Reprint from Journal of Marine Research, Volume 34, 4, 1976.

Forced baroclinic ocean motions. I. The linear equatorial unbounded case

by Mark A. Cane[1] and E. S. Sarachik[2]

ABSTRACT[3]

A method is developed for calculating the response of an unbounded inviscid ocean to wind stress and thermal forcings. Although emphasis is on equatorial baroclinic motions, the mathematical technique is first illustrated in detail for the motions described by the similar but simpler barotropic vorticity equation. This serves to clarify the significance of the asymptotic approximations made for the baroclinic planetary modes. We describe in detail response to forcings switched on at $t = 0$ and steady thereafter and that are independent of longitude or step functions of longitude. The response to a zonal wind stress grows linearly in time in the vicinity of the equator. The response to a meridional wind stress tends to be less equatorially confined and exhibits secular growth only at a forcing discontinuity, such a discontinuity acting as a line source of vorticity. Unlike a zonal wind stress, a meridional wind stress cannot excite equatorial Kelvin waves. A buoyancy source has less of a tendency to excite gravity waves than a wind stress, though the response to such a source qualitatively resembles that to a zonal wind. In a subsequent paper, the effects of boundaries are treated using the methods discussed in the present work.

1. Introduction

An important consequence of the vanishing of the vertical component of the Coriolis term is that equatorial motions have time scales which are very much shorter than those of mid-latitude motions: the baroclinic time scale is weeks at the equator, compared to years at mid-latitudes. The most impressive instance of this short time scale is the reversal in direction of the Somali Current within a month of the onset of the Southwest Monsoon (Leetmaa, 1973). In general, time-dependent oceanic motions with time scales longer than a few days have received relatively little attention. Equatorial regions are rewarding areas for the study of such time variations because of the rapidity of the ocean's response to atmospheric forcings and the large seasonal component of equatorial wind systems.

1. Goddard Institute for Space Studies, Goddard Space Flight Center, NASA, New York, N.Y., 10025, U.S.A.

2. Center for Earth & Planetary Physics, Harvard University, Cambridge, Mass., 02138, U.S.A.

3. Some of the results reported in this paper were presented at the Woods Hole Summer Geophysical Fluid Dynamics Program, see Cane (1974).

In this paper we are concerned with calculating the linear, inviscid, time-dependent baroclinic response to switched on atmospheric forcings with long time and space scales. We have in mind the monsoonal systems which vary seasonally and whose spatial scales are on the order of 1000 km or more. Since equatorial currents are relatively swift and narrow, nonlinear advective effects will usually be significant. Neverthless, in the real ocean the linear model is useful because it provides insight into the more realistic nonlinear case which will be discussed in future papers of this series (see also Cane, 1975). In particular, it provides information about time and space scales of the oceanic response. Moreover, there are a number of important special situations where the linear theory may be expected to apply directly. For example, it should be applicable to the initial response of the Indian Ocean to the onset of the monsoon.

The equations appropriate to our study are the linear inviscid shallow water equations (Veronis, 1963a, b).

$$\frac{\partial u}{\partial t} - fv + g\frac{\partial h}{\partial x} = F \tag{1a}$$

$$\frac{\partial v}{\partial t} + fu + g\frac{\partial h}{\partial y} = G \tag{1b}$$

$$\frac{\partial h}{\partial t} + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = Q\ , \tag{1c}$$

where $f = \beta y$ on an equatorial β plane.

The symbols on the left-hand side of these equations have their conventional meanings: x is zonal distance, y is meridional distance; u and v are the velocity components in the x and y directions respectively; h is the deviation of the ocean depth from its mean value H; g is the acceleration of gravity; $\beta = 2\Omega/a$, where Ω is the rate of rotation and a the radius of the earth; and subscripts denote differentiation. In the parlance of tidal theory Eqs. (1) describe the horizontal structure of the vertical mode with equivalent depth H. The equivalent depths for the equatorial baroclinic modes are on the order of .5 m or less (Moore and Philander, 1976). For such shallow oceans the equatorial beta plane is an excellent approximation to the more correct spherical geometry in the sense that the solutions to the unforced version of (1) are close to the eigenfunctions on a sphere (Lindzen, 1967; Longuet-Higgins, 1968).

The forcing functions which appear on the right-hand sides of Eqs. (1) are the projections of total atmospheric forcing onto the baroclinic mode of equivalent depth H. F and G are the projections of the zonal and meridional components of the wind stress, respectively, while Q is the projection of a buoyancy source (e.g., a heating function). A method for calculating these projections is given in detail in the Appendix of Lighthill (1969). This method models the wind stress as a body force uniformly distributed over the depth of the surface mixed layer.

There is a natural scaling for the system (1) chosen by the dynamics of the motions. Specifically, no nondimensional numbers will appear on the left-hand side of the nondimensional equatorial form of (1) if the length-scale L and time-scale T are given by

$$L = (gH/\beta^2)^{\frac{1}{4}}$$
$$T = (\beta L)^{-1} = (gH\,\beta^2)^{-\frac{1}{4}} \ . \tag{2}$$

L is referred to as the equatorial radius of deformation. Recall that the usual definition of the Rossby radius of definition $L_R(y)$ for a shallow water system (Rossby, 1938) is

$$L_R(y) = (gH)^{\frac{1}{2}}/f(y) \ ,$$

where $f(y)$ is the Coriolis parameter at latitude y.

Since near the equator $f = \beta y$, it follows that

$$L = L_R(L) = (gH)^{\frac{1}{2}}/\beta L \ ,$$

which says that the equatorial radius of deformation L is just the usual Rossby radius evaluated at a latitude equal to L. As the equator is approached from higher latitudes the radius of deformation increases until a latitude equal to the radius of deformation is reached. The radius of deformation may be interpreted as the scale over which the motions of an ocean of depth H rotating at a rate $f/2$ adjust to a geostrophic equilibrium. The limit on its size as the equator is approached says that motions at $y < L$ must feel the effect of the Coriolis parameter at latitude $y = L$ if they attempt to adjust over a distance $L_R(y) > L$.

It may be seen from (2) that $T = f(L)^{-1}$; that is, T is the inverse of the Coriolis parameter at the equatorial radius of deformation. For $H = .5$ m, $L \approx 320$ km and $T \approx 1.6$ days. Wave speeds are scaled by

$$LT^{-1} = \beta L_R{}^2 = (gH)^{\frac{1}{2}} \approx 2.2 \text{ m sec}^{-1} \ .$$

This speed, which is external gravity wave speed for an ocean of depth H, scales both the equatorial gravity waves and the equatorial Rossby waves. For mid-latitude motions the gravity wave speed is still $(gH)^{\frac{1}{2}}$, but the Rossby wave speeds are scaled by the local value of $\beta L_R{}^2$. At 45° $\beta L_R{}^2 \approx .8$ cm sec^{-1}. The difference in equatorial and extraequatorial time-scales for planetary motions may thus be traced to the order of magnitude difference in the radii of deformation. For Rossby waves at any latitude the principal restoring force is the gradient of planetary vorticity, β. This varies little between the equator and mid-latitudes. The inertia of a baroclinic Rossby wave depends on $L_R{}^{-2}$ so the mid-latitude waves move along more slowly.

We can see all this by deriving a single exact equation for v from the homogeneous form of (1) [Rattray, 1964; Moore, 1968]. Applying the equatorial scaling this is, for $u, v, h \sim \exp i(k\,x\text{-}\omega t)$

$$v_{yy} + \left[\left(\omega^2 - k^2 - \frac{k}{\omega}\right) - y^2\right] v = 0 \ . \tag{3}$$

The solutions that go to zero at $y = \pm \infty$ are

$$v_n = \psi_n(y) = \pi^{-\frac{1}{4}} (2^n n!)^{-\frac{1}{2}} e^{-y^2/2} H_n(y) \tag{4a}$$

where the H_n's are the usual Hermite polynomials and with

$$\omega^2 - k^2 - \frac{k}{\omega} = 2n + 1 \ . \tag{4b}$$

Equation (3) has a turning point at $y = (2n + 1)^{\frac{1}{2}}$ (dimensionally, $y = (2n + 1)^{\frac{1}{2}} L$). The mid-latitude analog of (3), equatorially scaled, is

$$v_{yy} + \left[\left(\omega^2 - k^2 - \frac{k}{\omega}\right) - \epsilon\right] v = 0 \tag{5}$$

where $\epsilon = \dfrac{f_0{}^2}{c\beta}$ and f_0 is the mid-latitude Coriolis parameter specific to where our mid-latitude beta plane is centered. Eq. (5) is derived from the homogeneous part of the exact equation (3) by simply replacing f by f_0.

Eq. (5) admits simple trigonometric solutions e^{ily} which, it should be noted, have no turning points. It is known from an analysis of the exact modes on a sphere (Longuet-Higgins, 1968, Munk and Phillips, 1968) that small k mid-latitude modes *do* have turning points at the local inertial latitude, i.e. the frequency equal to the local Coriolis frequency. We preserve this feature of the mid-latitude modes by working on an equatorial beta plane, but at high mode number $2n + 1 \cong \epsilon$.

Since ϵ is essentially the Lamb parameter $\left(\dfrac{4\Omega^2 a^2}{gH}\right)^{\frac{1}{2}}$, it has the value of about 100 for the first baroclinic mode. We can therefore either think of mid-latitude modes as equatorial modes with $n = 0(100)$ so that the turning latitudes lie in mid-latitudes, or equivalently as equatorial modes of low mode number whose scaling is $T_m = (c\beta)^{-\frac{1}{2}} \epsilon^{-\frac{1}{2}}$, $L_m = \epsilon^{-\frac{1}{2}} \sqrt{\dfrac{c}{\beta}}$. In either case, the mid-latitude time scale is an order of magnitude slower, and the baroclinic radius of deformation an order of magnitude smaller, than the equatorial. Mid-latitude velocities are then two orders of magnitude smaller than equatorial, namely $\epsilon^{-1} c$ instead of c. In what follows, we will concentrate on equatorial modes—a simple transposition to high n will yield valid baroclinic mid-latitude results.

The plan of the remainder of this paper is as follows. In Sec. 2 we will review the complete equatorial free wave solutions. Sec. 3 deals with a barotropic analog of the forced problem in which the approximations to be used in the full problem are explored. The general unbounded forced problem is formulated in Sec. 4 and the general method of solution is given. The method is illustrated by giving explicit solutions for certain piecewise constant forcings in Sec. 5 and Sec. 6. Sec. 7 completes and summarizes our results.

Further papers in this series will use the methods and results of this paper to deal with responses in bounded basins, a numerical model to illustrate the transition to nonlinearity, and finally a numerical simulation of a highly nonlinear feature: the equatorial undercurrent.

2. Free wave solutions

The free wave solutions of (1) have been discussed in the literature (Matsuno, 1966; Blandford, 1966; Moore and Philander, 1976). Since our method of solution to the forced problem makes use of the properties of the free wave solutions they will be quickly reviewed here. This will also serve to establish notation. We assume that Eqs. (1) have been scaled using the relations (2). Then the free wave solutions to the unforced version of (1) (i.e. $F = G = Q = 0$) for an infinite ocean with the boundary conditions

$$u, v, h \to 0 \text{ as } |y| \to \infty$$

may be written

$$(u, v, h) = \phi_{n,j}(k,y) \exp i[kx - \omega_{n,j}(k)\, t] \ . \tag{6}$$

As a rule, n indexes the meridional structure (it is analogous to the meridional wave number) and j, the wave type (inertia-gravity or Rossby). The subscript pairs (n, j) may take on the possible values

$$j = 1 \text{ for } n = -1$$

$$j = 1 \text{ or } 2 \text{ for } n = 0$$

$$j = 1, 2 \text{ or } 3 \text{ for } n > 0 \ .$$

We denote this set of possible values of (n,j) by J. For $n > 0$, the $\omega_{n,j}(k)$'s satisfy the dispersion relation

$$\omega_{n,j}^2 - k^2 - k/\omega_{n,j} = 2n + 1 \ . \tag{7a}$$

For a given n and k there are three real roots to this equation, indexed by $j = 1, 2$ or 3. For definiteness we distinguish among these by their values as $k \to 0$:

$$\omega_{n,1} \to (2n + 1)^{\frac{1}{2}},\ \omega_{n,2} \to -(2n + 1)^{\frac{1}{2}},\ \omega_{n,3} \to -k/(2n + 1) \ .$$

Then $j = 1$ and $j = 2$ label inertial-gravity waves with phase speeds to the east and west, respectively, while $j = 3$ labels the Rossby waves. When $n = 0$ the root $\omega = -k$ of (7a) must be rejected because the corresponding u and h functions become unbounded at infinity. The acceptable $n = 0$ mode is referred to as the mixed mode or Yanai wave. The dispersion relation (7a) simplifies to

$$\omega_{0,j} - \omega_{0,j}^{-1} = k \ . \tag{7b}$$

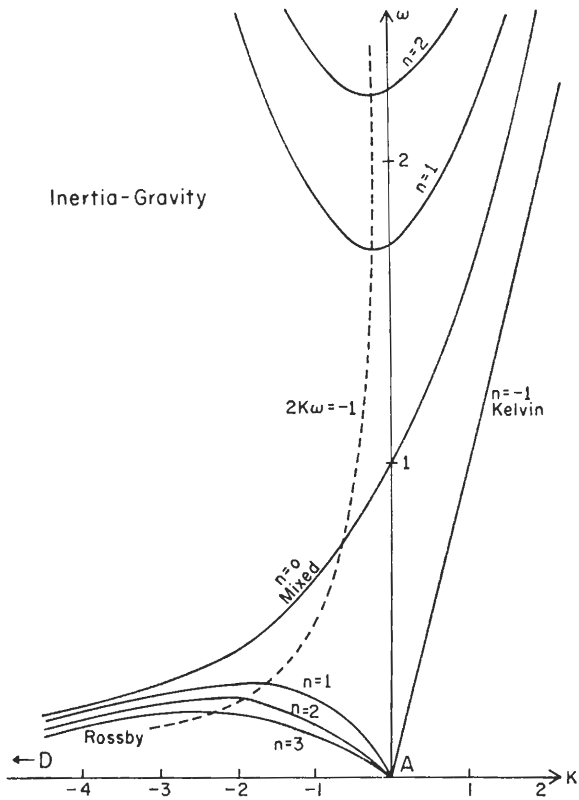


Figure 1. Dispersion relation for waves on an equatorial β plane for $\omega > 0$. The relation $\omega(k) = -\omega(-k)$ may be used for $\omega < 0$. The points A and D are described in the text.

