9
Internal Waves and Small-Scale Processes

Walter Munk

9.1 Introduction

Gravity waves in the ocean's interior are as common as waves at the sea surface—perhaps even more so, for no one has ever reported an interior calm.

Typical scales for the internal waves are kilometers and hours. Amplitudes are remarkably large, of the order of 10 meters, and for that reason internal waves are not difficult to observe; in fact they are hard *not* to observe in any kind of systematic measurements conducted over the appropriate space-time scales. They show up also where they are not wanted: as short-period fluctuations in the vertical structure of temperature and salinity in intermittent hydrocasts.

I believe that Nansen (1902) was the first to report such fluctuations;[1] they were subsequently observed on major expeditions of the early nineteen hundreds: the *Michael Sars* expedition in 1910, the *Meteor* expeditions in 1927 and 1938, and the *Snellius* expedition in 1929-1930. [A comprehensive account is given in chapter 16 of Defant (1961a)]. In all of these observations the internal waves constitute an undersampled small-scale noise that is then "aliased" into the larger space-time scales that are the principal concern of classical oceanography.

From the very beginning, the fluctuations in the hydrocast profiles were properly attributed to internal waves. The earliest theory had preceded the observations by half a century. Stokes (1847) treated internal waves at the interface between a light fluid overlaying a heavy fluid, a somewhat minor extension of the theory of surface waves. The important extension to the case of a vertical mode structure in *continuously* stratified fluids goes back to Rayleigh (1883). But the discreteness in the vertical sampling by hydrocasts led to an interpretation in terms of just the few gravest modes, with the number of such modes increasing with the number of sample depths (giving j equations in j unknowns). And the discreteness in sampling time led to an interpretation in terms of just a few discrete frequencies, with emphasis on tidal frequencies.

The development of the bathythermograph in 1940 made it possible to repeat soundings at close intervals. Ufford (1947) employed three vessels from which bathythermograph lowerings were made at 2-minute intervals! In 1954, Stommel commenced three years of temperature observations offshore from Castle Harbor, Bermuda, initially at half-hour intervals, later at 5-minute intervals.[2] Starting in 1959, time series of isotherm depths were obtained at the Navy Electronics Laboratory (NEL) oceanographic tower off Mission Beach, California, using isotherm followers (Lafond, 1961) installed in a 200-m triangle (Cox, 1962).

By this time oceanographers had become familiar with the concepts of continuous spectra (long before routinely applied in the fields of optics and acoustics), and the spectral representation of surface waves had proven very useful. It became clear that internal waves, too, occupy a frequency continuum, over some six octaves extending from inertial to buoyant frequencies. [The high-frequency cutoff had been made explicit by Groen (1948).] With regard to the vertical modes, there is sufficient energy in the higher modes that for many purposes the discrete modal structure can be replaced by an equivalent three-dimensional continuum.

We have already referred to the measurements by Ufford and by Lafond at horizontally separated points. Simultaneous current measurements at vertically separated points go back to 1930 (Ekman and Helland-Hansen, 1931). In all these papers there is an expression of dismay concerning the lack of resemblance between measurements at such small spatial separations of oscillations with such long periods. I believe (from discussions with Ekman in 1949) that this lack of coherence was the reason why Ekman postponed for 23 years (until one year before his death) the publication of "Results of a Cruise on Board the 'Armauer Hansen' in 1930 under the Leadership of Björn Helland-Hansen" (Ekman, 1953). But the decorrelation distance is just the reciprocal of the bandwidth; waves separated in wavenumber by more than Δk interfere destructively at separations exceeding $(\Delta k)^{-1}$. The small observed coherences are simply an indication of a large bandwidth.

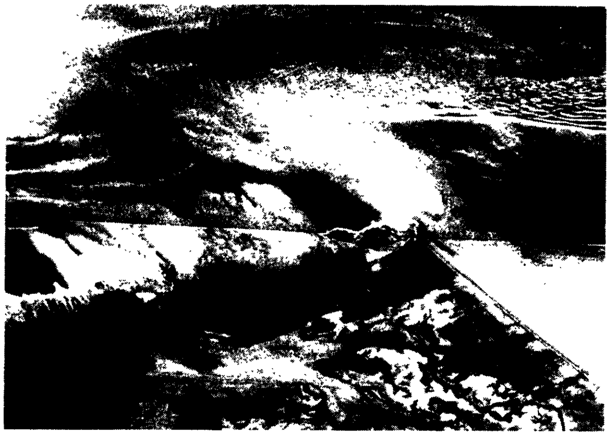
The search for an analytic spectral model to describe the internal current and temperature fluctuations goes back over many years, prompted by the remarkable success of Phillips's (1958) saturation spectrum for surface waves. I shall mention only the work of Murphy and Lord (1965), who mounted temperature sensors in an unmanned submarine at great depth. They found some evidence for a spectrum depending on scalar wavenumber as $k^{-5/3}$, which they interpreted as the inertial subrange of homogeneous, isotropic turbulence. But the inertial subrange is probably not applicable (except perhaps at very small scales), and the fluctuations are certainly not homogeneous and not isotropic.

Briscoe (1975a) has written a very readable account of developments in the early 1970s. The interpretation of multipoint coherences in terms of bandwidth was the key for a model specturm proposed by Garrett and Munk (1972b). The synthesis was purely empirical, apart from being guided by dimensional considerations and by not violating gross requirements for the finiteness of certain fundamental physical properties. Subsequently, the model served as a convenient "strawman" for a wide variety of moored, towed and "dropped" experiments, and had to be promptly modified [Garrett and Munk (1975), which became known as GM75 in the spirit of planned obsolescence]. There have been further modifications [see a review paper by Garrett and Munk (1979)]; the most recent version is summarized at the end of this chapter.

The best modern accounts on internal waves are by O. M. Phillips (1966b), Phillips (1977a), and Turner (1973a). Present views of the time and space scales of internal waves are based largely on densely sampled moored, towed, and dropped measurements. The pioneering work with moorings was done at site D in the western North Atlantic (Fofonoff, 1969; Webster, 1968). Horizontal tows of suspended thermistor chains (Lafond, 1963; Charnock, 1965) were followed by towed and self-propelled isotherm-following "fishes" (Katz, 1973; McKean and Ewart, 1974). Techniques for dropped measurements were developed along a number of lines: rapidly repeated soundings from the stable platform FLIP by Pinkel (1975), vertical profiling of currents from free-fall instruments by Sanford (1975) and Sanford, Drever, and Dunlap (1978), and vertical profiling of temperature from a self-contained yo-yoing capsule by Cairns and Williams (1976). The three-dimensional IWEX (internal wave experiment) array is the most ambitious to date (Briscoe, 1975b). These experiments have served to determine selected parameters of model spectra; none of them so far, not even IWEX, has been sufficiently complete for a straightforward and unambiguous transform into the multidimensional $(\omega, \mathbf{k})$-spectrum. The FLIP measurements come closest, giving an objective spectrum in the two dimensions ω, k_z, with fragmentary information on k_x, k_y. Otherwise only one-dimensional spectra can be evaluated from any single experiment, and one is back to model testing. Yet in spite of these observational shortcomings, there is now evidence for some degree of universality of internal wave spectra, suggesting that these spectra may be shaped by a saturation process (the interior equivalent of whitecaps), rather than by external generation processes.

Internal waves have surface manifestations consisting of alternate bands of roughened and smooth water (Ewing, 1950; Hughes, 1978), and these appear to be visible from satellites (figure 9.1). High-frequency sonar beams are a powerful tool for measuring internal wave related processes in the upper oceans (figures 9.2, 9.3). The probing of the deep ocean interior by acoustics is ultimately limited by scintillations due to internal waves (Flatté et al., 1979; Munk and Wunsch, 1979) just as the "diffraction-limited" telescope has its dimensions set by the small-scale variability in the upper atmosphere.

It will be seen that internal waves are a lively subject. The key is to find the connections between internal waves and other ocean processes. The discovery of ever finer scales, down to the scale of molecular processes,



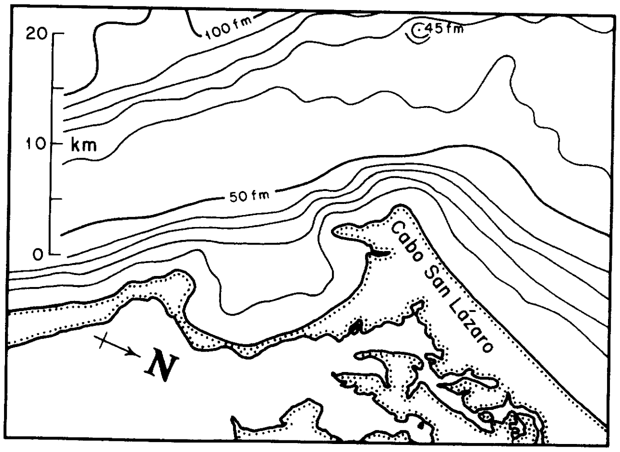


Figure 9.1 SEASAT synthetic aperture radar image off Cabo San Lázaro, Baja California (24°48′N, 112°18′W) taken on 7 July 1978. Scale of image nearly matches that of bathymetric area. The pattern in the right top area is most likely formed by internal waves coming into the 50 fathom line. (I am indebted to R. Bernstein for this figure.)

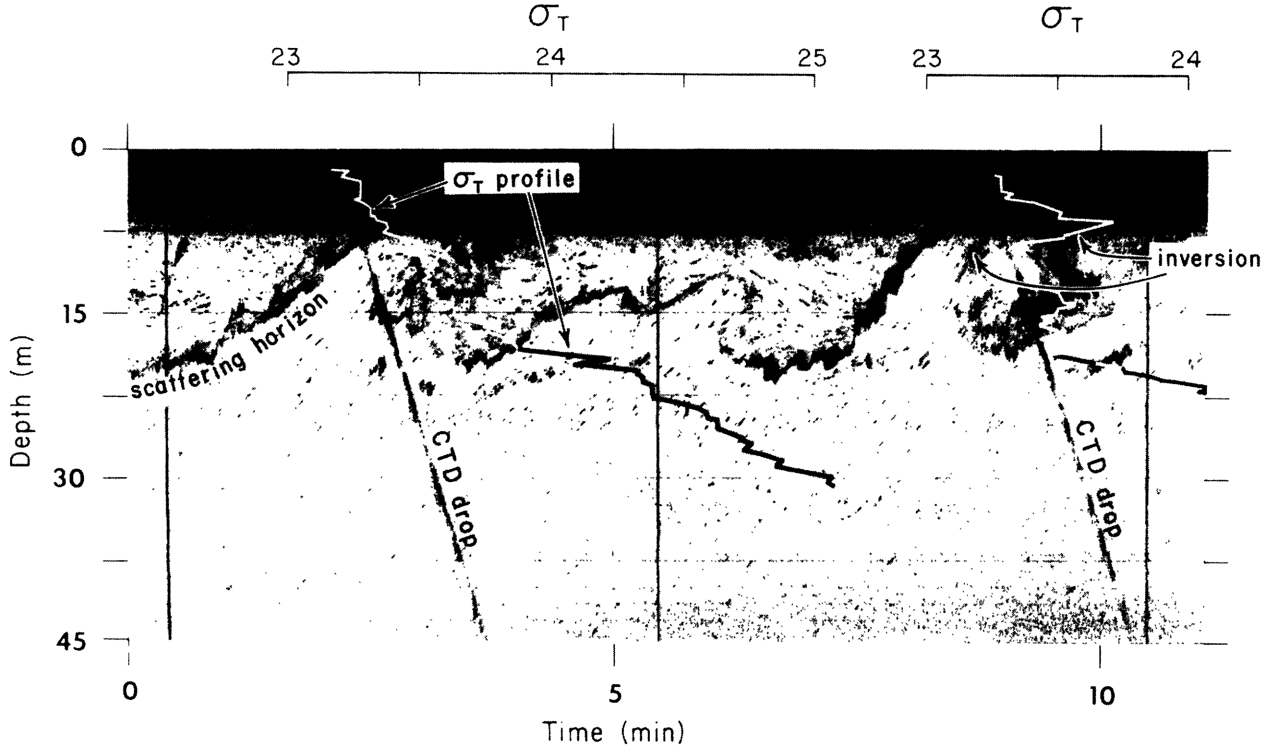


Figure 9.2 The water column is insonified with a narrow downward sonar beam of 200 kHz (wavelength 0.75 cm). The dark band is presumably a back-scattering layer convoluted by shear instabilities. In a number of places the instabilities have created density inversions. This is confirmed by the two σ_t-profiles. The acoustic reflection from the sinking CTD along the steeply slanting lines shows the depth-time history of the σ_t-profiles. The profiling sound source was suspended from a drifting ship. The horizontal distance between overturning events was estimated to be 60–70 m. (I am indebted to Marshall Orr of Woods Hole Oceanographic Institution for this figure; see Haury, Briscoe, and Orr, 1979.)

has been a continuing surprise to the oceanographic community for 40 years. Classical hydrographic casts employed reversing (Nansen) bottles typically at 100-m intervals in the upper oceans beneath the thermocline, and half-kilometer intervals at abyssal depths. Only the gross features can be so resolved. Modern sounding instruments (BT, STD, CTD) demonstrated a temperature and salinity[3] fine structure down to meter scales. An early clue to microstructure was the steppy traces on the smoked slides of bathythermographs. These steps were usually attributed to "stylus stiction," and the instruments suitably repaired.

Free-fall apparatus sinking slowly ($\sim 0.1\ \mathrm{m\,s^{-1}}$) and employing small, rapid-response (~ 0.01 s) transducers, subsequently resolved the structure down to centimeter scales and beyond. The evolving terminology

gross structure:	larger than 100 m vertical
fine structure:	1 m to 100 m vertical
microstructure:	less than 1 m vertical

is then largely based on what could be resolved in a given epoch (see chapter 14). The fine-structure measurements of temperature and salinity owe much of their success to the evolution of the CTD (Brown, 1974). The pioneering microscale measurements were done by Woods (1968a) and by Cox and his collaborators (Gregg and Cox, 1972; Osborn and Cox, 1972). Measurements of velocity fine structure down to a few meters have been accomplished by Sanford (1975) and Sanford, Drever, and Dunlap (1978). Osborn (1974, 1980) has resolved the velocity microstructure between 40 and 4 cm. Evidently velocity and temperature structure have now been adequately resolved right down to the scales for which molecular processes become dominant. At these scales the dissipation of energy and mean-square temperature gradients is directly proportional to the *molecular* coefficients of viscosity and thermal diffusivity. The dissipation scale for salinity is even smaller (the haline diffusivity is much smaller than the thermal diffusivity) and has not been adequately resolved. The time is drawing near when we shall record the entire fine structure and microstructure scales of temperature, salinity and currents [and hence of the buoyancy frequency $N(z)$ and of Richardson number $Ri(z)$] from a single free-fall apparatus.

Perhaps the discovery of very fine scales could have been anticipated. There is an overall ocean balance between the generation and dissipation of mean-square gradients. Eckart (1948) refers to the balancing processes as *stirring* and *mixing*. Garrett (1979) has put it succinctly: "Fluctuations in ocean temperature produced by surface heating and cooling, and in salinity due to evaporation, precipitation, run-off and freezing, are stirred into the ocean by permanent current systems and large scale eddies." Mixing ultimately occurs through dissipation by "molecular action on small-scale irregularities produced by a variety of processes." The microstructure (where the mean-square gradients largely reside) are then a vital component of ocean dynamics. This leaves open the question whether mixing is important throughout the ocean, or whether it is concentrated at ocean boundaries and internal fronts, or in intense currents (an extensive discussion may be found in chapter 8).

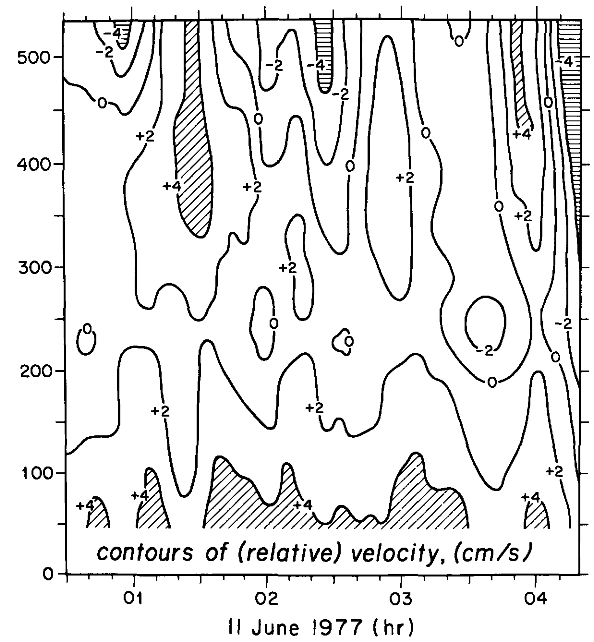


Figure 9.3 Measurements of Doppler vs. range were made at 2-minute intervals with a quasi-horizontal 88-kHz sound beam mounted on FLIP at a depth of 87 m. Bands of alternating positive and negative Doppler (in velocity contours) are the result of back scatter from particles drifting toward and away from the sound source (the mean drift has been removed). The velocities are almost certainly associated with internal wave-orbital motion. The range-rate of positive or negative bands gives the appropriate projection of phase velocity. The measurements are somewhat equivalent to successive horizontal tows at 3000 knots! (I am indebted to Robert Pinkel of Scripps Institution of Oceanography for this figure.)

What are the connections between internal waves and small-scale ocean structure? Is internal wave breaking associated with ocean microstructure? Is there an associated flux of heat and salt, and hence buoyancy? Does the presence of internal waves in a shear flow lead to an enhanced momentum flux, which can be parameterized in the form of an eddy viscosity? What are the processes of internal-wave generation and decay? I feel that we are close to having these puzzles fall into place (recognizing that oceanographic "breakthroughs" are apt to take a decade), and I am uncomfortable with attempting a survey at this time.

