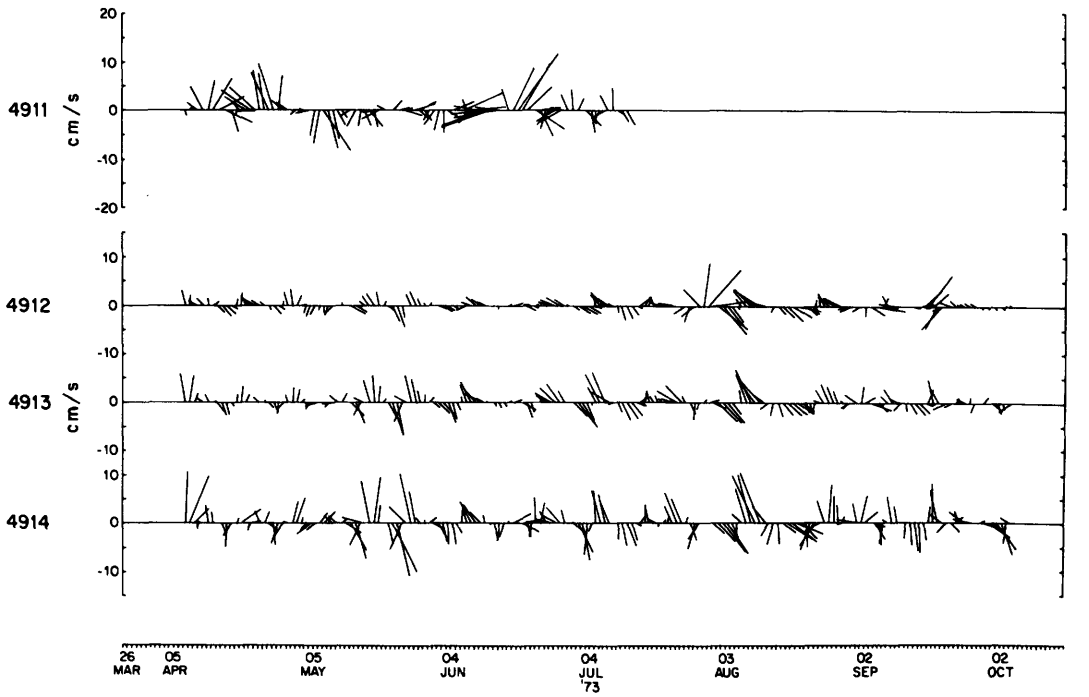


Figure 10.17B A high-passed version of figure 10.17A. The lower layers are now dominated by fast-bottom intensified oscillations. (Rhines, 1977.)

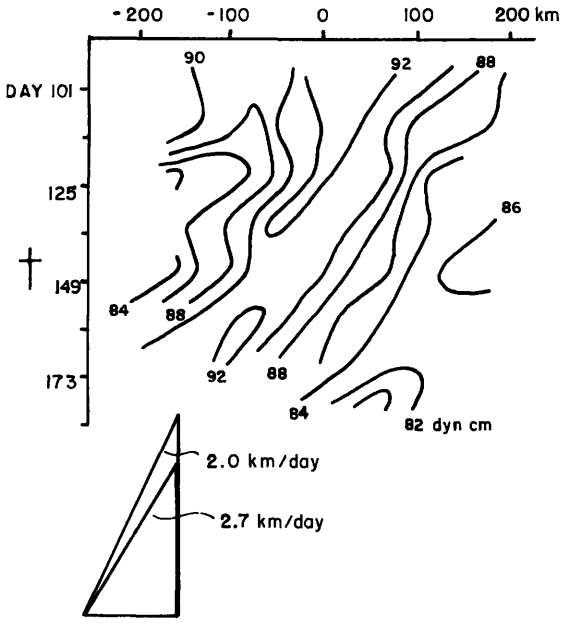


Figure 10.18 Time-longitude plot of 501-1500-db dynamic height along 28'N from MODE showing westward propagation of "thermocline eddies" but at a rate significantly slower than that observed in 1500-m currents (figure 10.6). (Rhines, 1977.)

as the boundary-value problem governing periodic [exp(−$i\sigma t$)] second-class (and low-frequency first-class internal) waves over the linear beach. Ou (1979) saw that the affine transformation

$$z = \hat{z}(f_0^2 - \sigma^2)^{1/2}/N_0 \tag{10.121}$$

followed by a rotation of coordinates from $(x, y, \hat{z})$ to (x', y, z') such that the beach $z = -ax$ now becomes $z' = 0$ leads to

$$\frac{\partial^2 p}{\partial z'^2} + \frac{\partial^2 p}{\partial x'^2} - k^2 p = 0, \tag{10.122}$$

$$\frac{\partial p}{\partial n} = \frac{kf_0}{\sigma}\frac{a'}{(1 + a'^2)^{1/2}} p \qquad \text{at} \quad z' = 0, \tag{10.123}$$

$$\frac{\partial p}{\partial n} = 0 \qquad \text{at} \quad z' = a'x' \tag{10.124}$$

for motions periodic in y[exp(iky)], and furthermore that this is *exactly* Ursell's (1952) problem just turned upside down; (10.122), (10.123), and (10.124) correspond to (10.93), (10.94), and (10.97). The transformed beach slope a' is given by

$$a' = aN(f_0^2 - \sigma^2)^{-1/2}. \tag{10.125}$$

For second-class waves

$$\sigma^2 < f_0^2, \tag{10.126}$$

a' is real and all of Ursell's (1952) results are immediately available. There are thus coastally trapped waves whose dispersion relation is

$$\sigma = f_0 a'/\{(1 + a'^2)^{1/2} \sin[2n + 1)\tan^{-1}a']\} \tag{10.127}$$

as well as a continuum of bottom-trapped waves that have the form

$$P = \cos(lx' + \text{phase})\exp[-i\sigma t + iky - (l^2 + k^2)^{1/2}z']$$

as $x' \to \infty$ and that have their dispersion relation included in

$$\sigma = -N \sin(\tan^{-1} a)\left[\frac{k}{(l^2 + k^2)^{1/2}}\right]. \tag{10.128}$$

The latter are just the bottom-trapped waves (10.110)–(10.112) of Rhines (1970).

Equation (10.128) is their dispersion relation in a half-plane bounded by the sloping bottom (Rhines, 1970). The frequency may be either sub- or super-inertial. The coastally trapped waves have $\sigma > N \sin(\tan^{-1} a)$ [by analogy with the fact that for Ursell's (1952) edge waves $\sigma^2 < gk$] and there are a finite number of them:

$$n = 1, 2, \ldots < [\pi/(4 \tan^{-1} a') - 1/2]. \tag{10.129}$$

Note that $n = 0$ would imply $\sigma = f_0$ but this does not solve the full equations (10.30) and associated boundary conditions.

In the limit of decreasing slope a', (10.129) allows ever more modes, and the dispersion relation (10.127) for coastally trapped waves simplifies to

$$\sigma = -f_0(2n + 1). \tag{10.130}$$

This is the low-frequency, short-wavelength second-class limit of the barotropic dispersion relation (10.98). We thus identify Ou's coastally trapped modes as the stratified analog of the already familiar refractively trapped second-class topographic Rossby waves.

Over the linear beach, then, stratification limits the number of second-class refractively trapped topographic waves and opens up a new continuum of bottom-trapped waves. Figure 10.19 compares barotropic and baroclinic dispersion relations when the ocean surface is rigid. Further results are given by Ou (1979).

Suppose now that the linear beach terminates in a flat bottom of depth D_0, as in (10.102). Ou's (1979) transformation allows us to deal efficiently with the stratified problem. Figure 10.20 summarizes the boundary-value problem and its alteration by Ou's transformation into an equivalent problem in deep-water waves (figure 10.20D). In this latter problem, the deep-water continuum that existed for the linear beach must now be quantized into an infinite family of modes by repeated reflection between the shoaling and the

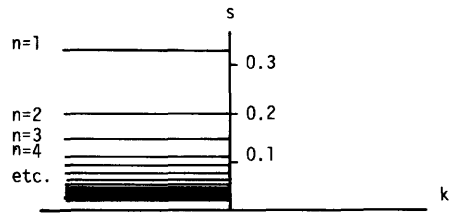


Figure 10.19A The dispersion relation (10.130), viz., $s = -1/(2n + 1)$ for topographic Rossby waves refractively trapped over the linear beach $D = -ax$ beneath rotating homogeneous fluid. There are an infinite number $n = 1,2,...$ of trapped modes.

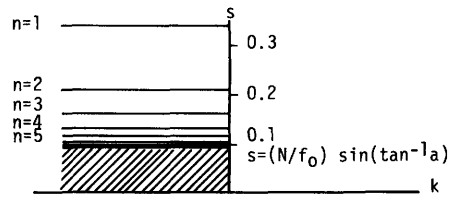


Figure 10.19B The dispersion relation (10.127)

$$s = a'/\{(1 + a'^2)^{1/2} \sin[(2n + 1)\tan^{-1} a']\},$$

where $a' = aN_0/[f_0(1-s^2)^{1/2}]$, for topographic Rossby waves refractively trapped over the linear beach $D = -ax$ beneath uniformly stratified (buoyancy frequency N_0) rotating fluid. There are a finite number $n = 1,2,... < [\pi/(4\tan^{-1}a') - \frac{1}{2}]$ of trapped modes all with frequencies $s > N_0 f_0^{-1} \sin[\tan^{-1} a]$. At lower frequencies a continuum of bottom-trapped modes reflected from $x = \infty$ exists. Sketch is for $N_0 = 10f_0$, $a = 10^{-2}$.

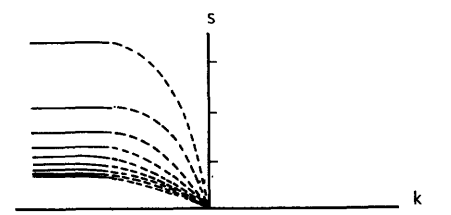


Figure 10.19C Sketch of dispersion relation for topographic Rossby waves in rotating stratified fluid over a linear beach that terminates in a uniform-depth ocean. The continuum of bottom-trapped modes that existed over the semi-infinite beach is quantized. All dispersion curves pass through $s = 0$ as $k \to 0$.

overhanging coasts. Low-mode edge waves have small amplitude at the overhanging coast and are not much affected by it. But higher-mode edge waves have appreciable amplitude at the overhanging coast and they blend smoothly into the infinite family of modes made up of waves repeatedly reflected between beach and the overhanging coast. All these results have direct analogs in the original stratified problem (figure 10.20A). There are a number of second-class topographic waves that are refractively trapped near the coast and that have decayed to very small amplitudes at the seaward termination of the beach. The continuum of bottom-trapped waves present over the unending linear beach is replaced by an infinite family of bottom-trapped waves reflected repeatedly between the coast and the seaward termination of the beach. In the special case $a = \infty$ of a perpendicular coast, their dispersion relations are easily seen to be

$$\sigma = ND_0 k/n\pi \tag{10.131}$$

and they then correspond to ordinary internal Kelvin waves (in this case there are no refractively trapped modes). Other shelf geometries invite similar treatment. Figures 10.20E and 10.20F show the equivalent deep-water wave problem for a step shelf. Deep-water waves on surface 1 of figure 10.20F that are short enough that their particle displacements at the level of surface 2 are negligible correspond in the stratified problem (figure 10.20E) to internal Kelvin waves trapped against the coast. Deep-water waves on surface 2 of figure 10.20F correspond to baroclinic counterparts of the double Kelvin wave that, in homogeneous fluid, may be trapped along a discontinuity in depth (Longuet-Higgins, 1968b; see also section 10.4.6).

In general, if the equivalent deep-water wave problem has the waveguide-like dispersion relations $\sigma^2/g = F_n[H,\gamma,k]$, $n = 1, 2, \ldots$, with H,γ as defined in figure 10.20D, then the corresponding stratified shelf problem must have the dispersion relation

$$\frac{kf}{\sigma} - \frac{a'}{(1 + a'^2)^{1/2}} = F_n[\{(D_0 a'/a)^2 + D_0 a^{-2}\}^{1/2}, \tan^{-1} a', k] \tag{10.132}$$

so that for all the dispersion curves $\sigma \to 0$ as $k \to 0$. The dispersion relation must thus qualitatively look like figure 10.19C. Wang and Mooers (1976) have generalized the problem numerically to more complex shelf profiles with nonuniform stratification $N(z)$.

Ou's (1979) transformation is most useful for second-class motions (10.125) because for them the transformed coordinates are real. For first-class motions with $\sigma^2 > f_0^2$ it still produces Laplace's equation but now z' is imaginary. Wunsch (1969) has nonetheless been able to use it to discuss first-class internal waves obliquely incident on the linear beach without rota-

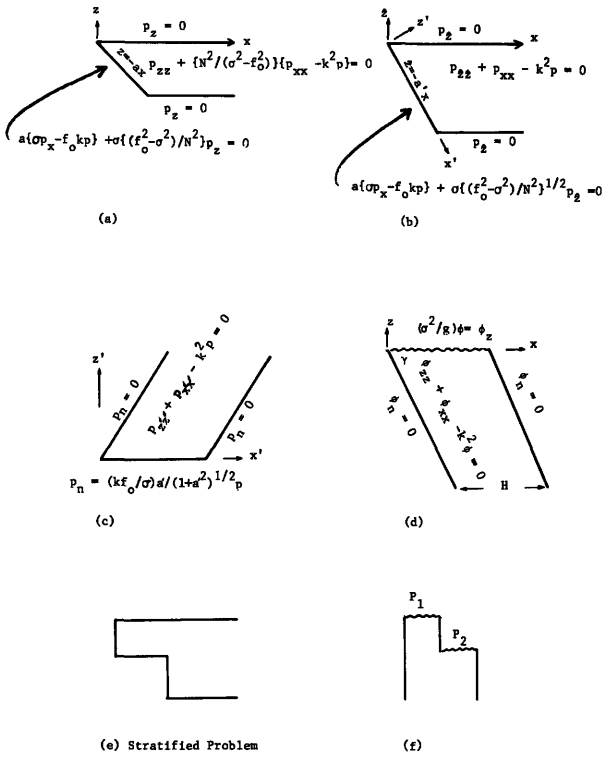


Figure 10.20 The deep-water surface-wave analog (d,f) of two shelf problems involving topographic Rossby waves in uniformly stratified rotating fluid: (a) stratified problem; (b) result of affine transformation; (c) result of rotation; (d) equivalent deep-water problem (velocity potential ϕ); (e) stratified problem; (f) equivalent deep-water problem (atmospheric pressure P_1 must be maintained lower than P_2 for physical realizability).

tion. This is the stratified analog of Eckart's (1951) nonrotating LSW study of waves over a sloping beach (section 10.4.6).

For beach slopes much smaller than the slope (σ/N) of (low-frequency) internal wave characteristics (10.43), Wunsch thus found that internal waves are refracted just like surface gravity waves by the shoaling relief and that refractively trapped edge modes occur. From the dispersion relation

$$\left(\frac{n\pi}{D_0}\right)^2 = \frac{N^2 - \sigma^2}{\sigma^2}(l^2 + k^2)$$

for plane internal waves of the form

$$w = \sin\left(\frac{n\pi z}{D_0}\right)\exp(-i\sigma t + ilx + iky)$$

over a uniform bottom D_0, l must ultimately become imaginary if D_0 is allowed to grow parametrically offshore while n and k are held fixed. One would therefore expect a WKB treatment of internal waves over gently shoaling relief to result in refraction and refractive trapping *provided* that the mode number n does not change, i.e., provided that the relief does not scatter energy from one mode into others. Constancy of n is indeed a feature of Wunsch's solutions but it cannot be expected to hold for more abrupt relief, especially if the relief slope exceeds the characteristic slope. If the relief couples modes efficiently, then scattering into higher modes allows l to remain real even in deep water far from shore so that energy is not refractively trapped near the coast. In principle, scattering into internal modes thus even destroys the perfect trapping of long surface gravity waves predicted by LSW theory over a step shelf, but in practice appreciable trapping is often observed. The efficiency of mode coupling depends both on the relief and on the vertical profile $N(z)$ of the buoyancy frequency, so that a general result for internal waves is difficult to formulate.

10.4.8 Free Oscillations of Ocean Basins

Finding the free oscillations allowed by LTE in rotating ocean basins is difficult even in the f-plane (section 10.4.2). Platzman (1975, 1978) has developed powerful numerical techniques for finding the natural frequencies and associated flow fields of free oscillations allowed by LTE in basins of realistic shape and bottom relief. The general classification of free oscillations into first- and second-class modes characteristic of the idealized cases discussed in sections 10.4.2 and 10.4.5 (effectively for a global basin) persists in Platzman's (1975) calculations. For a basin composed of Atlantic and Indian Oceans, there are 14 free oscillations with periods between 10 and 25 hours. Some of these are very close to the diurnal and semidiurnal tidal periods, and all of them, being within a few percentage points of equipartition of kinetic and potential energies, are first-class modes. There are also free oscillations of much longer period, for which potential energy is only about 10% or even less of kinetic energy; they are second-class modes.

I know of no extratidal peaks in open-ocean sea-level records that correspond to these free oscillations. There is some evidence in tidal admittances for the excitation of free modes but the resonances are evidently not very sharp (see section 10.5.1). Munk, Bryan, and Zetler (private communication) have searched without success for the intertidal coherence of sea level across the Atlantic that the broad spatial scale of these modes implies. The modes are evidently very highly damped.