For definiteness take $\omega_{01} > 0$; then $\omega_{02} < 0$. We have labeled the equatorial Kelvin wave by $n = -1$. Its dispersion relation is simply

$$\omega_{-1} = k \ . \tag{7c}$$

(We drop the redundant second subscript.) The dispersion relations (7) are displayed in Fig. 1 for $\omega > 0$; since $\omega(-k) = -\omega(k)$, the values for negative ω may be obtained by reflecting the graph through the origin.

The vector functions $\boldsymbol{\phi}_{n,j}(k,y)$ specify the meridional structure of u, v, and h for each wave. First define three vector functions of y only:

$$\begin{aligned} \mathbf{W}_n(y) &= (y\,\psi_n(y),\, 0,\, -d\,\psi_n(y)/dy) \\ \mathbf{V}_n(y) &= (0,\, \psi_n(y),\, 0) \\ \mathbf{M}_n(y) &= (-d\,\psi_n(y)/dy,\, 0,\, y\,\psi_n(y)) \end{aligned} \tag{8}$$

where ψ_n is the nth (normalized) Hermite function defined by (4a). For $n \geqq 0$

$$\boldsymbol{\phi}_{n,j}(k,y) = \omega_{n,j}(k)\,\mathbf{W}_n(y) + k\,\mathbf{M}_n(y) - i(\omega_{n,j}^2(k) - k^2)\,\mathbf{V}_n(y) \ . \tag{9}$$

For the Kelvin wave, $n = -1$,

$$\boldsymbol{\phi}_{-1}(k,y) = \mathbf{M}_{-1}(y) \equiv \mathbf{W}_{-1}(y) = 2^{-\frac{1}{2}}\,(\psi_0(y), 0, \psi_0(y)) \ . \tag{10}$$

Having established our notation, we wish to note some of the characteristics of these solutions important to what follows. (See Moore and Philander, 1976, for a more complete account). The higher frequency branches in Fig. 1 are the inertia-gravity waves, the lower frequency curves for $n > 0$ are Rossby waves. The Rossby waves all have a westward phase velocity. The dotted line $2\,k\omega = -1$ divides those waves with eastward group velocity from those with westward group velocity. For a given zonal wavenumber k the larger the n the smaller the group velocity. For the Rossby waves v and h are in approximate geostrophic balance for large k, while as $k \to 0$, u and h approach geostrophic balance. The large k Rossby waves are approximately nondivergent. The mixed mode ($n = 0$) behaves like a Rossby wave for small wavelength waves with westward phase speed; it behaves like a gravity wave for $k > 0$. Both the Kelvin wave and the mixed mode have eastward group velocity for all wavelengths. The Kelvin wave is purely hyperbolic while all other wave solutions to (1) are dispersive.

3. Forced barotropic response

We have discussed the free solutions and the dispersion relations for the equatorial beta plane equations. The kinematics of these waves are quite involved and make a direct attack on the initial value problem almost byzantine in its complexity. To illuminate our important points in a simpler context, we will, in this section, examine a problem which exhibits many of the same features of the equatorial beta plane problem. The model chosen is the barotropic vorticity equation. It can be derived from the single v equation by making two assumptions viz. the time scale is long compared to a half pendulum day, and the depth is large enough so that the rotational Froude number $\left(\dfrac{f^2\,L^2}{gH}\right)$ is no larger than order unity. (See Anderson and Gill, 1975, for a more extended discussion). The resulting equation can be written in terms of a stream function ψ such that $v = \psi_x$, $u = -\psi_y$:

$$(\nabla^2 - L_R^{-2})\,\psi_t + \beta\,\psi_x = C(x,y;t) \ , \tag{11}$$

where $C(x,y;t) = F_y - G_x$ is the curl of the wind stress and L_R is the radius of deformation.

We will confine ourselves to longitudinal forcing of the form e^{ily} since this case gives a purely zonal propagation, in analogy to the baroclinic equatorial case. The equation we will look at is therefore

$$\psi_{xxt} - m^2\,\psi_t + \beta\,\psi_x = C(x;t) \tag{12}$$

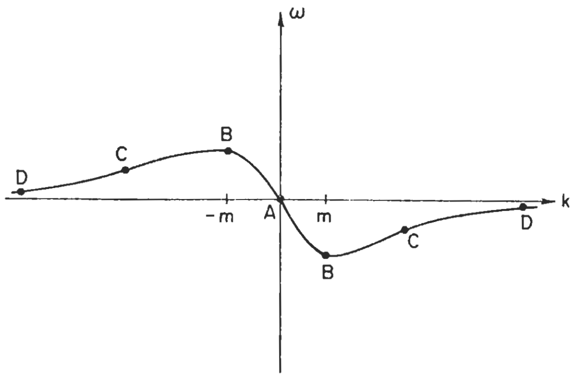


Figure 2. The barotropic dispersion relation $\omega = -\beta k(k^2 + m^2)^{-1}$. The points A, B, C, D are described in the text.

where $m^2 = l^2 + L_R^{-2}$. [Note that m^2 corresponds to the $2n + 1$ of Eq. (4b)]. North-South propagation can be recovered by Fourier analyzing the y dependence of a forcing, solving for each l, and summing the response.

We will be particularly interested in forcing functions switched on in time, in the form $H(t)$, where H is the Heaviside step function. Previous studies have examined periodic time behavior (Veronis and Stommel, 1956), and impulsive time behavior $\delta(t)$ (Veronis, 1958, exactly; and Longuet-Higgins, 1965, approximately). Dickinson (1969) has examined the general problem of impulsive and switched-on sources in a mostly atmospheric context. The seminal paper of Lighthill (1969) has treated many of these same problems: it is to clarify the approximations made in that paper that this section is written. After much of this work was completed, the paper of Anderson and Gill (1975) appeared. Many of the features to be described below can be observed in their numerical simulation although, as with Lighthill's paper, the nature of the approximations they made in describing their simulations is here clarified.

Eq. (12) is dispersive and admits free wave solutions of the form $e^{i(kx-\omega t)}$ if the dispersion relation

$$\omega = -\beta k(k^2 + m^2)^{-1} \tag{13}$$

is obeyed.

Fig. 2 shows a plot of the dispersion relation. Note how similar this curve is to the baroclinic Rossby wave dispersion relations shown in Fig. 1. Since we will be interested in $H(t)$ forcing, we might expect the zero frequency points A and D to have special significance—we will see that this is true. The point B $\left(\text{coordinates } k = \pm m, \omega = \pm \dfrac{\beta}{2m}\right)$ is a point of zero group velocity ($\omega'(\pm m) = 0$). Point C

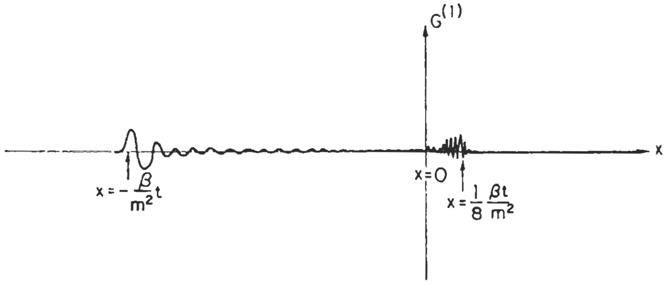


Figure 3. Sketch of asymptotic barotropic response to $\delta(x)\,\delta(t)$ forcing. Not to scale.

$\left(k = \pm \dfrac{\sqrt{3}\beta}{4m}\right)$ is of a maximum eastward group velocity $\left(\omega''\ (\pm\sqrt{3}m) = 0,\right.$ $\left.\omega'\ (\pm\ 3m) = \tfrac{1}{8}\dfrac{\beta}{m^2}\right)$ A is, in addition to being a point of zero frequency, also a point of maximum *westward* group velocity $-\beta/m^2$. We will see that it is this dual property that A possesses, being both a point of zero frequency and an inflection point of the dispersion relation, that accounts for its asymptotic dominance.

Longuet-Higgins (1965) has investigated the long time behavior of the impulse Green function

$$G_{xxt}{}^{(1)} - m^2\, G_t{}^{(1)} + \beta\, G_x{}^{(1)} = \delta(x)\,\delta(t) \tag{14}$$

by looking at the asymptotic properties of the Fourier transformed solution. Asymptotically the response may be summarized by the sketch in Fig. 3. The response arises primarily from the inflection points ("caustics") of the dispersion relation. After a very long time, the $t^{-\frac{1}{2}}$ wake will die away and nothing will remain. This is in contrast to the f plane analog in which a geostrophically balanced flow over a radius of deformation would remain.

The switched-on forcing is quite different even on the β plane because at long time, a steady response endures. The Green function for this case satisfies

$$G_{xxt}{}^{(2)} - m^2\, G_t{}^{(2)} + \beta\, G_x{}^{(2)} = \delta(x)\, H(t)\ . \tag{15}$$

An immediate solution may be obtained by Fourier transforming in space and Laplace transforming in time, with initial condition $G^{(2)} = 0$ for $t < 0$. The resulting transform denoted by the argument, p, is

$$G^{(2)}\ (k; p) = \frac{1}{p}\,\frac{1}{(-k^2 - m^2)\,p + ik\,\beta}\ . \tag{16}$$

There are two interesting and seemingly incompatible limits to this equation. If $k \to 0$,

$$G^{(2)}(k,p) \to (mp)^{-2} \text{ and } G^{(2)}\ (x,t) \to -\,t\,m^{-2}\,H(t)\ . \tag{17a}$$

If, however, we take the long time limit $p \to 0$ first,

$$G^{(2)}(kp) \to (ik\,\beta\,p)^{-1} \text{ and } G^{(2)}(x,t) \to \beta^{-1} H(x)\,H(t) \ . \tag{17b}$$

Both limits (17a) and (17b) are valid long-time solutions. The first represents secular growth due to resonant forcing at $k = 0$, $\omega = 0$ and the second represents the steady (we will call it Sverdrup) long-term forced response. In terms of the original equation (16), the first solution has $-m^2\,G_t{}^{(2)}$, and the second $\beta\,G_x{}^{(2)}$, being balanced by the forcing. We will first present the correct asymptotic expansions and than a heuristic argument that will explain the asymptotics and the transition from (17a) to (17b).

The complete solution to (15) with explicit time dependence is expressed in terms of the inverse Fourier transform as

$$G^{(2)}(x,t) = I_1{}^{(2)} + I_2{}^{(2)} \ , \tag{18a}$$

where

$$I_1{}^{(2)} = (2\pi)^{-1} \int_{-\infty}^{\infty} (ik\beta)^{-1} \exp[ikx]\,dk = \beta^{-1} H(x) \tag{18b}$$

and

$$I_2{}^{(2)} = -\,(2\pi)^{-1} \int_{-\infty}^{\infty} (ik\beta)^{-1} \exp[it\,\chi(k)]\,dk \tag{18c}$$

with

$$\chi(k) = k\xi + \beta\,k\,(k^2 + m^2)^{-1} \ .$$

Two immediate comments can be made about $G^{(2)}$. First, that an exact solution can be obtained by first Laplace transforming (15) in t and suitably inverting the transform, as in the method of Veronis (1958). The exact solution is a complicated double integral from which it is impossible to extract much useful information, just as in Veronis' case. (The interested reader may apply to the authors for this solution—it is not worth giving here). The second comment is that it appears at first sight that the integral (18c) is easily performed since the integral has three isolated singularities (the pole at $k = 0$ and essential singularities at $k = \pm\,im$), and the contour can be closed at infinity. The residues at the essential singularities, however, are only expressible as multiple infinite sums and it has proven impossible to extract any useful information from the integral this way.