Forty years ago, internal waves played the role of an attractive nuisance: attractive for their analytical elegance and their accessibility to a variety of experimental methods, a nuisance for their interference with what was then considered the principal task of physical oceanography, namely, charting the "mean" density field. Twenty years from now I expect that internal waves will be recognized as being intimately involved with the vertical fluxes of heat, salt, and momentum, and so to provide a vital link in the understanding of the mean fields of mass and motion in the oceans.

9.1.1 Preview of This Chapter

We start with the traditional case of a two-layer ocean, followed by a discussion of continuous stratification: constant buoyancy frequency N, N decreasing with depth, a maximum N (thermocline), a double maximum. Conditions are greatly altered in the presence of quite moderate current shears. Short (compliant) internal waves have phase velocities that are generally slower than the orbital currents associated with the long (intrinsic) internal waves, and thus are subject to critical layer processes. There is further nonlinear coupling by various resonant interactions.

Ocean fine structure is usually the result of internal-wave straining, but in some regions the fine structure is dominated by intrusive processes. Microstructure is concentrated in patches and may be the residue of internal wave breaking. Little is known about the breaking of internal waves. Evidently, there are two limiting forms of instability leading to breaking: advective instability and shear instability.

The chapter ends with an attempt to estimate the probability of wave breaking, and of the gross vertical mixing and energy dissipation associated with these highly intermittent events. An important fact is that the Richardson number associated with the internal wave field is of order 1. Similarly the wave field is within a small numerical factor of advective instability. Doubling the mean internal wave energy can lead to a large increase in the occurrence of breaking events; halving the wave energy could reduce the probability of breaking to very low levels. This would have the effect of maintaining the energy level of internal waves within narrow limits, as observed. But the analysis is based on some questionable assumptions, and the principal message is that we do not understand the problem.

9.2 Layered Ocean

We start with the conventional discussion of internal waves at the boundary between two fluids of different density. The configuration has perhaps some application to the problem of long internal waves in the thermocline, and of short internal waves in a stepwise fine structure.

Following Phillips (1977a), this can be treated as a limiting case of a density transition from ρ_u above $z = -h$ to ρ_l beneath $z = -h$, with a transition thickness δh (figure 9.4). The vertical displacement $\zeta(z)$ has a peak at the transition, and the horizontal velocity $u(z)$ changes sign, forming a discontinuity (vortex sheet) in the limit $\delta h \rightarrow 0$. For the second mode (not shown), $\zeta(z)$ changes sign within the transition layer and $u(z)$ changes sign twice; this becomes unphysical in the limit $\delta h \rightarrow 0$. For higher modes the discontinuities are even more pathological, and so a two-layer ocean is associated with only the gravest internal mode.

For the subsequent discussion it is helpful to give a sketch of how the dependent variables are usually derived and related. The unknowns are u,v,w,p (after eliminating the density perturbation), where p is the departure from hydrostatic pressure. The four unknowns are determined by the equations of motion and continuity (assuming incompressibility). The linearized x,y equations of motion are written in the traditional f-plane; for the vertical equation it is now standard [since the work of Eckart (1960)] to display the density stratification in terms of the buoyancy (or Brunt-Väisälä) frequency

$$N(z) = \left\{ -\frac{g}{\rho}\left[\frac{d\rho}{dz} - \left(\frac{d\rho}{dz}\right)_{\text{adiabatic}}\right]\right\}^{1/2}, \tag{9.1}$$

thus giving

$$\frac{\partial w}{\partial t} = \frac{1}{\rho_0}\frac{\partial p}{\partial z} - N^2\zeta = 0.$$

The last term will be recognized as the buoyancy force $-g\,\delta\rho/\rho_0$ of a particle displaced upwards by an amount $\zeta = \int w\,dt$.

For propagating waves of the form $\zeta(z)\exp i(kx - \omega t)$ the equations can be combined (Phillips, 1977a, §5.2 and §5.7) into

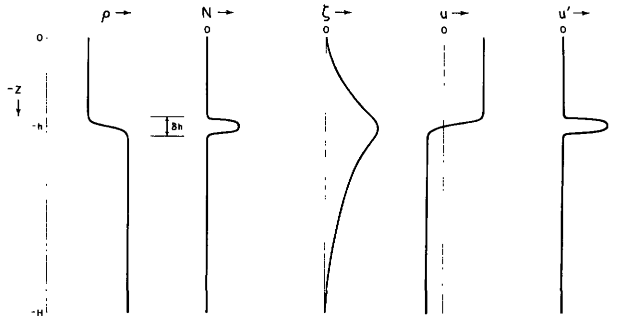


Figure 9.4 A sharp density transition from ρ_u to ρ_l takes place between the depths $-z = h - \frac{1}{2}\delta h$ and $-z = h + \frac{1}{2}\delta h$. This is associated with a delta-like peak in buoyancy frequency $N(z)$. Amplitudes of vertical displacement $\zeta(z)$, horizontal velocity $u(z)$, and shear $u'(z) = du/dz$ are sketched for the gravest internal wave mode.

$$\frac{d^2\zeta}{dz^2} + k^2\frac{N^2(z) - \omega^2}{\omega^2 - f^2}\,\zeta = 0. \tag{9.2}$$

The linearized boundary conditions are $\zeta = 0$ at the surface and bottom.

A simple case is that of $f = 0$ and $N = 0$ outside the transition layer. We have then

$$\zeta_u = A \sinh kz,$$

$$\zeta_l = B \sinh k(z + H),$$

above and below the transition layer, respectively. The constants A and B are determined by patching the vertical displacement at the transition layer:

$$\zeta_u = \zeta_l = a \qquad \text{at} \quad z = -h.$$

The dispersion relation is found by integrating Eq. (9.2) across the transition layer:

$$\begin{aligned}\zeta_u' - \zeta_l' &= -k^2\zeta\int_{\delta h} dz\frac{N^2(z) - \omega^2}{\omega^2}\\ &\approx -\frac{k^2\zeta}{\omega^2}\left(g\frac{\delta\rho}{\rho} - \omega^2\delta h\right) \qquad \text{at} \quad z = -h,\end{aligned}$$

where $\zeta' \equiv d\zeta/dz$. In the limit of small $k\,\delta h$, that is, for waves long compared to the transition thickness, the foregoing equations lead to the dispersion relation

$$\omega^2 = \frac{g(\delta\rho/\rho)k}{\coth kh + \coth k(H - h)}.$$

For a lower layer that is deep as compared to a wavelength, the denominator becomes $\coth k(h - H) + 1$. If the upper layer is also deep, it becomes $1 + 1$, and

$$\omega^2 = \frac{1}{2}gk\frac{\delta\rho}{\rho} = \frac{1}{2}gk\frac{\rho_l - \rho_u}{\frac{1}{2}(\rho_l + \rho_u)}.$$

As $\rho_u \to 0$, $\omega^2 \to gk$, which is the familiar expression for surface waves in deep water.

The case of principal interest here is that of an isolated density transition $\delta\rho \ll \rho$ and $k\,\delta h \ll 1$. Then $\omega^2 = \frac{1}{2}gk\,\delta\rho/\rho$. The vertical displacement is a maximum at the transition and dies off with distance δz from the transition as $a\exp(-k|\delta z|)$.

A question of interest is the variation of Richardson number across the transition layer. We know from the work of Miles and Howard [see Miles (1963)] that for a transition $\rho(z)$ and a *steady* $u(z)$ of the kind shown in figure 9.4, the flow becomes unstable to disturbances of length scale δh if $Ri < \frac{1}{4}$. I find it convenient to refer to the *root-reciprocal* Richardson number

$$Ri^{-1/2} = |u'/N|,$$

so that large values imply large instabilities (as for Reynolds numbers); the critical value is $|u'/N| = 2$. One would think offhand that $|u'/N|$ is a minimum at the transition where N reaches a maximum, but just the opposite is true. To prove this, we use the condition of incompressibility, $iku(z) - i\omega\zeta' = 0$, and equation (9.2) to obtain

$$u'(z) = \frac{\omega}{k}\zeta'' = -\frac{N^2(z) - \omega^2}{\omega}\,ka, \tag{9.3}$$

and so $u' \sim N^2$ for small ω/N; accordingly u'/N varies as N. Thus the layers of largest gravitational stability (largest N) are also the layers of largest shear instability (largest $|u'/N|$).

9.3 Continuously Stratified Ocean

The simplest case is that of constant N. The solution to (9.2) is

$$\zeta(z) = a\sin mz, \qquad m^2 = k^2\frac{N^2 - \omega^2}{\omega^2 - f^2} \tag{9.4}$$

with m so chosen that ζ vanishes at $z = -H$. Solving for ω^2,

$$\omega_j^2 = \frac{k^2N^2 + m_j^2f^2}{m_j^2 + k^2}, \qquad m_jH = j\pi, \qquad j = 1,2,.... \tag{9.5}$$

This dispersion relation is plotted in figure 9.5. The vertical displacements for the first and third mode are shown in figure 9.6. Very high modes (and the ocean is full of them) in the deep interior are many wavelengths removed from the boundaries, and we can expect the waves to be insensitive to the precise configuration of top and bottom. The discrete dispersion $\omega_j(k)$ is then replaced by an equivalent continuous dispersion $\omega(k, m)$.

The standard expressions for the particle velocities u, w and the group velocities $\mathbf{c}_g$ with components $\partial\omega/\partial k$, $\partial\omega/\partial m$ as functions of the propagation vector

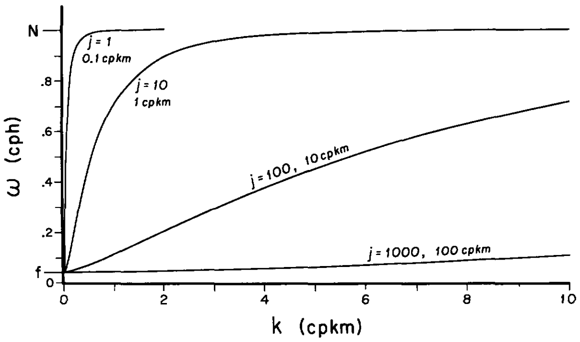


Figure 9.5 The dispersion $\omega_j(k)$ [equation (9.5)], for modes $j = 1, 10, 100, 1000$, corresponding to vertical wavenumbers $m = 0.1, 1, 10, 100$ cpkm in an ocean of depth 5 km. The inertial frequency is taken at $f = 0.0417$ cph (1 cpd), and the buoyancy frequency at $N = 1$ cph.

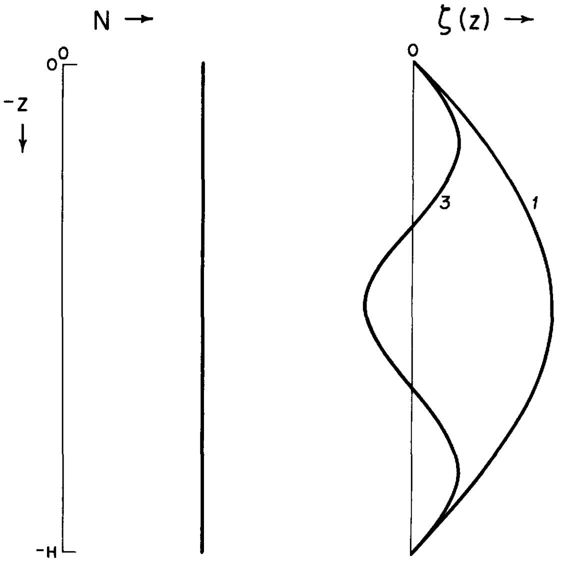


Figure 9.6 Vertical displacements $\zeta(z)$ in a constant-N ocean, for modes $j = 1$ and $j = 3$.

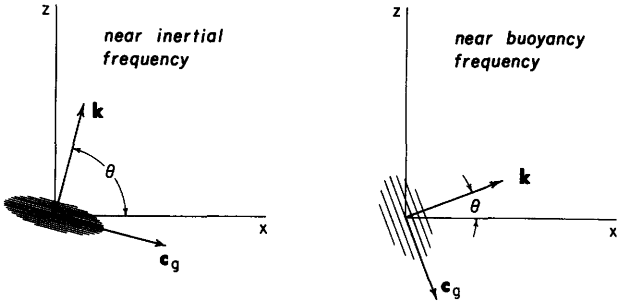


Figure 9.7 The wavenumber vector $\mathbf{k} = (k, m)$ and group velocity $\mathbf{c}_g$ near the inertial frequency $(\omega = f + \epsilon)$ and near the buoyancy frequency $(\omega = N - \epsilon)$, respectively. A packet of wave energy is projected on the (x, z)-plane. Crests and troughs in the wave packet are in a plane normal to $\mathbf{k}$, and travel with phase velocity $\mathbf{c}$ in the direction $\mathbf{k}$. The wave packet travels with group velocity $\mathbf{c}_g$ at right angles to $\mathbf{k}$, thus sliding sideways along the crests and troughs. The particle velocity $\mathbf{u}$ (not shown) is in the planes at right angles to $\mathbf{k}$.

$\mathbf{k} = (k, m)$ are easy to derive, but hard to visualize. Consider a wave packet (figure 9.7) with crests and troughs along planes normal to the paper and inclined with respect to the (x, z)-axis as shown. The phase velocity is in the direction $\mathbf{k}$ normal to the crests, but the group velocity $\mathbf{c}_g$ is parallel to the crests, and the wave packet slides sideways. $\mathbf{k}$ is inclined to the horizontal by

$$\tan\theta = \frac{m}{k} = \left(\frac{N^2 - \omega^2}{\omega^2 - f^2}\right)^{1/2} \tag{9.6a}$$

and so the angle is steep for inertial waves $(\omega = f + \epsilon)$ and flat for buoyancy waves $(\omega = N - \epsilon)$. The energy packet is propagated horizontally for inertial waves, and vertically for buoyancy waves, but the group velocity goes to zero at both limits.

The flow $\mathbf{u} = (u, w)$ takes place in the plane of the crest and troughs. For inertial waves, particles move in horizontal circles. The orbits become increasingly elliptical with increasing frequency, and for buoyancy waves the particle orbits are linear along the z-axis, in the direction of $\mathbf{c}_g$. The wavenumber $\mathbf{k}$ is always normal to both $\mathbf{c}_g$ and $\mathbf{u}$. [The nonlinear field accelerations $(\mathbf{u} \cdot \nabla)\mathbf{u}$ vanish for an isolated elementary wave train, leading to the curiosity that the linear solution is an *exact* solution.] Readers who find it difficult to visualize (or believe) these geometric relations should refer to the beautiful laboratory demonstrations of Mowbray and Rarity (1967).

It is not surprising, then, that internal waves will do unexpected things when reflected from sloping boundaries. The important property is that the inclination θ relative to the x-axis depends only on frequency [equation (9.6a)]. Since frequency is conserved upon reflection, incident and reflected θ must be symmetric with respect to a level surface rather than with respect to the reflecting surface. At the same time the flow $\mathbf{u}$ for the combined incident and reflected wave must be parallel to the reflecting boundary. For a given ω, there is a special angle for which the orbital flow is parallel to the boundary. This requires that the boundary be inclined at a slope

$$\tan\beta(z) = \tan(90^\circ - \theta) = \left[\frac{\omega^2 - f^2}{N^2(z) - \omega^2}\right]^{1/2}. \tag{9.6b}$$

It can be shown that for slopes steeper than β, the energy of "shoreward" traveling internal waves is reflected "seaward": for slopes of less than β, the energy is forward reflected. Repeated reflections in a wedge-shaped region such as the ocean on the continental slope can lead to an accumulation of energy at ever smaller scales (Wunsch, 1969). For a given slope, we can expect an amplification of the internal waves at the frequency ω determined by (9.6b). Wunsch (1972b) has suggested that a peak in the spectrum of temperature fluctuations measured southeast of Bermuda could be so explained. Pertinent values are N = 2.6 cph, f = 0.045 cph, and $\beta \approx 13°$. Equation (9.6b) gives ω = 0.59 cph, in agreement with the observed spectral peak at 0.5 cph.

9.4 Turning Depths and Turning Latitudes

Figure 9.8 shows the situation for an ocean with variable $N(z)$. For frequencies that are less than N throughout the water column, the displacements are similar to those for constant N (figure 9.6) except that the positions of the maxima and zeros are displaced somewhat upward, and that the relative amplitudes are somewhat larger at depth. The important modification occurs for frequencies that exceed $N(z)$ somewhere within the water column. At the depths z_T where $\omega = N(z_T)$, called turning depths, we have the situation shown to the right in figure 9.8. Equation (9.2) is locally of the form $\zeta'' + z\zeta = 0$ where z is now a rescaled vertical coordinate relative to z_T. The solution (called an *Airy function*) has an inflection point at the turning depth (here z = 0), is oscillatory above the turning depth, and is exponentially damped beneath. The amplitudes are somewhat larger just above the turning depth than at greater distance, but nothing very dramatic happens.

The refraction of a propagating wave packet is illustrated in figure 9.9. As the packet moves into depths of diminishing $N(z)$ the crests and troughs turn steeper, and the direction of energy propagation becomes more nearly vertical. The waves are totally reflected at the turning depth z_T where $\omega = N(z_T)$. Modal solutions $\zeta_j(z) \times \exp i(kx - \omega_j t)$ with $\zeta_j(z)$ as illustrated in figure 9.8 can be regarded as formed by superposition of propagating waves with equal upward and downward energy transport. The wave energy remains trapped between the surface and the turning depth.