10.5 The Ocean Surface Tide

10.5.1 Why Ocean Tides Are of Scientific Interest

The physical motivation for studying and augmenting the global ensemble of ocean-tide records has expanded enormously since Laplace's time. In this section I have tried to sketch the motivating ideas without getting

involved in the details of theoretical models; some of these receive attention in subsequent sections.

Certain of the ancients knew a great deal about tides [see, e.g., Darwin's (1911a) summary of classical references], but the first extant reduction of observations made explicitly for predictive purposes may be the table of "flod at london brigge" due to Wallingford who died as Abbot of St. Alban's in 1213 (Sager, 1955). Making practical tide predictions was probably *the* preoccupation of observers for the next 500 years.

In 1683, Flamsteed (Sager, 1955) produced a table of high waters for London Bridge as well as, in the following year, corrections making it applicable to other English ports. Darwin (1911a) quotes Whewell's description, written in 1837, of how successors to Flamsteed's tables were produced:

> The course . . . would have been to ascertain by an analysis of long series of observations, the effects of changes in the time of transit, parallax, and the declination of the moon and thus to obtain the laws of phenomena.
>
> . . . Though this was not the course followed by mathematical theorists, it was really pursued by those who practically calculated tide tables. . . . Liverpool London, and other places had their tables, constructed by undivulged methods . . . handed down from father to son.
>
> . . . The Liverpool tide tables . . . were deduced by a clergyman named Holden, from observations made at that port . . . for above twenty years, day and night. Holden's tables, founded on four years of these observations, were remarkably accurate.
>
> At length men of science began to perceive that such calculations were part of their business. . . . Mr. Lubbock . . . , finding that regular tide observations had been made at the London docks from 1795, . . . took nineteen years of these . . . and caused them to be analyzed. . . . In a very few years the tables thus produced by an open and scientific process were more exact than those which resulted from any of the secrets.

Quite aside from its proprietary aspects, Darwin (1911b) explicitly notes the synthetic nature of this process; it at least conceptually represents "the oscillation of the sea by a single mathematical expression" provided by Bernoulli in 1738 for an inertialess ocean (the equilibrium tide), by Laplace for a global ocean obeying Newton's laws of motion, and assumed to exist for actual oceans even if too complex to represent in simple form.

Kelvin, in about 1870 (Darwin, 1911b) introduced the harmonic method, which Darwin (1911b) calls "analytic" because synthesis of the entire tide into one dynamically derived form is abandoned and instead the tide at any given place is regarded as a sum of harmonic oscillations whose frequencies are determined from astronomy (section 10.2) but whose amplitudes and phases must be determined from analysis of *in situ* sea-level observations. Prediction is then carried out by recombining the harmonic oscillations at future times.

Kelvin's suggested procedure was made feasible by the introduction of recording tide gauges in which the motion of a float in a well, insulated from short-period waves but otherwise freely connected with the sea, drives a pencil up and down a paper wrapped on a drum rotated by clockwork, thus producing a continuous plot of sea level versus time. [Darwin (1911a) describes contemporary instruments.] Harmonic analysis of this record *at relatively few astronomically determined frequencies* was feasible by judicious sampling and manual calculation. The recombination of harmonics at future times was then carried out mechanically by means of a series of pulleys, movable at frequencies corresponding to the astronomical ones, that drove a pencil over a paper wrapped on a drum rotated by clockwork, thus ultimately providing a plot of predicted sea level versus time. The design of such a machine was due to Kelvin, and elaborations were in regular use until the mid-1960s (Zetler, 1978).

Even before Kelvin's introduction of the harmonic method, Lubbock and Whewell (Darwin, 1911b) had begun to combine observations at different ports into cotidal maps showing the geographical variation of sea level associated with tides. Thus Airy in 1845 gave a chart (modified by Berghaus in 1891) of locations of high water at different times of day in the North Sea (figure 10.21). Concerning this, Darwin (1911b) remarks, "It will be noticed that between Yarmouth and Holland the cotidal lines cross one another. Such an intersection of lines is in general impossible; it is indeed only possible if there is a region in which the water neither rises nor falls. . . . A set of observations by Captain Hewitt, R.N. made in 1840 appears to prove the existence of a region of this kind." This is probably the first recorded observation of an amphidromic point.

But hourly maps of high-water locations change throughout the month. Kelvin's harmonic analysis decomposed the tide into harmonic components for which a single cotidal map, with cotidal lines drawn at fractions of the period of the component, can represent the entire spatial variation of that component forever. Time series at thousands of ports may thus be reduced to a handful of global maps that are ideal summaries of observations for comparison with solutions of LTE forced by the different harmonic components of the ATGF. Kelvin's abandonment of the "synthetic" viewpoint thus in effect provided the means for its reinstatement.

Of the handful of such maps constructed empirically for global tides, Dietrich's (1944) are perhaps the most widely quoted. Villain (1952) gives an extensive discussion of the observations leading to his global M_2 cotidal map (figure 10.22). Much modern tidal research has consisted of attempts to apply the principles of

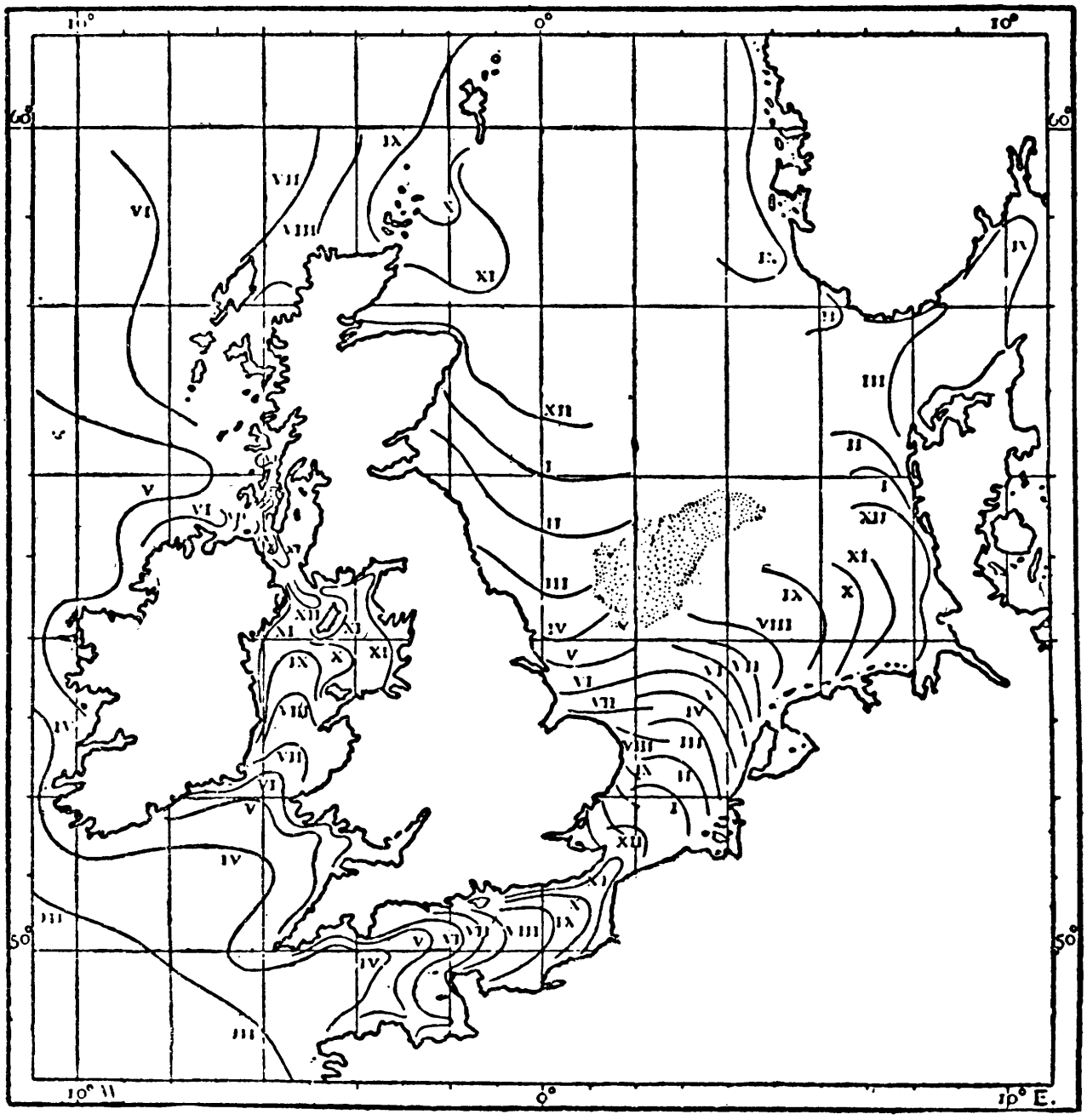


Figure 10.21 Airy's chart of cotidal lines in British seas. (Darwin, 1911b.)

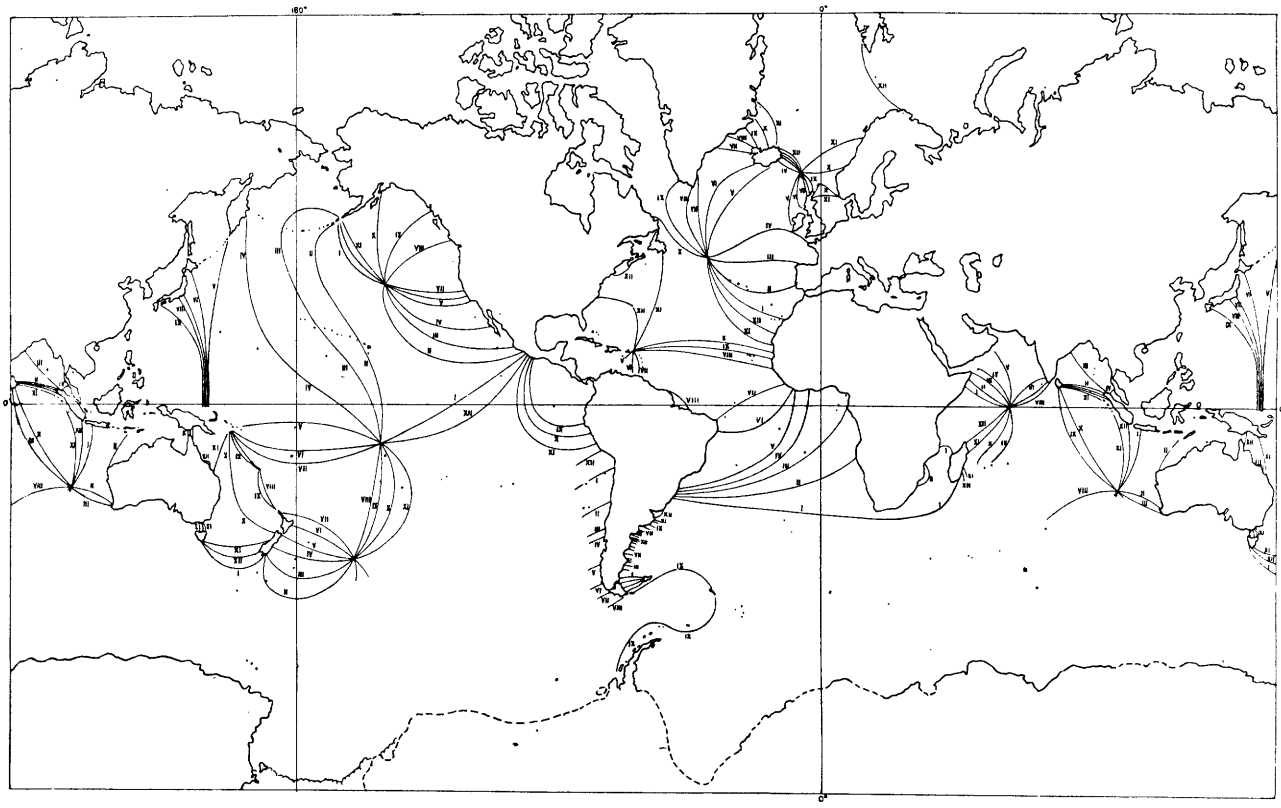


Figure 10.22 Cotidal lines for M_2 (in lunar hours relative to moon's transit over Greenwich). (Villain, 1952.)

dynamics to reproduce and hence to "explain" the global distribution of tides as suggested by such empirical maps. But the degree of success achieved to date, as well as insight into the variation of the dynamics of tides over the globe, has required thinking about how the response of the ocean would change if tidal frequencies could be varied. Since they cannot be, this implies comparing global tidal maps at different tidal periods. The origins of this viewpoint are found in the work of Munk and Cartwright (1966), who were enabled by the advent of modern computers to analyze 19 years of hourly tide readings at Honolulu and Newlyn "without astronomical prejudice as to what frequencies are present and what are not, thus allowing for background noise."

Their work has been influential in two very general ways quite apart from the improvement in tide prediction that it introduced. First, it provided a clear distinction between sea-level fluctuations due to TGF and those of similar period due to nontidal agents, a distinction crucial in establishing the significance of any geophysical interpretation of all but the strongest constituents of ocean or solid-earth tides. Second, it introduced the idea of oceanic admittance, the (possibly complex) ratio between ocean response and forcing, as a continuous function of frequency that can be estimated from tidal observations and that summarizes the dynamic response of the ocean to time-variable forcing in a manner easily related to the properties of free solutions of LTE by an expansion in eigenfunctions.

If the ocean had many sharp resonances within the frequency bands spanning the three species, the tidal admittance would have amplitude peaks and rapid phase shifts. Typical deep-sea admittances tend to be smooth across a species but are far from constant. Admittance curves for the Coral Sea (Webb, 1974) and at Bermuda (Wunsch, 1972c) are shown in figures 10.23A and 10.23B, respectively. The Coral Sea admittance is unusual in its very sharp sudden variation between M_2 and S_2, apparently showing the existence of a sharp local resonance. The amplitude of the Bermuda admittance rises smoothly, by 400% toward lower frequencies over the semidiurnal band; Wunsch's result is consistent with Platzman's (1975) prediction of an Atlantic resonance of roughly 14-hour period, but one appreciably broadened by dissipation.

Smoothness of the admittance across tidal bands was anticipated by Munk and Cartwright (1966) in their "credo of smoothness": "We do not believe, nor will we tolerate, the existence of very sharp resonance peaks." In part, this credo had its origin in the prevailing beliefs, since then largely confirmed, that ocean

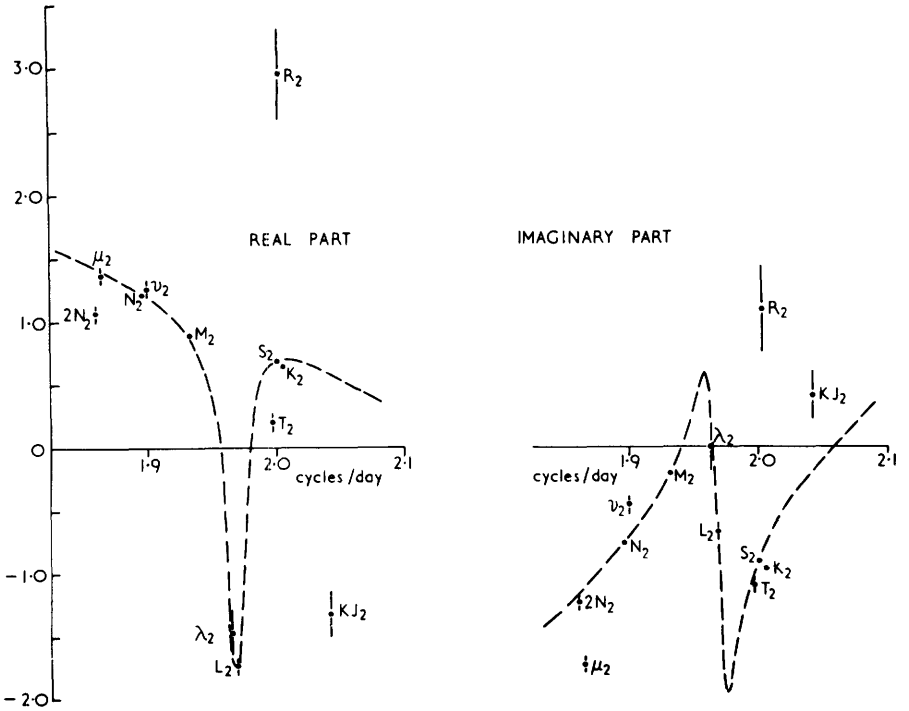


Figure 10.23A The real and imaginary parts of the response function at Cairns (16°55'S, 145°47'E), showing a resonance in the Coral Sea. (Webb, 1974.)