We therefore proceed to analyze the integral (18c) by approximate methods. The short time behavior can be obtained by expanding the integral in powers of t, the first term yielding

$$G^{(2)}(x,t) \to -\,t\,e^{-m\,|x|} \text{ as } t \to 0 \ .$$

We note that the initial response grows linearly with t arbitrarily far from the forcing. The source of this curious behavior is that the approximations leading to (16) have idealized the ocean as infinitely deep. There are therefore infinitely fast waves ("precursors") which allow a small growth before any planetary wave response can have sizeable amplitude.

The large t response $\left(t >> \frac{m}{\beta}\right)$ is obtained by applying the method of stationary phase to the integral (18c) for constant values of $\xi = \frac{x}{t}$. For almost all ξ, the method will yield a $t^{-\frac{1}{2}}$ behavior. Anything larger than $t^{-\frac{1}{2}}$ will then clearly dominate the asymptotic behavior.

It proves illuminating to consider $\xi = \frac{x}{t} = \epsilon$ where ϵ is very small. This arises, for example, in considering the long-time behavior at a fixed point in space.

The wavelengths contributing to the stationary phase solutions are solutions of $\omega'(k_i) = \epsilon$ and are found to be $k = \pm\, m$, with no restrictions on x, and $k = \pm \sqrt{\frac{\beta}{\epsilon}}$ when $x > 0$. We expect the behavior at a fixed point to be dominated by the zero group velocity points and a glance at Fig. 2 shows that these are indeed points B and D. The restriction $x > 0$ attached to point D simply means that the group velocity, no matter how small, is always positive, so that for $t > 0$ only points to the east of the forcing point $x = 0$ will be affected.

The contribution from the points B on the dispersion curve is easily found to be

$$I_2{}^{(2)}\,(B) \rightarrow \sqrt{\frac{2}{\pi}}\left(\frac{m}{\beta^3\, t}\right)^{\frac{1}{2}} \cos\left(mx + \frac{\beta t}{2m} + \frac{\pi}{4}\right) \tag{19a}$$

while the contribution from the points D, $k = \pm\, (\beta\epsilon^{-1})^{\frac{1}{2}}$, is

$$I_2{}^{(2)}\,(D) \rightarrow -\frac{1}{\sqrt{\pi}}\,\frac{1}{\beta}\,\frac{1}{(x\beta t)^{\frac{1}{4}}} \cos\left(2\sqrt{\beta\, x\, t} - \frac{\pi}{4}\right). \tag{19b}$$

Note that the contribution from point D goes as $t^{-\frac{1}{4}}$ and is therefore expected to make a major asymptotic contribution. Note, also, that the limit $x \rightarrow 0$ in (19b) is not defined. The source of this difficulty can be found by looking at (18c) for extremely large k:

$$I^{(2)}\,(D) \rightarrow -\frac{1}{2\pi\beta}\int_{-\infty}^{\infty} \frac{1}{ik}\; e^{\,it\left(k\epsilon + \frac{\beta}{k}\right)}\, dk\,.$$

This integral can be evaluated exactly by changing variables, expanding the exponential in Bessel functions and then evaluating by contour integration at the pole. The result is an ordinary Bessel function of zero order (this solution first appears in Rossby, 1945):

$$I_2{}^{(2)}\,(D) \rightarrow -\frac{1}{\beta}\, J_0\,(2\sqrt{\beta x t}) \tag{19c}$$

and this is the uniform asymptotic result for t large, $x \geqq 0$.

The large t expansion of (19c) is identical to the stationary phase result (19b). The local wave number is $\sqrt{\frac{\beta t}{x}}$ and so increases with time, and the local frequency is $\sqrt{\frac{\beta x}{t}}$ which decreases with time. The width of the main peak about $x = 0$ is $0\left(\frac{1}{t}\right)$.

The asymptotic response due to the inflection point C of the dispersion curve (maximum eastward group velocity) can be uniformly expanded near the wave front $\delta_C \equiv \xi - \frac{1}{8}\beta/m^2 = 0$ as

$$I_2^{(2)}(C) \rightarrow -\frac{2}{\sqrt{3}\,m}\left(\frac{32}{at}\right)^{\frac{1}{3}} Ai\left[\left(\frac{32\,t^2}{a}\right)^{\frac{1}{3}} \delta_C\right] \sin\left[\sqrt{3}\left(mx + \frac{\beta t}{4m}\right)\right] \tag{20}$$

where $a \equiv 3\,\beta\,m^{-4}$. This behaves as $t^{-\frac{1}{3}}$ at the wavefront, $t^{-\frac{1}{2}}$ behind the wavefront (to blend in with the general residual $t^{-\frac{1}{2}}$ decaying motion), and e^{-t} ahead of the wavefront.

The major response to the switch-on forcing arises from the point of maximum westward group velocity $\left(-\frac{\beta}{m^2}\right)$, point A in Fig. 2. The method of Chester, Friedman, and Ursell, 1957; (see also Friedman, 1959) breaks down in this case because $k = 0$ is a pole of the integrand as well as being a double zero of the phase. This problem can be avoided by formally differentiating $I_2^{(2)}$ with respect to ξ and uniformly approximating the derivative near the wavefront $\delta_A = \xi + \frac{\beta}{m^2} = 0$:

$$\frac{\partial I_2^{(2)}}{\partial \xi}(A) \rightarrow -\frac{t^{\frac{2}{3}}}{\beta a^{\frac{1}{3}}} Ai\left[\frac{t^{\frac{2}{3}}}{a^{\frac{1}{3}}}\delta_A\right]. \tag{21}$$

Asymptotic series can be integrated term by term so that we are assured that the integral of (21) is indeed asymptotic to $I_2^{(2)}(A)$:

$$I_2^{(2)}(A) \rightarrow -\frac{1}{\beta}\int_0^{t^{\frac{2}{3}}a^{-1}\delta_A} Ai(-z)\,dz - \frac{1}{3\beta}. \tag{22}$$

The response (22) is plotted in Fig. 4. As the front propagates, its slope steepens as $t^{\frac{2}{3}}$ and its width narrows as $t^{-\frac{2}{3}}$. The wake consists of the $-1/\beta$ steady response plus wiggles ("postcursors") which decay as $t^{-\frac{1}{2}}$. At very long times the response (22) essentially (to within wiggles) becomes a step function about $x = \frac{\beta}{m^2}t$ so that

$$G^{(2)} \rightarrow \frac{1}{\beta}\left[H(x) - H\left(x + \frac{\beta}{m^2}t\right)\right] - \frac{1}{\beta}J_0\left(2\sqrt{\beta x t}\right) + O(t^{-\frac{1}{2}}). \tag{23}$$

The higher order contributions come from the previously derived (19c) and (20)

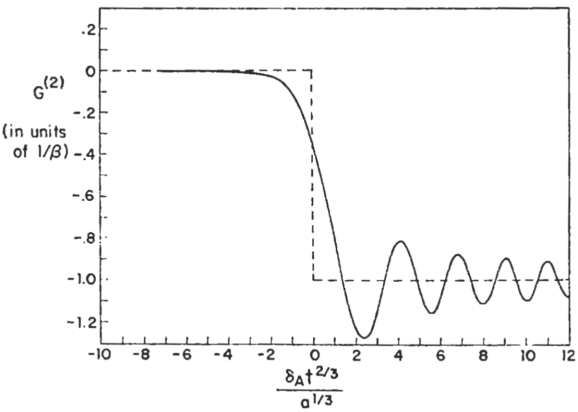


Figure 4. Behavior near the hyperbolic wave front $\delta_A = \xi + \beta t/m^2 = 0$. The dashed line is $H(\delta_A)$.

with the $t^{-\frac{1}{2}}$ behavior of the Bessel function response (19c) formally being next most important. A sketch of the total asymptotic ($t >> m/\beta$) response is given in Fig. 5.

We may interpret these results for $G^{(2)}$ using an argument implicit in the work by Anderson and Gill (1975). If we scale the time derivative in the exact Eq. (15) by $\frac{1}{T}$, the equation should go asymptotically as

$$\frac{G_{xx}{}^{(2)}}{T} - \frac{m^2\, G^{(2)}}{T} + \beta\, G_x{}^{(2)} = \delta(x)\, H(t)\ . \tag{24}$$

The large T limit reduces the order of the equation from second to first (in space) so that we may appeal to the methods of singular perturbation theory. We therefore first look at the "interior" equation

$$-\, m^2\, G_t{}^{(2,H)} + \beta\, G_x{}^{(2,H)} = \delta(x)\, H(t)\ . \tag{25}$$

Eq. (25) is purely hyperbolic (hence the superscript H) and supports discontinuities on characteristics. The solution to (25) is

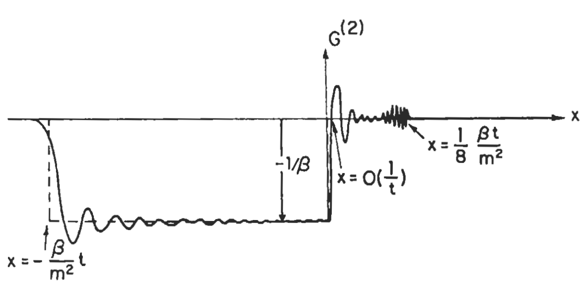


Figure 5. Sketch of asymptotic barotropic response to $\delta(x)\, H(t)$ forcing. Not to scale.

$$G^{(2,\mathrm{H})} = \frac{1}{\beta}\left[H(x) - H\left(x + \frac{\beta}{m^2} t \right)\right] \tag{26}$$

which is the behavior already found in (23).

It should be noted that the approximate integral that led to the asymptotic results (21) and (22) can be written

$$I_2{}^{(2)}(A) = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{ik\beta} \exp\left\{ it\left(\xi + \frac{\beta}{m^2}\right) k - itak^3 \right\} dk \ .$$

The k^3 term in the exponential gives all the asymptotic dispersion. If it were absent the integral would simply be $-\frac{1}{\beta} H\left(x + \frac{\beta}{m^2} t \right)$. It is in this sense that the integral of the Airy function is the (asymptotically) dispersed form of the step function, and it is in this asymptotic sense that the approximations of Lighthill (1969) and Anderson and Gill (1975) leading directly to the step function must be understood. Similarly, the Airy function solution to (14) is the dispersed form of the corresponding solution to the hyperbolic part of Eq. (14), $\delta\left(x + \frac{\beta}{m^2} t \right)$. The effects of dispersion are more pronounced in the $\delta(t)$ forcing case in the sense that the Airy function resembles the delta function less than the integral of the Airy function resembles the step function. This difference is due to the lack of a sustained forcing in the $\delta(t)$ case: all the response simply disperses and propagates away, while in the $H(t)$ case a steady response remains.

The full Eq. (15) with forcing $\delta(x)\ H(t)$ has the following jump conditions at the forcing point:

$$[G^{(2)}] = 0,\ [G_{xt}{}^{(2)}] - m^2\,[G_t{}^{(2)}] = 1 \ , \tag{27}$$

where the bracket indicates the difference of $G^{(2)}$ at $0+$ and $0-$. The hyperbolic solution (26) satisfies the jump conditions

$$[G^{(2,H)}] = -\frac{1}{\beta}\ ,\ [G_t{}^{(2,\mathrm{H})}] = [G_{xt}{}^{(2,H)}] = 0 \tag{28}$$

which clearly violate the correct conditions (27).

A boundary layer correction to the interior solution must be added to satisfy the correct jump conditions. The boundary layer equation for (24) is

$$G_{xxt}{}^{(2,J)} + \beta\, G_x{}^{(2,J)} = 0 \ . \tag{29}$$

The boundary layer width is expected to be $O(t^{-1})$. The exact solution to (29) is a sum of Bessel function terms of the form

$$\left(\frac{\beta x}{t}\right)^{n/2} J_n\left(2\sqrt{\beta\, x\, t}\right) \text{ for } -\infty < n < \infty \ .$$

It can easily be seen that $n = 0$ is the correct choice. The correct correction

$$G^{(2,J)} = -\frac{1}{\beta} J_0 (2\sqrt{\beta x t}) \tag{30}$$

satisfies the jump conditions

$$[G^{(2,J)}] = -\frac{1}{\beta}, [G_{xt}^{(2,J)}] - m^2 [G_t^{(2,J)}] = 1 \tag{31}$$

and does have a peak of width $O(t^{-1})$.

The sum

$$G^{(2)} = G^{(2,H)} + G^{(2,J)} = \frac{1}{\beta}\left[H(x) - H\left(x + \frac{\beta}{m^2} t \right)\right] - \frac{1}{\beta} J_0 \left(2\sqrt{\beta x t}\right) \tag{32}$$

gives a complete solution satisfying the correct jump conditions at $x = 0$. If $G^{(2,H)}$ is understood as the limit of (22), it is smooth about the wavefront. Understood this way, (32) is the correct asymptotic interior-plus-boundary-layer solution to (15) and agrees with the dominant asymptotic behavior previously derived in (23).

The effects of a distributed forcing can be understood by considering the equation

$$G_{xxt}^{(3,4)} - m^2 G_t^{(3,4)} + \beta G_x^{(3,4)} = H(\pm x) H(t) \tag{33}$$

where the superscript 3 will refer to the $H(x)$ forcing and 4 to $H(-x)$.