The common situation for the deep ocean is the main thermocline associated with a maximum in $N(z)$. Internal waves with frequencies less than this maximum are in a waveguide contained between upper and lower turning depths. For relatively high (but still trapped) frequencies the sea surface and bottom boundaries play a negligible role, and the wave solutions can be written in a simple form (Eriksen, 1978). The bottom boundary condition (9.5) for a constant-N ocean, e.g., $m_jH = j\pi$, j = 1,2,..., is replaced in the WKB approximation by

$$m_j b = j\pi \left(\frac{N^2 - \omega^2}{N_0^2 - \omega^2}\right)^{1/2} \approx j\pi N/N_0, \tag{9.7}$$

where b is a representative thermocline (or stratification) scale. Equation (9.7) assures an exponential attenuation outside the waveguide. For the case of a double peak in $N(z)$ with maxima N_1 and N_2, the internal wave energy is concentrated first at one thermocline, then

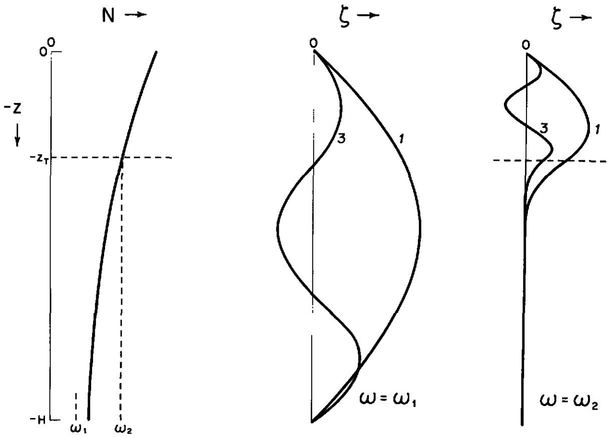


Figure 9.8 Vertical displacements $\zeta(z)$ in a variable-N ocean, for modes j = 1 and j = 3. ω_1 is taken to be less than $N(z)$ at all depths. ω_2 is less than $N(z)$ in the upper oceans above $z = -z_p$ only.

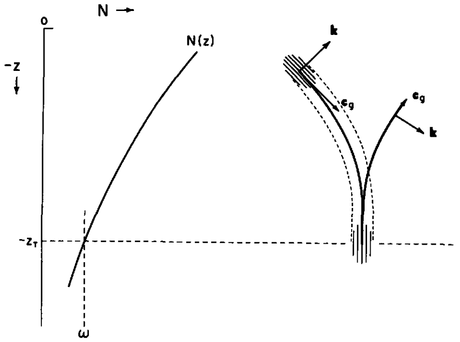


Figure 9.9 Propagation of a wave packet in a variable-$N(z)$ ocean without shear (U = constant). The turning depth z_T occurs when $\omega = N(z_T)$.

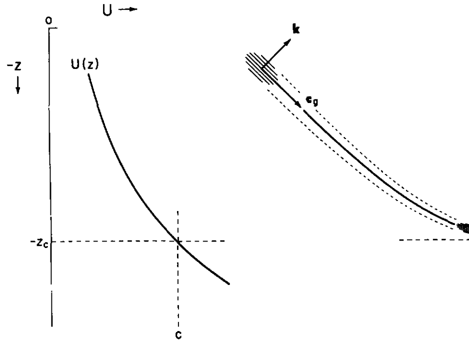


Figure 9.10 Propagation of a wave packet in a constant-N ocean with shear. The critical depth z_C occurs where $U = c(z_C)$.

the other, migrating up and down with a frequency $|N_1 - N_2|$ (Eckart, 1961). This is similar to the behavior of two loosely coupled oscillators. The quantum-mechanical analogy is that of two potential minima and the penetration of the potential barrier between them.

There is a close analogy between the constant- and variable-N ocean, and the constant- and variable-f ocean (the f-plane and β-plane approximations). For a fixed ω, the condition $\omega = f = 2\Omega \sin \phi_T$ determines the turning latitude ϕ_T. Eastward-propagating internal gravity waves have solutions of the form $\eta(y)\zeta(z)\exp i(kx - \omega t)$. The equation governing the local north-south variation is (Munk and Phillips, 1968)

$$\eta'' + y\eta = 0, \qquad \eta'' = d^2\eta/dy^2,$$

where y is the poleward distance (properly scaled) from the turning latitude. This is in close analogy with the up-down variation near the turning depth, which is governed by

$$\zeta'' + z\zeta = 0, \qquad \zeta'' = d^2\zeta/dz^2.$$

Thus $\eta(y)$ varies from an oscillatory to an exponentially damped behavior as one goes poleward across the turning latitude. Poleward-traveling wave packets are reflected at the turning latitude.

From an inspection of figure 9.7, it is seen that the roles of horizontal and vertical displacements are interchanged in the $N(z)$ and $f(y)$ turning points. In the $N(z)$ case the motion is purely vertical; in the $f(y)$ case the motion is purely horizontal (with circular polarization).

It has already been noted that nothing dramatic is observed in the spectrum of vertical displacement (or potential energy) near $\omega = N$—only a moderate enhancement, which can be reconciled to the behavior of the Airy function (Desaubies, 1975; Cairns and Williams, 1976). Similarly we might expect only a moderate enhancement in the spectrum of horizontal motion (or kinetic energy) near $\omega = f$. In fact, the spectrum is observed to peak sharply. If the horizontal motion is written as a sum of rotary components (Gonella, 1972), it is found that the peak is associated with negative rotation (clockwise in the northern hemisphere).

I have made a parallel derivation of the spectra at the two turning points (figure 9.11), assuming horizontally isotropic wave propagation within the entire equatorial waveguide. It turns out that the buoyancy peak is in fact much smaller than the inertial peak at moderate latitudes. But at very low latitude the inertial peak vanishes. This is in accord with the equatorial observations by Eriksen (1980). Fu (1980) gives an interesting discussion of the relative contributions to the spectral peak at the local inertial frequency $\omega = f_{\text{local}}$ from two processes: (1) *local* generation of resonant inertial

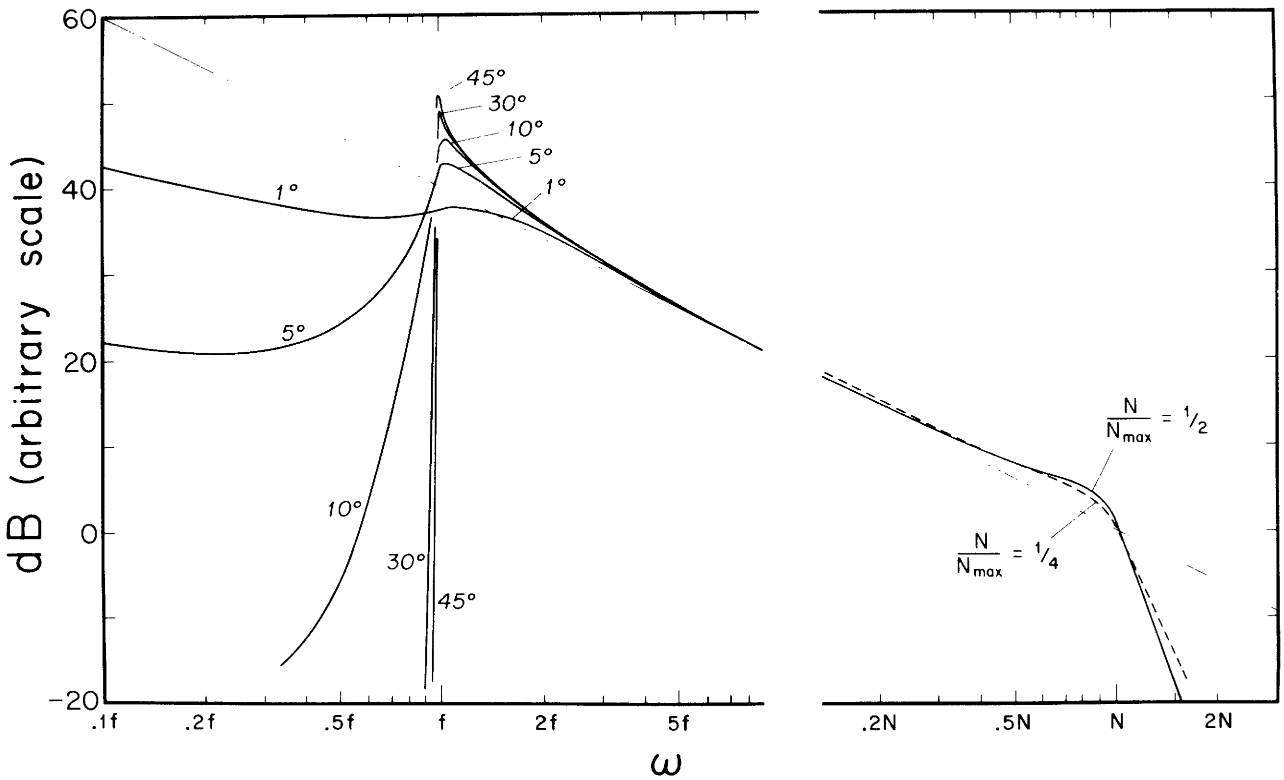


Figure 9.11 Enhancement of the kinetic-energy spectrum (left) and of the potential-energy spectrum (right) at the inertial and buoyancy frequencies, respectively. The inertial spectrum is drawn for latitudes 1°, 5°, 10°, 30°, 45°. The buoyancy spectrum is drawn for two depths, corresponding to $N = \frac{1}{2}, \frac{1}{4}$ times the maximum buoyancy frequency.

waves $\omega = f_{\text{local}}$; and (2) *remote* generation of waves of the same frequency $\omega = f_{\text{local}}$ at lower latitudes (where $f < f_{\text{local}}$). Figure 9.11 is drawn for case 2 under the assumption that the equatorial waveguide is filled with horizontally isotropic, freely propagating radiation. Take the curve marked 30°, say. Then for $\omega > f$ a station at lat. 30° is within the equatorial waveguide; for $\omega < f$ the spectrum is the result of evanescent extensions from a waveguide bounded by lower latitudes. Over rough topography and in regions of strong surface forcing, the case can be made for local generation of the inertial peak. It would seem that the buoyancy peak at mid-depth must always be associated with remote generation.

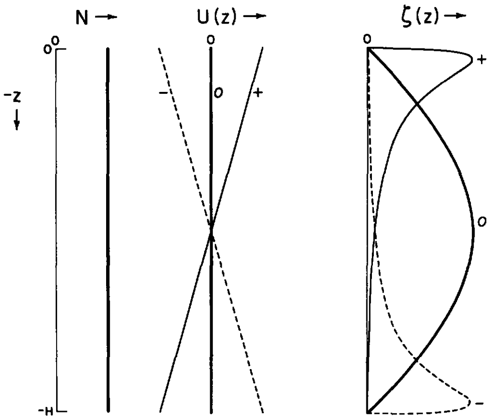


Figure 9.12 First mode vertical displacements $\zeta(z)$ in a Couette flow (constant U' and constant N), for $U'/N = 0, \pm 1$. Waves move from left to right, and U is positive in the direction of wave propagation. (Thorpe, 1978c.)

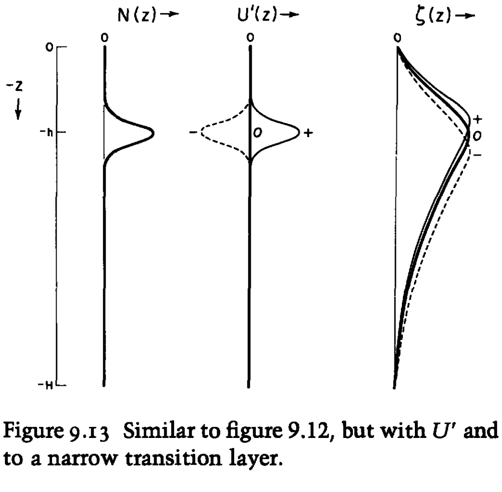


Figure 9.13 Similar to figure 9.12, but with U' and N confined to a narrow transition layer.

9.5 Shear

Internal waves are greatly modified by an underlying shear flow.[4] A variable $U(z)$ can have a more traumatic effect on internal waves than a variable $N(z)$. For ready comparison with figure 9.9 showing the effect of a variable $N(z)$ on a traveling wave packet, we have sketched in figure 9.10 the situation for a wave packet traveling in the direction of an increasing $U(z)$. As the wave packet approaches the "critical depth" z_C where the phase velocity (in a fixed frame of reference) equals the mean flow, $c = U(z_C)$, the vertical wavenumber increases without limit (as will be demonstrated).

For the present purpose we might as well avoid additional complexities by setting $f = 0$. The theoretical starting point is the replacement of ∂_t by $\partial_t + U\partial_x + w\,\partial_z$ in the linearized equations of motion. The result is the Taylor-Goldstein equation [Phillips (1977a, p. 248)]:

$$\frac{d^2\zeta}{dz^2} + \left(\frac{N^2}{(U-c)^2} - \frac{U''}{U-c} - k^2\right)\zeta = 0, \qquad U'' = \frac{d^2U}{dz^2}, \tag{9.8}$$

where c is the phase velocity in a fixed reference frame. (This reduces to

$$\frac{d^2\zeta}{dz^2} + k^2\frac{N^2-\omega^2}{\omega^2}\zeta = 0 \tag{9.9}$$

for $U = 0$.) The singularity at the critical depth where $U = c$ is in contrast with the smooth turning-point transition at $N = \omega$; this is the analytic manifestation of the relative severity of the effect of a variable $U(z)$ versus that of a variable $N(z)$.

Thorpe (1978c) has computed the wave function $\zeta(z)$ for (1) the case of constant N and U' and (2) the case where N and U' are confined to a narrow transition layer. The results are shown in figures 9.12 and 9.13. The profiles are noticeably distorted relative to the case of zero shear, with the largest amplitudes displaced toward the level at which the mean speed (in the direction of wave propagation) is the greatest. Finite-amplitude waves have been examined for the case 2. Where there is a forward[5] flow in the upper level (including the limiting case of zero flow), the waves have narrow crests and flat troughs, like surface waves; with backward flow in the upper layer, the waves have flat crests and narrow troughs. Wave breaking is discussed later.

9.5.1 Critical Layer Processes[6]

The pioneering work is by Bretherton (1966c), and by Booker and Bretherton (1967). Critical layers have been associated with the occurrence of clear-air turbulence; their possible role with regard to internal waves in the oceans has not been given adequate attention.

Following Phillips (1977a), let

$$\omega_F = kU + \omega, \qquad \frac{\omega}{N} = \frac{k}{(k^2 + m^2)^{1/2}} = \cos\theta \qquad (9.10)$$

designate the frequency in a *fixed* reference frame. $U(z)$ is the *mean current* relative to this fixed frame, and $\omega_F - kU = \omega$ is the *intrinsic* frequency [as in (9.5)], as it would be measured from a reference frame drifting with the mean current $U(z)$.

Bretherton (1966c) has given the WKB solutions for waves in an ocean of constant N and slowly varying U. [It is important to note the simplification to (9.8) when $U'' = 0$ at the critical layer.] Near the criticial layer depth z_C, the magnitudes of w, u, and of the vertical displacement ζ vary as

$$w \sim |z - z_C|^{1/2}, \qquad u \sim |z - z_C|^{-1/2}, \qquad \zeta \sim |z - z_C|^{-1/2}.$$

The quantities ω_F and k are constant in this problem, but m and ω are not. The vertical wavenumber increases, whereas the intrinsic frequency decreases as a wave packet approaches its critical layer:

$$m \sim |z - z_C|^{-1}, \qquad \omega \sim |z - z_C|.$$

A sketch of the trajectory is given in figure 9.10. Waves are refracted by the shear and develop large vertical displacements ζ (even though $w \to 0$), large horizontal velocities u, and very large induced vertical shears u'. This has implications for the dissipation and breaking of internal waves.

For $Ri > \frac{1}{4}$, Booker and Bretherton (1967) derived an energy transmission coefficient

$$\rho = \exp(-2\pi\sqrt{Ri - \tfrac{1}{4}}). \qquad (9.11)$$

In the usual case, $U' \ll 2N$ so that $Ri \gg \frac{1}{4}$ and ρ is small. This is interpreted as wave energy and momentum being absorbed by the mean flow at z_C. As $Ri \to \infty$, $\rho \to 0$, consistent with the WKB prediction of Bretherton (1966c) that a wave packet approaches but never reaches the critical layer.

The small coefficient of transmission for Richardson numbers commonly found in the ocean implies that the critical layer inhibits the vertical transfer of wave energy. This effect has been verified in the laboratory experiments of Bretherton, Hazel, Thorpe, and Wood (1967). When rotation is introduced, the energy and momentum delivered to the mean flow may alternatively be transferred from high-frequency to low-frequency waves (if the time scales are appropriate). Thus it is possible that some sort of pumping mechanism may exist for getting energy into, for example, the high-mode, quasi-inertial internal waves. This mechanism can be compared with McComas and Bretherton's (1977) parametric instability, a weakly nonlinear interaction (section 9.6).

The work of Bretherton and of Booker and Bretherton has prompted a great number of critical-layer studies. One of the most interesting extensions was done by Jones (1968). Whereas Booker and Bretherton found the critical layer to be an absorber, not a reflector, when $Ri > \frac{1}{4}$, Jones found that reflection from the critical layer is possible when $Ri < \frac{1}{4}$; in fact, the reflected wave amplitude can exceed that of the incident wave. Jones called these waves "overreflected," their energy being enhanced at the expense of the mean flow. This is illustrated in figure 9.14, based on a solution for a hyperbolic-tangent profile intended to display the results of linear theory. Transmission and reflection ratios at $z = \pm\infty$ were derived using definitions of wave energy density appropriate to moving media. "Overtransmission" as well as overreflection occurs at very small Richardson numbers, with the internal waves gaining energy from the mean flow on both counts.