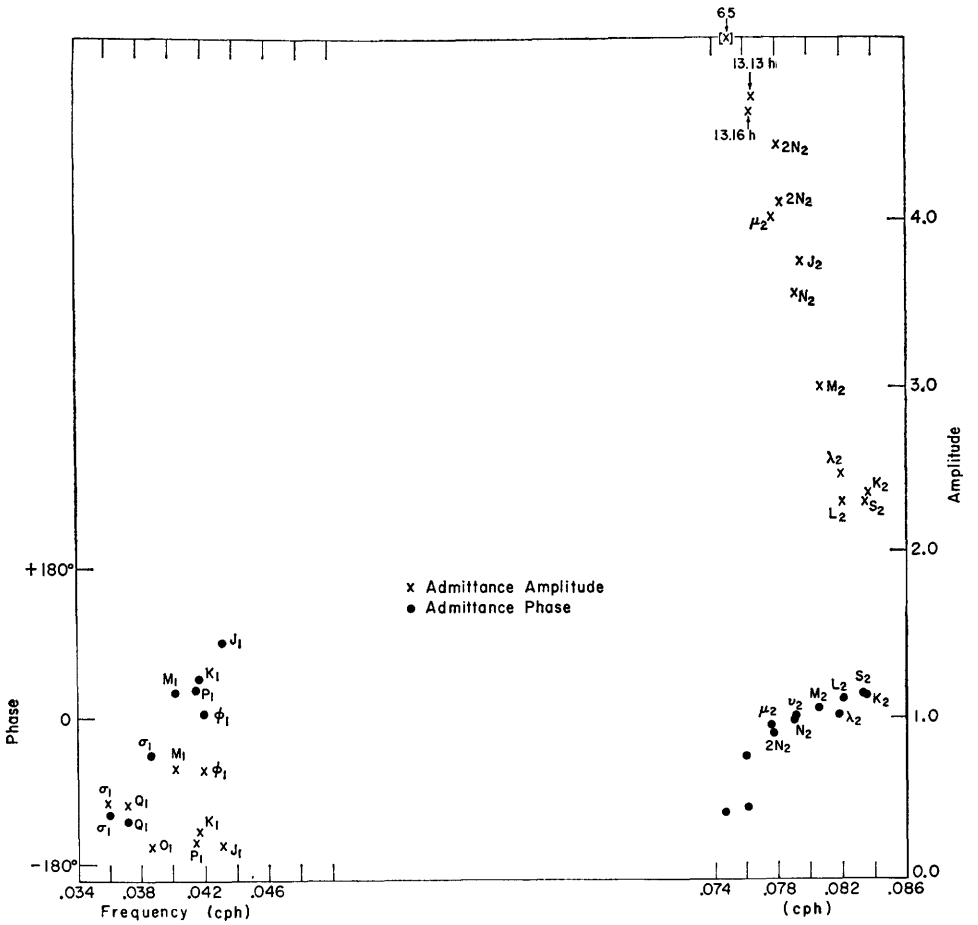


Figure 10.23B Amplitude and phase of the admittance at Bermuda. (Wunsch, 1972c.)

tides must be of rather low Q. Evaluation of Q requires knowledge of the total energy E stored in the tide as well as the rate $\dot{E}$ at which it is dissipated; then

$$Q = \frac{2\pi E}{\dot{E}T} .$$

Estimation of the stored energy E was difficult before modern numerical solutions of LTE because of the open-ocean detail required. Earliest estimates assumed the tide to be in equilibrium; allowance for (likely) equipartition between potential and kinetic energy and for the area of the oceans led to an estimate of 5.6×10^{16} J for M_2 (Garrett and Munk, 1971). I interpolated coastal M_2 harmonic constants over the globe by solving LTE with these as boundary values and thus obtained (Hendershott, 1972) an estimate of 7.29×10^{17} J. But my kinetic energy was over twice my potential energy and I now believe this to have been a numerical artifact, especially since Platzman's (1975) near-tidal normal modes are within a few percentage points of equipartition. My estimate should thus be revised to 5.14×10^{17} J. Parke and Hendershott (1980) improved the interpolation by taking island data into account and found 2.68×10^{17} J (assuming equipartition).

The estimation of $\dot{E}$ historically has been of importance in cosmology. Halley in 1695 first discovered that the apparent position of the moon is *not* that predicted by (frictionless) Newtonian mechanics. The discrepancy is real; Munk (1968) outlines ultimately unsuccessful attempts to resolve it by appealing to perturbation of the moon's orbit by changes in the earth's orbit around the sun. Much of the discrepancy is now believed to be due to tidal friction. As Immanuel Kant noted in 1754, tidal friction must slow the earth's daily rate of rotation; this alone gives rise to an apparent perturbation of the moon's mean longitude. By the conservation of angular momentum, the moon's angular velocity about the earth-moon center of mass is also altered and (for the present prograde rotation of the moon about the earth) the moon recedes from the earth (by about 6 cm yr^{-1}; Cartwright, 1977). Müller (1976) reviews astronomical data both ancient (eclipse observations) and modern, and analyzes them simultaneously to estimate $\dot{n}_t$ (the tidal acceleration of the moon's longitude n), $\dot{\Omega}/\Omega$ (the observed apparent acceleration of the earth's rotation frequency Ω), $\dot{\Omega}/\Omega_{TNT}$ (the nontidal part of $\dot{\Omega}/\Omega$), and $\dot{G}/G$ (the possible rate of change of the gravitational constant G). He finds

$$\dot{n}_t = -27.2 \pm 1.7'' \text{ cy}^{-2},$$

$$\dot{\Omega}/\Omega = -22.6 \pm 1.1 \times 10^{-11} \text{ yr}^{-1},$$

the latter corresponding to a lengthening of day of 2.0×10^{-3} s cy^{-1}. If he assumes $\dot{G}/G = 0$, then $\dot{\Omega}/\Omega_{TNT}$ becomes $9.2 \pm 2.5 \times 10^{-11}$ yr^{-1}, a sizable portion of $\dot{\Omega}/\Omega$ that demands geophysical explanation. Various cosmological theories have $\dot{G}/G$ of order 5×10^{-11} yr^{-1}; $\dot{\Omega}/\Omega_{TNT}$ then becomes zero with an uncertainty of order 5×10^{-11} yr^{-1}. "It appears that either we really have a (non zero) cosmological constant $\dot{G}/G$ consistent with the Hubble constant, or we have a significant $\dot{\Omega}/\Omega_{TNT}$" [Müller (1976)].

Lambeck (1975) gives expressions for the tidally induced rates of change of the semimajor axis a of the moon's orbit, of its eccentricity e, and of its orbital inclination i in terms of a spherical harmonic decomposition of the ocean tide ζ_0. For semidiurnal tides only the second harmonic is important. Once these rates of change have been estimated, then $\dot{n}_t$ and $(\dot{\Omega}/\Omega)_t$ (i.e., the tidal part of $\dot{\Omega}/\Omega$) may be estimated from, respectively, Kepler's law (Cartwright, 1977, equation 8.3) and from the conservation of angular momentum (Lambeck, 1977, equations 2). $\dot{n}_t$ and $(\dot{\Omega}/\Omega)_t$ imply a rate $\dot{E}_t$ of tidal energy dissipation in the earth-moon system (Lambeck, 1977, equations 3). Using global calculations of ζ_0 for the M_2 ocean tide by Bogdanov and Magarik (1967), by Pekeris and Accad (1969) and by myself (Hendershott, 1972), Lambeck (1977) thus estimates for M_2

$$\dot{n}_t = -27.8 \pm 3'' \text{ cy}^{-} \qquad \text{(his table 7)},$$

$$(\dot{\Omega}/\Omega)_t = -25.8 \times 10^{-11} \text{ yr}^{-1} \qquad \text{(his table 8)},$$

$$\dot{E}_t = 3.35 \times 10^{19} \text{ erg s}^{-1} \qquad \text{(his equation 3b)}.$$

Since his work, new M_2 calculations by Accad and Pekeris (1978) and by Parke and Hendershott (1980) have appeared. These calculations include ocean self-attraction and loading (section 10.5.3) and are not unrealistically resonant. Accad and Pekeris (1978) directly evaluate the flow of M_2 energy out of the numerical ocean and obtain 2.44-2.79 $\times 10^{19}$ erg s^{-1}. Parke and Hendershott (1980) evaluate the rate $\langle \dot{W} \rangle$ at which the M_2 tide generating forces (potential Γ) and ocean floor (solid earth M_2 tide δ) do work on the ocean averaged ($\langle\ \rangle$) over a tidal period

$$\langle \dot{W} \rangle = \int\int_{\text{ocean}} \left(\rho \left\langle \Gamma \frac{\partial \zeta_0}{\partial t} \right\rangle + \rho g \left\langle \zeta_0 \frac{\partial \delta}{\partial t} \right\rangle \right) dA$$

(Hendershott, 1972) and obtain 2.22×10^{19} erg s^{-1}. All this work is lost in tidal friction. If these results are taken as an improved estimate $\dot{E}'_t$ of E_t for the M_2 tide

$$\dot{E}'_t = 2.2\text{-}2.8 \times 10^{19} \text{ erg s}^{-1},$$

then Lambeck's (1975) procedure would yield

$$\dot{n}'_t = -(18.3\text{-}23.2)'' \text{ cy}^{-2},$$

$$(\dot{\Omega}/\Omega)'_t = -(16.9\text{-}21.6) \times 10^{-11} \text{ yr}^{-1}$$

for M_2. If we retain unaltered Lambeck's (1977) estimate of the contribution $\Delta\dot{n}$ and $\Delta(\dot{\Omega}/\Omega)$ of all remaining tides to $\dot{n}_t$ and $(\dot{\Omega}/\Omega)_t$,

$\Delta\dot{n} = -3.1''\ \mathrm{cy}^{-2}$,

$\Delta(\dot{\Omega}/\Omega) = -6.9 \times 10^{-11}\ \mathrm{yr}^{-1}$,

then we obtain the revised estimates for all tides:

$n_t = (21.4\text{–}26.3)''\ \mathrm{cy}^{-2}$,

$(\dot{\Omega}/\Omega) = -(23.8\text{–}28.5) \times 10^{-11}\ \mathrm{yr}^{-1}$.

These are to be compared with Müller's (1976) estimates from astronomical data:

$\dot{n}_t = -27.2''\ \mathrm{cy}^{-2}$

and

$(\dot{\Omega}/\Omega)_t = -(13.4\text{–}22.6) \times 10^{-11}\ \mathrm{yr}^{-1}$

for

$\dot{G}/G = -(0\text{–}6.9) \times 10^{-11}\ \mathrm{yr}^{-1}$.

The comparison is worst if $\dot{G}/G$ is taken zero and becomes rather good if $\dot{G}/G$ is allowed to differ from zero.

There is thus some interest in estimating $\dot{E}$ for ocean tides but, as indicated above, results differ significantly depending on details of the estimation procedure. The estimates referred to above (except for that of Accad and Pekeris, 1978) essentially use global cotidal maps to find the part of the ocean tide in phase with the tide generating forces. The resulting rate of working $\dot{W}$ is then attributed to friction without having to localize it anywhere. Indeed, the long waves making up the tide transmit energy over the globe so readily that we may expect no correlation between where the moon and sun work hardest on the sea and where the energy thus put into the sea is dissipated.

It may be that little of that dissipation occurs in the open ocean. Taylor (1920) estimated tidal friction in the Irish Sea and showed that most of the energy thus lost comes from the adjacent deep ocean with little direct input due to local working by moon and sun. His methods were extended to the world's coasts and marginal seas by Jeffreys (1921), Heiskanen (1921), and Miller (1966). Miller finds $\dot{E} = 0.7\text{–}2.5 \times 10^{19}\ \mathrm{erg\,s}^{-1}$, two-thirds of which occurs in the Bering Sea, the Sea of Okhotsk, the seas north of Australia, the seas surrounding the British Isles, the Patagonian shelf, and Hudson Bay. This is below all but the most recent estimates of $\langle\dot{W}\rangle$. It should be, by perhaps 10%, because of open-ocean internal tidal dissipation not consistently or completely taken into account (section 10.6). It is now difficult to say whether or not the difference indicates an important omission of some dissipative mechanism.

Additional information about tidal dissipation is contained in the width of conjectured or observed peaks in the admittance amplitude and in shifts in phase of the admittance from one constituent to another. Thus the width of the amplitude-response curve at Bermuda (figure 10.23B; Wunsch, 1972c) suggests a local Q exceeding about 5. Garrett and Munk (1971) surveyed the difference in admittance phase between M_2 and S_2 (the age of the tide) and concluded that worldwide semidiurnal tides had a Q of order 25. Webb (1974) argued that such age-derived estimates of Q primarily reflect localized resonances. It is thus difficult to compare such results with the global Q, for M_2, with a Q of 17 emerging from the most recent cotidal chart of Parke and Hendershott (1980).

Astronomical and oceanographic interest in the amount and geographical distribution of tidal friction constitutes one of the principle modern motivations for studying ocean tides. The other principle motivation is the need, by solid-earth tidalists (Farrell, 1979) and satellite geodesists (Marsh, Martin, McCarthy, and Chovitz, 1980) for a very accurate map of the global distribution of ocean tides. Significant improvement of the most recent numerical maps is going to require extensive new observations.

The technology of deep-ocean pressure sensors suitable for gathering pelagic tide records was pioneered by Eyriés (1968), F. E. Snodgrass (1968) and Filloux (1969). The latest compilation of such results (Cartwright, Zetler, and Hamon, 1979) summarizes harmonic constants for 108 sites irregularly distributed around the world. Cartwright (1977) reviews the history and considerable accomplishments of pelagic tide recording but concludes that economic and political difficulties as well as rapidly evolving research priorities make it an unlikely method for detailed global tide mapping.

Several alternative methods are beginning to be studied. Given a sufficient number of measurements of the solid-earth tide, it is possible to construct estimates of the ocean tide that (in part) generated the solid-earth tide. But high precision earth-tide measurements are needed, and ocean tides in the vicinity of coastal earth-tide stations must be accurately known in order to perceive global ocean tide contributions (Farrell, 1979). Kuo and Jachens (1977) document attempts along these lines.

Satellites may be employed to study ocean tides in two ways. First, the periodic tidal deformation of earth and ocean results in significant perturbation in the orbits of close satellites (Cazenave, Daillet, and Lambeck, 1977). The lowest-order spherical-harmonic components of the tide are the most accessible by this method. It therefore complements the second possibility, direct measurement of satellite-to-sea-surface altitude. The greatest obstacle to extraction of ocean tides from such altimetry is not the error in the altitude measurement but rather the error in our knowledge of where the satellite is relative to the center of the earth. This "tracking" or "orbit" error is greatest at a spatial

scale corresponding to the earth's circumference but decreases rapidly at smaller spatial scales. It probably makes the large-scale features of ocean tides inaccessible from the GEOS-3 altimetry-data set. But smaller-scale tidal systems appear to be directly observable from the later SEASAT-1 altimetry. Parke has permitted me to reproduce (figure 10.24) his recovery of tides along the Patagonian shelf from SEASAT-1 altimetry (Parke, 1980) as an example.

Determining the combination of ocean-tide gauge data (coastal, island, and pelagic), of earth-tide data, of satellite-orbit perturbations, and of satellite altimetry optimal for mapping ocean tides and localizing their dissipation is now perhaps the outstanding theoretical problem in ocean tides.

10.5.2 Partial Models of Ocean Tides

Introduction In his George Darwin lecture "The Tides of the Atlantic Ocean," Proudman (1944) stated, "I shall mainly be concerned with the discovery of the distribution of tides over the open Atlantic Ocean, by the application of the principles of dynamics."

This was, of course, Laplace's goal for global tides. From Laplace's time until now, many researchers have pursued this goal with dogged persistence by solving LTE with astronomical forcing for oceans having shape and relief sufficiently idealized that existing methods of solution could produce an evaluable answer. With hindsight, the properties of these solutions may be appreciated by regarding the solution as eigenfunction expansions in which the various eigenfunctions $Z_n(\phi,\theta)\exp[-i\sigma_n t]$ or free oscillations allowed by LTE have the properties summarized in section 10.4. The frequency σ_n of oscillation is the most natural eigenparameter, but the eigenfunction expansion $\zeta(\phi,\theta,t) = \Sigma_n a_n Z_n \exp[-\sigma_T t]$ for a tide forced at frequency σ_T is *not* of the usual form in which (in the absence of dissipation) $a_n \sim (\sigma_T^2 - \sigma_n^2)^{-1}$. If however, for a given frequency σ_T of forcing, the inverse Δ^{-1} of the mean depth

$$\Delta \equiv (4\pi)^{-1} \iint D(\phi,\theta)\cos\theta\, d\theta\, d\phi$$

[$D(\phi,\theta) = 0$ on land] is regarded as the eigenparameter with "resonant" depths Δ_n, then $a_n \sim (\Delta^{-1} - \Delta_n^{-1})^{-1}$. Nothing restricts Δ_n to positive values. Indeed, negative-depth modes having $\Delta_n < 0$ often exist and may be important in the eignefunction expansion of forced solutions. This evidently was pointed out first by Lindzen (1967) for atmospheric tides.

Direct numerical solution of LTE in realistically shaped basins may be viewed as summation of this eigenfunction expansion, and has gone some distance toward attaining Proudman's stated goal. But Proudman's George Darwin lecture marked an important break with the sequence of dynamic studies that have since culminated in modern numerical solutions.