The solution to (33) is

$$G^{(3,4)} (x,t) = I_1^{(3,4)} + I_2^{(3,4)}$$

where

$$I_1^{(3,4)} = \mp \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{k^2 \beta} e^{ikx} dk = \frac{1}{\beta} x H(\pm x) \tag{34a}$$

and

$$I_2^{(3,4)} = \pm \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{k^2 \beta} e^{it\,\chi(k)} dk \ . \tag{34b}$$

Care must be exercised in the interpretation of (34a,b). The poles are to be slightly displaced into the positive and negative imaginary plane in performing the integration in order to guarantee the existance of the integrals.

The stationary phase evaluation of (34b) for $\xi = \epsilon$ (small) yields

$$I_2^{(3,4)} (B) = -\sqrt{\frac{2}{\pi}} (\beta^3 mt)^{-\frac{1}{2}} \cos\left(mx + \frac{\beta t}{2m} - \frac{\pi}{4}\right) \tag{35a}$$

from $k = \pm m$, and

$$I_2^{(3,4)} (D) = \pm \frac{1}{\beta}\sqrt{\frac{x}{\beta t}} J_1 (2\sqrt{\beta x t}) \rightarrow$$

(35b)

$$\pm \frac{1}{\sqrt{\pi}} \frac{1}{\beta} \left(\frac{x}{\beta^3 t^3} \right)^{\frac{1}{4}} \quad \cos\left(2\sqrt{\beta x t} + \frac{\pi}{4} \right)$$

from $k = \pm \sqrt{\dfrac{\beta}{\epsilon}}$. Also, as before, we find

$$I_2^{(3,4)}\,(C) = \pm \frac{2}{3\,\beta\, m^2} \left(\frac{32}{at} \right)^{\frac{1}{3}} \quad Ai\left[\left(\frac{32}{a} \right)^{\frac{1}{3}} \; t^{\frac{2}{3}}\, \delta_C \right] \tag{35c}$$

$$\times \cos\left(\sqrt{3} \left(mx + \frac{\beta t}{4m} \right) \right)$$

near the wavefront $\delta_C = 0$. For the hyperbolic wavefront $\delta_A = 0$, the integral (34b) has a double pole at the point $k = 0$. A partial integration gives

$$I_2^{(3,4)} = \mp \frac{\left(x + \dfrac{\beta}{m^2}\, t \right)}{2\pi\beta} \int_{-\infty}^{\infty} \frac{1}{ik}\, e^{it\chi(k)}\, dk$$

and if care is taken in keeping the locations of the pole (or equivalently the contour) straight, the integrals may be done as before.

$$I_2^{(3)}\,(A) = -\frac{1}{\beta} \left(x + \frac{\beta}{m^2}\, t \right) \left[\int_0^{\delta_A a^{-\frac{1}{3}} t^{\frac{2}{3}}} Ai(-z)dz + \tfrac{1}{3} \right] \tag{35d}$$

$$\rightarrow -\frac{1}{\beta} \left(x + \frac{\beta}{m^2}\, t \right) H \left(x + \frac{\beta}{m^2}\, t \right).$$

$$I_2^{(4)}\,(A) = -\frac{1}{\beta} \left(x + \frac{\beta}{m^2}\, t \right) \left[-\int_0^{\delta_A a^{-\frac{1}{3}} t^{\frac{2}{3}}} Ai(-z)dz + \tfrac{2}{3} \right] \tag{35e}$$

$$\rightarrow -\frac{1}{\beta} \left(x + \frac{\beta}{m^2}\, t \right) H \left(-\left(x + \frac{\beta}{m^2}\, t \right) \right).$$

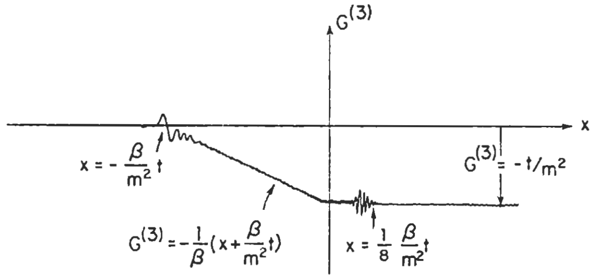
Sketches of the response $G^{(3)}$ and $G^{(4)}$ are given in Fig. 6.

Two points are important in these distributed forcing cases. The first is that the hyperbolic response now by itself satisfies the jump conditions at $x = 0$:

$$[G^{(3,4)}] = [G_{xt}^{(3,4)}] = 0$$

$$[G_{xxt}^{(3,4)}] + \beta[G_x^{(3,4)}] - m^2\, [G_t^{(3,4)}] = \pm\, 1$$

so that no boundary layer correction is needed at $x = 0$. This is why the Bessel response (35b) is of order $t^{-\frac{3}{4}}$ and therefore below the $t^{-\frac{1}{3}}$ background noise in the system. The Bessel boundary layer correction is only needed, to $0(t^{-\frac{1}{3}})$, at a forcing discontinuity when the hyperbolic *response* is discontinuous at the discontinuity.



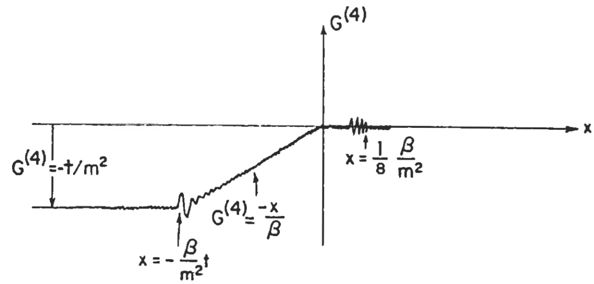


Figure 6. (a) Sketch of asymptotic barotropic response to $H(x)$ $H(t)$ forcing. Not to scale. (b) Same for $H(-x)$ $H(t)$ forcing.

The second point involves the secular growth, $-t/m^2$, that shows up in both $G^{(3)}$ and $G^{(4)}$. We can now explain the two possible long-time limits we found previously in Eq. (17). The essential explanation of secular growth is that it is the resonant response to forcing at $\omega = 0$, $k = 0$. That the ultimate frequency is $\omega = 0$ is guaranteed by the nature of the forcing $H(t)$. The $k = 0$ is harder to achieve; in fact it is the existence of forcing discontinuities that prevent $k = 0$ from being achieved. If, at a point x, at a time t, there is also an equal forcing at point $x' = x + \frac{\beta}{m^2}t$, the point x will have no way of knowing that there is a forcing boundary somewhere out to the east of x'. The growth will then be secular in t. As soon as the point x' falls beyond a forcing region, the secular growth will cease. Thus, in Fig. 6a and 6b, there is secular growth for all x at time t such that $x + \frac{\beta}{m^2}t$ *also* lies in the forcing region and the growth ceases when the condition no longer holds. The explanation of the transition from secular growth to Sverdrup for a switched-on forcing is in essence related to our previous result that the steady response is signalled by the hyperbolic group velocity $-\beta/m^2$.

4. Forced baroclinic response: general formulation

In this section we give a general formulation for the solution of the forced shallow water equations on an equatorial beta plane. Our approach is along the lines suggested by Matsuno (1966). The first task is to separate out the meridional dependence by an eigenfunction expansion. In doing so we rely on the information about the free solutions that was reviewed in Sec. 2. Once this first task has been accomplished we are left with the problem of finding the temporal and longitudinal structure of the response. This latter task is simplified by making use of the knowledge gained from the asymptotic analysis of Sec. 3.

If, in the shallow water equations (1a,b,c) we take $f = \beta y$ and apply the equatorial scaling (2) we arrive at the equations

$$u_t - yv + h_x = F \tag{36a}$$

$$v_t + yu + h_y = G \tag{36b}$$

$$h_t + u_x + v_y = Q \ . \tag{36c}$$

These may be written in the compact form

$$\frac{\partial \mathbf{u}^T}{\partial t} + \Omega\ \mathbf{u}^T = \mathbf{F}^T$$

where $\mathbf{u} \equiv (u, v, h)$ and $\mathbf{F} \equiv (F, G, Q)$. Superscript T indicates transpose and Ω is an operator depending only on the spatial variables x and y. Fourier transform $\mathbf{u}$ and $\mathbf{F}$ from (x, y, t) space to (k, y, t) space by applying the operator $\int_{-\infty}^{\infty} (\ \ )\ e^{-ikx}\, dx$ to each component. Then

$$\frac{\partial \mathbf{u}^T(k,y,t)}{\partial t} + \Omega(k,y)\ \mathbf{u}^T(k,y,t) = \mathbf{F}^T(k,y,t)$$

where

$$\Omega(k,y) = \begin{bmatrix} 0 & -y & ik \\ y & 0 & \dfrac{\partial}{\partial y} \\ ik & \dfrac{\partial}{\partial y} & 0 \end{bmatrix} .$$

It now follows immediately that the free wave solutions (6) ff. yield the vector eigenfunctions of $\Omega(k,y)$; i.e.,

$$\Omega(k,y)\ \boldsymbol{\phi}_{n,j}{}^T(k,y) = i\omega_{n,j}(k)\ \boldsymbol{\phi}_{n,j}{}^T(k,y)$$

where the eigenvalues $i\omega_{n,j}$ are given by the free wave dispersion relations (7).

We now define a scalar product by

$$([u_a, v_a, h_a], [u_b, v_b, h_b]) \equiv \int_{-\infty}^{+\infty} (u_a{}^* u_b + v_a{}^* v_b + h_a{}^* h_b)\, dy \tag{37}$$

where * denotes complex conjugate. In the Appendix it is shown that the eigenfunctions $\phi_{n,j}(k,y)$ are orthogonal and complete. This means that any vector forcing may be expanded in the $\phi_{n,j}$'s if its components may be expanded in Hermite functions. As a general rule, a function may be represented as a convergent series of Hermite functions if it is square integrable in the interval $(-\infty, +\infty)$. Questions of convergence make for some nice mathematical problems, but in view of our purpose such questions may be circumvented. We are concerned with ocean basins in equatorial regions of limited latitudinal extent. The form of the forcing function (or the response) beyond the limits of the basin should make no difference to the basin response so the forcing may always be taken to go to zero sufficiently rapidly as $|y| \to \infty$. For example, any physically reasonable forcing may be multiplied by exp $(-b^2 y^2)$, $b^2 << 1$ to guarantee convergence without changing its value near the equator; the projection of this forcing onto the modes with n small will be unchanged (since these modes have small amplitude away from the equator).

Explicitly, expand

$$e^{-b^2 y^2} = \sum_{n=0}^{\infty} f_n \psi_n(y)$$

where the coefficients are $f_n = \int_{-\infty}^{\infty} e^{-b^2 y^2} \psi_n(y)\, dy$. The integral is easily performed—only even coefficients enter

$$f_{2m} = \frac{\pi^{\frac{1}{4}}(2(2m)!)^{\frac{1}{2}}}{2^m\, m!} \frac{(1-2\, b^2)^m}{(1+2\, b^2)^{m+1}} \quad .$$

For a forcing extending far beyond L, $b^2 << 1$ so that

$$f_{2m} \sim (1 - (4m+2)b^2)$$

and we conclude that only those modes that have their turning points lying within the half width of the forcing will be significantly excited. For $b^2(4m+2) << 1$, the result is identical to the result for a constant forcing.

The fact that modes with n large might be affected by this alteration is an indication of the fact that these infinite beta plane modes are not the eigenfunctions for a bounded basin. (The correct modes involve the parabolic cylinder functions which give $v = 0$ at the zonal walls). Those modes which have their turning latitudes equatorward of the latitudes bounding the basin will be essentially unaffected by the walls. For an ocean bounded at $\pm$ 15° with a baroclinic radius of deformation of 300 km this means those modes with $n \leqq 12$. Higher modes must be corrected by considering the effects of walls at a finite distance from the equator. Such changes will make little difference near the equator where the amplitude of these modes is small. In summary, since our problem is to calculate the equatorial response, we needn't concern ourselves much with questions of convergence or the influence of northern and southern boundaries. The chief exception to this statement is the possibility of fast moving boundary trapped modes which may turn the corners at the bounding meridians and propagate into the equatorial region (e.g., coastal Kelvin waves; see Moore, 1968).