We shall now consider the condition for critical-layer absorption. Let ω_1 designate the intrinsic frequency of a wave packet at some depth z_1 with a mean flow U_1 in the direction of wave propagation. According to (9.10),

$$\omega_F = kU_1 + \omega_1.$$

Let U increase to some value U_2 at z_2. Then since ω_F and k are conserved along the trajectory of the wave packet,

$$\omega_F = kU_2 + \omega_2.$$

For the special case that z_2 is to be a critical depth, we have $\omega_F = kU_2$, hence $\omega_2 = 0$, and so

$$\omega_1/k = U_2 - U_1.$$

The vertical wavenumber of internal waves is given by the dispersion relation

$$m = k\left(\frac{N^2 - \omega^2}{\omega^2 - f^2}\right)^{1/2}$$

$$\approx kN/\omega \qquad \text{for} \quad f \ll \omega \ll N. \qquad (9.12)$$

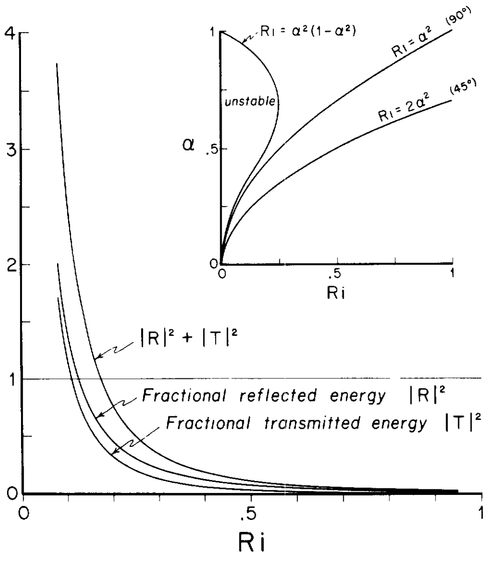


Figure 9.14 Fractional internal wave energy reflected and transmitted through a mean shear flow $U = U_0 \tanh(z/d)$ at constant N, as a function of the minimum Richardson number $Ri = N^2d^2/U_0^2$. Internal wave energy is lost to the mean flow for $|R|^2 + |T|^2 < 1$, or $Ri > 0.18$; internal wave energy is gained from the mean flow for $Ri < 0.18$. The plot is drawn for $Ri = 2\alpha^2$, where $\alpha = kd$ is the dimensionless horizontal wavenumber. This corresponds to a wave packet traveling at an inclination of 45° at $z = \pm\infty$. ($Ri = \alpha^2$ corresponds to the limiting case of vertical group velocity to $\pm\infty$.) (I am indebted to D. Broutman for this figure.)

For critical absorption within the interval Δz over which the mean flow varies by ΔU, we replace ω/k by ΔU, and obtain the critical vertical wavenumber

$$m_C = N/\Delta U. \tag{9.13}$$

R. Weller (personal communication) has analyzed a month of current measurements off California for the expected *difference* $\Delta U = |U_2 - U_1|$ in a velocity component (either of the two components) at two levels separated by $\Delta z = |z_2 - z_1|$. The observations are, of course, widely scattered, but the following values give representative magnitudes:

Δz in m	0	10	25	50	100	
ΔU in cm s^{-1}	0	4	7	10	15	(upper 100 m)
	0	7	10	12	15	(100–300 m depth)

For $\Delta U = 10$ cm s^{-1} and $N = 0.01$ s^{-1} (6 cph), (9.13) gives $m_C = 10^{-3}$ cm^{-1} (16 cpkm). Internal waves with vertical wavelengths of less than 60 m are subject to critical-layer interactions.

A large fraction of the measured velocity difference ΔU can be ascribed to the flow field $u(z)$ of the internal waves themselves, and deduced from the model spectra. The expected velocity difference increases to $\sqrt{2}$ times the rms value as the separation increases to the vertical coherence scale, which is of order 100 m. Here most of the contribution comes from low frequencies and low wavenumbers. I am tempted to interpret ΔU for critical layer processes as rms u from the internal waves themselves. The internal wave spectrum is then divided into two parts: (1) the *intrinsic* part $m < m_C$, which contains most of the energy, and (2) the *compliant* part $m > m_C$, which is greatly modified by interaction with the intrinsic flow field. The phase speed for critical reflection is

$$c_C = \text{rms}\, u, \tag{9.14}$$

and the critical wavenumber is

$$m_C = N/\text{rms}\, u. \tag{9.15}$$

There is the separate question whether the internal waves at the critical layer will be underreflected, just reflected, or overreflected, and this depends on the ambient Richardson number. In the underreflected case there is a flux of energy from the compliant to the intrinsic waves. In the overreflected case the flow is the other way. For an equilibrium configuration, one may want to look for a transmission coefficient ρ near unity, and the exponential behavior of $\rho(Ri)$ will then set narrow bounds to the ambient spectrum. But this gets us into deep speculation, and had better be left to the end of this chapter.

9.6 Resonant Interactions

Up to this point the only interactions considered are those associated with critical layers. In the literature the focus has been on the resonant interaction of wave triads, using linearized perturbation theory. There are two ways in which critical layer interactions differ from resonant interactions: (1) compliant waves of *any* wavenumber and *any* frequency are modified, as long as c equals u somewhere in the water column; and (2) the modification is apt to be large (the ratio u/c being a very measure of nonlinearity). For the wave triads, the interaction is (1) limited to *specific* wavenumbers and frequencies, and (2) assumed to be small in the perturbation treatment.[7] To borrow some words of O. M. Phillips (1966b), the contrast is between the "strong, promiscuous interactions" in the critical layer and the "weak, selective interactions" of the triads.

The conditions for resonance are

$$\mathbf{k}_1 \pm \mathbf{k}_2 = \mathbf{k}_3, \qquad \omega_1 \pm \omega_2 = \omega_3,$$

where $\mathbf{k}_i = (k_i, l_i, m_i)$, and all frequencies satisfy the dispersion relation $\omega_i(\mathbf{k}_i)$. Resonant interactions are well demonstrated in laboratory experiments. For a transition layer (as in figure 9.4), Davis and Acrivos (1967) have found that a first-order propagating mode, which alternately raises and lowers the transition layer, was unstable to resonant interactions, leading to a rapid growth of a second-order mode, which alternately thickens and thins the transition layer like a propagating link sausage. Martin, Simmons, and Wunsch (1972) have demonstrated a variety of resonant triads for a constant-N stratification.

Among the infinity of possible resonant interactions, McComas and Bretherton (1977) have been able to identify three distinct classes that dominate the computed energy transfer under typical ocean conditions. Figure 9.15 shows the interacting propagation vectors in (k, m)-space. The associated frequencies ω are uniquely determined by the tilt of the vectors, in accordance with (9.4). Inertial frequencies (between f and $2f$, say) correspond to very steep vectors, buoyancy frequencies (between $\frac{1}{2}N$ and N) to flat vectors, as shown.

Elastic scattering tends to equalize upward and downward energy fluxes for all but inertial frequencies. Suppose that $\mathbf{k}_3$ is associated with waves generated near the sea surface propagating energy downward (at right angles to $\mathbf{k}_3$, as in figure 9.7). These are scattered into $\mathbf{k}_1$, with the property $m_1 = -m_3$, until the upward energy flux associated with $\mathbf{k}_1$ balances the downward flux by $\mathbf{k}_3$. The interaction involves a near-intertial wave $\mathbf{k}_2$ with the property $m_2 \approx 2m_3$. (The reader will be reminded of Bragg scattering from waves having half the wavelength of the incident and back-scattered radiation.) Similarly, for bottom-generated $\mathbf{k}_1$ waves with upward energy fluxes, elastic scattering will transfer energy into $\mathbf{k}_3$ waves.

Induced diffusion tends to fill in any sharp cutoffs at high wavenumber. The interaction is between two neighboring wave vectors of high wavenumber and frequency, $\mathbf{k}_1$ and $\mathbf{k}_3$, and a low-frequency low-wavenumber vector $\mathbf{k}_2$. Suppose the $\mathbf{k}_2$ waves are highly energetic, and that the wave spectrum drops sharply for wavenumbers just exceeding $|\mathbf{k}_3|$, such as $|\mathbf{k}_1|$. This interaction leads to a diffusion of action (energy/ω) into the low region beyond $|\mathbf{k}_3|$, thus causing $\mathbf{k}_1$ to grow at the expense of $\mathbf{k}_2$.

Parametric subharmonic instability transfers energy from low wavenumbers $\mathbf{k}_2$ to high wavenumbers $\mathbf{k}_1$ of half the frequency, $\omega_1 = \frac{1}{2}\omega_2$, ultimately pushing energy into the inertial band at high vertical wavenumber. The interaction involves two waves $\mathbf{k}_1$ and $\mathbf{k}_3$ of nearly opposite wavenumbers and nearly equal frequencies. The periodic tilting of the isopycnals by $\mathbf{k}_2$ varies the buoyancy frequency at twice the frequency of $\mathbf{k}_1$ and $\mathbf{k}_3$. (The reader will be reminded of the response of a pendulum whose support is vertically oscillated at twice the natural frequency.)

The relaxation (or interaction) time is the ratio of the energy density at a particular wavenumber to the net energy flux to (or from) this wavenumber. The result depends, therefore, on the assumed spectrum. For representative ocean conditions, McComas (in preparation) finds the relaxation time for elastic scattering to be extremely short, of the order of a period, and so up- and downgoing energy flux should be in balance. This result does not apply to inertial frequencies, consistent with observations by Leaman and Sanford (1975) of a downward flux at these frequencies. The relaxation time for induced diffusion is typically a fraction of a period! (This is beyond the assumption of the perturbation treatment.) Any spectral bump is quickly wiped out. The conclusion is that the resonant interactions impose strong restraints on the possible shapes of stable spectra.

In a challenging paper, Cox and Johnson (1979) have drawn a distinction between radiative and diffusive transports of internal wave energy. In the examples cited so far, energy in wave packets is radiated at group velocity in the direction of the group velocity. But suppose that wave-wave interactions randomize the direction of the group velocity. Then eventually the wave energy is spread by diffusion rather than radiation. The relevant diffusivity is $\kappa = \frac{1}{3}\langle c_g^2 \rangle \tau$, where τ is the relaxation time of the nonlinear interactions. Cox and Johnson have estimated energy diffusivities and momentum diffusivities (viscosities); they find that beyond 100 km from a source, diffusive spreading is apt to dominate over radiative spreading. There is an interesting analogy to crystals, where it is known that energy associated with thermal agitation is spread by diffusion rather than by radiation. The explanation lies in the anharmonic restoring forces between molecules, which bring about wave-wave scattering at room temperatures with relaxation times in the nanoseconds.

9.7 Breaking

This is the most important and least understood aspect of our survey. Longuet-Higgins has mounted a broadly based fundamental attack on the dynamics of breaking surface waves, starting with Longuet-Higgins and Fox (1977), and this will yield some insight into the internal-wave problems. At the present time we depend on laboratory experiments with the interpretation of the results sometimes aided by theoretical considerations.

Figure 9.16 is a cartoon of the various stages in an experiment performed by Thorpe (1978b). A density transition layer is established in a long rectangular tube. An internal wave maker generates waves of the first vertical mode. Before the waves have reached the far end of the tube, the tube is tilted through a small angle to induce a slowly accelerating shear flow. The underlying profiles of density, shear, and vertical displacement correspond roughly to the situation in figure 9.13.

For relatively steep waves in a weak positive[8] shear, the waves have sharpened crests. At the position of the crest, the density profile has been translated upward and steepened (B_1). There is significant wave energy loss in this development (Thorpe, 1978c, figure 10).

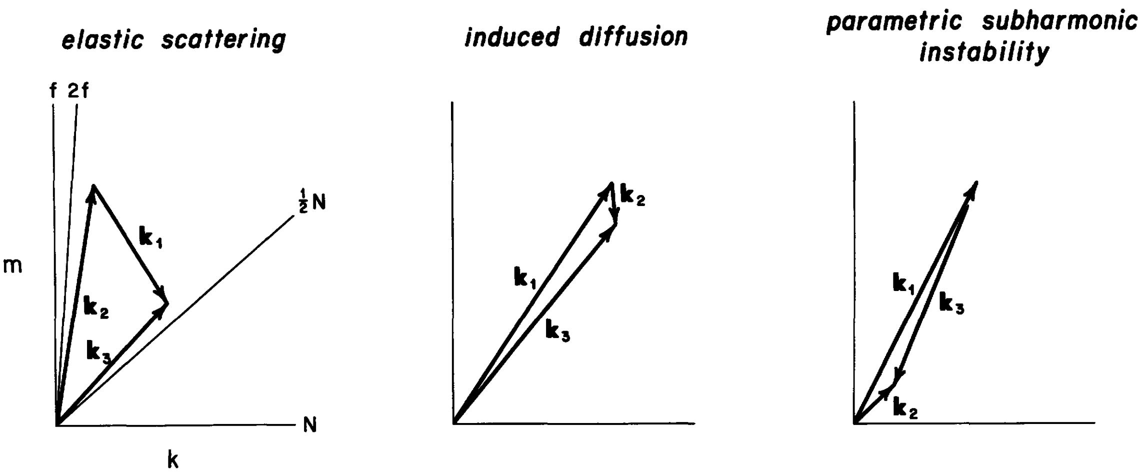


Figure 9.15 Resonant triads for three limiting classes of interaction, according to McComas and Bretherton (1977). The propagation vectors are drawn in (k, m)-space. Radial lines designate the tilt of the $\mathbf{k}$ vectors for $\omega = f, 2f, \frac{1}{2}N, n$. taking $N = 24f$.

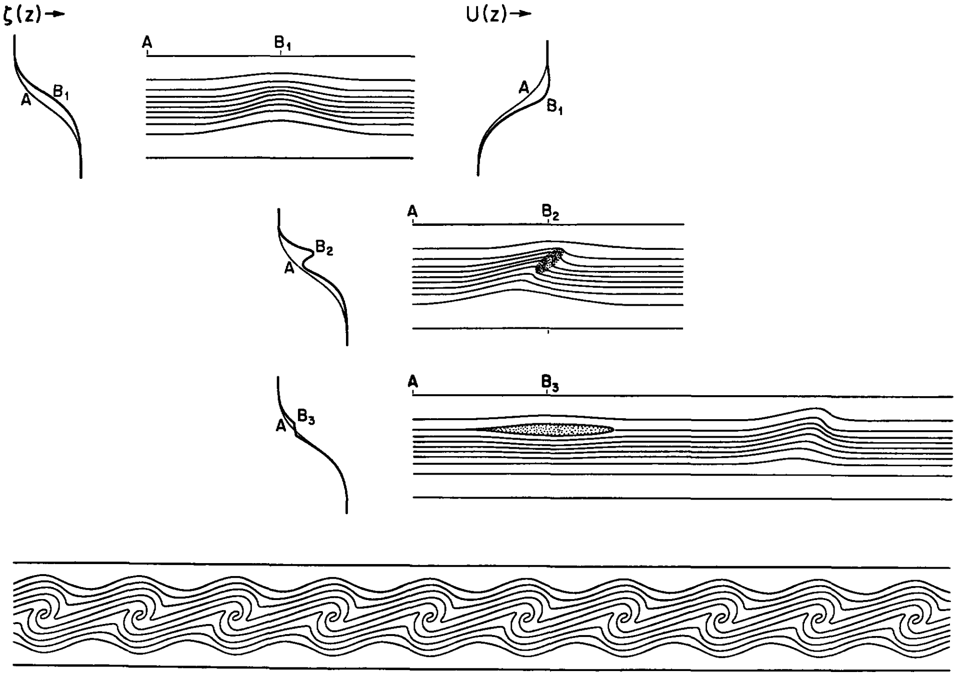


Figure 9.16 Cartoon for various stages of Thorpe's experiment. The early stages lead to the development of *advective* instability (upper three sketches), and the final stage to *shear* instability (bottom). Waves are traveling from left to right; the mean flow is forward (in the direction of wave propagation) above the density transition layer and backward below the transition layer. The density profiles along the indicated vertical sections are shown to the left; a velocity profile is shown to the top right (thin lines give the undisturbed profiles).

With increased positive shear, or with increased time, the particles at the crest accelerate, the isopycnal wave front becomes momentarily vertical, and a jet of fluid moves forward of the crest (B_2). The resulting density inversion gives rise to a Rayleigh-Taylor instability, forming a turbulent patch (shaded) whose turbulent energy is irretrievably lost to the organized wave motion. The turbulent patch becomes fairly well mixed, and introduces a steplike feature into the density profile (B_3). The patch spreads horizontally under the influence of the ambient stratification, forming *blini,* or pancakes. The detailed dynamics are complicated (Barenblatt and Monin, 1979); it is possible that in the oceans the spreading of the patches is eventually retarded by geostrophic confinement.

In Thorpe's laboratory experiment, the later stages of horizontal spreading are interrupted by the sudden formation of billows that grow rapidly, extracting energy from the mean shear flow (bottom of figure). Their wave length is quite short, only several times the thickness of the transition layer.

Hence, Thorpe (1978b, 1979) distinguishes between two types of instability leading to internal wave breaking. In the case of *advective instability,* breaking grows out of existing large-amplitude internal waves: more precisely, waves associated with steep isopycnal slopes. Eventually the particles in the crest are advected forward of the crest, leading to a local density inversion with the potential for a Rayleigh-Taylor instability. Advective instability can take place in the absence of ambient shear, though it is advanced by shear. The second type is induced *shear instability* (Kelvin-Helmholtz instability in the limit of an abrupt density transition), and can take place even in the absence of any (finite) wave disturbance, but is catalyzed by an existing wave background.

The two types appear as end points on a stability curve in slope-shear space, constructed by Thorpe (1978b, 1979) from theory and experiment (figure 9.17). Under the conditions described by the author, internal waves on a transition layer are unstable if their slope exceeds 0.34 in the absence of ambient shear, and if the shear exceeds $2N$ in the absence of slope. Away from the end points, there is advective instability modified by shear, and shear (K-H) instability modified by advection. The stability curve for the transition profile is not symmetric, implying that (under the prescribed geometry) negative shear delays instability.