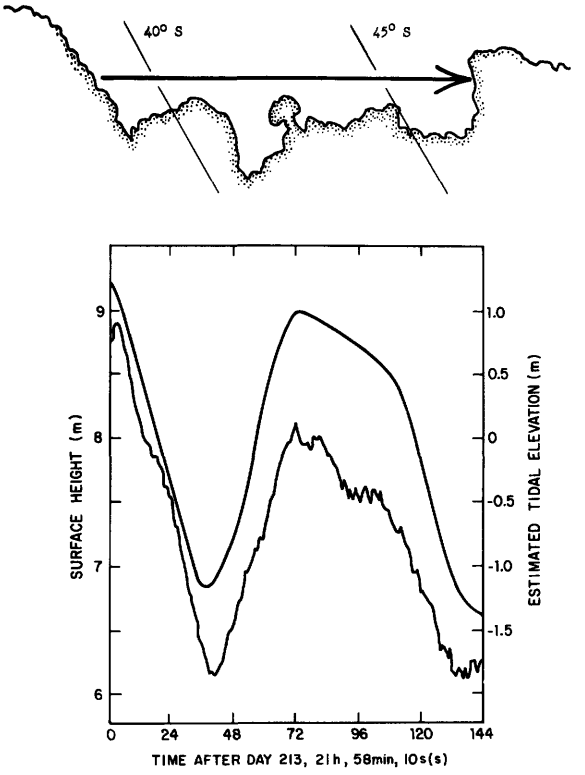


Figure 10.24 SEASAT altimeter record (wiggly line) and a reconstruction (smooth lines) from coastal harmonic constants of nearshore Patagonian shelf tides at the subsatellite point for the SEASAT pass whose path is shown in the upper panel. (Parke, 1980.)

Rather than solving LTE for Atlantic tides, Proudman computed free and forced M_2 solutions of LTE for a portion of the Atlantic and fitted their sum to observations. Subsequent studies carried out in the same spirit but for more simple continental-shelf and marginal-sea geometries have provided dynamically understandable rationalizations for the distribution of tides in these regions and have led to a reappraisal of both observations and of global solutions of LTE. Discussion of these matters occupies the remainder of this section.

Tides in the Gulf of California Godin (1965) and Hendershott and Speranza (1971) noted that (10.29) is satisfied for all the Poincaré channel modes $n = 1,2,\ldots$ in many of the world's long and narrow marginal seas. In these, then, all Poincaré modes are evanescent so that the tide away from the ends of the basin must be mainly a sum of two oppositely traveling Kelvin waves, usually of unequal amplitude. Friction in the basin (or a net rate of working on the tide-generating body by tides in the basin) will make the outgoing Kelvin wave of lower amplitude than the incoming one and will shift amphidromic points (at which the two Kelvin waves interfere destructively) toward the "outgoing" coast.

Figure 10.25 illustrates application of these ideas to the M_2 tide in the Gulf of California. The westward displacement of the amphidromes points to substantial dissipation in the upper reaches of the Gulf. But this Kelvin wave fit does not well represent the tide there. On the basis of his extensive network of tide-gauge observations, Filloux (1973a) was able to estimate the tidal prism and mass transport for six sections along the length of the Gulf and could thus directly evaluate stored energy and energy flux along the Gulf, and energy flux from the moon into dissipation. About 10% of the energy entering the mouth from the Pacific (4.7×10^{16} erg s^{-1}) is lost as the Gulf M_2 tide works on the moon; the remainder is dissipated frictionally (over 80% northward of the islands in figure 10.25).

Elementary considerations suggest that the Gulf of California has a resonance fairly close to the semidiurnal tidal frequency. Filloux (1973a) estimates a Q of about 13 for the thus nearly resonant M_2 tide. Stock (1976) constructed a finite-difference model of Gulf tides using a very fine (10-km) mesh. His solutions effectively sum both of the Kelvin waves and all the evanescent Poincaré modes as well as allowing for their distortion by the irregular shape of the basin. He included dissipative effects and specified the elevation across the mouth of the Gulf in accordance with observations. His model is resonant at about 1.8 cpd with a Q sufficiently high that different discretizations of the problem, all a priori equally reasonable, can give very different Gulf tides. He found it necessary to force his model to have a realistic resonant frequency—fixed by arbitrarily varying the mean depth—before it would produce realistic cotidal maps (figure 10.26). Once this had been done, he found small but nevertheless significant sensitivity of the solution to the localization of dissipation; the solution agreeing best with Filloux's data was that in which most of the dissipation took place around the islands in the upper portion of the Gulf.

The Boundary-Value Problem for Marginal Sea Tides The Gulf of California is one of many marginal seas that connect with the global ocean across a relatively small mouth. Dynamic models of tides in such regions have generally been constructed by solving LTE in the region subject to the condition that the elevation

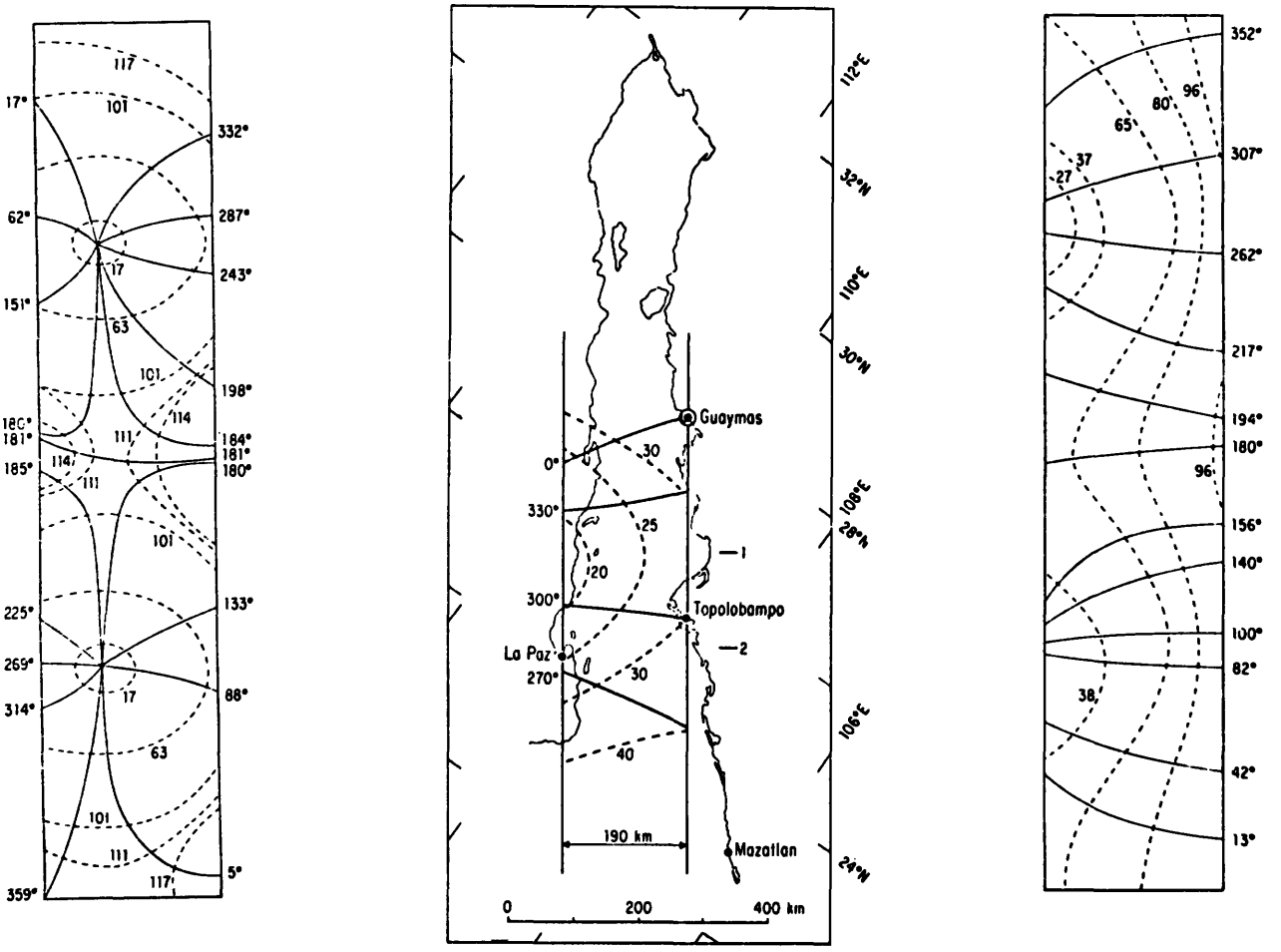


Figure 10.25 Left and right panels are co-oscillating tides in a rectangular gulf with little (left panel) and much (right panel) absorption at upper boundary. Center panel is a Kelvin wave fit to M_2 as observed in the Gulf of California.

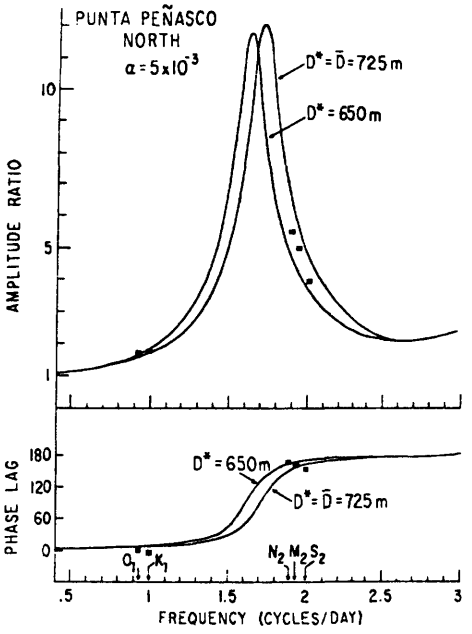


Figure 10.26 A comparison of the tidal response at Punta Penasco (solid squares) with the tide at the mouth of the Gulf of California for two numerical models with different mean depths.

across the open mouth is equal to that actually observed. This is disadvantageous for two reasons. First of all, it eliminates damping of the marginal sea tide by radiation into the deep sea; second, it results in solutions that cannot predict the effects of changes in basin geometry (i.e., installation of causeways, etc.) on the tides because the tide across the open mouth is not allowed to respond to them.

Garrett (1974) pointed out that in many cases these difficulties may be resolved partially by allowing the marginal sea to radiate into an idealized deep sea. For a given constituent, suppose that, with forcing included and all other boundary conditions (i.e., no mass flux through coasts) satisfied, the mass flux $\alpha\delta(S - S')$ normal to the mouth (across which distance is measured by S) would result in the tide $\zeta_G(S) + \alpha K_G(S,S')$ across the mouth when the marginal-sea problem is solved and would result in $\zeta_D(S) + \alpha K_D(S,S')$ when the deep-sea problem is solved. The tides $\zeta_G(S)$ and $\zeta_D(S)$ are thus those that would result just inside and just outside across the mouth if it were closed by an imaginary impermeable barrier. In the real world, the mass flux $U(S)$ across the mouth is fixed by the necessity that its incorporation into either the marginal-sea or the deep-sea problem give the same tide $\zeta_M(S)$ across the mouth:

$$\zeta_M(S) = \zeta_G(S) + \int U(S')K_G(S,S')dS'$$
$$= \zeta_D(S) + \int U(S')K_D(S,S')dS. \qquad (10.133)$$

The latter half of this relation is an integral equation to be solved for $U(S)$. Once $U(S)$ has been found, the problem for the marginal-sea tide is well posed. If the deep sea is, for example, idealized as an infinite half-plane ocean, then $K_D(S,S')$ can be constructed by imposing the radiation condition far from the mouth. The boundary condition across the mouth for marginal sea tides, the specified mass flux $U(S)$, will thus incorporate radiative damping into the solution for marginal-sea tides. Garrett (1974) has discussed limiting cases of (10.133). Garrett and Greenberg (1977) have used the method to discuss possible perturbations of tides by construction of a tidal power station in the Bay of Fundy.

$U(S)$ as given by (10.133) is also the correct marginal-sea boundary condition for models of deep-ocean tides. Its application could allow optimal coupling of finely resolved marginal-sea models to more coarsely resolved global ones, but the methodology requires further development.

Continental Shelf Tides When the tide progresses parallel to a fairly long, straight continental shelf, then the free waves of section 10.4.6 are natural ones in terms of which to expect an economical representation of the tide. Munk, Snodgrass, and Wimbush (1970) analyzed California coastal tides in this way. In addition to the free waves capable of propagating energy at tidal frequencies, they introduced a forced wave to take local working by TGF into account. For the M_2 tide, the Kelvin wave, the single representative member of the Poincaré continuum, and the forced wave have coastal amplitudes of 54, 16, and 4 cm, respectively. The coastal tide is dominated by the northward-propagating Kelvin mode, but further at sea the modes unexpectedly combine to yield an amphidrome (figure 10.27) whose existence was subsequently confirmed by Irish, Munk, and Snodgrass (1971). For the K_1 California tide, the corresponding amplitudes are 21, 24, and 9 cm; the Kelvin wave is not nearly as important. Platzman (1979) has shown how this local representation is related to the properties of eigensolutions of LTE for the world ocean.

The California coast is too low in latitude for second-class shelf modes (section 10.4.6) to propagate energy at tidal frequencies. At higher latitudes, however, Cartwright (1969) has found evidence of their excitation; strong diurnal tidal currents without correspondingly great diurnal surface tides. At very low latitudes, low-mode edge waves could be resonant at tidal frequencies. Stock (private communication) has applied these ideas to the west coast of South America and to the Patagonian shelf. Geometrical difficulties prevent quantitative results in the latter case but the qualitative prediction that the coastally dominant Kelvin

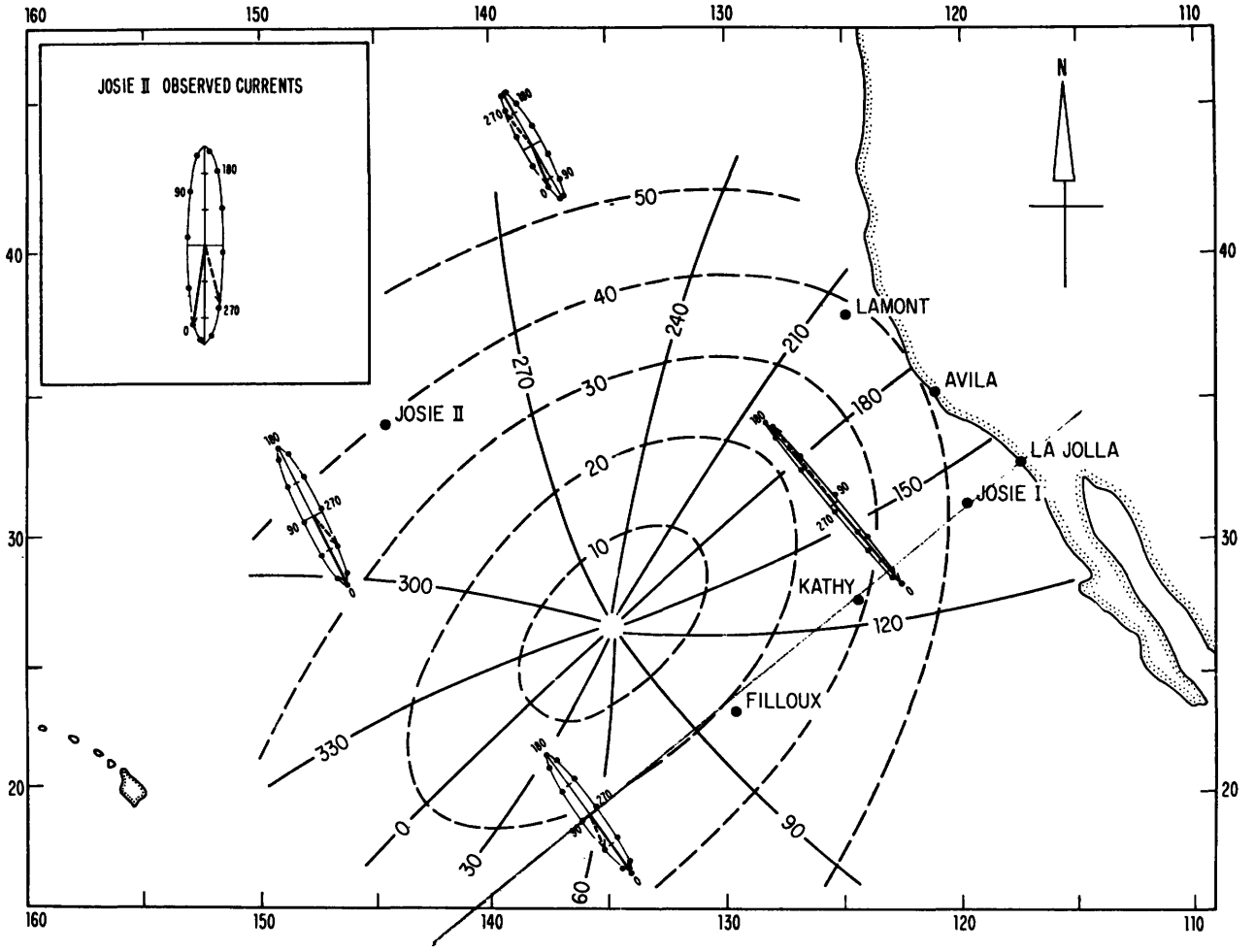


Figure 10.27 M_2 cotidal chart from Munk, Snodgrass, and Wimbush (1970) (amplitudes in cm, phases relative to moon's transit over Greenwich). Ellipses show computed currents at ellipse center (ticks on ellipse axis correspond to 1 cm s^{-1}). Modal fit was to coastal stations plus Josie I, Kathy, and Filloux. Subsequent observations at Josie II confirmed phase shift across predicted amphidrome (Irish, Munk, and Snodgrass, 1971.)

mode decays by e^{-1} across the broad and shallow Patagonian shelf and that the low speed of long-wave propagation over the shallow shelf so compresses the length scale of the tides that a complex system of several amphidromes fits over the shelf are nonetheless important.