The completeness of the eigenfunctions means that for any (physically interesting) forcing function we may write

$$\mathbf{F}(k,y,t) = \sum_{(n,j)\in J} b_{n,j}(k,t)\boldsymbol{\phi}_{n,j}(k,y) \tag{38}$$

where J is again the set of permissible subscripts. The $b_{n,j}$ may be computed as follows. Since the $\boldsymbol{\phi}_{n,j}$ are orthogonal it follows from (38) that

$$b_{n,j} = N_{n,j}^{-1}\,(\boldsymbol{\phi}_{n,j}, \mathbf{F}) \tag{39}$$

where

$$N_{n,j}(k) \equiv (\boldsymbol{\phi}_{n,j}, \boldsymbol{\phi}_{n,j}) = (2n+1)(\omega_{n,j}^2 + k^2) + 2k\omega_{n,j}^2 + (\omega_{n,j}^2 - k^2)^2 \tag{40}$$

Then with the notation $(A)_n = \int_{-\infty}^{+\infty} A\,\psi_n\,dy$,

$$b_{-1}(k,t) = 2^{-\frac{1}{2}}\,(F + Q)_0 \ , \tag{41a}$$

and for $n \geqq 0$

$$b_{n,j}(k,t) = N_{n,j}^{-1}(k)\,\{\omega_{n,j}\,d_n(k,t) + k\,e_n(k,t) + i(\omega_{n,j}^2 - k^2)\,g_n(k,t)\} \ , \tag{41b}$$

where

$$d_n(k,t) = (yF + Q_y)_n \tag{42a}$$

$$e_n(k,t) = (F_y + yQ)_n \tag{42b}$$

and

$$g_n(k,t) = (G)_n \ . \tag{42c}$$

Once the $b_{n,j}$ has been obtained, one may proceed in the manner usual with eigenfunction expansions:

Let

$$\mathbf{u} = \sum_{(n,j)\in J} a_{n,j}(k,t)\,\boldsymbol{\phi}_{n,j}(k,y) \ . \tag{43}$$

Then

$$\frac{\partial a_{n,j}}{\partial t} + i\,\omega_{n,j}\,a_{n,j} = b_{n,j} \text{ for all } (n,j) \in J \ ,$$

or

$$a_{n,j}(k,t) = a_{n,j}(k,0)\,e^{-i\,\omega_{n,j}\,t} + \int_0^t b_{n,j}(k,s)\,e^{i\,\omega_{n,j}(s-t)}\,ds \ . \tag{44}$$

This equation is familiar from the linear harmonic oscillator problem. If the initial conditions are that $\mathbf{u} = 0$ at $t = 0$ and the forcing is at a single frequency σ so that $b_{n,j}(k,t) = b_{n,j}(k)e^{-i\,\sigma\,t}$, then

$$a_{n,j}(k,t) = \frac{b_{n,j}(k)}{i(\omega_{n,j} - \sigma)} \; [e^{-i\sigma t} - e^{-i\omega_{n,j}t}] \cdot$$

As with the linear oscillator, the first term in the square brackets has the same time behavior as the forcing, while the second is the free wave response needed to satisfy the initial conditions. Clearly, the closer the forcing frequency is to the natural frequency the larger the response. At resonance $\sigma = \omega_{n,j}$ and $a_{n,j} = tb_{n,j}(k)$—secular growth. For a steady forcing $\sigma = 0$ so that

$$a_{n,j}(k,t) = \frac{b_{n,j}(k)}{i\,\omega_{n,j}(k)} \; [1 - e^{-i\omega_{n,j}t}] \; . \tag{45}$$

In a formal sense the problem of finding the ocean's response to an arbitrary forcing is now solved—one need only invert the necessary Fourier transforms. That is

$$\mathbf{u}(x,y,t) = [u(x,y,t),\, v(x,y,t),\, h(x,y,t)] = \sum_{(n,j)\in J} \mathbf{u}_{n,j}(x,y,t)$$

where

$$\mathbf{u}_{n,j}(x,y,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} a_{n,j}(k,t)\, \boldsymbol{\phi}_{n,j}(k,y)\, e^{ikx}\, dk \; . \tag{46}$$

As a rule the integrals in (46) are very difficult to evaluate, primarily because both $a_{n,j}$ and $\boldsymbol{\phi}_{n,j}$ depend on $\omega_{n,j}$ and the latter has a very complicated dependence on k [see Eq. (7a)]. Some simplifications are clearly in order. To begin with, we consider only the case where F is switched on but otherwise steady and the initial conditions are $u = v = h = 0$. This amounts to seeking the response to a step function in time; the response to other time structures may be found by a convolution. In this case (45) applies so that (46) becomes

$$\mathbf{u}_{n,j} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{b_{n,j}(k)}{i\omega_{n,j}(k)} \; [1 - e^{-i\omega_{n,j}(k)t}] \, e^{ikx}\, \boldsymbol{\phi}_{n,j}(k,y) dk \; . \tag{47}$$

Let us for the moment restrict ourselves to *forcings* $\mathbf{F}$ which do not have singularities which contribute to the integral (47), except perhaps at $k = 0$; we thus focus on the nature of the *response*. Among the forcings with this property are those with appreciable amplitude only at long wavelengths, $|k| << 1$ (cf. Lighthill, 1969). This is a useful parameter range to consider because the baroclinic length scale L is small compared to the scale of the atmospheric forcings. Since the Kelvin wave is nondispersive ($\omega = k$) (47) is easily calculated for $n = -1$. It will be sufficient for our purposes to evaluate the asymptotically less important inertia-gravity ($j = 1{,}2$) versions of (47) by using the long wave approximations

$$\omega_{n,1} \approx (2n + 1)^{\frac{1}{2}},\; \omega_{n,2} \approx -(2n + 1)^{\frac{1}{2}} \; .$$

The integrals for the Rossby mode case are still complicated but we will obtain approximate asymptotic solutions using the results found in Sec. 3. [Note that for

Rossby waves the dispersion relation (7a) is well approximated by $\omega_{n,3} = -k/(k^2 + 2n + 1)$ which has the same form as the barotropic dispersion relation (13).] In particular, only the hyperbolic part of the dispersion relation

$$\omega_{n,3} \approx -k(2n+1)^{-1} \tag{48}$$

is needed to give the dominant asymptotic response. It must be remembered that any step function response should be interpreted with its appropriate Airy function smoothing as in Fig. 4. [Compare Eqs. (22) and (23); also see Eq. (35d)]. Also, in the infinitely deep barotropic ocean the precursors can travel infinitely fast, whereas the baroclinic system has a finite depth H so that no signal will travel faster than $(gH)^{\frac{1}{2}}$ (=1 in nondimensional units). We also note that because of the way that $b_{n,3}$ and $\boldsymbol{\phi}_{n,3}$ depend on k the directly forced part of the integral (47) has a more complicated form that the barotropic analogs, Eqs. (18b), (34a), considered in Sec. 3. [The directly forced part arises from the "1" in the square brackets in (47) as opposed to the $e^{-i\omega_{n,3}t}$ term which gives the free motions.] However, it is still the case that the only contribution to this part of (47) comes from the pole at $k = 0$ so that (48) may be used.

Hence the dominant asymptotic response that arises from point A ($\omega = k = 0$) may be obtained by using (48) in (47). We may write $b_{n,3} \approx k^{-1}\, r_n(k)$ where r_n depends only on the vector of forcings $\mathbf{F}$. From (41)

$$r_n(k) = e_n(k) - (2n+1)^{-1}\, d_n(k) \ . \tag{49a}$$

In addition, $\boldsymbol{\phi}_{n,3} \approx k\, \mathbf{R}_n + i\, k^2\, (2n+1)^{-1}\, \mathbf{V}_n$, where

$$\mathbf{R}_n(y) = [4n(n+1)]^{-1}\, [(2n+1)\, \mathbf{M}_n - \mathbf{W}_n] \tag{49b}$$

is a multiple of the $k = 0$ Rossby mode with $v \equiv 0$ and u and h in geostrophic balance. Eq. (47) now becomes, with $\xi = x + (2n+1)^{-1}\, t$

$$\begin{aligned} \mathbf{u}_{n,3}\,(x,y,t) &= 2\pi \int_{-\infty}^{+\infty} \frac{r_n(k)}{ik}\, [e^{ik\xi} - e^{ikx}] \\ &\quad \times \{(2n+1)\, \mathbf{R}_n(y) + ik\, \mathbf{V}_n(y)\}\, dk \\ &= (2n+1)\, \mathbf{R}_n(y) \int_x^{\xi} r_n(x')\, dx' + \mathbf{V}_n(y)\, [r_n(\xi) - r_n(x)] \ . \end{aligned} \tag{50}$$

Since the long wave assumption is equivalent to asserting that $|\partial\mathbf{F}/\partial x| << |\mathbf{F}|$ it follows that $\partial r_n/\partial x << r_n$. Hence the long wave Rossby modes described by (50) have v components which are much smaller than their u and h components. Further, these modes have u and h in geostrophic balance. Eq. (50) also shows that the total response $\mathbf{u}_{n,3}$ at a point x consists of a locally forced part and a part that has propagated in from a point $x' = x + t(2n+1)^{-1}$. The latter is a synthesis of free waves that arises from the need to satisfy the initial conditions (recall the discussion preceding Eq. (45)).

5. The response to x-independent forcings

The procedure used in obtaining (50) is equivalent to obtaining the leading asymptotic behavior of (47) from the pole at $k = 0$. If $\mathbf{F}$ is truly independent of x then the $b_{n,j}$ are all proportional to $\delta(k)$ and the solution is exact. It may be written as a sum

$$\mathbf{u} = \mathbf{u}_K + \sum_{n=0}^{\infty} \mathbf{u}_{n,I} + \sum_{n=1}^{\infty} \mathbf{u}_{n,A} \ . \tag{51a}$$

where, with $m = (2n + 1)^{\frac{1}{2}}$

$$\mathbf{u}_{n,I} = \mathbf{u}_{n,1} + \mathbf{u}_{n,2} = [m^{-3}\, d_n \sin mt + m^{-2}\, g_n(1 - \cos mt)]\mathbf{W}_n(y)$$

$$+ [m^{-2}\, d_n (\cos mt - 1) + m^{-1}\, g_n \sin mt]\mathbf{V}_n(y) \tag{51b}$$

$$\mathbf{u}_{n,A} = \mathbf{u}_{n,3} = t\, r_n\, \mathbf{R}_n(y) \tag{51c}$$

$$\mathbf{u}_K = \mathbf{u}_{-1} = t\, b_{-1}\, \mathbf{W}_{-1}(y) \ . \tag{51d}$$

The subscripts I, and K denote inertia-gravity and Kelvin modes, respectively, while A denotes the Rossby mode contribution from point A. (We exploit the one-to-one correspondence between the eigenfunctions $\boldsymbol{\phi}_{n,j}$ and the free waves $\boldsymbol{\phi}_{n,j} \exp i(kx - \omega_{n,j}\, t)$ to carry over the free wave nomenclature.) The solution is of course qualitatively similar to the approximate long wave solution. The inertia-gravity modes of the latter are exactly like (51b) except that d_n and g_n are functions of x; the similarities in the Rossby modes are evident from a comparison of (51c) and (50); the Kelvin mode would have $\int_{x-t}^{x} b_{-1}(x')dx'$ in place of tb_{-1} in (51d).

We now wish to consider the solution (51) in some detail. Suppose first that the forcing consists solely of an east-west wind stress (i.e., $F = F(y)$; $G = Q = 0$). The response consists of secularly growing u and h fields, plus a steady v component:

$$(u, v, h) = (t\, U(y), V(y), t\, H(y)) \ . \tag{52}$$

In addition, there is a series of inertia-gravity waves which are required to satisfy the initial condition $v = 0$. From a mathematical point of view the solution is best explained in terms of Eq. (45) and the dispersion diagram Fig. 1. The forcing function has zero frequency and zonal wavenumber so it lies at the origin of the dispersion diagram. This is a point of resonance for the Rossby and Kelvin waves resulting in a secularly growing solution. The discussion following (50) shows that this secularly growing part of the solution may be viewed as the sum of a locally forced part which goes like $-x$, and a propagating free wave part (required by the initial conditions) which goes like $x + t\,(2n+1)^{-1}$. This point of view clearly reveals the way in which the secular growth depends on the assumption that the ocean is unbounded. The propagating part of the solution which arrives at the point x at time t

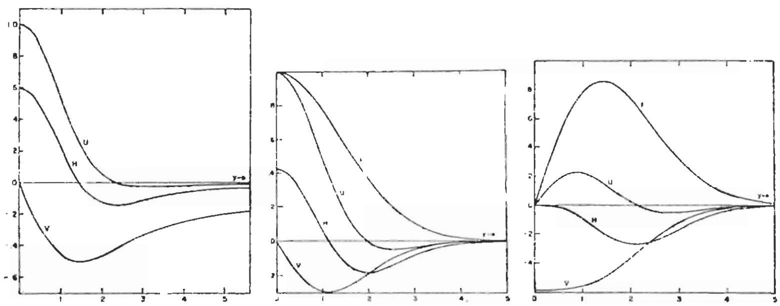


Figure 7. Unbounded baroclinic response to a westerly wind $F \neq 0$, $G = Q = 0$. (a) $F(y) = 1$; (b) $F(y) = \exp(-y^2/4)$; (c) $F(y) = y \exp(-y^2/4)$.

started at $x' = x + t\,(2n+1)^{-1}$ at $t = 0$; as $t \to \infty$ $|x'| \to \infty$. Secular growth is maintained by waves which propagate in from infinity. A boundary at $x'' < x'$ would cut off the source of these waves and prevent this growth from continuing.