The essential feature of advective instability is that the particle speed at the crest eventually exceeds the wave speed. The stability curve in figure 9.16 has been constructed from $u_{\text{crest}} = c$ (carrying the theory to third order in wave slope). This is in fair agreement with experiment. From a similar point of view, Orlanski and Bryan (1969) had previously derived the required critical amplitude for advective instability in the oceans, and have checked their analysis with numerical experiments. They conclude that more than enough internal wave energy exists for this type of instability to occur. They also conclude that conditions favor advective instability over shear instability, by the following very simple argument. From (9.12),

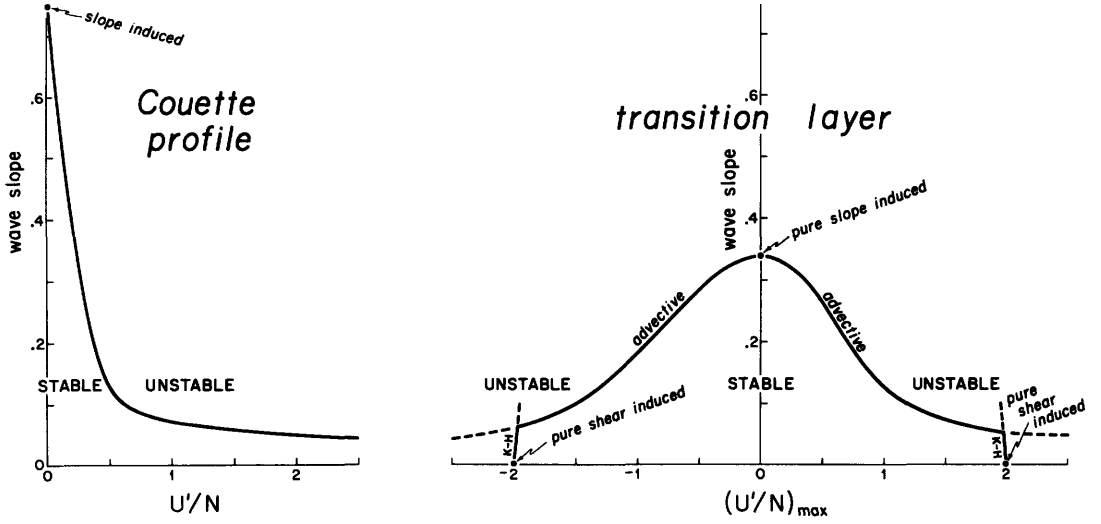


Figure 9.17 Stability diagram for internal waves in a shear flow for a Couette profile (as in figure 9.12) and for a transition layer (as in figure 9.13). Slope is defined as $\pi \cdot$wave height/wavelength. The ordinate is U'/N (or $\pm Ri^{-1/2}$), with U in the direction of wave propagation. For the transition profile, U' and N^2 are proportional to $\text{sech}^2(z - h)$, and hence $U'/N \sim \text{sech}(z - h)$ has its maximum value at the transition $z = h$. The curves are drawn for the specific dimensions described by Thorpe (1978a, 1979).

$$\begin{aligned}\frac{mu}{N} &= \frac{u}{\omega/k}\left(\frac{1-\omega^2/N^2}{1-f^2/\omega^2}\right)^{1/2}\\ &\approx \frac{u}{\omega/k} \qquad \text{for} \quad f \ll \omega \ll N.\end{aligned} \tag{9.16}$$

But $mu/N = Ri^{-1/2} = 2$ for shear instability, and $u/(\omega/k) = 1$ for advective instability. The linearized[9] treatment then says that waves have to be twice as high to be shear unstable than to be advectively unstable. The trouble with this argument is that it is limited to the self-shear of an elementary wave train, and does not take into account the imposed ambient shear (possibly due to other components of the internal wave spectrum). Thorpe's stability plot (figure 9.17) shows that the instabilities can go either way, depending on wave slope and the ambient shear.

McEwan (1973) has generated breaking internal waves in the laboratory by crossing two internal wave beams from separate sources. He finds that the breaking is associated with localized, abruptly appearing intensification in density gradient and shear. These "traumata" persist and spread, and become the locus of incipient turbulence. True turbulent disorder was always preceded by the sudden and widespread occurrence of the traumata.

In some further laboratory experiments with breaking internal waves, McEwan (personal communication) has estimated separately the work done in generating internal waves (allowing for wall friction), and the fraction of this work going into mixing, e.g., going into the increase in the potential energy of the mean stratification. The remaining energy is dissipated into heat. McEwan finds that something less than $\frac{1}{4}$ of the input energy goes into mixing, in support of an estimate by Thorpe (1973b). (In the ocean, the mixing of salt and heat may proceed at different rates because of the disparity in the diffusivities.) Thompson (1980) argues that this ratio is, in fact, the critical Richardson number.

All of this points toward a strong connection between breaking internal waves and the microstructure of density and velocity. Evidently, breaking internal waves can modify a density profile, reducing gradients in turbulent patches and sharpening them elsewhere. This can lead to a steppy fine structure. But we have shown that internal wave shear is concentrated at the steps, thus producing conditions for shear instability, and renewed breaking. This is like the chicken and the egg: which comes first?

9.8 Ocean Fine Structure and Microstructure

Measurements by Gregg (1975) off Cabo San Lucas and in the North Pacific gyre (figure 9.18 and table 9.1) speak for great geographic variability in the mixing processes. (This is apart from the *local* patchiness in microstructure even in regions of strong mixing.) Three water masses intermingle off Cabo San Lucas: the saline outflow from the Gulf of California, the relatively fresh waters being brought in from the northwest by the California Current, and Equatorial Water of intermediate salinity from the eastern tropical Pacific. MR6 remains in Equatorial Water. MR7 is from a shallower drop taken the next day within a few kilometers of MR6. Here we see the intermingling of the three water masses, each jostling for a level appropriate to its density.

Temperature inversions (negative dT/dz) are generally balanced by positive salinity gradients, so that the density increases with depth, and N^2 is positive. The temperature inversions have typical vertical scales of 5 m, with a step structure (e.g., just beneath feature D) attributed to the *diffusive* regime of double diffusion. The underside of temperature inversions (just above E) is often characterized by strong salinity inversions (positive dS/dz), and by prominent microstructure attributed to the *fingering* regime of double diffusion. Double-diffusive processes can be very important locally; they are discussed by J. S. Turner in Chapter 8.

Occasional density inversions (such as at 13 m depth in MR7) are accompanied by intense microstructure. These inversions are very local, and they disappear in a plot of 3-m averages. We are tempted to attribute the density inversions and associated intermittent microstructure to internal wave breaking.

MR7 is a good example of *intrusive* fine structure. Stommel and Fedorov (1967) gave the first discussion of such features based on their measurements near Timor and Mindanao. At the bottom of a well-mixed layer they found a pronounced temperature inversion (balanced by high salinity) that could be traced for 200 km! Evidently the warm saline water was formed 1 or 2 months earlier over the Australian continental shelf at a distance of 500 km, sliding down along an isopycnal surface. The thickness of the inversion layer varied from 20 to 40 m. Beneath the inversion layer, a number of warm, saline lamina of typically 5-m thickness could be traced over 5 km. All these features are associated with horizontal pressure gradients that must be geostrophically balanced. The authors made some calculations of the rate of lamina spreading associated with frictional dissipation in Ekman spirals above and beneath the lamina boundaries. Once the lamina are thinner than 1 m, they are swiftly conducted away. I refer the reader to Stommel and Fedorov's stimulating discussion.

Table 9.1 summarizes some statistical parameters. For comparison we have included MSR4 from the midgyre of the central North Pacific (Gregg, Cox, and Hacker, 1973). The three stations MR7, MR6, and MSR4 characterize strongly intrusive, weakly intrusive

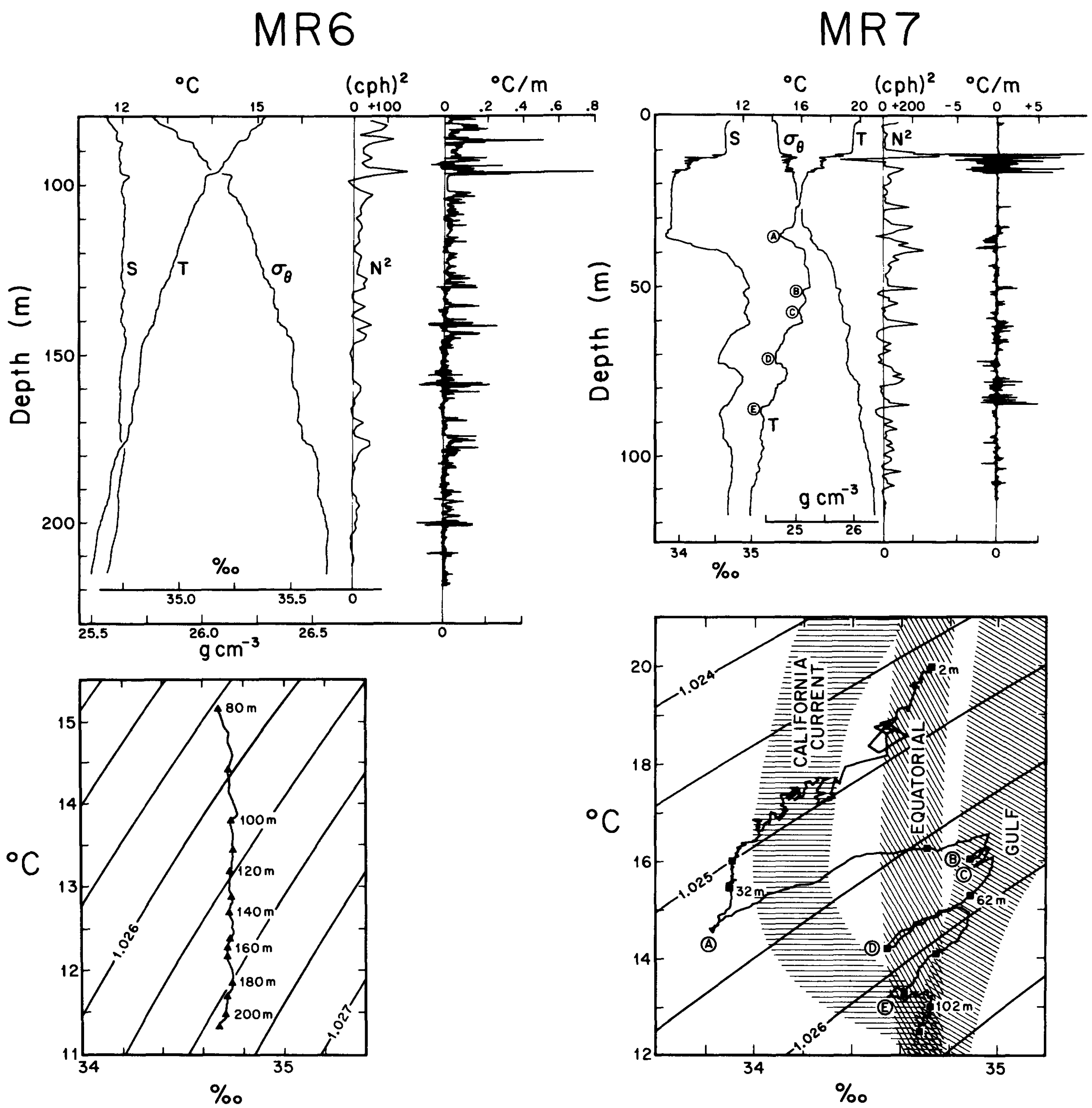


Figure 9.18 The water structure at two stations 60 km southwest of Cabo San Lucas (the southern tip of Baja California), and the associated *T-S* diagrams (from Gregg, 1975). The measurements have been processed to give the fine structure of *S*, *T*, σ_θ, *N* and the microstructure of *dT*/*dz*. Note differences in scale. The *cuspy T-S* diagram for MR7 is an indication of intrusive fine structure.

Table 9.1 Variances and Spectra of Vertical Gradients in the Ocean Fine Structure and Microstructure at Two Stations off Cabo San Lucas (MR7 and MR6) and in the Mid-Gyre of the Central North Pacific (MSR4)

	MR7			MR6			MSR4[a]		
	Strongly intrusive						Nonintrusive		
$\langle (\partial_z T - \partial_z \bar{T})^2 \rangle$ in $\left(\frac{°\mathrm{C}}{\mathrm{m}}\right)^2$		26			9×10^{-3}			1×10^{-3}	
$\langle \partial_z T \rangle^2$ in $\left(\frac{°\mathrm{C}}{\mathrm{m}}\right)^2$		4×10^{-3}			8×10^{-4}			6×10^{-4}	
Thermal diffusivity in $\mathrm{cm^2\,s^{-1}}$		8[b]			0.015			0.002	
Cycles per meter	0.1	1	10	0.1	1	10	0.1	1	10
Spectrum of									
$\partial_z T$ in $\left(\frac{°\mathrm{C}}{\mathrm{m}}\right)^2\Big/\mathrm{cpm}$	2×10^{-1}	1×10^{-2}	1×10^{-2}	2×10^{-3}	2×10^{-4}	7×10^{-5}	7×10^{-4}	1×10^{-4}	3×10^{-5}
$\partial_z S$ in $\left(\frac{‰}{\mathrm{m}}\right)^2\Big/\mathrm{cpm}$	2×10^{-2}	1×10^{-3}	8×10^{-4}	5×10^{-5}	1×10^{-5}	?	4×10^{-6}	8×10^{-6}	2×10^{-5}
$\partial_z \rho$ in $\left(\frac{\mathrm{g\,cm^{-3}}}{\mathrm{m}}\right)^2\Big/\mathrm{cpm}$	2×10^{-9}	7×10^{-10}	7×10^{-10}	2×10^{-10}	2×10^{-11}	?	2×10^{-11}	1×10^{-11}	9×10^{-12}
Spectrum of									
$a\,\partial_z T$ in $\left(\frac{\mathrm{g\,cm^{-3}}}{\mathrm{m}}\right)^2\Big/\mathrm{cpm}$	2×10^{-8}	1×10^{-9}	1×10^{-9}	2×10^{-10}	2×10^{-11}	?	3×10^{-11}	4×10^{-12}	1×10^{-12}
$b\,\partial_z S$ in $\left(\frac{\mathrm{g\,cm^{-3}}}{\mathrm{m}}\right)^2\Big/\mathrm{cpm}$	1×10^{-8}	0.6×10^{-9}	0.5×10^{-9}	0.3×10^{-10}	0.6×10^{-11}	?	0.3×10^{-11}	5×10^{-12}	13×10^{-12}
$\partial_z \rho$ in $\left(\frac{\mathrm{g\,cm^{-3}}}{\mathrm{m}}\right)^2\Big/\mathrm{cpm}$	0.2×10^{-8}	0.7×10^{-9}	0.7×10^{-9}	2×10^{-10}	2×10^{-11}	?	2×10^{-11}	10×10^{-12}	9×10^{-12}

a. MSR4 is not necessarily representative for the mid-gyre; subsequent cruises have given larger mean-square gradients.

b. The vertical heat flux for MR7 can probably not be modeled by an eddy coefficient (Gregg, 1975).

and nonintrusive situations, respectively. The conclusions are: (1) The ratio of the mean-square gradient to the mean gradient squared (the "Cox number") for temperature is highly variable, from 5000 at MR7 to 2 in the mid-gyre. Under certain assumptions (Osborn and Cox, 1972), the eddy diffusivity is the molecular diffusivity times this ratio, giving values all the way from 8 to 0.002 $\mathrm{cm^2\,s^{-1}}$ (but see the footnote to table 9.1). The canonical value of 1 $\mathrm{cm^2\,s^{-1}}$ [for which I am partly responsible (Munk, 1966)] is of no use locally. (2) Spectral levels in vertical gradients diminish with increasing vertical wavenumber up to 1 cpm, and then level off. (3) The relative contributions to the density-gradient spectrum has been estimated from $\partial_z \rho = -a\,\partial_z T + b\,\partial_z S$, with $a = 1.7 \times 10^{-4}\ \mathrm{g\,cm^{-3}\,(°C)^{-1}}$, $b = 8 \times 10^{-4}\ \mathrm{g\,cm^{-3}\,(‰)^{-1}}$. For MR7 at 0.1 cpm, the measured density gradient is much smaller than that inferred from either temperature alone or salinity alone. This is consistent with the near cancellation between temperature and salinity for intrusive features. (4) At higher wavenumbers for MR7, and at all wavenumbers of MR6 and MSR4, the density gradient spectrum is of the same order as that inferred from temperature or salinity alone, thus implying the dominance of internal waves.

Probability densities of the temperature gradients are highly non-Gaussian with an enormous flatness factor (138 for MR7, 55 for MR6) attesting to the patchiness (Gregg, 1975). The construction of meaningful ensemble averages in a highly intermittent environment (space and time) is an important task for the future.

To return now to internal waves, we can distinguish between two quite different effects on vertical profiles: (1) an (irreversible) microstructure and fine structure associated with intermittent internal wave breaking, and (2) a (reversible) fine structure due to the vertical straining of an otherwise smooth profile by internal waves of short vertical wavelength. The reversible contribution to fine structure by internal waves was first noticed by Lazier (1973b) and Garrett (1973). How are we to distinguish it from diffusive (and other irreversible) fine structure?

Let δT and δS designate departures in (potential) temperature and salinity from some long-time or long-distance averages $T_0(z)$, $S_0(z)$ at the same depth. Then

$$\delta\rho = -a\,\delta T + b\,\delta S$$

is the associated density departure, with $a(T, S, p)$ and $b(T, S, p)$ designating the (positive) coefficients of thermal expansion and haline contraction. Take first the case of an intrusion only (figure 9.19). If it is totally compensated,

$$\delta\rho = 0,$$

and if it is not totally compensated, it soon will be (in a time of order N^{-1}). It follows that any vertical displacement of the isopycnals is not intrusive but due to a vertical displacement of the water, which we associate with internal waves (see next section). The vertical displacement ζ can be found from conservation of potential density:

$$\rho(z + \zeta) = \rho_0(z).$$

For the case of internal waves only, conservation of potential temperature and salinity give

$$T(z + \zeta) = T_0(z), \qquad S(z + \zeta) = S_0(z),$$

and the ζ-values from the three preceding equations should be the same:

$$\zeta_\rho = \zeta_T = \zeta_S. \tag{9.17}$$

Then in general, ζ_ρ gives the vertical displacement by internal waves, and $\zeta_T - \zeta_\rho$ and $\zeta_S - \zeta_\rho$ are measures of intrusive activity.