On all the shelves so far mentioned, the tide advances parallel to the shelf so that decomposition into modes traveling parallel to the coast is natural. But not all shelf tides are of this nature. Redfield (1958) has summarized observations of United States east coast continental shelf tides (figure 10.28). There the salient features are a very close correspondence between local shelf width and the coastal amplitude and phase of the tide. Tides are nearly coincident over the entire length and width (Beardsley et al., 1977) of the shelf, in marked contrast with the California case.

Island Modification of Tides Island tide records have been prized in working out the distribution of open-ocean tides not only because of their open-ocean location but also because they have been supposed more representative of adjacent open-ocean tides than are coastal records.

Nevertheless, they are not entirely so. Tsunami travel-time charts suggest that tides in island lagoons may be delayed by as much as 20 minutes; harmonic constants for open-ocean tide charts correspondingly may need revision (Parke and Hendershott, 1980). Pelagic records (section 10.5.1) do not, of course, present this problem.

Diffractive effects near island chains may result in appreciable local modification of the tides. Larsen (1977) has studied the diffraction of an open-ocean plane wave of tidal frequency by an elliptical island (intended to model the Hawaiian Island plateau). A typical cotidal chart is shown in figure 10.29. Diffraction alters the time of high water by as much as an hour.

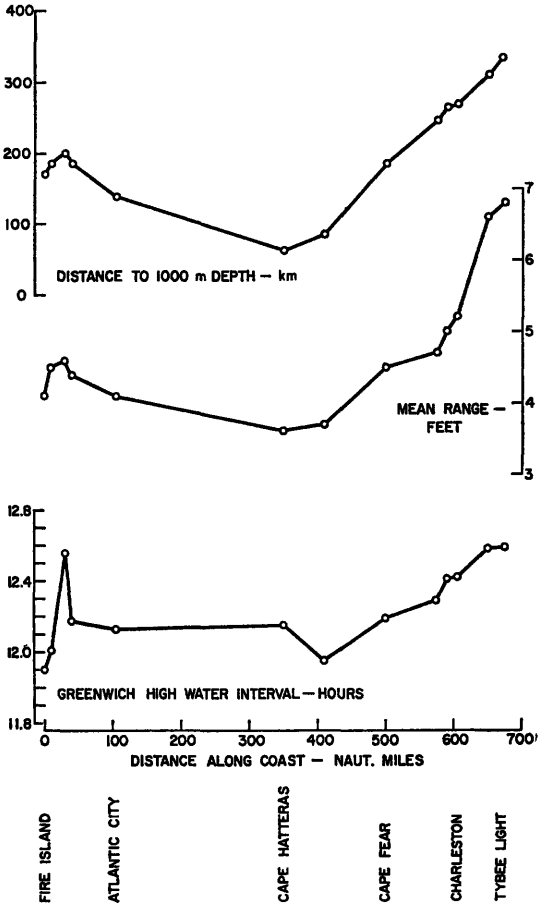


Figure 10.28 Distance from shore to 1000 m depth contour, mean coastal tidal range, and Greenwich high-water interval for selected outlying stations along the eastern coast of the United States. (Redfield, 1958.)

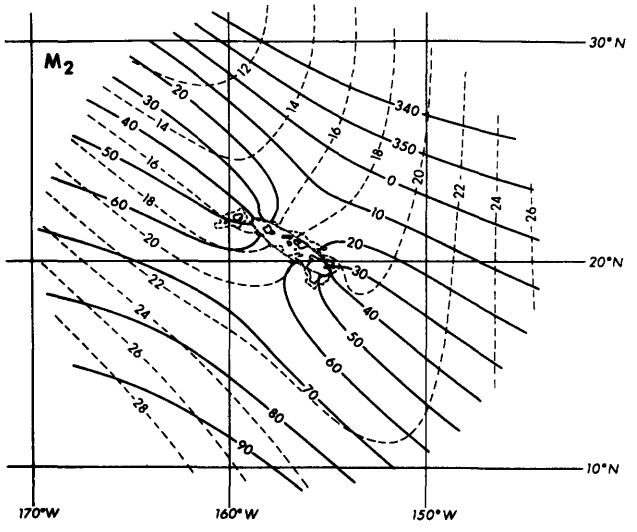


Figure 10.29 Theoretical cotidal chart for an M_2 plane wave in a uniformly rotating ocean of 5000 m depth incident from the northeast on an elliptical island modeling the Hawaiian Chain. (Larsen, 1977.)

10.5.3 Global Tidal Models

The shape of the world's oceans is so complicated that realistic solutions of LTE must be numerical. Pioneering studies were made by Hansen (1949) and by Rossiter (1958). The first global solution was presented by Pekeris and Dishon at the 1961 IUGG Assembly in Helsinki. I have reviewed subsequent developments elsewhere (Hendershott and Munk, 1970; Hendershott, 1973, 1977) and so will not attempt a comprehensive discussion.

Generally, numerical tidalists have solved (often by time-stepping) the forced LTE (10.5) with adjoined dissipative terms, and taking the numerical coasts as impermeable, or else they have solved the elliptic elevation equation [obtained by eliminating the velocities from LTE (10.5) without dissipative terms] for individual constituents (most often M_2) with elevation at the numerical coast somehow specified from actual coastal observations. Combinations of these approaches have also been employed.

The first procedure yields solutions that may be thought of as a weighted sum of the dissipative analogs of Platzman's (1975) normal modes (section 10.4.8). Neither mass nor energy flows across the numerical coast. If the dissipation is modeled accurately (a matter of real concern since the smallest feasible global mesh spacing of about 1° cannot adequately resolve many marginal-sea and shelf tides), then such models should have fairly realistic admittances.

The second procedure attempts to circumvent this difficulty by allowing most or all dissipation to occur beyond the numerical coasts in regions that thus do not have to be resolved. It yields solutions that may be thought of as a tide reproducing the prescribed coastal tide plus a superposition of eigensolutions [of LTE (10.5) or of the elevation equation] that have vanishing elevation at the numerical coast. These eigensolutions have no simple oceanic counterparts since their coastal boundary condition does not require vanishing coastal normal velocity. The full solution satisfies the forced LTE and reproduces the prescribed coastal tide but also generally does not have vanishing normal velocity at the numerical coasts. Consequently there may be at any instant a net flow of water through the numerical coastline, and the flux of energy (averaged over a tidal period) through the numerical coastline need not be zero.

This flux of energy through the numerical coast is a realistic feature since the numerical coast is not intended to model the actual coast but, instead, crudely models the seaward edges of the world's marginal seas and shelves. The same is true of the mass flux, although, in using the solution to estimate ocean-tide perturbations of gravity, etc., the water that thus flows through the numerical coast must somehow be taken into account (Farrell, 1972b). Perhaps the greatest

drawback of the second procedure is the possibly resonant forcing of the unphysical zero-coastal-elevation eigensolutions. This can cause the model to have a very unrealistic admittance even though it is in principle capable of correctly reproducing all constituents. In practice, it often causes the model to be unrealistically sensitive to the way in which discretization of the equations or of the basin has been carried out. Thus Parke and Hendershott (1980) encountered resonances in solving for semidiurnal constituents by the second procedure and were forced to appeal to island observations in the manner described below in order to obtain realistic results. They encountered no similar resonances when solving for the diurnal K_1 constituent, perhaps because the artificial coastal condition filters out the Kelvin-like modes that could be resonant at subinertial frequencies in the f-plane (section 10.4.2). All these remarks also apply to marginal-sea-tide models (section 10.5.2): when the elevation at the connection to the open ocean is specified *ab initio*.

These two procedures and variants of them have resulted in global solutions (most for M_2) that show good qualitative agreement (Hendershott, 1973, 1977). The most recent published global models are by Zahel (1970), Parke and Hendershott (1980), and Accad and Pekeris (1978). I know of new calculations by Zahel, by Estes, and by Schwiderski (Parke, 1979) as well, but have not been able to examine them in detail. When all have been published, a careful comparison of these models with one another, with island and pelagic tidal data, with gravity data, and with tidal perturbations of satellite orbits ought to be carried out.

All the most recent solutions include effects of ocean loading and self-attraction (section 10.3). Many of them have been published since Cartwright (1977) and I (Hendershott, 1977) reviewed the tidal problem. The varying methods of solution may be summarized by abbreviating LTE (10.5) or the elevation equation as in Hendershott (1977):

$$\mathscr{L}[\zeta_0] = \mathscr{L}'[\iint G\zeta_0] + \mathscr{L}'[(1 + k_2 - h_2)U_2/g]. \qquad (10.134)$$

Here U_2 is the tide-generating potential (a second-order spherical harmonic) for a given constituent, (k_2, h_2) are Love numbers (section 10.3), $\mathscr{L}$ and $\mathscr{L}'$ are operators elliptic in space with $\mathscr{L}$ representing LTE (10.5) or the elevation equation, and $\iint G\zeta_0$ abbreviates the global convolution expressing effects of loading and self-attraction as in (10.14).

I attempted to solve (10.134) for M_2 using the second procedure iteratively,

$$\mathscr{L}[\zeta_0^{(i+1)}] = \mathscr{L}'[\iint G\zeta_0^{(i)}] + \mathscr{L}'[(1 + k_2 - h_2)U_2/g], \qquad (10.135)$$

(Hendershott, 1972) but the iteration did not look as though it would converge. Gordeev, Kagan, and Polyakov (1977) found that inclusion of dissipation could result in convergence. Parke (1978) used the iterates $\zeta_0^{(i)}$ as a basis set for a least-squares solution $\hat{\zeta}_0$ of (10.134) of the form

$$\hat{\zeta}_0 = \Sigma A_i \zeta_0^{(i)} \qquad (10.136)$$

in which the A_i are found by solving

$$\frac{\partial}{\partial A_i}\{E \equiv \int\int_{\text{ocean}} |\mathscr{L}(\hat{\zeta}_0) - \mathscr{L}'[\iint G\hat{\zeta}_0]$$

$$- \mathscr{L}'[(1 + k_2 - h_2)U_2/g]|^2\} = 0. \qquad (10.137)$$

He obtained solutions that evidently were quite accurate [E as defined in (10.137) was small], but their realism was marred by the unphysical resonances of the second procedure. Parke and Hendershott (1980) therefore effectively adjusted the locations of these resonances to yield realistic global results by getting the A_i from a least-squares fit of (10.136) to island and pelagic observations.

Accad and Pekeris (1978) noticed that $\iint G\zeta_0^{(i)}$ was very similar to $\zeta_0^{(i+1)}$. They therefore put

$$\iint G\zeta_0^{(i)} = K\zeta_0^{(i)} + \iint \Delta\zeta_0^{(i)}, \qquad (10.138)$$

where K is a constant evaluated empirically at each iteration by

$$K = \iint[\zeta_0^{(i)*} \iint G\zeta_0^{(i)}] / \iint [\zeta_0^{(i)*}\zeta_0^{(i)}] \qquad (10.139)$$

and then iterated not (10.135) but

$$\mathscr{L}[\zeta_0^{(i+1)}] - K\mathscr{L}'[\zeta_0^{(i+1)}]$$

$$= \mathscr{L}'[\iint \Delta\zeta_0^{(i)}] + \mathscr{L}'[(1+k_2 - h_2)U_2/g]. \qquad (10.140)$$

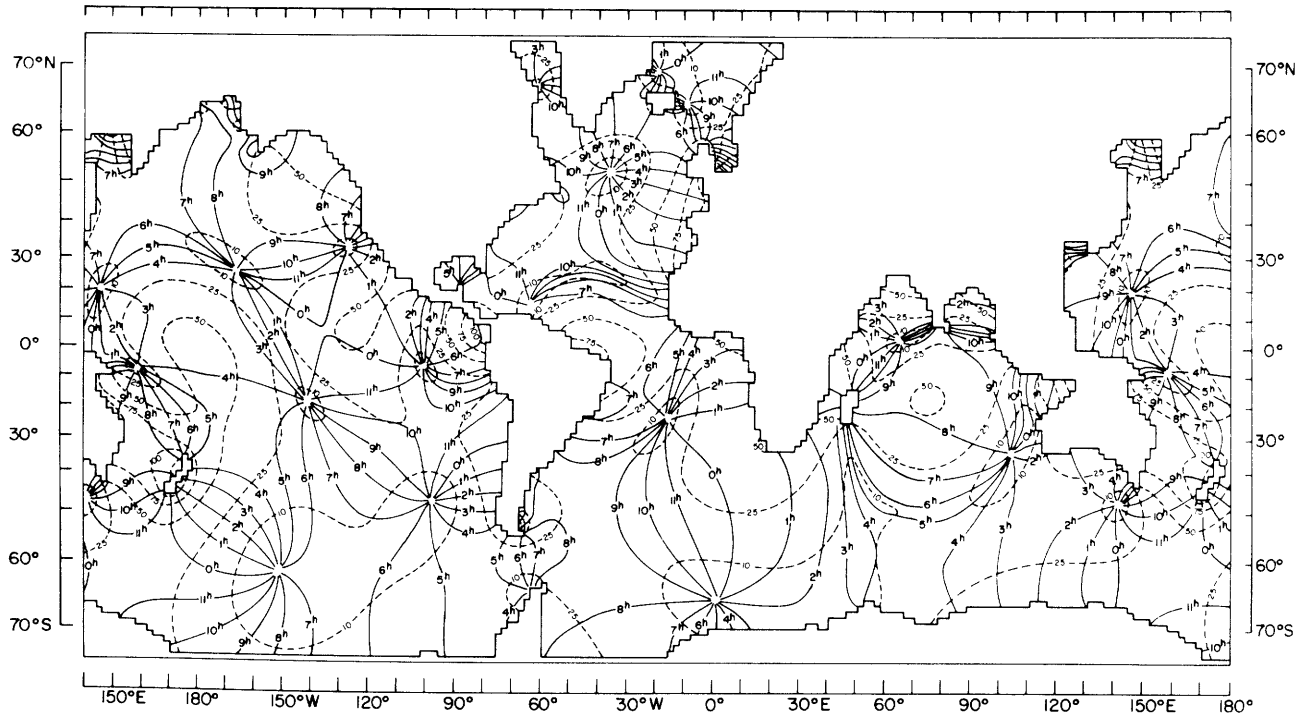
This greatly accelerated the slow convergence of (10.135), presumably already established by dissipation in their calculations.

Figure 10.30 shows two M_2 global cotidal maps of Accad and Pekeris (1978), which differ only in the inclusion of the convolution terms $\iint G\zeta_0$. These terms do not result in an order of magnitude alteration of the computed tide but their effects are large enough that they must be included in any dynamically consistent model aiming at more than order-of-magnitude correctness. These solutions and others like them are obtained solely from a knowledge of the tidal potential and are, in that sense, as close as modern investigators have come to attaining Laplace's original goal.

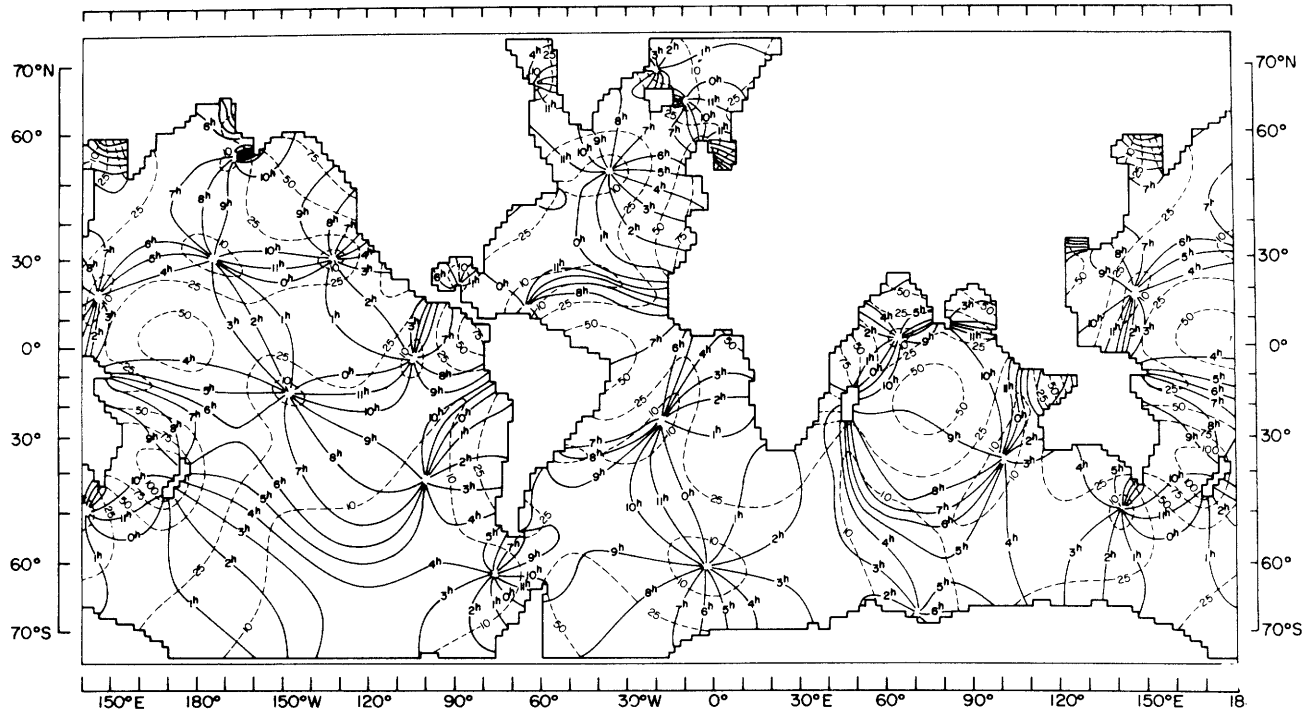
10.6 Internal Tides

10.6.1 Introduction

Internal tides have long been recognized as internal waves somehow excited at or near tidal periods. Their potential as a source of error in hydrographic casts seems to have been recognized since their earliest re-



(10.30A)



(10.30B)

Figure 10.30 Two theoretical calculations of the global M_2 tide obtained solely from a knowledge of the astronomical tide-generating forces (A) and differing only in the inclusion (B) of the effects of loading and self-attraction. (Accad and Pekeris, 1978.)

ported observation by Nansen (1902). Because constant-depth internal-wave modes are almost orthogonal to the ATGF (they would be completely so if the sea surface were rigid and the ATGF exactly depth independent) it has always been difficult to see why internal tides exist at all. The work of Zeilon (1911, 1912) appears to be the precursor of the now generally accepted explanation—energy is scattered from surface to internal tides by bottom roughness—but there has been a history of controversy. The lack of correlation between internal tides at points separated vertically by O (100 m) or horizontally by O (100 km) puzzled early observers. Subsequent observations showed semidiurnal and diurnal internal tides to be narrow-band processes each with a finite band width $\Delta\sigma$ of order several cycles per month. This property manifests itself both in a decay of spatial coherence of internal tides over a length associated with the spread of spatial wavenumbers corresponding to $\Delta\sigma$ and in temporal intermittency over times $\Delta\sigma^{-1}$, as well as in a corresponding lack of coherence with either the surface tide or the ATGF. Typical observations are shown in figure 10.31.