The steady part $v = V(y)$ is the forced response of the inertia-gravity modes at $k = 0$ (not on resonance), while the oscillating part is made up of inertia-gravity waves with $k = 0$. At the equator the Coriolis term is absent and the wind stress causes a steady acceleration in the direction of wind $U(0) = F(0)$. Note that x-independent zonal winds off the equator do not contribute to the zonal current at the equator. Note too that U and H of (52) are in geostrophic balance so that dH/dy is zero at the equator. As a general rule, the time growing part of the response tends to be more equatorially confined than is the (smooth) zonal wind system that forces it, while the steady v field asymptotes to the wind drift value $-F(y)/y$ as y increases and the Coriolis balance becomes dominant. An illustration of these features is given in Fig. 7. (The solution for the case $F = 1$ was first obtained by Yoshida, 1959). A somewhat special instance of these general rules arises when F is the simplest continuous odd function, i.e. $F(y) = y$. In this case $U = H = 0$ and $V = -1$ everywhere.

The response to a purely meridional wind stress ($G = G(y)$, $F = Q = 0$) is very different, consisting of steady u and h components and a series of inertia-gravity waves of zero zonal wavenumber which are required to satisfy $u = h = 0$ at $t = 0$. There is no steady (or other non-oscillating) v component. Extra-equatorially, the steady part of the solutions $U_s(y)$, $H_s(y)$ tends to approach the wind drift if G is smooth; i.e.

$$\text{as } |y| \to \infty \;\; U_s(y) \to G(y)/y,\; H_s(y) \to 0 \;.$$

At the equator the Coriolis term vanishes and the wind stress is balanced by the "sea-surface setup"—that is, by dH_s/dy. U_s need not vanish at the equator even if the stress does. The solution is of course not geostrophic. Mathematically speaking, the response comes from the inertia-gravity modes at the points on the axes $k = 0$

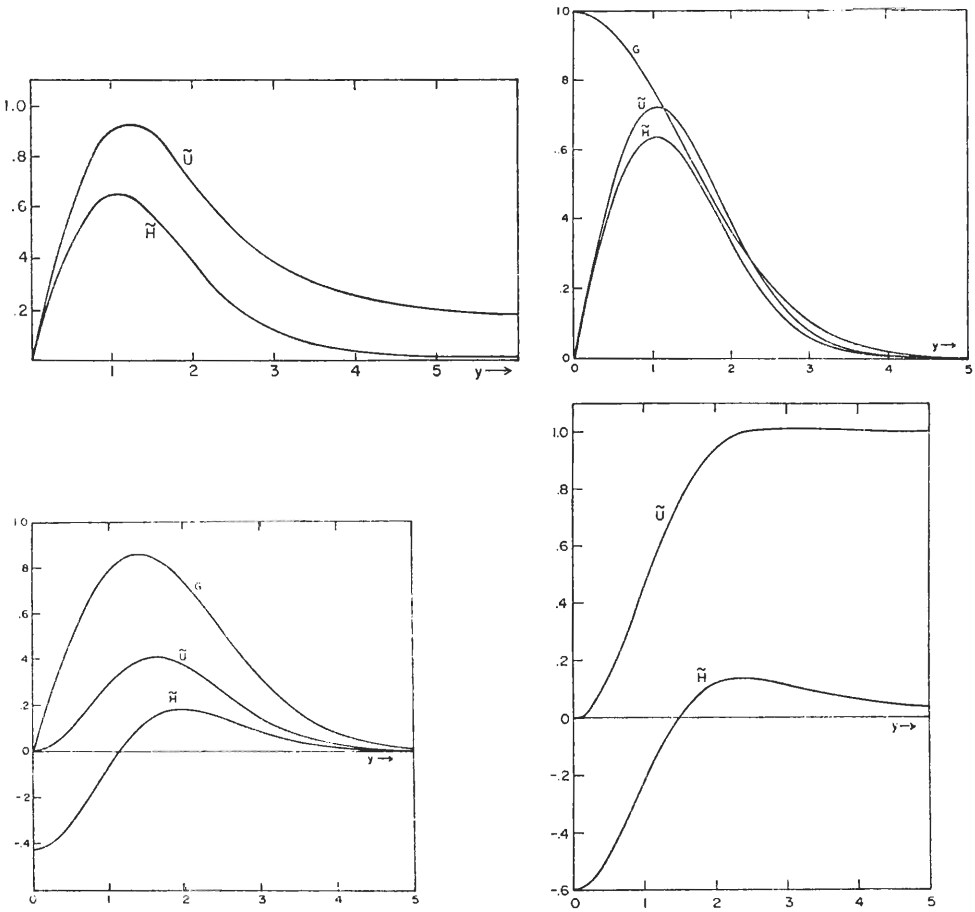


Figure 8. Unbounded baroclinic response to a southerly wind $F = 0$, $G \neq 0$, $Q = 0$. (a) $G(y) = 1$; (b) $G(y) = \exp(-y^2/4)$; (c) $G(y) = y \exp(-y^2/4)$; (d) $G(y) = y$.

of Fig. 1. While the forcing is again at $k = 0$, $\omega = 0$, there is no resonant response in the Rossby and Kelvin modes because these modes have no meridional component at $k = 0$. Fig. 8 shows $U_s(y)$ and $H_s(y)$ for a number of cases. $U_s = 0$ at the equator and asymptotes to $1/y$ as $y \to \infty$; these constraints determine its general shape.

With a meridional wind stress the only meridional motions are those associated with the inertia-gravity waves. These are the motions which accomplish the redistribution of mass required to set up a non-zero $H_s(y)$. The partition of energy between the steady part of the solution U_s, H_s and the inertia-gravity waves thus depends solely on the initial conditions; for $u = v = h = 0$ at $t = 0$, the energy is equally divided. By contrast the amount of energy going into the secularly growing solution for $F \neq 0$ is independent of the initial conditions. The difference between the two cases is reflected in the fact that if one tries a solution of the form (52) in Eqs. (1) for $G = Q = 0$ and $F = F(y)$ the solution is determined; whereas with $F = Q = 0$ and $G = G(y)$ choosing the forms $U = U_s\,(y)$, $H = H_s\,(y)$, $v = 0$ leads to an indeterminate set of equations.

Finally, we remark that the response to only a heating function forcing ($F = G = 0$, $Q = Q(y)$) has the same general components as the case of a zonal wind stress; that is, a form like (52) plus inertia-gravity waves. Of course, on a less superficial level of description, it is very different. For example, the response to $Q = 1$ is simply $u = v = 0$ and $h = t$; no gravity waves are excited.

From (42a) and (51b) it may be seen that the amplitude of the gravity wave part of the response depends on the gradients of the external heating Q_y and the latitude times the zonal wind stress yF. Hence a large-scale heat source excites little gravity wave response, while a large scale zonal momentum source has a large gravity wave response. This is consistent with the well-known results for a mid-latitude f-plane or β-plane (e.g., Blumen, 1972). In addition, consideration of the coefficients in (51) shows that the response to a large-scale heating source tends to be no more equatorially confined than the source itself; the case $Q \equiv 1$ is the simplest example of this.

Before considering forcings which vary with x we wish to consider briefly a problem closer to the classical Rossby geostrophic adjustment problem: the response when the forcing is impulsive in time $[\mathbf{F}(y)\ \delta\ (t)]$. Since our solution is for a step function forcing, and since the left-hand side of (36) is independent of time, the response to an impulse forcing may be obtained by differentiating the solution (51) with respect to t. Letting $m = (2n + 1)^{\frac{1}{2}}$ the result is

$$\begin{aligned} \mathbf{u} = b_{-1}\,\mathbf{W}_{-1} + \sum_{n=1}^{\infty} r_n\,\mathbf{R}_n \\ + \sum_{n=0}^{\infty} \{[m^{-2}\,d_n \cos mt + m^{-1}\,g_n \sin mt]\,\mathbf{W}_n \\ + [-\,m^{-1}\,d_n \sin mt + g_n \cos mt]\,\mathbf{V}_n\}\ . \end{aligned}$$

As compared with the longer time forcing, the gravity wave response is enhanced and the planetary mode response is reduced. This is to be expected, since the impulsive forcing has, in effect, a higher frequency. If the forcing consists only of a meridional wind stress, there is no response in steady or slowly varying planetary modes; only free inertia-gravity waves are excited. In this case $(\mathbf{u}_{n,I},\ \mathbf{u}_{n,I}) = g_n{}^2$ so that the energy in the nth mode depends only on the projection of the forcing function onto ψ_n and not on the nature of the medium; in particular, the more equatorially confined modes are not preferentially excited. The response to a zonal wind stress F or a heat source Q consists of free inertia-gravity waves together with a steady response $(u, v, h) = (U(y), O, H(y))$ where U and H are the same functions which appear in (52). The latter have been discussed above. The nth inertia-gravity wave has $(2n+1)^{\frac{1}{2}}$ times the amplitude of the nth wave in response to the step function forcing, but it is still true that the amplitude depends on $yF + dQ/dy$. These waves are excited by a large scale momentum impulse, but are barely present in response to a smooth large scale heat source. (The response to $Q \equiv 1 \times \delta(t)$ is simply $h = 1$ for $t > 0$). Again, this is qualitatively similar to the mid-latitude result.

6. The response to step function forcings

In this section we solve for the response to a forcing which is again turned on at $t = 0$ and steady thereafter but whose zonal dependence is also a step function; i.e.

$$\mathbf{F}(x, y, t) = \mathbf{F}(y)\, H(t)\, H(x) \ . \tag{53}$$

(The response to a forcing with arbitrary zonal structure could be obtained by a convolution). The solution of this section provides information about the effect of switching off a forcing in space that will prove very useful in understanding the effects of meridional boundaries. (See Cane and Sarachik, 1976). The solution may be obtained by applying the methods of Sec. 3 and 4. In doing so we note that for a meridional wind stress G proportional to $H(x)$ the curl of the wind stress is proportional to $\delta(x)$. This promotes the asymptotic importance of contributions arising from the points D where $k \to -\infty$ and $\omega \approx -k^{-1} \to 0$ [cf. Eqs. (19c) and (35b)]. These contributions may be calculated by noting that as $k \to -\infty$

$$\frac{b_{n,3}}{i\omega_{n,3}}\, \boldsymbol{\phi}_{n,3} \approx g_n \{\mathbf{M}_n(y) + ik\, \mathbf{V}_n(y)\}$$

so that the contribution to (47) from point D is

$$\mathbf{u}_{n,3}(D) \approx -(2\pi)^{-1} \left[\mathbf{M}_n(y) + \mathbf{V}_n(y) \frac{\partial}{\partial x} \right] \int_{-\infty}^{+\infty} g_n\, e^{i(kx + t/k)}\, dk \ . \tag{54}$$

Hence, with $g_n(x) = H(x)$ so $g_n(k) = (ik)^{-1}$

$$\mathbf{u}_{n,3}(D) \approx -\left[\mathbf{M}_n(y) + \mathbf{V}_n(y) \frac{\partial}{\partial x} \right] J_0(2\sqrt{xt})$$

$$= -\mathbf{M}_n(y)\, J_0\,(2\sqrt{xt}) + \mathbf{V}_n(y) \left(\frac{t}{x}\right)^{\frac{1}{2}} J_1\,(2\sqrt{xt}) \ . \tag{55}$$

The similarity of (55) to (19c) is even more striking when one realizes that $\psi = \psi_n(y)\, J_0\,(2\sqrt{xt})$ is a stream function for $\mathbf{u}_{n,3}(D)$ with $(u, v, h) = (-\psi_y, \psi_x, y\psi)$. We note further that (55) can hold only for $0 < x < C_n t$ where C_n is the maximum eastward group velocity, i.e. the value of $\frac{\partial \omega}{\partial k}$ at point C where

$$C_n = \max \frac{\partial \omega}{\partial k} \approx [8(2n+1)]^{-1} \ . \tag{56}$$

In interpreting the asymptotic solution (55) we recognize that there is no discontinuity in the full solution because the asymptotic term (55) "blends" in with the asymptotically less important contribution arising at the wavefront where $x = C_n t$ (cf. Eqs. 20, 35c).

By making use of previous results, especially (50) and (55), we are now able to write down the total response to the forcing (53). Before doing so we digress to indicate another method of obtaining the solution. Apply the x-independent solution (51) for $x > 0$ and take its Laplace transform in time, thus going from the time to the frequency domain. Then find the free solutions needed to eliminate the discontinuities in u and h at $x = 0$. Finally, transform back to the time domain. (For details see Cane, 1975). This procedure is similar to that used by Lighthill (1969) and later by Anderson and Rowlands (1976a, 1976b) to calculate reflections at boundaries.