Figure 9.20 shows the situation in (T, S)-space. In the combined case, a projection parallel to the isopycnals can separate the two effects. For constant a and b, it is convenient to introduce a family of lines that are orthogonal to the lines of equal potential density (Veronis, 1972). They are here designated by π, for "spiciness" (hot and salty[10]), and they give a measure of the strength of the intrusion. The construction in (ρ, π)-space has some convenient properties. If the x- and y-axes are scaled in equivalent density units, bS and aT, then

$$\delta\rho = -aT + bS, \qquad \delta\pi = aT + bS. \tag{9.18}$$

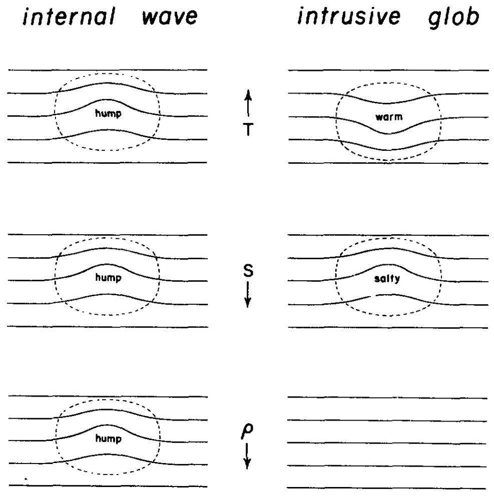


Figure 9.19 Contours of potential temperature, salinity and potential density in a vertical section (x, z) for an internal wave hump and a compensated warm and salty intrusive glob.

Figure 9.21 shows plots of the inferred vertical displacements in an area 200 miles southwest of San Diego (Johnson, Cox, and Gallagher, 1978). ζ_T, ζ_S, and ζ_ρ should all be alike for the case of a fine structure due to internal waves only [equation (9.17)], and this turns out to be the case down to a depth of 225 m. There is a broad intrusion between 225 and 260 m, and a narrow intrusion at 275 m. From a spectral analysis it was found that internal waves dominated the fine structure for all vertical scales that could be resolved, that is, down to 5 m.

The displacement spectrum in vertical wavenumber m steepens from approximately m^{-2} for $m < m_u$ to m^{-3} for $m > m_u$, with m_u near 0.6m^{-1} (~0.1 cpm). This kink appears to be a common feature in temperature spectra (Gregg, 1977; Hayes, 1978), and is most clearly portrayed in the temperature-gradient spectra (figure 9.22). A similar steepening is found in the spectrum of currents and current shear, but at a somewhat lower vertical wavenumber (Hogg, Katz, and Sanford, 1978).

A free-fall instrument called the "camel" for measuring the velocity microstructure has been developed by Osborn (1974; see chapter 14). Figure 9.23 presents measurements in the Atlantic Equatorial Undercurrent during the GATE experiment (Crawford and Osborn, 1980). The most intense microstructure of temperature and current was found above the velocity core. The microstructure in the core was weak and intermittent. Moderately intensive microstructure was found below the core, near the base of the thermocline. This is shown in great detail in figure 9.24. As an example, between 81 and 82 m there is an active temperature microstructure with positive and negative $\partial_z T$, accompanied by an active velocity microstructure. Similar evidence is found in horizontal tows, as for example in the upper part of figure 9.25 (Gibson, Schedvin, and Washburn, personal communication; see Gibson, 1980). The important conclusion is that velocity microstructure and small-scale temperature inversions must be closely linked, for one is not found without the other.

Occasionally one encounters patches of temperature microstructure without velocity microstructure. The inference is that these patches are the remains of a mixing event for which the velocity microstructure has decayed (fossil turbulence). Examples are found in the vertical profiles (figure 9.24 between 69 and 70 m depth), and in the horizontal tows (figure 9.25, bottom). But for the vertical profiles the temperature microstructure is here limited to only positive $\partial_z T$; the authors suggest that this might be a peculiarity of the core (velocity and salt) of the undercurrent.

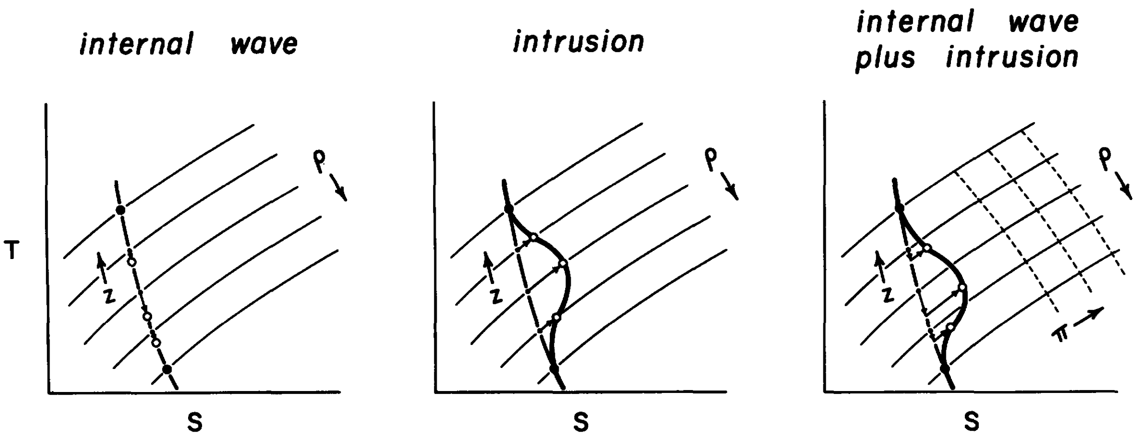


Figure 9.20 T-S relations for an internal wave hump and a compensated intrusive glob. The dots (●) correspond to the undisturbed positions of the five contours in figure 9.19. The open circles (○) give the positions through the center of the disturbance. The "isospiceness" lines (constant π) are orthogonal to the isopycnals (constant ρ).

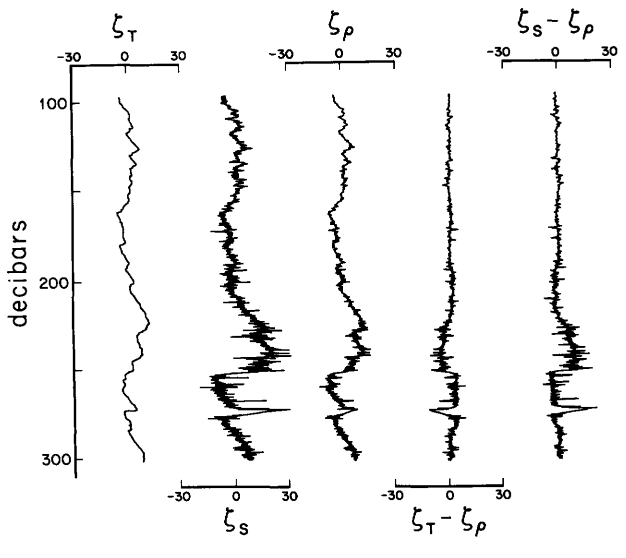


Figure 9.21 Displacement profiles in meters as inferred from the temperature, salinity, and density profiles. These should be alike for internal wave produced fine structure. The mean T and S gradients were of opposite sign (as in figure 9.19), hence the opposite signs of $\zeta_T - \zeta_\rho$ and $\zeta_S - \zeta_\rho$. (Johnson, Cox, and Gallagher, 1978.)

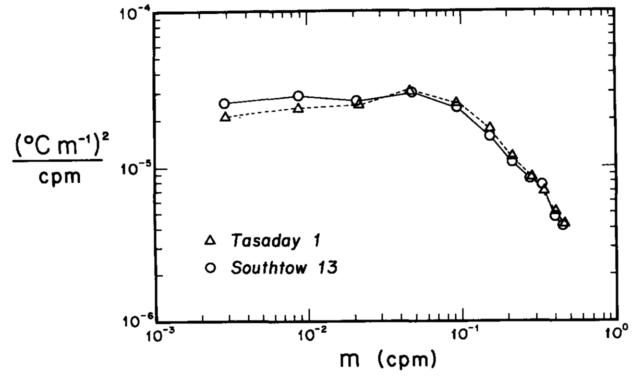


Figure 9.22 Temperature gradient spectra for two stations in the North Pacific. (Gregg, 1977.)

Crawford and Osborn have calculated the dissipation rates. Typical values just beneath the core of the undercurrent range from $\epsilon = 10^{-4}$ to 10^{-3} cm^2 s^{-3} (10^{-5} to 10^{-4} W m^{-3}). (Measurements away from the equator fall within the same limits.) But Belyaev, Lubimtzev, and Ozmidov (1975) obtain dissipation rates in the area of the undercurrent from horizontal tows that are higher by two orders of magnitude.[11]

The question of the relative magnitudes of vertical and horizontal scales has been examined by Hacker (1973) from a comparison of wing tip and nose temperatures of a rotating free-fall instrument. The horizontal separation is 1.7 m. The two records are coherent for vertical wavelengths down to 1 m. At smaller wavelengths the analysis is made difficult by the random tilts (5° rms) associated with internal waves. By selectively analyzing depth ranges of small tilts, Elliott and Oakey (1975) found coherence over a horizontal spacing of 0.5-m down to 10-cm vertical wavelengths. The conclusion is that anisotropy extends beyond the fine structure into the microstructure, perhaps as far as the dissipation scale ($\sim$1 cm).

The picture that emerges is one of a fine structure that is usually dominated by internal wave straining and is fairly uniform, in contrast to a microstructure that is extremely patchy and variable even in the mean. Patches of temperature microstructure without velocity microstructure ("fossil turbulence") evidently mark the demise of internal waves that had previously broken.

9.9 An Inconclusive Discussion

Is there a connection between internal wave activity, dissipation, and buoyancy flux? What is the explanation for the seeming steadiness of the internal wave field? Having gone this far, I cannot refrain from continuing with some speculation. The reader is encouraged to go no further (if he has gotten this far).

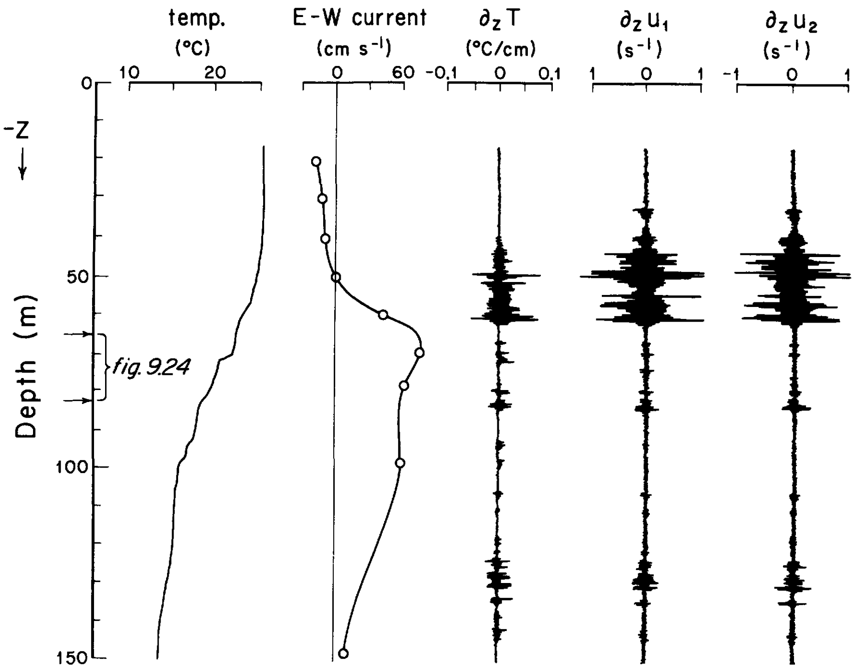


Figure 9.23 Temperature and velocity microstructure in the Atlantic Equatorial Undercurrent at 0°18′S, 28°01′W. (Crawford and Osborn, 1980.) The large scale-current profile was measured by J. Bruce. The region between 65 and 82 m is shown on an enlarged scale in figure 9.24.

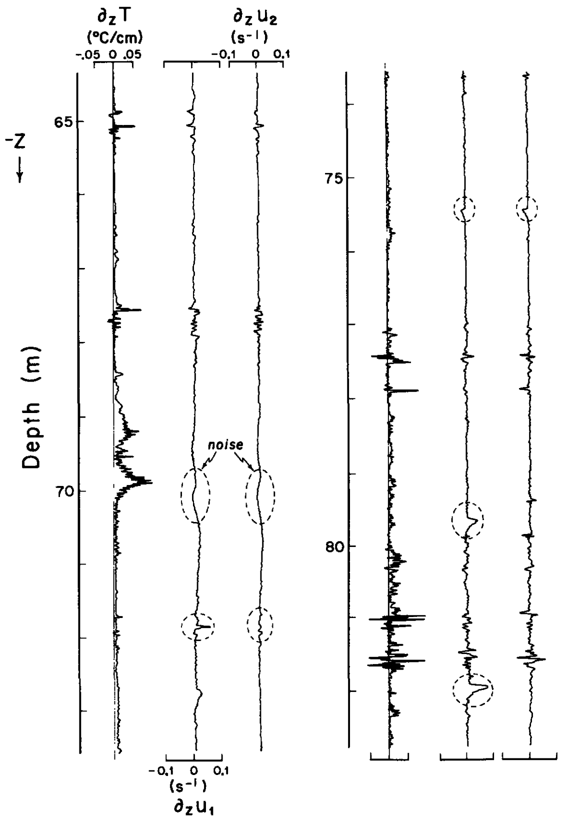


Figure 9.24 An enlarged section of the microstructure profile shown in figure 9.23. The encircled features have been attributed to various sources of instrumental noise.

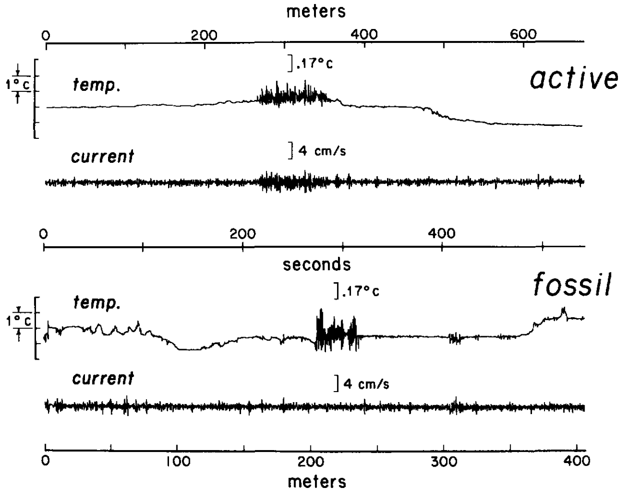


Figure 9.25 Active and fossil turbulence from towed body measurements during the mixed-layer experiment (MILE) in September 1977 near ocean station PAPA. The body was towed in the seasonal thermocline at a depth of 33 m; note the horizontal temperature change by about 0.3°C across the patches. The 1°C scale (left) refers to frequencies $f < 1$ H_z; the 0.17°C scale is for 1-12 H_z.

(I am indebted to C. Gibson, J. Schedvin, and L. Washburn for permission to show these measurements. See also Gibson, 1980.)

We shall need some quantitative guidance for internal wave intensities. For that purpose I shall use the model spectrum of Garrett and Munk (1972b, 1975, 1979), with slight modifications. The model fails near the ocean boundaries[12] (Pinkel, 1975). The spectrum was developed on the basis of rank empiricism, with no trace of underlying theory. But it has since gained some respectability by the theoretical findings of Watson and collaborators that the shape of the GM spectrum is stable to nonlinear interactions, except for the lowest modes and near-inertial frequencies [Meiss, Pomphrey, and Watson (1979); Pomphrey, Meiss, and Watson (1980); see also McComas (1977)].

9.9.1 Model Spectrum GM79

The internal wave energy is assumed to be equally distributed in all horizontal directions, so that only a single horizontal wavenumber, $k = (k_1^2 + k_2^2)^{1/2}$, is used. Upward and downward energy flux are taken as equal. The spectra of vertical displacement, horizontal velocity, and energy per unit mass are[13]

$$F_\zeta(\omega,j) = b^2 N_0 N^{-1}(\omega^2 - f^2)\omega^{-2}E(\omega,j), \tag{9.19}$$

$$F_u(\omega,j) = F_{u_1} + F_{u_2} = b^2 N_0 N(\omega^2 + f^2)\omega^{-2}E(\omega,j), \tag{9.20}$$

$$F_e(\omega,j) = \tfrac{1}{2}(F_u + N^2 F_\zeta) = b^2 N_0 N E(\omega,j), \tag{9.21}$$

where j is the vertical mode number, $b \approx 1.3$ km the e-folding scale of $N(z)$, with $N_0 \approx 5.2 \times 10^{-3}\,\mathrm{s}^{-1}$ (3 cph) the surface-extrapolated buoyancy frequency and $f = 7.3 \times 10^{-5}\,\mathrm{s}^{-1}$ the Coriolis frequency at lat. 30°. We can ignore F_w compared to F_u. At high frequencies, $\omega \gg f$, kinetic and potential energy densities are equal: $\tfrac{1}{2}F_u = \tfrac{1}{2}N^2 F_\zeta$. $E(\omega,j)$ is a dimensionless energy density that is factored as follows:

$$E(\omega,j) = B(\omega)\cdot H(j)\cdot E,$$

$$B(\omega) = 2\pi^{-1} f\omega^{-1}(\omega^2 - f^2)^{-1/2}, \qquad \int_f^{N(z)} B(\omega)\,d\omega = 1,$$

$$H(j) = \frac{(j^2 + j_*^2)^{-1}}{\sum_1^\infty (j^2 + j_*^2)^{-1}}, \qquad \sum_{j=1}^{\infty} H(j) = 1.$$

The factor $(\omega^2 - f^2)^{-1/2}$ in the expression for $B(\omega)$ is a crude attempt to allow for the peak at the inertial turning frequency (see figure 9.11); $j_* = 3$ is a mode scale number, and E is the internal wave "energy parameter." We set

$$E = 6.3 \times 10^{-5} \quad \text{(dimensionless)}. \tag{9.22}$$

There is a surprising universality[14] to the value of E (mostly within a factor of two).