10.6.2 Generation Mechanisms

Zeilon (1934) carried out laboratory experiments showing that a step in bottom relief could excite internal waves in a two-layer fluid when a surface tidal wave passed overhead. Two-layer models are attractive analytically because each layer is governed by a well-posed boundary-value problem; such experiments have been studied theoretically by Rattray (1960) and many others.

Haurwitz (1950) and Defant (1950) noticed that in the f-plane both the horizontal wavelength and the phase speed of plane internal waves grow very large as $\sigma \to f_0$ [(10.23e) with $D_* = D_n$]. Resonance with the ATGF might thus be possible near the inertial latitudes corresponding to tidal frequencies. But the equatorial β-plane solutions (section 10.4.5) (even though only qualitatively applicable at tidal inertial latitudes) show that this apparent possibility of resonance is an artifact of the f-plane, which provides WKB solutions of LTE, and so cannot be applied at the inertial latitudes.

Miles (1974a) has shown that the Coriolis terms customarily neglected in the traditional approximation scatter barotropic energy into baroclinic modes (section 10.3). Observations of internal tides (section 10.6.3) appear to favor bottom relief as the primary scatterer, but this may be because steep bottom relief is spatially localized whereas the "extra" Coriolis terms are smoothly distributed over the globe. Further theoretical work is needed to suggest more informative observations.

For a continuously stratified ocean, Cox and Sandstrom (1962) calculated the rate of energy flow from surface to internal tides due to single scattering from small-amplitude, uniformly distributed, open-ocean-bottom roughness $\epsilon D_1(x, y)$ [where $\epsilon \ll 1$, D_1 is $O(1)$]. Their calculation is most succinctly summarized by specializing to one-dimensional relief $D_1(x)$ and constant buoyancy frequency N_0. If the incident surface tidal-velocity field is idealized as $U\exp(-i\sigma_T t)$, with no space dependence, then the singly scattered internal-tide field $u^{(1)}$ is obtained by solving (10.45)

$$\frac{\partial^2 w^{(1)}}{\partial z^2} - \left(\frac{N_0^2}{\sigma_T^2 - f_0^2}\right)\frac{\partial^2 w^{(1)}}{\partial x^2} = 0 \tag{10.141}$$

and

$$\frac{\partial u^{(1)}}{\partial x} + \frac{\partial w^{(1)}}{\partial z} = 0 \tag{10.142}$$

subject to

$$w^{(1)} = 0 \quad \text{at} \quad z = 0, \tag{10.143}$$

$$w^{(1)} = \epsilon U\,\partial D_1/\partial x \quad \text{at} \quad z = -(D_* + \epsilon D_1) \tag{10.144}$$

plus a radiation condition as $|x| \to \infty$.

Equation (10.143) idealizes the free surface as rigid (adequate for internal waves); (10.144) is the $O(\epsilon)$ expansion about the mean relief $z = -D_*$ of the condition (10.15) of zero normal flow at the actual relief:

$$w = u\frac{\partial}{\partial x}(D_* + \epsilon D_1) \quad \text{at} \quad z = -D_* - \epsilon D_1. \tag{10.145}$$

The solution of (10.141)–(10.144) for $w^{(1)}$ is

$$w_1(x,z) = \int_{-\infty}^{\infty} \overline{w}_1(l,z)\exp(ilx)\,dl, \tag{10.146}$$

where

$$\overline{w}_1(l,z) = \overline{(\epsilon U\,\partial D_1/\partial x)}\,\frac{\sin[lzN_0/(\sigma_T^2 - f_0^2)^{1/2}]}{\sin[-lD_*N_0/(\sigma_T^2 - f_0^2)^{1/2}]} \tag{10.147}$$

with $\overline{(\epsilon U\,\partial D_1/\partial x)}$ defined as the Fourier transform of $(\epsilon U\,\partial D_1/\partial x)$. The integrand of (10.14.6) has simple poles at $[-lD_*N_0/(\sigma_T^2 - f_0^2)^{1/2}] = n\pi$, i.e., at horizontal wavenumbers l satisfying the internal wave dispersion relation $\sigma_T^2 - f_0^2 = gD_n l^2$, $D_n = N_0^2 D_*^2/gn^2\pi^2$. Equatorward of the tidal inertial latitude, $\sigma_T^2 < f_0^2$, so that each pole is real and corresponds to an internal wave traveling away from the scattering relief. Poleward of the tidal inertial latitude, $\sigma_T^2 < f_0^2$, so that each pole is imaginary and the corresponding internal mode decays exponentially away from the scattering roughness without carrying energy away. The sum of all evanescent modes also decays in the vertical away from the scattering relief. Wunsch (1975) reports the existence of observations showing this evanescent behavior for diurnal internal tides.

When $\sigma_T^2 > f_0^2$, (10.141) is hyperbolic in (x, t) with characteristic slope $(\sigma_T^2 - f_0^2)^{1/2}/N_0$. Baines (1971) solved

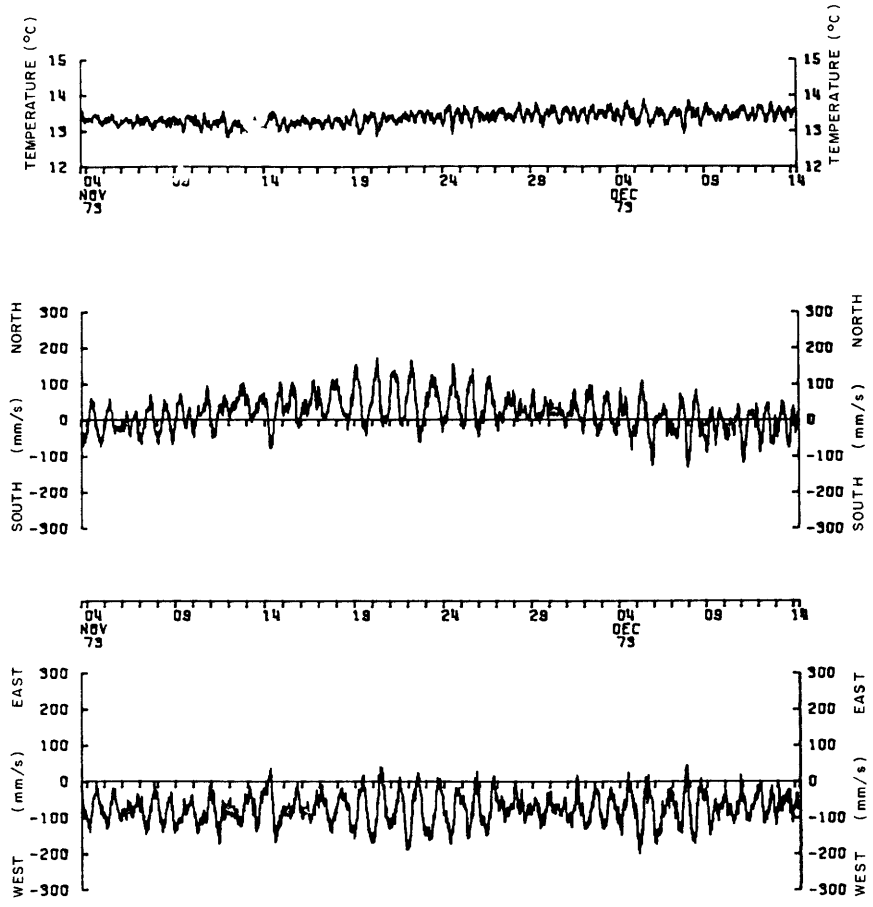


Figure 10.31A Time series of temperature and velocity at the IWEX mooring (Hatteras Abyssal Plain) at 640 m depth (Briscoe, 1975b).

(10.141) and (10.145) exactly, by the method of characteristics, thus eliminating the restriction to weakly sloping relief. The analytical novelty of his work was the imposition of the radiation condition on the characteristic form

$$F(x - Rz) + G(x + Rz), \qquad R = N_0/(\sigma_T^2 - f_0^2)^{1/2}$$

of the solutions of (10.141) by, for example, choosing

$$F(x) = \int_0^\infty e^{ilx}\overline{F(l)}dl$$

so that $F(x \pm Rz)\exp(-i\sigma_T t)$ contains only outgoing plane waves. Laboratory work (Sandstrom, 1969) and analysis (Wunsch, 1969) showed that when bottom and tidal characteristic slopes coincide, the near-bottom motion is strongly intensified. Wunsch and Hendry (1972) show evidence for such intensification over the continental slope south of Cape Cod (figure 10.32).

The general possibility that diurnal tides enhance diurnal inertial motion by some mechanism has been suggested by Ekman (1931), Reid (1962) and Knauss (1962b). I (Hendershott, 1973) estimated the amplitude of motion if the mechanism is scattering of surface tides into internal tides by open-ocean bottom roughness, but obtained a result sufficiently small that it would not stand out noticeably against the high level of inertial motion found at all mid-latitudes (Munk and Phillips, 1968).

Thus far, the discussion is in terms of linearly scattered linear waves. Bell (1975) considers the formation of internal lee waves on periodically varying barotropic tidal currents. This process could generate a complex spectrum of internal waves even with a monochromatic surface tide. What actually occurs when laboratory or ocean stratified flow passes over relief is complicated. In Massachusetts Bay, Halpern (1971) has observed that tidal flow over a ridge generates a thermal front that propagates away as a highly nonlinear internal wavetrain (Lee and Beardsley, 1974) or as an internal bore. Such bores are commonly observed along the Southern California coast (Winant, 1979).

Maxworthy (1979) emphasizes the importance of the collapse of the stirred region that initially develops over relief in the subsequent generation of laboratory internal waves. The relative importance of all these processes near the sea floor is unknown. If separation

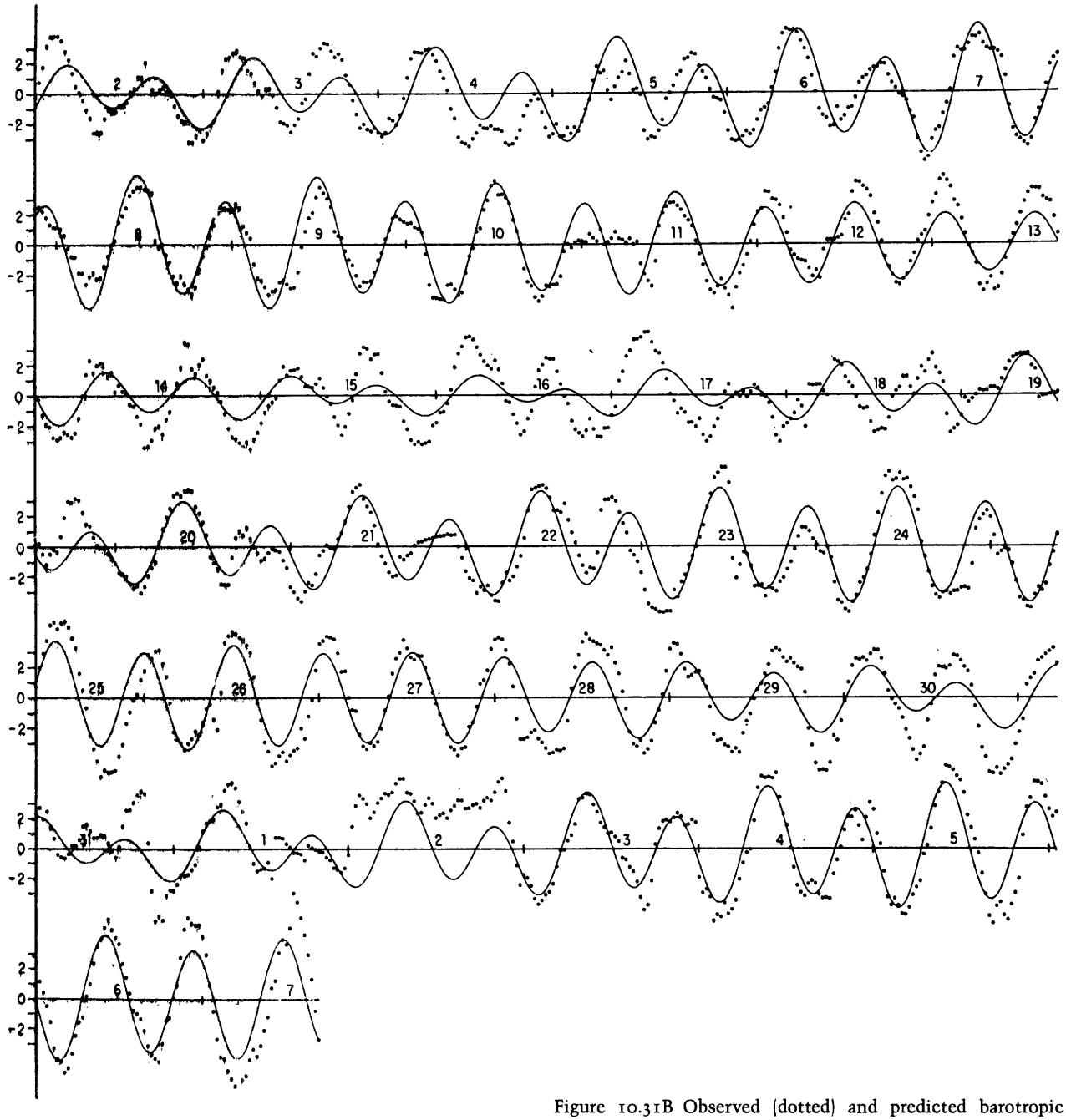


Figure 10.31B Observed (dotted) and predicted barotropic (solid) longshore bottom velocity at Josie I (figure 10.27) off the southern California coast. (Munk, Snodgrass, and Wimbush, 1970.)

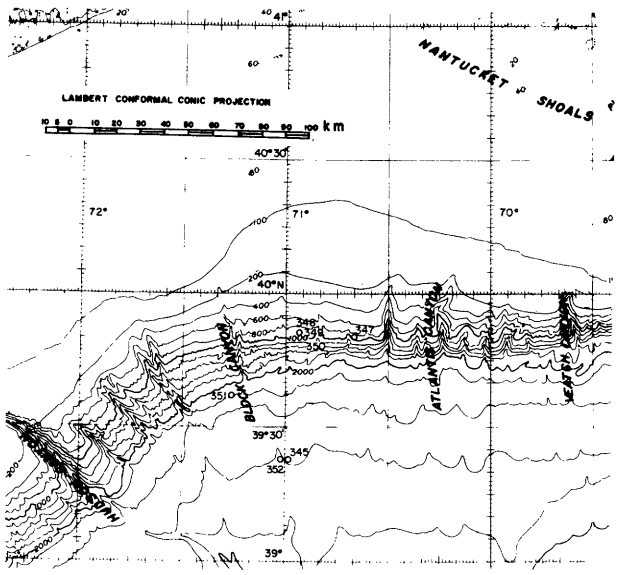


Figure 10.32A Current meter mooring positions and shelf topography. (Wunsch and Hendry, 1972.)