The (leading asymptotic terms of the) response to the step function forcing (53) may be written as a sum

$$\mathbf{U} = \mathbf{U}_K{}^{(1)} + \mathbf{U}_K{}^{(2)} + \sum_{n=0}^{\infty} \{\mathbf{U}_{n,I}{}^{(1)} + \mathbf{U}_{n,I}{}^{(2)} + \mathbf{U}_{n,I}{}^{(3)}\}$$

$$+ \mathbf{U}_0{}^{(3)} + \sum_{n=1}^{\infty} \{\mathbf{U}_{n,A}{}^{(1)} + \mathbf{U}_{n,A}{}^{(2)} + \mathbf{U}_{n,A}{}^{(3)} + \mathbf{U}_{n,D}{}^{(3)}\} \tag{57}$$

where, with $m = (2n + 1)^{\frac{1}{2}}$,

$$\mathbf{U}_K{}^{(1)} = \mathbf{U}_K\, H(x),\ \mathbf{U}_{n,I}{}^{(1)} = \mathbf{U}_{n,I}\, H(x),\ \mathbf{U}_{n,A}{}^{(1)} = \mathbf{U}_{n,A}\, H(x) \tag{58a}$$

$$\mathbf{U}_{n,I}{}^{(2)} = -d_n\, T\,(0; x; \tfrac{1}{2}\, m^{-2}\, t)\, \{m^{-3} \sin mt\, \mathbf{W}_n + m^{-2} \cos mt\, \mathbf{V}_n\} \tag{58b}$$

$$\mathbf{U}_{n,I}{}^{(3)} = g_n\, T\,(0; x; \tfrac{1}{2}\, m^{-2}\, t)\, \{m^{-2} \cos mt\, \mathbf{W}_n + m^{-1} \cos mt\, \mathbf{V}_n\} \tag{58c}$$

$$\mathbf{U}_K{}^{(2)} = -b_{-1}\, T\,(0; x; t)\, (t-x)\, \mathbf{W}_{-1} \tag{58d}$$

$$\mathbf{U}_{n,A}{}^{(2)} = r_n\, T\,(-m^{-2}\, t; x; 0)\, \{[t + m^2\, x]\, \mathbf{R}_n + \mathbf{V}_n\} \tag{58e}$$

$$\mathbf{U}_{n,A}{}^{(3)} = -g_n \left[\frac{2n+1}{4n(n+1)}\right] T(-m^{-2}\, t; x; 0)\, \mathbf{R}_n \tag{58f}$$

$$\mathbf{U}_0{}^{(3)} = g_0\, T\,(0; x; t) \Big\{ -J_0[2(x\{t-x\}^{\frac{1}{2}}]\, \mathbf{M}_0 + \left[\frac{t-x}{x}\right]^{\frac{1}{2}} \quad J_1\, [2(x\{t-x\})^{\frac{1}{2}}]\, \mathbf{V}_0 \tag{58g}$$

$$\mathbf{U}_{n,D}{}^{(3)} = g_n\, T\,(0; x; C_n\, t) \Big\{ -J_0\, (2\sqrt{xt})\, \mathbf{M}_n + \left[\frac{t}{x}\right]^{\frac{1}{2}} \quad J_1\, (2\sqrt{xt})\, \mathbf{V}_n \Big\}\ . \tag{58h}$$

The vectors without superscripts appearing in (58a) are the x-independent solutions (51). $T(x_L; x; x_R) = H(x - x_L)\, H(x_R - x)$ is the "top hat" function. C_n is the maximum eastward group velocity [cf. Eq. (56)]. The terms superscripted with a "1" are the forced solution to the x-independent problem, applied for $x > 0$. The other terms are free solutions required to satisfy the jump conditions at $x = 0$. Those superscripted with a "2" are needed if $F(y)$ or $Q(y)$ is non-zero while those with a "3" are required if $G(y)$ is non-zero. As before, the subscripts K and I are used to denote Kelvin and inertia-gravity modes, respectively, while A denotes the long wave, and D the short wave, Rossby mode contributions. Note that the mixed mode ($n = 0$) appears in the solution both in the role of a gravity wave at high frequencies and in the role of a Rossby wave for low frequencies.

The presence of the function T in (58b)-(58h) shows that all the free wave solutions exist only in a region bounded by $x = 0$ where the forcing switches on, and $x = C_g\, t$, the farthest point to which a disturbance traveling at the group velocity C_g will have propagated in time t. In the interior of the region each of the terms given by Eqs. (58b)-(58g) is a solution to the unforced Eqs. (36); that is, a free solution of the shallow water equations. This is almost true of (58h) as well but here the complete solution is a sum of terms of the form $[x/t]^{v/2}\, J_v\, (2\sqrt{xt})$; the form (58h) is clearly the leading term in the large t expansion. Furthermore, the total solution (58) does satisfy the proper jump conditions at $x = 0$.

It might then appear that this solution, with the exception noted in the previous paragraph, is not only the large t asymptotic solution, but is in fact the complete solution. This is almost true. However, some of the free solutions fail to satisfy Eqs. (36) at the wave front where $x = C_n\, t$. Consider the term $\mathbf{U}_{n,A}{}^{(2)}$ defined by (58e), for which $C_g = -\,(2n + 1)^{-1}$. To the left of $x = -t(2n + 1)^{-1}$ the v component of this term is zero, while to the right of this point $v = r_n\, \psi_n(y)$. The equations do not admit this discontinuity. More formally, if (58e) is substituted into the Eqs. (36) with $\mathbf{F} = 0$ there is a remainder $r_n\, \psi_n(y)\, \delta(x + t\,(2n{+}1)^{-1})$ in the second of these equations. This term arises from taking the derivative of $H(t(2n + 1)^{-1} + x)$ when $\partial v/\partial t$ is calculated. If (36) were a pure, nondispersive hyperbolic system this discontinuity at $x = C_g t$ would be admissible. As it is, precursors exist ahead of the wave front at $x = C_g t$ which have the effect of smoothing the discontinuity. These precursors can have a velocity equal to the fastest signal in the system (a velocity of one in our scaling). The point $x = C_g t$ which travels with the group velocity still defines a wave front because it marks the farthest penetration of a signal with substantial amplitude (see Brillouin, 1960, Chap. 3). We have discussed these issues in greater detail in Sec. 3 in connection with the qualitatively similar barotropic vorticity equation. The point to be emphasized here is that (57) is an excellent approximation to the exact solution.

A zonal wind stress or heating source also gives rise to short wave Rossby terms of the form

$$\mathbf{U}_{n,D}{}^{(2)} = -e_n\, T(0; x; C_n t) \left\{ \left(\frac{x}{t} \right)^{\frac{1}{2}} J_1\,(2\sqrt{xt})\, \mathbf{M}_n + J_0\,(2\sqrt{xt})\, \mathbf{V}_n \right\} . \tag{59}$$

Such terms are asymptotically of lower order than the terms retained in (57). [Compare the barotropic example Eq. (35b)]. Note that the u and h components of (59) go to zero at $x = 0$ so no discontinuities in these components are introduced. The v component, while also asymptotically small for $x >> 0$ becomes $-e_n \psi_n$ at $x = 0$ and thus in a narrow region $0 < x \lesssim 0(t^{-1})$ may be comparable to the v component of the retained terms in (57).

We now wish to show how (57) may be obtained in a simpler and more intuitive manner. We again take the solution to the x-independent problem given by (51) to apply for $x > 0$; this gives (58a) determining $\mathbf{U}^{(1)}$. We then seek the free solutions needed to make the jumps in u and h zero at $x = 0$. First consider the inertia-gravity wave terms of $\mathbf{U}^{(1)}$. These are free waves with zero east-west wavenumber; they lie on the $k = 0$ axis and the upper curves of the dispersion diagram, Fig. 1. The nth such mode propagates eastward with group velocity $C_g = [2(2n + 1)]^{-1}$ so that the motion present at a point x at time t must have originated at a point $x_0 = x - C_g t$ at $t = 0$. No motion was excited at points $x_0 < 0$ (since the forcing is present only for $x > 0$). So at time t there will be no nth inertia-gravity mode present at $x < C_g t$. Since (58a) specifies the existence of these modes for all $x > 0$, Eqs. (58b) and (58c) are needed to cancel them in the region $0 < x < C_g t$.

Now consider the planetary wave part of $\mathbf{U}^{(2)}$ defined by (58d,e). As remarked above, each Rossby and Kelvin mode piece of the solution $\mathbf{U}^{(1)}$ was forced on resonance at $\omega = 0$; $k = 0$ and may be viewed as consisting of a locally forced part varying like $C_g^{-1}\, x$ and a propagating part which goes like $t - C_g^{-1}\, x$. Only the latter violates the jump conditions. But since each such propagating part is a free solution of Eqs. (36), the jump conditions can be matched by considering how these modes propagate through $x = 0$. The nth Rossby mode may be thought of as a synthesis of Rossby waves with amplitude $\delta\,(k)$. It has a group velocity of magnitude $(2n + 1)^{-1}$ to the west. Each such mode continues to propagate westward beyond $x = 0$, so we must add these propagating solutions for $x < 0$; this gives (58e). Like the $k = 0$ inertia-gravity waves the Kelvin wave propagates eastward; its group velocity is 1. Since the forcing extends only as far to the west as $x = 0$ no Kelvin waves originate at points $x < 0$. The term $\mathbf{U}_K{}^{(2)}$ of (58d) serves to cancel the propagating part of the Kelvin response at points $x < C_g t = t$.

The steady forced response to a meridional wind stress does not have a simple relation to free wave solutions. As a result the terms (58f, g, h) needed to correct for the jump it introduces are more complicated than the other corrections. These terms are syntheses of free waves with u and h components, that are independent of t at $x = 0$. Eq. (58f) describes a non-dispersive long wave form which rapidly propagates westward. Eq. (58h) is a synthesis of the short dispersive waves which lie in the lower left-hand corner of the dispersion diagram ($\omega << 1, -k >> 1$). All of these $n > 0$ eastward propagating modes are essentially trapped to the discontinuity at $x = 0$; they have very low group velocities, so the "disturbance" moves away from $x = 0$ very slowly. The mixed mode (Eq. 58g) also has most of its energy trapped at the boundary, although there is appreciable amplitude at the rapidly travelling wave front at $x = t$. Since each of the modes (58g, h) has its v and h components in geostrophic balance, the entire meriodional current is geostrophic. The amplitude of the v component of each mode grows linearly in time at $x = 0$; the mode also becomes thinner in time. Thus the vorticity of the flow increases in response to the constant input of vorticity from the wind stress.

A few additional remarks about the general shape of the solution (57) may be made. As noted above inertia-gravity waves with meridional index n are absent in the (expanding) region $x < t/(4n + 2)$. Eqs. (58a) and (58d) show that the Kelvin mode is absent for $x < 0$ and has amplitude proportional to x for $0 < x < t$ and to t for $x > t$. Similarly, Eqs. (58a) and (58e) show that the nth Rossby mode generated in response to a zonal wind stress F or a heat source Q is absent for $x < x_n \equiv -t/(2n + 1)$, has a non-zero v component in the region bounded by $x = 0$ and $x = x_n$, and its u and h components have amplitude proportional to $t-(2n + 1)\ x$ for $x_n < x < 0$ and proportional to t for $x > 0$. Thus at a fixed point $x < 0$ the more equatorially confined modes (i.e., those with small n) are enhanced relative to the less equatorially confined ones. The more equatorially confined modes are also present over a large area since they propagate more rapidly. In the previous section we noted the response to an x-independent F or Q forcing would have $u = 0$ at the equator if $F = 0$ there. With a forcing which is non-zero to the right of $x = 0$, u may be non-zero at the equator between $x = -t/3$ and $x = t$, even if there is no wind stress at the equator. The features described in this paragraph are shown schematically in Fig. 9.

The westward propagating modes (58f) that are present if there is a non-zero meridional wind stress have no v component. Inertia-gravity waves aside, the only meridional flow generated by the step function G forcing is the ever thinning current just to the right of $x = 0$; this is the sum of the modes (58g, h). Each of these modes has v growing secularly at $x = 0$ while u and h are constant. Geostrophy in the downstream direction is maintained by having the current become thinner at a rate which makes the cross-stream height gradient increase linearly in time at $x = 0$. The zonal flow (exclusive of inertia-gravity waves) is the same as in the x-independent case to the right of $x = 0$; Eq. (58f) shows that while the more equatorially confined modes extend the furthest to the west of $x = 0$, their amplitudes are not enhanced relative to the less confined modes. The features described in this paragraph are illustrated in Fig. 10.