The transfer into (ω,k)- or (ω,m)-space is accomplished by setting $F(\omega,j)\,\delta j = F(\omega,k)\,dk = F(\omega,m)\,dm$, with

$$m = k\left(\frac{N^2 - \omega^2}{\omega^2 - f^2}\right)^{1/2} = \pi b^{-1}\left(\frac{N^2 - \omega^2}{N_0^2 - \omega^2}\right)^{1/2} j \tag{9.23a}$$

for a slowly varying $N(z)$, in accord with the WKB approximation [equations (9.4) and (9.7)]. For most purposes we can ignore the situation near the buoyancy turning frequency,[15] so that

$$m \approx kN(\omega^2 - f^2)^{-1/2} \approx \pi b^{-1}(N/N_0)j. \tag{9.23b}$$

For the sake of simplicity, the energy spectrum has been factored into $B(\omega)\cdot H(j)$. But there is evidence from Pinkel (1975) and from the IWEX measurements (Müller, Olbers, and Willebrand, 1978) that there is relatively more energy in the low modes at high frequency, and this could account for the astounding vertical coherences found by Pinkel in the upper 400 m at high frequencies.

I have no doubt that further discrepancies will be found; still, I believe that the model can now give useful quantitative estimates. For example, according to "Fofonoff's rule" (he disclaims ownership), the mean-square current within a 1-cph band centered at 1 cph is $1\,\mathrm{cm^2\,s^{-2}}$; this compares to $F_u(1\ \mathrm{cph}) = 0.8\,\mathrm{cm^2\,s^{-2}\,(cph)^{-1}}$ from (9.20). This agreement is not an accident, of course, the GM model having been based, in part, on the site D measurements (Fofonoff, 1969).

The mean-square quantitites are likewise in accord with the usual experience. From (9.19), (9.20), and (9.21),

$$\begin{aligned} \langle \zeta^2 \rangle &= \int d\omega \sum F_\zeta(\omega,j) \\ &= \tfrac{1}{2} b^2 E N_0 N^{-1} = 53(N/N_0)^{-1}\,\mathrm{m}^2, \\ \langle u^2 \rangle &= \langle u_1^2 \rangle + \langle u_2^2 \rangle = \int d\omega \sum F_u(\omega,j) \\ &= \tfrac{3}{2} b^2 E N_0 N = 44(N/N_0)\,\mathrm{cm^2\,s^{-2}}, \end{aligned} \tag{9.24}$$

$$\begin{aligned} \hat{E}(z) &= \int d\omega \sum F_e(\omega,j) \\ &= b^2 E N_0 N = 30(N/N_0)\,\mathrm{cm^2\,s^{-2}}, \end{aligned}$$

giving 7 m for the rms vertical displacement and $7\,\mathrm{cm\,s^{-1}}$ for the rms current in the upper oceans beneath the mixed layer. The energy can be written alternatively $\hat{E}(z) = \tfrac{1}{2}[\langle u^2 \rangle + N^2\langle \zeta^2 \rangle]$ so that the *total* kinetic energy is three times the total potential energy in the GM model. $\rho\hat{E}(z)$ is the energy per unit volume; the energy per unit area is

$$\begin{aligned} \rho\hat{E} &= \int \rho\hat{E}(z)\,dz = \rho b^2 E N_0 \int N\,dz = \rho b^2 E N_0 \int b\,dN \\ &\approx \rho b^3 N_0^2 E \\ &= 3.8 \times 10^6\,\mathrm{erg\,cm^{-2}} = 3800\,\mathrm{J\,m^{-2}}, \end{aligned}$$

using $b^{-1} = N^{-1}\cdot dN/dz$ as definition for the e-folding scale b.

9.9.2 Universality

It has turned out, quite unexpectedly, that the intensities are remarkably uniform in space and time.

Wunsch (1976) has made a deliberate attempt to find systematic deviations for a variety of deep water locations in the North Atlantic, with the purpose of identifying sources and sinks of internal wave energy. Using the frequency band $\frac{1}{6}$ to $\frac{1}{4}$ cph as a standard, the only clear deviations he could find were associated with topographic features, particularly Muir Seamount, and even these were inconspicuous at short distance. In a further study (Wunsch and Webb, 1979) some evidence is presented for deviations on the equator and in regions of high mean shear.

Figure 9.26 shows a continuing spectral display over an 18-day period, and this is found consistent with a stationary Gaussian process (Cairns and Williams, 1976). [The mean distribution is in accord with the equation (9.19) for $F_\zeta(\omega) = \Sigma F_\zeta(\omega,j)$, setting $\Sigma H(j) = 1$.] These observations were taken during 21 days of mild to moderate winds. Davis (personal communication) has recorded currents in the seasonal thermocline over a 19-day interval with two periods of heavy winds (figure 9.27). The first event is followed in about 2 days by an increase in mean-square currents, the second event in somewhat less time. Energy enhancement is by a factor of three or less. Johnson, Cox, and Gallagher (1978) found a temporarily elevated spectral level on a windy day. Following these events, the intensities rapidly relax to their normal state.

9.9.3 Generation

The observed growth times are consistent with a theory for the generation of internal waves by resonant interaction with surface waves (Brekhovskikh, Goncharov Kurtepov, and Naugol'nykh, 1972; Watson, West, and Cohen, 1976).

But there are other means of generating internal waves. Garrett (1979) has reviewed a variety of contenders, all of which fall (surprisingly) into the right order of magnitude. For reference, he takes $7 \times 10^{-3}\,\mathrm{W\,m^{-2}}$ for the internal wave dissipation (corresponding to a relaxation time of one week). Globally, this amounts to 2 TW (terawatts: tera = 10^{12}). The total loss of energy of the earth-moon system is known from the moon's orbit to be 4 TW, mostly by tidal dissipation in the oceans. It is not impossible that surface tides pump significant amounts of energy into internal wave motions via internal tides (cf. chapter 10). Other contenders are surface forcing by traveling fluctuations of wind stress and buoyancy flux, currents over bottom topography, and extraction from the mean current shear. There is no problem with supplying internal waves with 2 TW of power; the problem is rather to eliminate some of the potential donors.

9.9.4 Instability

We can now derive some numerical estimates for a variety of instability parameters. The spectrum $F_{u'/N}$ of

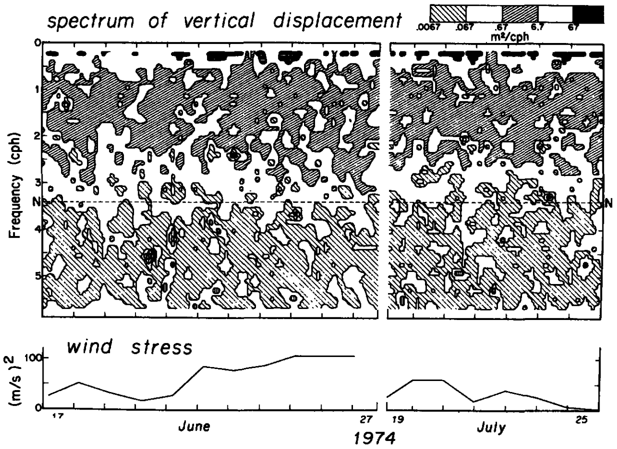


Figure 9.26 Time-frequency display of $F_\zeta(\omega)$ from MISERY 1 and MISERY 3. (Cairns and Williams, 1976.) $\zeta(t)$ is the depth of the 6.60° isotherm (at a mean depth of 350 m) in a location 800 km offshore of San Diego, California, measured with a yo-yoing midwater capsule. The squared wind (bottom) shows light winds at the start and end of the experiments.

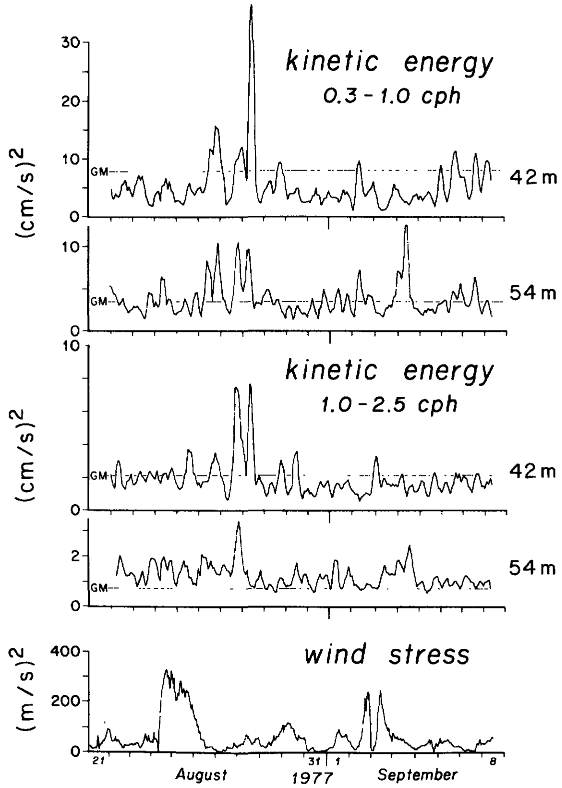


Figure 9.27 Kinetic energy $\langle u^2 \rangle = \langle u_1^2 \rangle + \langle u_2^2 \rangle$ during the mixed-layer experiment (MILE) on station PAPA (50°N, 145°W) at 42 and 54 m depths. The upper two plots refer to a frequency band of 0.3 to 1.0 cph, the next two plots to a band 1.0 to 2.5 cph. The squared wind (bottom) shows two episodes of large wind stress. The GM model levels are indicated [using $N = 0.023, 0.0096\ \mathrm{s}^{-1}$ (13, 5.5 cph) at 42, 54 m], though the model is not really applicable to such shallow depths and sharp N-gradients. (I am indebted to R. Davis for permission to use his measurements.)

the reciprocal Richardson number is defined by

$$Ri^{-1} = \langle (u')^2 \rangle / N^2 = \int d\omega \sum_j F_{u'/N}(\omega, j),$$

where $\langle (u')^2 \rangle = \langle (\partial_z u_1)^2 \rangle + \langle (\partial_z u_2)^2 \rangle = m^2 \langle u^2 \rangle$. But $m = \pi b^{-1}(N/N_0)j$ (9.23b) and so

$$F_{u'/N} = 2\pi E(N/N_0) f(\omega^2 + f^2)\omega^{-3}(\omega^2 - f^2)^{-1/2} j^2 H(j). \qquad (9.25)$$

The principal contribution comes from the inertial frequencies. Performing the ω-integration,

$$Ri^{-1} = \tfrac{3}{2}\pi^2 E(N/N_0) \sum j^2 H(j). \qquad (9.26)$$

We now perform the mode summations (subscript u for upper)

$$\sum_{j=1}^{j_u} j^2 (j^2 + j_*^2)^{-1} \approx j_u,$$

$$\sum_{j=1}^{j_u} (j^2 + j_*^2)^{-1} \approx \tfrac{1}{2} j_*^{-2} (\pi j_* - 1) \equiv J^{-1} = 0.47$$

for $j_u \gg j_* = 3$. The spectrum $F_{u'/N}(j)$ is white (except for the lowest few mode numbers) and Ri^{-1} depends on the choice of the upper cutoff j_u. In terms of the limiting vertical wavenumber $m_u = \pi b^{-1}(N/N_0) j_u$ we have, finally,

$$Ri^{-1} = \tfrac{3}{2}\pi J E m_u b. \qquad (9.27)$$

If we identify m_u with the kink (figure 9.22) at $0.6m^{-1}$ (0.1 cpm), then $Ri^{-1} = 0.52$. If the spectrum is extended beyond the break, with a slope m^{-1} to some new upper limit $m_{uu} = 10 m_u$ (say), then $Ri^{-1} = 0.52(1 + \ln m_{uu}/m_u) = 1.72$. Hogg, Katz, and Sanford (1978) find Ri^{-1} near 0.5 in the open ocean near Bermuda, including m up to 0.2 cpm. An interesting scatter plot has been produced by Eriksen (1978) and is shown in figure 9.28. The conclusion is that the internal-wave shear field is associated with Richardson numbers of order 1:

$$Ri^{-1} = \text{order}(1). \qquad (9.28)$$

We can proceed in a similar manner with regard to advective instability. The simplest generalization of the earlier discussion (section 9.7) is to derive the spectral decomposition of $\langle u^2/c^2 \rangle$, $c = \omega/k$. (Would $\langle u^2 \rangle / \langle c^2 \rangle$ be better?) The equations of continuity and of dispersion (away from the turning frequencies) can be written

$$k\xi + m\zeta = 0, \quad c = \omega/k = N/m,$$

so that

$$\frac{u}{c} = \frac{mu}{mc} = \frac{u'}{N}.$$

Similarly, for horizontal and vertical strain,

$$\partial\xi/\partial x = -\partial\zeta/\partial z = ik\xi = (k/\omega)u = u/c.$$

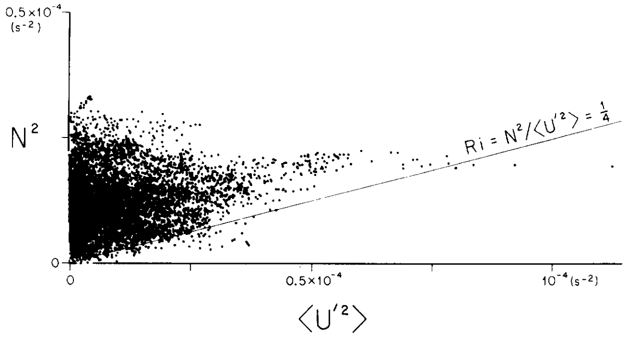


Figure 9.28 A scatter plot of squared shear (over 6.3 m vertical separation) versus N^2 (over 7.1 m) of estimates made every 40 s for 78 hours. (Eriksen, 1978.) Eriksen finds that Ri rarely falls below the critical value $\frac{1}{4}$, and that $\phi = \tan^{-1} Ri$ is uniformly distributed for ϕ greater than $\tan^{-1}(\frac{1}{4})$.

Thus shear, advection, and longitudinal strains all have similar conditions for instability, and we can write

$$\langle \phi^2 \rangle = C E m_u b \qquad (9.29)$$

for any of these, without having to go into gruesome details. [Integrations yield $C = 2.5, 5.8, 3.3, 3.3$ for $u'/2N$, u/c, $\partial_x \xi$, $\partial_z \zeta$, respectively; this is a spectrally weighted version of the Orlanski and Bryan argument [equation (9.16)] that advective instability is the most likely to occur.]

9.9.5 Compliant Wave Cutoff

The instability condition $\langle \phi^2 \rangle = \text{order}(1)$ is an argument for a universal value of the product $E m_u$. To account for a universal E we need some additional condition.

I propose that the upper cutoff m_u is related to the transition at $c_c \approx N/m_c = \text{rms}\, u$ from the intrinsic to the compliant parts of the internal wave spectrum:

$$m_u = C' m_c = C' N / \text{rms}\, u, \qquad (9.30)$$

where C' is a constant of order (1). It stands to reason that the strongly interacting high wavenumbers have a different spectral form from the intrinsic waves. If we identify m_u with the kink (figure 9.22) at 0.6 m^{-1} (0.1 cpm), then for $N = 0.01\ \text{s}^{-1}$ and $\text{rms}\, u = 10\ \text{cm}\,\text{s}^{-1}$ this gives $C' = 6$ (somewhat large for comfort).

There is an equivalent way of postulating the upper cutoff. The ϕ-spectrum is white up to some limit which is the reciprocal of the vertical extent Δ of the smallest ϕ-features. One might suppose this vertical extent to be some given fraction of the rms *amplitude* of the internal waves. (White caps occupy some fraction of the surface wave crests; the distance between crests is not a critical factor.) Using the foregoing numbers, we can write

$$\begin{aligned} \Delta &= m_u^{-1} = 1.37 (C')^{-1} \text{rms}\, \zeta \\ &= 0.23\, \text{rms}\, \zeta = 1.7\ \text{m}. \end{aligned} \qquad (9.31)$$

From the expressions (9.24) for $\langle \zeta^2 \rangle$ and $\langle u^2 \rangle$, either condition (9.30) or condition (9.31) leads to an energy parameter

$$E = [\langle \phi^2 \rangle / CC']^2 N_0/N. \tag{9.32}$$

For $C = 5$, $C' = 6$, $E = 10^{-3} \langle \phi^2 \rangle^2 N_0/N$. Then with $\langle \phi^2 \rangle$ a moderate fraction of its critical value 1, we can recover the numerical value $E = 6 \times 10^{-5}$. This is not to say that E has been calculated from first principles; it is only to say that acceptable values for the various coefficients lead to a small numerical value of the dimensionless energy parameter, as observed.

9.9.6 Dissipation

For small numerical values of the instability parameter $\langle \phi^2 \rangle$ we are in a regime of *sparse* instabilities in space and time (such as incipient whitecaps in light winds). When $\langle \phi^2 \rangle$ is near 1 the probability for instabilities is high. For an a priori estimate of $\langle \phi^2 \rangle$ we require (1) a model to relate $\langle \phi^2 \rangle$ to the internal wave energy dissipation, and (2) an estimate of the rate of dissipation (or generation for a given steady state). This is essentially the procedure followed by Longuet-Higgins (1969a) in his stimulating attempt to interpret the Phillips saturation constant for surface waves.