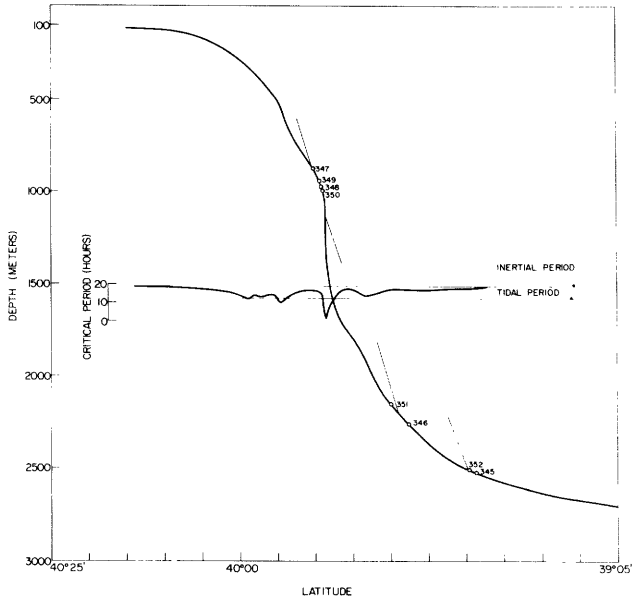


Figure 10.32B Profile of topography through the array along the dashed line of figure 10.32B with mooring positions indicated. Several internal wave characteristics for the M_2 tide are shown, and the critical period (at which internal-wave characteristics are locally tangent to the relief) is plotted across the profile. (Wunsch and Hendry, 1972.)

over abyssal relief does occur in tidal currents, it could contribute to abyssal mixing by helping to form the near-bottom laminae observed by Armi and Millard (1976). Wunsch (1970) made a somewhat similar suggestion based on laboratory studies.

Besides this potential complexity of generation, the medium through which the internal tide moves is strongly inhomogeneous in space and time. The overall result is the complicated and irregularly fluctuating internal tide observed. Still, away from generation regions, some features of the linear theory shine through.

10.6.3 Observations

In linear theory, breaks in the slope of the relief and extended regions where that slope coincides with a tidal characteristic slope make themselves felt in the body of the ocean as narrow-beam disturbances concentrated along the characteristics (Rattray et al., 1969). The beams are typically narrow (figure 10.33) and their (characteristic) slope in the presence of mean currents varies both with local stratification and shear. This suggests that, especially near generation regions, the internal tide will have a complex spatial structure and that its amplitude at a given point may vary markedly as nearby stratification and mean flow change. Thus Hayes and Halpern (1976) document very large variability of semidiurnal internal tidal currents during a coastal upwelling event; they account for much of it by appealing to the deformation of characteristics as vertical and horizontal density gradients change during the upwelling. Regal and Wunsch (1973) find internal tidal currents at site D, over the continental slope south of Cape Cod, to be concentrated near the surface and there highly (and uncharacteristically) coherent

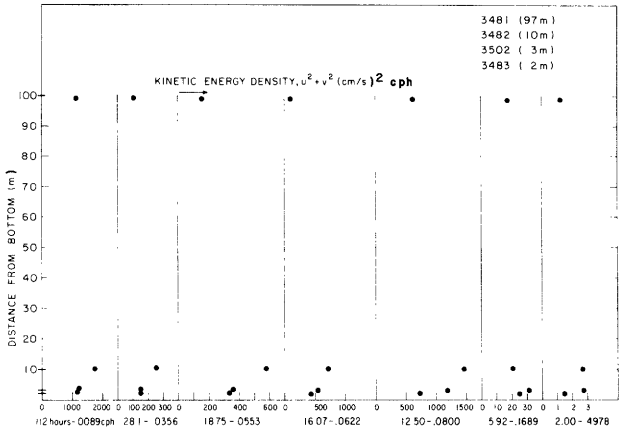


Figure 10.32C Vertical profiles of kinetic energy density (observed values are solid dots) for various periods over the slope where the tidal characteristic is locally tangent to the relief (moorings 347-350). (Wunsch and Hendry, 1972.)

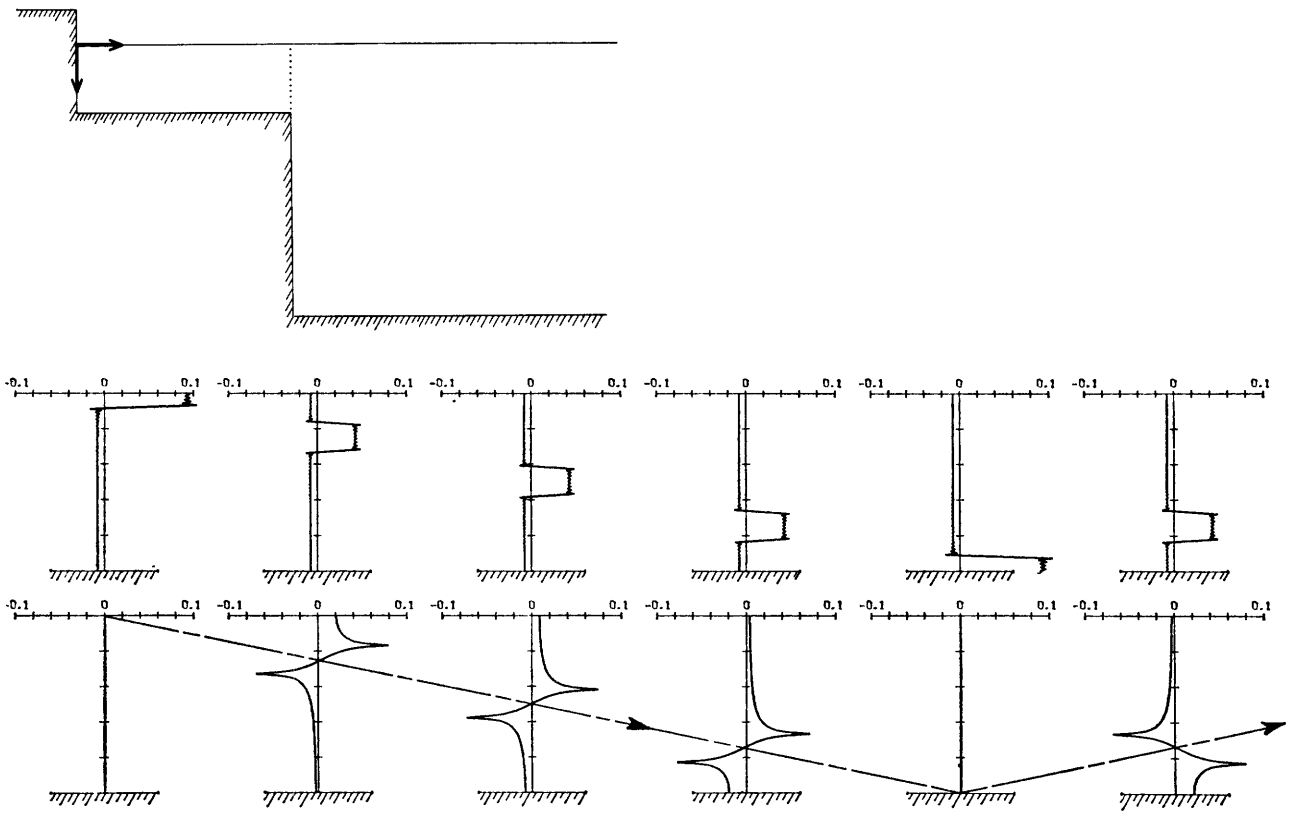


Figure 10.33 Depth distribution of horizontal internal tidal currents away from a step shelf (top) at two times (center and bottom) separated by a quarter-wave period. Ratio of deep sea to shelf depth is 12.5, characteristic slope is 1/715; current profiles begin at shelf edge and are separated horizontally by 1.8 of the deep-sea (first-mode) internal-tide wavelength. (Rattray et al., 1969.)

with the surface tide. The result is consistent with generation where tidal characteristics graze the slope perhaps 60 km to the north followed by propagation along characteristics that leave the region of tangency and bounce once off the ocean floor before passing through the near surface part of the water column at site D (figure 10.34). Observations at site L, some 500 km to the south, show no evidence of propagation along beams.

Beams are a coherent sum of many high internal modes. We intuitively expect that high modes are more rapidly degraded by whatever processes ultimately lead to dissipation than are low modes, and that they are more sensitive to medium motion and fluctuation than are low modes because they propagate so slowly. We thus do not expect beamlike features in the deep sea, and they are not observed. Instead, we expect a few low modes to dominate in a combination of arrivals from distant steep relief. These will have made their way through significant oceanic density fluctuations and through fluctuations of mean flows often at an appreciable fraction of internal-wave-phase speeds. The line spectrum characteristic of the ATGF and the surface tide will thus be so broadly smeared into semidiurnal and diurnal peaks that individual constituents or even the spring-neap cycle are at best very difficult (Hecht and Hughes, 1971) to perceive.

The most complete description of open-ocean internal tides is due to Hendry (1977), who used the western central Atlantic Mid-Ocean Dynamics Experiment (MODE) data. Figure 10.35 summarizes the results. M_2 tends to dominate semidiurnal temperature variance over the water column. Adjacent N_2 and S_2 variances are nearly equal, and the vertical variation of N_2, S_2 variance generally follows M_2 (with qualifications near the bottom). M_2 likewise dominates horizontal semidiurnal current variance over the water column. At subthermocline depths M_2 variance approaches estimates of the barotropic M_2 tidal current variance while S_2 and N_2 variances exceed their barotropic counterparts by an order of magnitude. All this suggests that much of what appears in the N_2 and S_2 bands has really been smeared out of M_2. The vertical distribution of variance is broader than the two WKB profiles $(\partial\bar{\theta}/\partial z)^2 N(z)^{-1}$ ($\bar{\theta}$ is mean potential temperature) and $N(z)$, for temperature and horizontal current variance, respectively. This indicates that the lowest vertical modes dominate. M_2 temperature variance has a coherence of about 0.7 with the ATGF in the upper thermocline while N_2 and S_2 are far less coherent with the

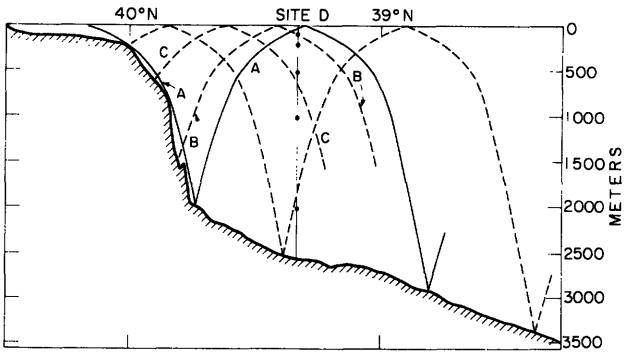


Figure 10.34A Profile of relief along 70°W (see figure 10.32 for local isobaths) together with selected semidiurnal characteristics passing near site D. (Regal and Wunsch, 1973.)

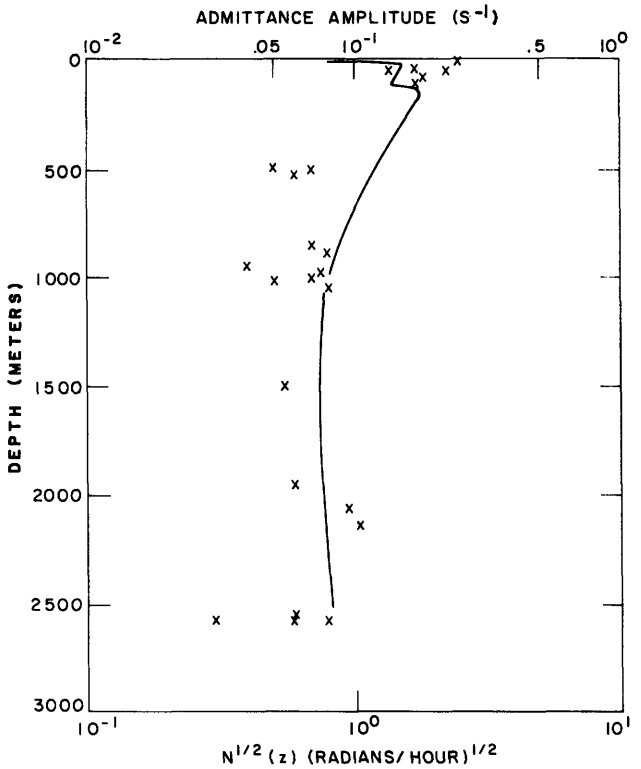


Figure 10.34B Admittance amplitude x for semidiurnal tidal currents together with buoyancy frequency $N(z)$ at site D. Near-surface admittances are strongly intensified; currents there are highly coherent with the surface tide. (Regal and Wunsch, 1973.)

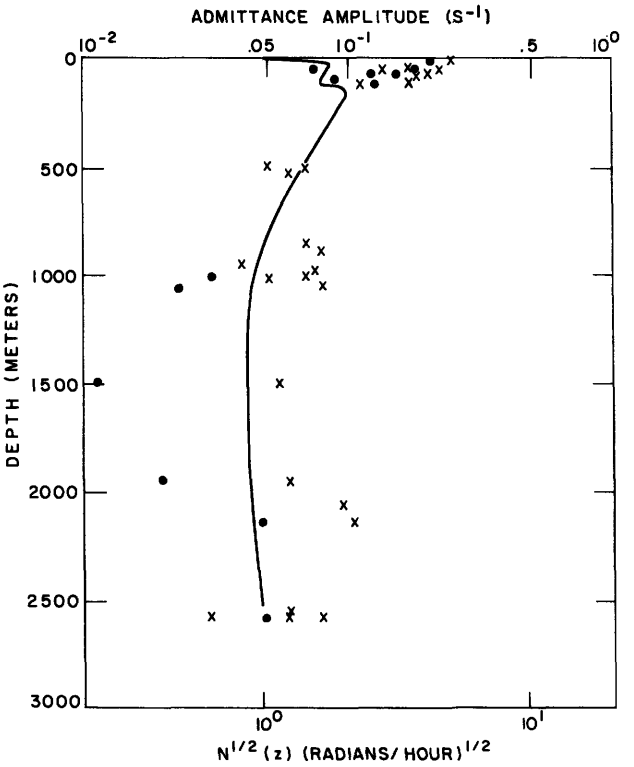


Figure 10.34C A similar display at site L, 599 km south of site D, shows no comparable surface intensification. (Regal and Wunsch, 1973.)

ATGF but yet not totally incoherent. The M_2 first internal mode dominates and propagates to the southeast; this plus the (significantly not random) phase lag between M_2 and S_2 in the thermocline (the age of the internal tide) point to the 700-km distant Blake escarpment as a generating region. Other discussions of open-ocean internal tides are consistent with the foregoing picture although necessarily based upon less extensive observations.

Wunsch (1975) reviews observations allowing estimation of the energy density (in units of ergs per squared centimeters) of internal tides and suggests that it is from 10 to 50% of the corresponding energy density of the barotropic tide, albeit with wide and unsystematic geographic variation.

10.6.4 Internal Tides and the Tidal Energy Budget

It thus appears that barotropic tides somehow give up energy to internal tides. Return scattering is probably unimportant. It is important to know the rate at which this energy transfer occurs because (section 10.5) the energy budget for global tides may not yet be closed. Wunsch (1975) reviews estimates arising from the various scattering theories outlined above (section 10.6.2); typical values are 0.5×10^{19} erg s^{-1} from deep-sea roughness [using the theory of Cox and Sandstrom as rediscussed by Munk (1966)], 6×10^{7} erg cm s^{-1} from continental shelves [using the theory of Baines (1974) and also from independent measurements by Wunsch and Hendry (1972)]. The latter value extrapolates to 5.6×10^{15} erg s^{-1} over the globe.

A bound on this estimate independent of scattering theories was pointed out by Wunsch (1975). Internal-tide energy densities E_I are order 0.1 to 0.5 times surface-tide energy densities E_S. Group velocities $\mathbf{c}_{gI}$ of internal waves are order $(D_n D_0)^{1/2} \simeq (N^2 D_0 / g n^2 \pi^2)^{1/2}$ times group velocities $\mathbf{c}_{gS}$ of long-surface gravity waves. If open-ocean tidal energy is radiated toward shallow seas (or any other dissipation region) at rates $\mathbf{c}_{gS} E_S$ and $\mathbf{c}_{gI} E_I$, then internal tides can never account for more than $O(10\%)$ of the total energy lost from surface tides.

Wunsch's (1975) discussion of the caveats to this result has not been substantially altered by subsequent developments. Nonlinear interactions that drain internal energy from the tidal bands to other frequencies and scales certainly do occur but their rates are not yet accurately estimable. Such rates as have been calculated [Garrett and Munk (1972a) calculated the energy loss due to internal wave breaking; McComas and Bretherton (1977) the time scale for the low-frequency part of the internal wave spectrum to evolve by resonant interactions]; they are small, but the problem is not closed.

10.6.5 Internal Tides and Ocean Stirring

Even if internal tides turn out to be a minor component of the global tidal-energy budget, they could be an im-

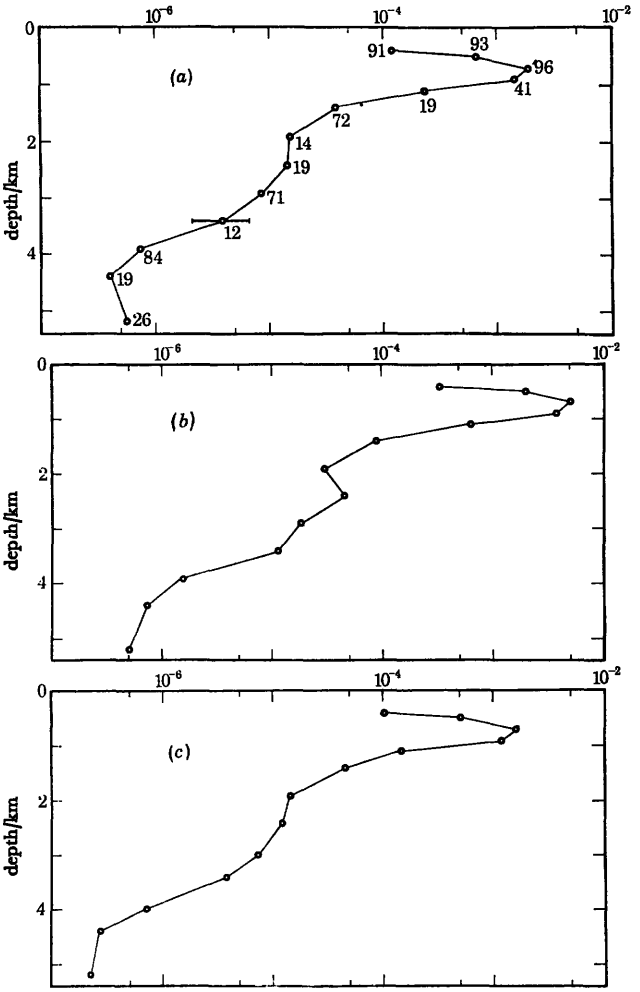


Figure 10.35A Vertical profile of squared temperature fluctuations in the (a) S_2 band (°C), averaged at depth levels over the entire array; the number of 15-day-piece lengths at each level is indicated. Vertical profile of average squared temperature fluctuations in the (b) M_2 band. Vertical profile of average squared temperature fluctuations in the (c) N_2 band. (Hendry, 1977.)

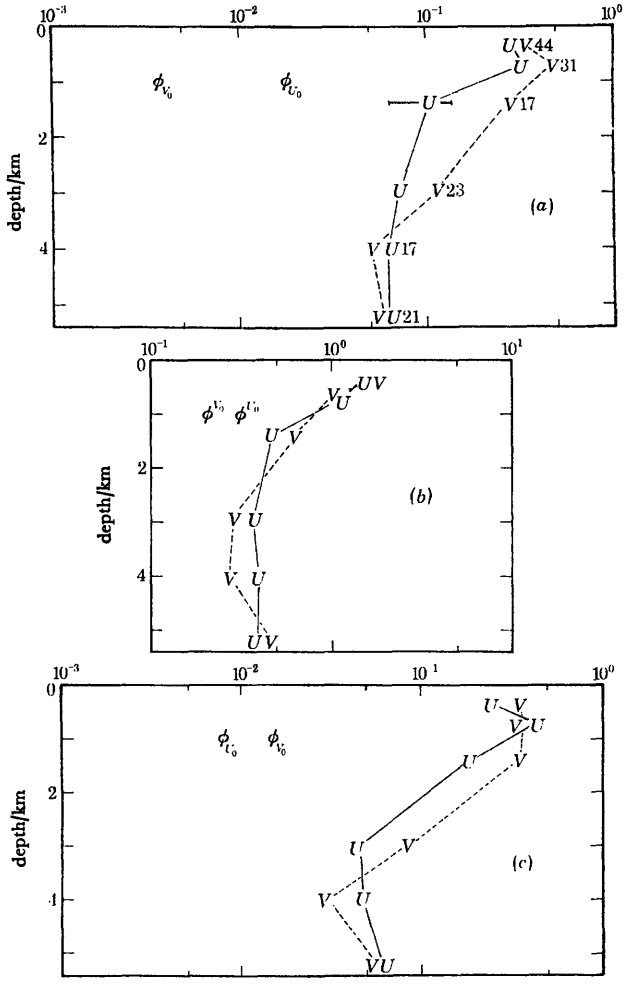


Figure 10.35B Vertical profile of squared horizontal current ($cm\,s^{-1}$) for U (east and V (north) in the (a) S_2 band, averaged at depth levels over the entire array; estimates of squared amplitude for the barotropic current components U and V are given, showing that the currents are dominanted by internal waves at all depths. Similar estimates for the (b) M_2 band currents. Here the deep currents are greatly influenced by the barotropic mode. Similar estimates are given for the (c) N_2 band; internal waves appear to dominate at all depths. (Hendry, 1977.)

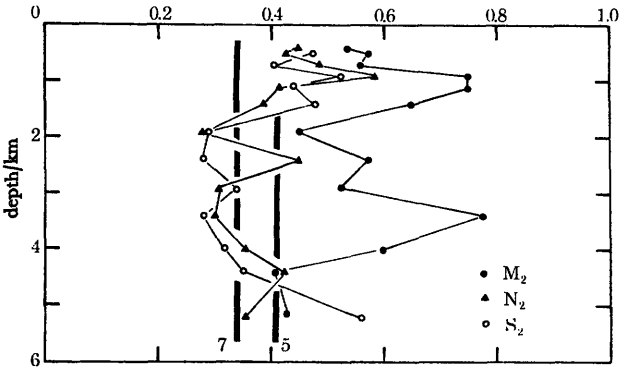


Figure 10.35C Vertical profile of average coherence amplitude of temperature fluctuations and the equilibrium tide for three semidiurnal frequency bands. The averages are taken over the whole array at depth levels, and include individual cases with both five and seven degrees of freedom. The expected values of coherence amplitude for zero true coherence are shown for each case, and while the central M_2 band shows a definite determinism, the adjacent frequency bands are much more dominated by randomly phased temperature fluctuations. (Hendry, 1977.)

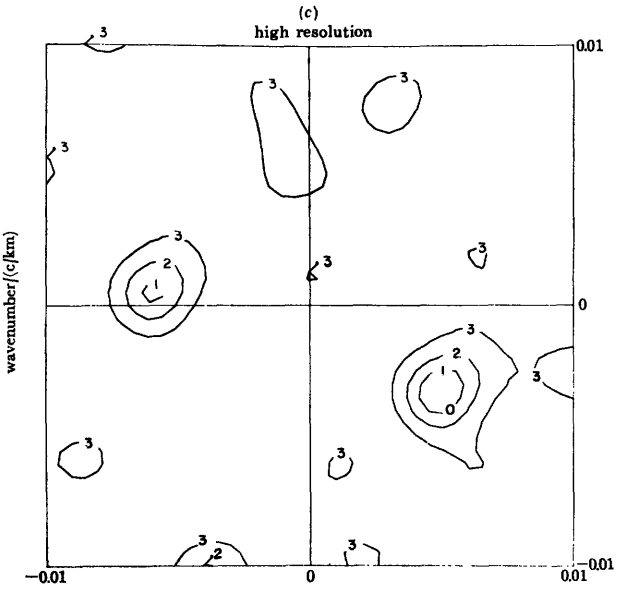


Figure 10.35D Conventional wavenumber spectrum of first-mode M_2 temperature fluctuations from MODE. The peak in the southeast quadrant has wavenumber 1/163 cpkm and represents a wave propagating from northwest to southeast. A secondary peak in the northwest quadrant is interpreted as an alias of the main peak. (Hendry, 1977.)

portant source of energy for ocean stirring and mixing. About 10% of the total tidal dissipation would be energetically adequate (Munk, 1966). The hypothesis is that internal tides somehow (by bottom turbulence, by nonlinear cascade) cause or enhance observed fine structure and microstructure events in which mixing is believed to be occurring. It thus is pertinent to examine observations for any correlation between tidal phenomena and smaller-scale events. Such a correlation could be temporal [with the intensity of small structure modulated at semidiurnal, diurnal, fortnightly (i.e., the spring–spring interval) or even longer tidal periods] or spatial (with small structure near generators different from that far away).

Although definitive studies have yet to be made, preliminary indications are that little such correlation exists. Cairns and Williams (1976) contour the spectrum of vertical displacement in a frequency–time plane for 17 days over the frequency band 0.2–6.0 cpd (figure 10.36) but see no modulation at tidal or fortnightly periods of any part of the spectrum. Wunsch (1976) finds no correlation between the overall spectral level of internal waves (specifically, the spectral intensity at 5-hour period of a model fitted to observed spectra) and the intensity of the internal tidal peak for observations from the western North Atlantic (figure 10.37).

The demonstration that tidal contributions to ocean mixing are significant will thus involve subtle measurements. Perhaps something may be learned by comparing internal waves, fine structure, and microstructure in the open ocean with their counterparts in the relatively tideless Mediterranean Sea or in the Great Lakes (see chapters 8 and 9).

10.7 Tidal Studies and the Rest of Oceanography

Although tidal studies were the first dynamic investigation of oceanic response to forcing, insight into wind- and thermohaline-driven ocean circulation developed largely independently of them. In the case of semidiurnal and diurnal tides, the reason is primarily dynamic. But the dynamics of long-period tides are likely to be much more like those of the wind-driven circulation (both steady and transient) than like those of semidiurnal and diurnal tides. It is possible that, if the long-period components of the ATGF had been large enough to make the long-period tides stand out recognizably above the low-frequency noise continuum, then the early tidalists might have recognized, in the low-frequency tides, features such as westward intensification also evident in the general circulation. They might then have been forced into the recognition that a linear superposition of (possibly damped) first-class waves could not account for long-period tides, as it seems able to do for semidiurnal and diurnal tides. As things are, however, long-period tides are so near to the noise level that their observation did not provide a global picture clear enough to force tidalists out of the semidiurnal–diurnal framework.

It was, in fact, insight into the problem of time-dependent wind-driven ocean circulation that led Wunsch (1967) to provide the modern view of long-period tides: a superposition of damped second-class waves (section 10.4.4) whose horizontal length scales are only $O(10^3$ km) and whose amplitudes and phases are likely to undergo substantial fluctuations in time on account of the overall time variability of the ocean currents through which they propagate. Laplace had supposed that a small amount of dissipation would bring the long-period tides into equilibrium, i.e., the geocentric sea surface ζ would be an equipotential of the total tide-generating potential Γ (Lamb, 1932, §217). The most recent elaboration of this view is by Agnew and Farrell (1978), who solved the integral equation

$$\zeta = \zeta_0 + \delta = \Gamma/g$$

[with δ (the solid-earth tide) and Γ given by (10.11) and (10.12) as functions both of the observed tide ζ_0 and of the long-period astronomical potential U_2] for equilibrium global-ocean fortnightly and monthly tides ζ_0 subject to the conservation of mass. Wunsch's (1967) analysis and dynamic model of the fortnightly tides suggest that they are not in equilbrium with either the astronomical potential U_2 or with the full potential Γ of Agnew and Farrell (1978). Their Pacific averaged admittance has magnitude 0.69 ± 0.02 relative to Γ with significant island-to-island variation. The Pacific averaged monthly tide admittance has magnitude 0.90 ± 0.05 relative to Γ, but island-to-island fluctuations vanish only by pushing individual island admittances to the very end of their error bands. Whether the kinds of dissipation and of nonlinear interaction between low-frequency motions that occur in the real ocean favor an equilibrium tidal response at sufficiently low frequencies is not yet known either observationally or theoretically. The 14-month pole tide is known to be significantly non-equilibrium in the shallow seas of Northern Europe (Miller and Wunsch, 1973) but nearly invisible elsewhere.

Observations of long-period tides have thus been too noisy to exert real influence on tidal studies and hence on dynamic oceanography. But the theoretical ideas emerging from study of the low-frequency solutions of LTE have been of great importance for dynamic oceanography (even though they are likely to be inadequate to model the full dynamics of the general circulation).

Steady (as opposed to wavelike) solutions of LTE obtained by active tidalists (Hough, 1897; Goldsbrough, 1933) had a profound effect on dynamic oceanography through the review by Stommel (1957b) of

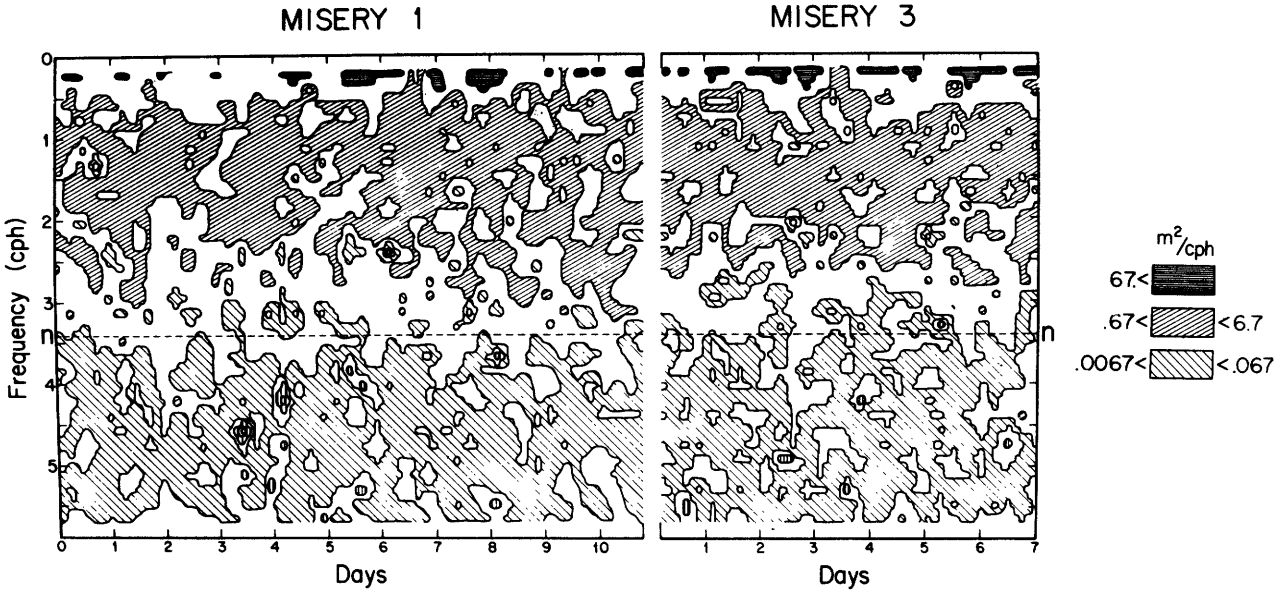


Figure 10.36 Contours of vertical-displacement spectral energy vs. elapsed time for the 6.60°C isotherm off the coast of southern California. Spatial estimates with 2 df are made for successive 5.8-h data segments and the results are contoured. (Cairns and Williams, 1976.)

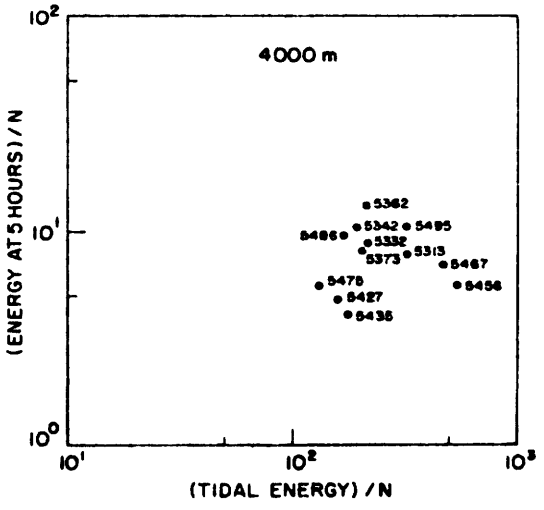


Figure 10.37 Five-hour internal wave energy at various locations vs. semidiurnal tidal energy. (Wunsch, 1976.)

ocean-current theory and through the work by Stommel and Arons (1960a,b) on abyssal circulation. Goldsbrough (1933) studied nonperiodic solutions of LTE driven by global patterns of evaporation and precipitation. The solutions are steady, provided that the precipitation–evaporation distribution vanishes when integrated along each parallel of latitude between basin boundaries. Stommel (1957b) pointed out that Ekman suction and blowing due to wind-stress convergence and divergence could effectively replace the evaporation–precipitation distribution, while the introduction of ageostrophic western boundary currents allowed the solutions to remain steady even when the integral constraint on the evaporation–distribution function was violated. The resulting flows display the main dynamic features of the theory of wind-driven circulation due to Sverdrup (1947), Stommel (1948), and Munk (1950). When the evaporation–precipitation function is viewed as modeling the high-latitude sinking of deep water and its mid-latitude subthermocline upwelling, the abyssal circulation theories of Stommel and Arons (1960a,b) result.

The seminal work on low-frequency second-class motions was the study (section 10.4.4) by Rossby and collaborators (1939), ironically inspired by meteorological rather than tidal studies. It led, through the studies by Veronis and Stommel (1956) and Lighthill (1969) of time-dependent motion generated by a fluctuating wind to the very different views of mid-latitudes and tropical transient circulation that prevail today (although, especially in mid-latitudes, linear dynamics

are now generally acknowledged to be inadequate for a full description; see chapter 5). Pedlosky (1965b) showed how the steady western boundary currents of Stommel (1948), Munk (1950), and Fofonoff (1954) could be viewed as Rossby waves reflected from the western boundary and either damped by friction or swept back toward the boundary by the interior flow that feeds the boundary current; Gates's (1968) numerical examples showed clearly the development of a frictional western boundary current as a group of short Rossby waves with seaward edge propagating away from the western boundary at the appropriate group velocity.

Modern interest in estimating the role of direct transient wind forcing in generating mesoscale oceanic variability (see chapter 11) calls for an up-to-date version of N. A. Phillips's (1966b) study of mid-latitude wind-generated Rossby waves using more realistic wind fields and taking into account new insight into the combined effects of bottom relief and stratification (section 10.4.7). Such a calculation would closely resemble a proper (linear) dynamic theory of long-period tides. But similar caveats apply to uncritically accepting either as representing an actual flow in the ocean.