Perhaps the simplest scheme is to relate the dissipation to the probability for $\phi > 1$.[16] The variance of ζ associated with $\phi > 1$ is $\langle \zeta^2 \rangle \cdot p\ (\phi > 1)$ for uncorrelated ζ and ϕ (as when $\phi = \partial_z \zeta$). Potential energy is proportional to $\langle \zeta^2 \rangle$; accordingly the rate of fractional energy dissipation can be written

$$\frac{1}{E}\frac{dE}{dt} = -\sigma p(\phi > 1), \tag{9.33}$$

where σ^{-1} is the characteristic interval during which the energy associated with $\phi > 1$ is lost to the organized wave field and renewed by generation processes. For a rough estimate (Garrett and Munk, 1972a),

$$\begin{aligned} \sigma^2 &= \pi^{-2} \int_f^N \omega^2 F_\phi(\omega) d\omega \Big/ \int_f^N F_\phi(\omega) d\omega \\ &\approx \pi^{-2} f N \end{aligned} \tag{9.34}$$

for any of the ϕ-spectra [such as (9.25)]. A Gaussian ϕ-distribution $p(\phi) = \pi^{-1/2} \beta \exp(-\beta \phi^2)$ leads to

$$p\ (\phi > 1) \approx \pi^{-1/2} \beta^{-1/2} \exp(-\beta), \qquad \beta \equiv \frac{1}{2\langle \phi^2 \rangle}, \tag{9.35}$$

provided β is large. In the upper ocean $\sigma = 20$ per day, and $\beta = 2.1, 3.7$, for relaxation times of 1 day, 1 week, respectively. The foregoing numerical values are not important; what is significant is that a tenfold increase in the rate of dissipation (and generation) is accompanied by only a threefold increase in β^2 (and hence in wave energy). Thus the energy level stays within rather narrow limits even though generation and dissipation processes may vary widely, particularly for large β. We propose for a "universality hypothesis" that the energy level responds only logarithmically to variable forcing. We shall examine this situation in more detail.

9.9.7 The Energy Balance

The differential equation of wave energy can be written

$$dE/dt = G(t) - D(t),$$

where $G(t)$ and $D(t)$ are the rates of energy generation and dissipation. We use the notation $\tilde{E}$, $\tilde{G}$, $\tilde{D}$ to represent the "normal" state of internal wave statistics. From (9.32) and (9.35)

$$E \sim \beta^{-2}(N_0/N), \qquad D \sim \beta^{-1/2} e^{-\beta},$$

with $\tilde{D} = \tilde{G}$. We define the relaxation time

$$\tilde{t} = \tilde{E}/\tilde{G};$$

accordingly $\tilde{t}^{-1}$ is the initial rate of decay for a wave field in equilibrium with $\tilde{G}$ if the generation is suddenly turned off. The differential energy equation can now be written

$$\begin{aligned} &\frac{d\mathscr{E}}{d\tau} + \mathscr{E}^{1/4} \exp[\tilde{\beta}(1 - \mathscr{E}^{-1/2})] = g(\tau), \\ &\tilde{\beta} = \tilde{E}^2 (N/N_0)^{-1/2}, \end{aligned} \tag{9.36}$$

where

$$\mathscr{E}(\tau) = E/\tilde{E}, \qquad g(\tau) = G/\tilde{G}, \qquad \tau = t/\tilde{t},$$

with

$$\langle \mathscr{E} \rangle = 1, \qquad \langle g \rangle = 1.$$

For large $\mathscr{E}$, the dissipation is $\mathscr{E}^{1/4} \exp \tilde{\beta}$ and thus large; for small $\mathscr{E}$, it is $\mathscr{E}^{1/4} \exp -(\tilde{\beta} \mathscr{E}^{-1/2})$ and thus very small.

The problem is to derive properties of the energy statistics for given generation statistics. (This is related to the fluctuation-dissipation theorem in the study of Brownian motion.) Two special solutions are easily found. For an equilibrium situation, (9.36) with $d\mathscr{E}/d\tau = 0$ gives the values $\mathscr{E}(g;\beta)$ in table 9.2. Departures from the normal state in E are much smaller than those in g, particularly at large β and for small g's.

To obtain some feeling for the nonlinear response time, let $g(\tau)$ go abruptly from 1 to g at time 0, and set $\mathscr{E}(\tau) = 1 + \epsilon(\tau)$, with $\epsilon \ll 1$ (but *not* $\epsilon \tilde{\beta} \ll 1$). Equation (9.36) becomes

$$dz/d\tau' + z^2 = gz, \qquad z = \exp(\tfrac{1}{2}\tilde{\beta}\epsilon), \qquad \tau' = \tfrac{1}{2}\tilde{\beta}\tau,$$

with the solution

$$\tau = \frac{2}{\tilde{\beta} g} \ln \frac{(g-1)z}{g-z}.$$

For the case $g \rightarrow 0$, $\tau' \rightarrow z^{-1} - 1$, and

$$d\epsilon/d\tau = 2\tilde{\beta}^{-1} d \ln z/d\tau \rightarrow -z, \tag{9.37}$$

Table 9.2 Energy Equilibrium $\mathscr{E}(g;\beta)$ from (9.36) for $d\mathscr{E}/d\tau = 0$ (top), Dimensionless Response Time $\tau_{1/2}$ (9.38) for an Abrupt Change in Generation from 1 to g (center), Relaxation Time $\tau_{1/2}$ (9.39) for a Change from g to 1 (bottom)[a]

		g (dB)				
		−10	−5	0	5	10
	Energy levels in dB					
$\tilde{\beta}$	2	−6.0	−3.4	0	4.9	13.0
	5	−3.0	−1.7	0	2.0	4.6
	10	−1.7	−0.9	0	1.0	2.1
	Response time (dimensionless)					
$\tilde{\beta}$	2	2.75	1.41	0.69	0.32	0.14
	5	1.10	0.56	0.28	0.13	0.06
	10	0.55	0.28	0.14	0.06	0.03
	Relaxation time (dimensionless)					
$\tilde{\beta}$	2	1.43	1.02	0.69	0.45	0.27
	5	0.57	0.41	0.28	0.18	0.11
	10	0.29	0.20	0.14	0.09	0.05

a. Generation g and energy $\mathscr{E}$ are in decibels relative to normal levels.

in accord with $d\epsilon/d\tau = -1$ for the initial rate of decay. As τ goes from 0 to ∞, z goes from 1 to g, and ϵ from 0 to $2\tilde{\beta}^{-1}\ln g$. Half energy response is for $\epsilon = \tilde{\beta}^{-1}\ln g$, $z = g^{1/2}$, and

$$\tau_{1/2} = 2\tilde{\beta}^{-1}g^{-1}\ln(1 + g^{1/2}). \tag{9.38}$$

For energy *relaxation,* going abruptly from g to 1 at time 0, the solution is

$$\tau = 2\tilde{\beta}^{-1}\ln\frac{(g-1)z}{g(z-1)}, \qquad \tau_{1/2} = 2\tilde{\beta}^{-1}\ln(1 + g^{-1/2}). \tag{9.39}$$

Half-times $\tau_{1/2}$ are given in table 9.2. For a linear system these would all be the same. Here the times are shorter for internal wave storms ($g > 1$) and longer for calms ($g < 1$), and this variation is more pronounced in response to a change in generation from a normal to a perturbed level than for relaxation back to normal generation. (The assumption of weak nonlinearity for computing τ is violated in the columns for $g = \pm 10$ dB.)

The imperceptible decay of internal wave intensities during relatively low winds (figure 9.26) is consistent with half the normal energy, and the rapid decay following a blow (figure 9.27) requires perhaps three times normal energy. If we can count on a storm or some other generation event to "top up" the internal wave energy once every hundred days, then we may expect the wave energies to remain generally within a factor of two.

I have paid no attention to depth dependence. The dissipation can be written

$$\hat{D} = \sigma(\pi\beta)^{-1/2}\exp(-\beta)\hat{E},$$

where $\rho\hat{D}(z)$, $\rho\hat{E}(z)$ are the dissipation and energy per unit volume, respectively. The dependence on depth is through

$$\sigma \sim n, \qquad \beta \sim n^{-1/2}, \qquad \hat{E} \sim n,$$

with $n(z) = N(z)/N_0$; hence

$$\hat{D} = \hat{D}_0\beta^{-9/2}\exp(\beta_0 - \beta), \qquad \hat{E} = \hat{E}_0\beta^{-2}.$$

Writing $dz = b\,dn/n = -2b\,d\beta/\beta$,

$$\int \hat{D}\,dz = 2b\hat{D}_0\int \beta^{-11/2}\exp(\beta_0 - \beta)\,d\beta,$$

$$\int \hat{E}\,dz = 2b\hat{E}_0\int \beta^{-3}\,d\beta.$$

Integrating from the surface ($\beta = \beta_0$) to the bottom ($\beta \approx \infty$), the integrals for large β_0 (as previously assumed) give

$$\int \hat{D}\,dz = 2b\hat{D}_0{\beta_0}^{-11/2}, \qquad \int \hat{E}\,dz = b\hat{E}_0\beta_0^{-2}$$

with an "integral relaxation time"

$$\tfrac{1}{2}\beta_0^{7/2}\hat{E}_0/\hat{D}_0 = \tfrac{1}{2}\pi^{1/2}\beta_0^4 e^{\beta_0}\sigma_0^{-1}. \tag{9.40}$$

Suppose the generation takes place in the upper few hundred meters, so that the dissipation in the interior ocean is compensated by downward radiation of internal wave energy. From a rotary decomposition of current profiles, Leaman (1976) estimates a downward flux of order 10^{-4} W m^{-2}. We compare this with the integrated dissipation beneath a scale depth $z = -b = -1.3$ km (where $n = e^{-1}$ and $\beta = \beta_0\sqrt{e}$):

$$\rho\int_{\beta_0\sqrt{e}}^{\infty} \hat{D}(z)\,dz = 2\pi^{-1/2}e^{-11/4}\beta_0^{-6}e^{-\beta_0}\sigma_0 N_0^2\rho b^3\tilde{E},$$

where $\tilde{E}$ is the "normal" dimensionless energy parameter. The result is $\beta_0 = 2.1$. The corresponding integral relaxation time [equation (9.40)] is 7 days. The normal surface relaxation time is

$$\tilde{t}_0 = \hat{E}_0/\hat{D}_0 = (\pi\beta_0)^{1/2}e^{\beta_0}\sigma_0^{-1} = 1.1 \text{ days}.$$

The dimensionless times in table 9.2 can be interpreted as shallow response times in days.

The strong dependence of dissipation on depth is an inherent feature of the proposed phenomenology.

9.9.8 Mixing

The balance between production and dissipation of turbulent energy can be represented by

$$\mu = \epsilon = \epsilon_p + \epsilon_k. \tag{9.41}$$

$\rho\epsilon_p$ is the rate of production of potential energy, e.g., the *buoyancy flux* $g\langle w'\rho'\rangle$, with the primes designating the fluctuating components; $\rho\epsilon_k$ is the dissipation of kinetic energy into heat. The fraction of work going into potential energy is the flux Richardson number Rf:

$$Rf = \epsilon_p/\mu = \epsilon_p/\epsilon. \tag{9.42}$$

$Rf < 1$ in order for $\epsilon_k > 0$. We have previously discussed the evidence that something less than $\frac{1}{4}$ of the kinetic energy dissipated appears as potential energy. Taking a typical value $\epsilon_p/\epsilon_k = \frac{1}{5}$ (Thorpe, 1973b, p. 749) gives $Rf = \frac{1}{6}$. The eddy flux of density can be written in terms of the eddy diffusivity A: $\langle w'\rho'\rangle = A \cdot d\rho/dz = AN^2(g/\rho)^{-1} = \rho\epsilon_p/g$, and so

$$A = \frac{\epsilon_p}{N^2} = \frac{\epsilon}{N^2} Rf = \frac{\epsilon_k}{N^2} \frac{Rf}{1 - Rf}. \tag{9.43}$$

Our procedure is to estimate ϵ_p from internal wave breaking, and to compute A and ϵ from (9.43), using $Rf = \frac{1}{6}$. There are, of course, other sources of turbulence; Osborn (1980) stresses the work done by the ambient turbulence in a mean shear: $\mu = \langle u'w'\rangle\, du/dz$. His procedure is to estimate ϵ from the measured mean-square shear, and to compute A from (9.43).

The random superposition of internal waves leads to the intermittent occurrence of "traumata"[17] associated with $\phi > 1$. The traumata are the locus of incipient turbulence, and quickly spread to some thickness Δ within which the average ϕ is reduced from 1 to about 0.9 (Thorpe, 1973b). Subsequently the patch continues to grow to some maximum thickness $\Delta_{\max}$ at the time the surrounding ϕ is largest, always keeping ϕ within the patch to 0.9.

The change of potential energy per unit surface area associated with perfect mixing over a depth Δ in a density gradient $d\rho/dz$ is $(1/12)g(d\rho/dz)\Delta^3 = (1/12)\rho N^2\Delta^3$. (Imperfect mixing just reduces the factor 1/12.) The average change of potential energy per unit time per unit volume is then

$$\rho\epsilon_p = \frac{1}{12}\rho N^2 \Delta^3 \nu,$$

where ν is the number of traumata per unit (t, z)-space. $\rho\epsilon_p$ equals the buoyancy flux $\rho N^2 A$ by definition of the eddy diffusivity A; hence

$$A = \frac{1}{12}\Delta^3\nu.$$

I have previously identified Δ with m_u^{-1} [equation (9.31)]. For ν we write

$$\nu = m'\sigma p(\phi > 1), \qquad m' = 0.2m_u, \qquad \sigma \approx \pi^{-1}(fN)^{1/2},$$

where m' is the rms spacial frequency derived from an equation analogous to (9.34). Putting all this together,

$$A \approx 10^{-2}\Delta^2/T, \qquad T \equiv \sigma^{-1}\beta^{1/2}e^{\beta}, \tag{9.44}$$

where T is the expected time interval between events over a distance $1/m'$. This is of similar form as the result of Stommel and Fedorov (1967), as is inevitable for what is, after all, a mixing-length theory.

The principal conclusion is a strong dependence of A on depth, and on any departures from normal generation. The numerical value for the normal state is $\tilde{A} = 10^{-2}\ \mathrm{cm^2\,s^{-1}}$, much lower than the global 1 $\mathrm{cm^2\,s^{-1}}$.

9.9.9 Saturation Spectra

There is an essential distinction between the usual formulation of turbulence and the saturation processes as here envisioned. We consider the regime of *sparse* instabilities in space and time (such as incipient whitecaps in light winds). Then the ϕ-field consists of scattered and uncorrelated spikes, and the ϕ-spectrum is accordingly white up to some limit that is the reciprocal of the vertical extent of the spikes. The dissipation is localized in physical space, and therefore broadly distributed in wavenumber space. In the usual turbulent situation, the dissipation is confined to a narrow (dissipation) region in wavenumber space, and spread in physical space.

A white spectrum in any of the ϕ-spectra, whether shear, advection or strain, implies an m^{-2} energy spectrum. The energy spectrum steepens (perhaps to m^{-3}) in the transition from the intrinsic to the compliant waves. Presumably the m^{-3} energy spectrum extends to the Ozmidov (or Richardson or Monin-Obukov) scale $m_0 = (N^3/\epsilon)^{1/2} \approx 4\ \mathrm{m}^{-1}$ (about 0.6 cpm), which is conveniently close to the definition of the microstructure boundary, then flattens out to the Kolmogorov dissipation scale $m_k = (\epsilon/\nu^3)^{1/4} \approx 3\ \mathrm{cm}^{-1}$ (0.5 cpcm), and finally cuts off exponentially (Gregg, Cox, and Hacker, 1973, figure 11). But such a description of "in-the-mean" scales may not be appropriate to a patchy environment, and is anyway beyond the scope of this survey.

9.10 Conclusion

I shall end as I started: the connection between internal waves and small scale processes—that is where the key is. I feel that we are close to having these pieces fall into place, and I am uncomfortable with having attempted a survey at this time.

Notes

1. They were found at Loch Ness at about the same time (Watson, 1904; Wedderburn, 1907).

2. The temperature measurements were made using a submarine cable from a recording Wheatstone bridge on the shore at Bermuda to two resistance thermometers offshore: both lay on the bottom, one at a depth of 50 m and the other at 500 m (Haurwitz, Stommel, and Munk, 1959).

3. Salinity is not directly measured, but has to be inferred from conductivity or sound speed, which are primarily responsive to temperature. This large temperature "correction" has been a source of some difficulty.

4. The notation $U' = \partial_z U$ refers to the ambient shear, and $u' = \partial_z u$ to the shear induced by the orbital wave motion. The distinction is not always so clear.

5. In general, *forward* flow refers to positive $\mathbf{c}\cdot(\mathbf{g} \times \operatorname{curl}\mathbf{U})$, where **g** and **U** are the vectors of gravity and ambient velocity, respectively.

6. I am indebted to D. Broutman for very considerable improvements of this section, and for the preparation of figure 9.14.

7. The linearized calculations indicated that they are in fact not small.

8. Backward breaking occurs for negative shear; this can be visualized by turning the figure upside down.

9. Frankignoul's (1972) treatment of *finite* amplitude waves [his equations (24) and (28)] lead to precisely the same result.

10. Garrett points out that a lot of laboratory experiments have been sweet-and-sour rather than spicy.

11. It has been suggested that vibration and temperature contamination contributes to the high values from the towed devices; it has also been suggested that the dropped devices have inadequate dynamic range to measure ϵ in the highly active patches where most of the dissipation takes place.

12. A normal-mode formulation is applicable near the boundaries (Watson, Siegmann, and Jacobson, 1977).

13. Frequencies are in rad s^{-1}, wavenumbers in rad m^{-1}. For comparison with computed spectra we sometimes include (in parenthesis) the values in cycles per hour (cph) and cycles per meter (cpm).

14. We note that $F_e(\omega \gg f) \propto E(\omega, j) \propto \omega^{-2} f E$. There is some evidence that the spectral energy density is independent of latitude (Wunsch and Webb, 1979; Eriksen, 1980), and we should probably replace fE by $N_0 E'$, with $E' = (f_{30^\circ}/N_0)E = 8.8 \times 10^{-7}$ the appropriate f-scaled energy parameter.

15. Desaubies (1973, 1975) explains the observed N-peak in the spectrum and the vertical coherence of vertical displacement.

16. This is related to the "intermittency index" evaluated by Thorpe (1977) from temperature inversions in Loch Ness.

17. "A disorderly state resulting from stress." This descriptive terminology is due to McEwan (1973).