7. Summary and discussion

If the equations of motion are linearized and viscous effects neglected, and the ocean is assumed to be flat-bottomed and stably stratified, then the vertical de-

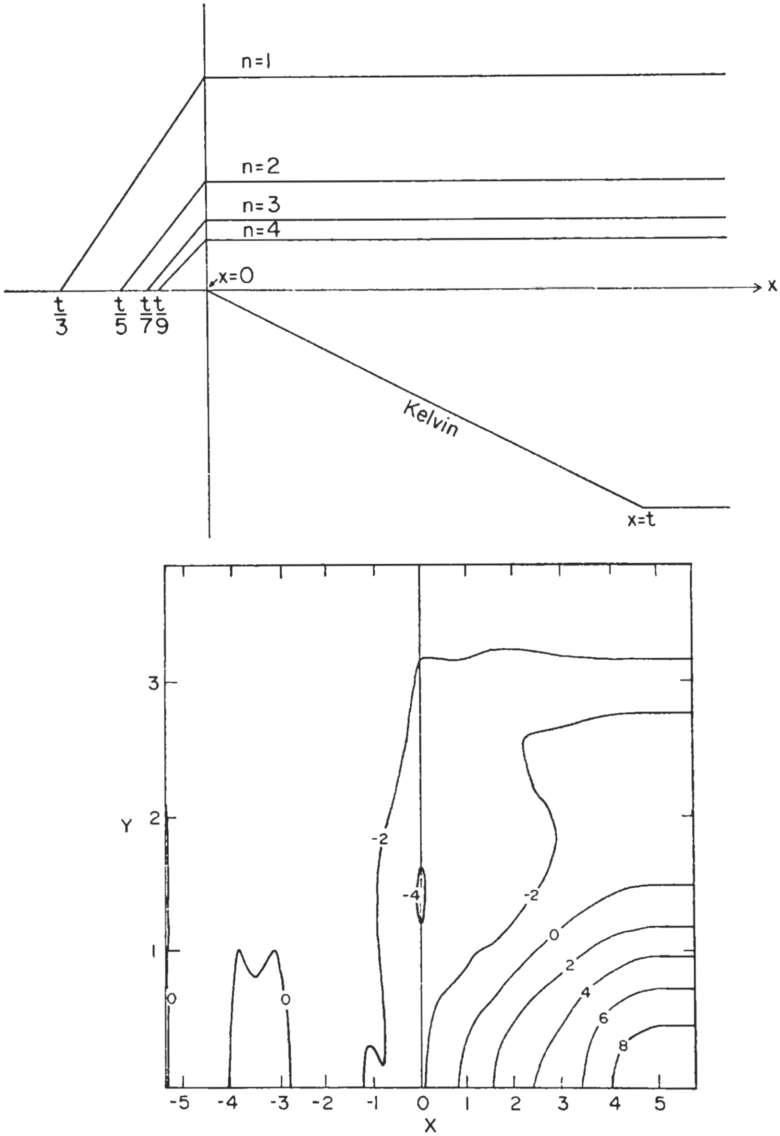


Figure 9. Unbounded baroclinic response to a step function westerly wind. (a) The longitudinal dependence of the low n modes. (b) The layer depth h at $t = 5$ for the case $F = H(x)$, $G = Q = 0$.

pendence may be analyzed into uncoupled vertical modes. The temporal and horizontal spatial dependence of each such mode is described by the linear inviscid shallow water equations (1). The subject of this work has been the solution of these equations for an unbounded ocean, with particular attention to motions forced in the vicinity of the equator.

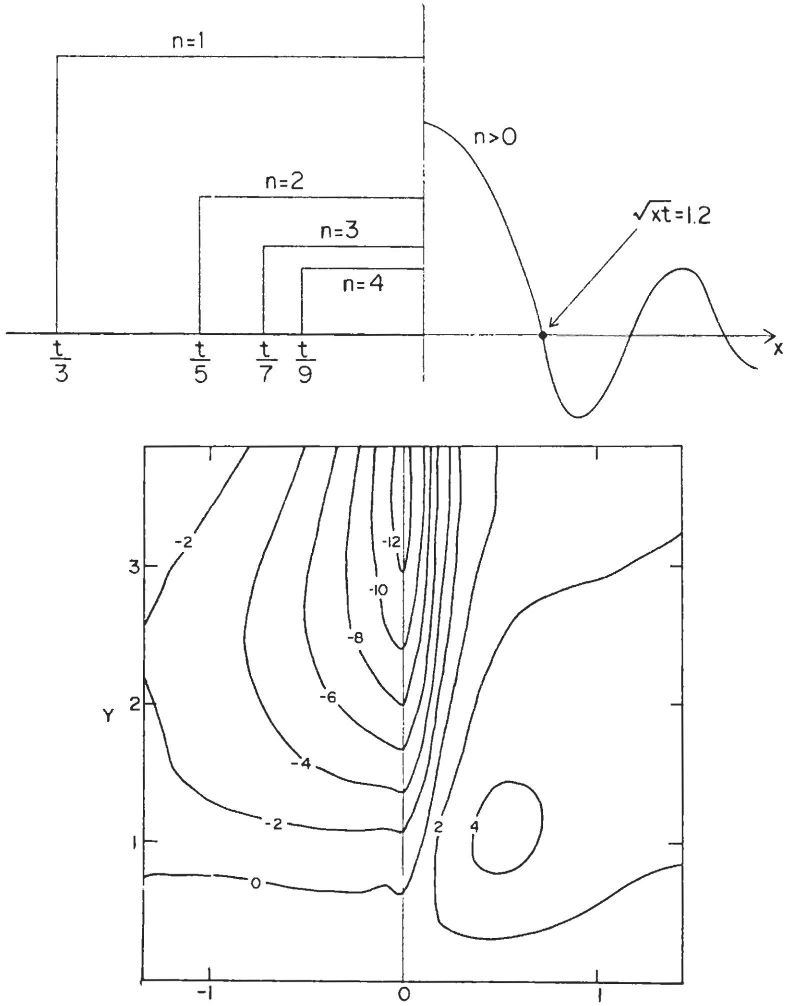


Figure 10. Unbounded baroclinic response to a step function southerly wind (a). The longitudinal dependence of the low n modes. (b) The layer depth h at $t = 9$ for the case $G = H(x)$, $F = Q = 0$.

Our method is similar to finding a Greens function. The equations are first Fourier transformed in the zonal (x) direction, after which the meridional structure is expressed as an eigenfunction expansion. The (x,t) dependence of each mode must then be determined. This problem is specialized by considering only forcings on a resting ocean that are switched on at $t = 0$ and remain steady thereafter. The solutions obtained are then appropriate for considering the response to forcings with time scales long compared to 2 days; for example, the seasonal variations of the wind systems in the tropics. For these time scales the largest response is in the Kelvin and Rossby planetary modes rather than the inertia-gravity modes. Since the Kelvin mode is nondispersive its (x,t) behavior is readily calculated; the response of the Rossby modes is far more complex as a consequence of their dispersive nature. Before considering the baroclinic response, an asymptotic analysis is made of the structurally similar nondivergent barotropic Rossby modes. It is shown that the asymptotically dominant response arises from the point A (Figs. 1,2) in the dispersion curve where $\omega = k = 0$. This asymptotic dominance is attributed to its dual role as both a point of inflection and a point of maximum (westward) group velocity. The response is well approximated by considering the modes to be purely hyperbolic. The more complete analysis provides an interpretation of the discontinuities in the response resulting from such an approximation. The second most significant contribution to the response arises from the points D ($\omega \approx k^{-1} \to 0$) where the modes are highly dispersive. In particular, the meridional velocity component arising from this source must be considered in the vicinity of small scale variations in the zonal structure of the forcing. General characteristics of the response are described for the cases where the forcing is independent of x (Sec. 5) and where it is a step function of x (Sec. 6). It is generally sufficient to consider only the dominant asymptotic contributions to the response arising from the points A and D in the (ω,k) plane, contributions that have a relatively simple form. For the barotropic case the complete solution is well approximated by (becomes asymptotic to) these parts of the total response for times $t >> m\ \beta^{-1} \approx$ ½ day for wavenumber $m =$ (1000 km)$^{-1}$. For the nth baroclinic mode the corresponding time is $(2n + 1)^{\frac{1}{2}}T$ where $T \lesssim 2$ days is the baroclinic equatorial time scale [Eq. (2)]. While the asymptotic time increases with the meridional wavenumber n the distance the nth mode travels before becoming asymptotic decreases with n since the propagation speed of the mode varies like $(2n + 1)^{-1}$.

It is important to stress the limitations of the model equations that we have solved. The equatorial regions of real oceans are characterized by swift, narrow currents so that neglect of the inertial terms in the equations of motion cannot be rigorously justified. For example, taking a typical undercurrent velocity (80 cm sec^{-1}) and half-width (150 km) and evaluating f at $y = L$ gives a Rossby number of O(1). Based on a value of vertical viscosity of 25 cm^2 sec^{-1} (as observed by Williams and Gibson, 1974) and a scale depth of 50m, viscous effects will become non-negligible in O(20 days). With these limitations in mind, the linear inviscid shallow water equations (1) remain a useful theoretical model. They are analytically tractable and can give us a qualitative understanding of the time scales and structures of the ocean's response to atmospheric forcing. The analytic solutions to (1) provide conceptual tools that have proven helpful in understanding the response of a viscous, fully nonlinear model ocean that can only be studied by numerical methods (see Cane, 1975).

Before knowledge of the linear response can be usefully applied to either the real ocean or numerical models, the role of boundaries must be considered. By making use of the approximations developed in the present work the effects of boundaries may be calculated by elementary methods. This is the subject of the sequel to this paper, Cane and Sarachik (1976).

Acknowledgments. The authors would like to thank Jule Charney who was the original source for this work and who made it all possible; and Dennis Moore who unstintingly shared his knowledge of equatorial waves with us in many conversations over the years. Conversations with R. S. Lindzen and K. K. Tung have been also very helpful. This work was supported by NSF Contract GA-28724X and NASA Grant NGR 22-009-727 at MIT, by NSF Grant GA-37116X at the 1974 Woods Hole Summer Geophysical Fluid Dynamics Program, and by NSF Grants GA-33990X and NSF-ATM-75-20156 at Harvard. This work was begun when one of us (E.S.S.) was an NSF Postdoctoral Fellow at MIT and completed when the other (M. A. C.) was an NAS/NRC Research Associate at GISS.

APPENDIX

Orthogonality and completeness of the eigenfunctions for the shallow water equations

(*a*) *Orthogonality*

From the orthogonality of the Hermite functions ψ_n, it follows that

$$(\phi_{n,j}(k,y),\ \phi_{m,l}(k,y)) = 0 \text{ if } n \neq m \ .$$

It remains to show that

$$A_{jl}^{n} \equiv (\phi_{n,j},\ \phi_{n,l}) = 0 \text{ if } j \neq l \ .$$

From the definition of the scalar product (37):

(i) $A_{jl}^{0} = 1 + \dfrac{1}{\omega_{01}\,\omega_{02}}$ for $j \neq l$; since ω_{01}, ω_{02} satisfy

$$\omega^2 - k\,\omega - 1 = 0$$

their product is -1 and

$$A_{jl}^{0} = 0 \text{ for } j \neq l \ .$$

(ii) For $n > 0$ $A_{jl}^{n} = \omega_j\,\omega_l\,[2n + 1 + \omega_j\,\omega_l]$

$$+\, k(\omega_l + \omega_j) + k^2\,[2n + 1 + k^2\,(\omega_l + \omega_j)^2 + 2\,\omega_l\,\omega_j]$$

(where we have written ω_j for $\omega_{nj}(k)$, etc.). Now ω_j and ω_l satisfy the dispersion relation (7a); let the third root be ω_*. Making use of the relations $\omega_j + \omega_l + \omega_* = 0$ and $\omega_l\,\omega_j\,\omega_* = k$ we obtain

$$A_{jl}^{n} = \frac{k}{\omega_*}\left[2n + 1 + \frac{k}{\omega_*}\right] - k\,\omega_* + k^2\left[2n + 1 + k^2 - \omega_*^2 + \frac{2k}{\omega_*}\right]$$

$$= \frac{k}{\omega_*}\left[2n + 1 + \frac{k}{\omega_*} - \omega_*^2\right] + k^2\left[\frac{k}{\omega_*}\right] = 0$$

since ω_* satisfies (7a).

(*b*) *Completeness*

We wish to show that if all the components of the vector $\mathbf{F} = (F, G, Q)$ have expansions of the form

$$\sum_{n=1}^{\infty} a_n(k,t)\, \psi_n(y)$$

then $\mathbf{F}$ has an expansion of the form $\sum b_{n,j}\, \Phi_{n,j}$.

It is sufficient to show that, for all n, $(1, 0, 1)\,\psi_n$, $(0, 1, 0)\,\psi_n$ and $(1, 0, -1)\,\psi_n$ have such expansions. From (8) and the recurrence relations for $y\psi_n$ and $d\psi_n/dy$ it follows that for $n \geqq 1$, the vectors $(1, 0, 1)\,\psi_{n+1}$, $(0, 1, 0)\,\psi_n$, $(1, 0, -1)\,\psi_{n-1}$ have expansions of the form $b_{n,1}\,\Phi_{n,1} + b_{n,2}\,\Phi_{n,2} + b_{n,3}\,\Phi_{n,3}$ if the matrix

$$\begin{bmatrix} \left[\frac{n+1}{2}\right]^{\frac{1}{2}} (\omega_1 + k) & \omega_1^2 - k^2 & \left[\frac{n}{2}\right]^{\frac{1}{2}} (\omega_1 - k) \\ \left[\frac{n+1}{2}\right]^{\frac{1}{2}} (\omega_2 + k) & \omega_2^2 - k^2 & \left[\frac{n}{2}\right]^{\frac{1}{2}} (\omega_2 - k) \\ \left[\frac{n+1}{2}\right]^{\frac{1}{2}} (\omega_3 + k) & \omega_3^2 - k^2 & \left[\frac{n}{2}\right]^{\frac{1}{2}} (\omega_3 - k) \end{bmatrix}$$

is non-singular. After some manipulation it may be seen that this is true if

$$\begin{bmatrix} 1 & \omega_1 & \omega_1^2 \\ 1 & \omega_2 & \omega_2^2 \\ 1 & \omega_3 & \omega_3^2 \end{bmatrix}$$

is nonsingular, or if $(\omega_1 - \omega_2)(\omega_1 - \omega_3)(\omega_2 - \omega_3) \neq 0$. This is equivalent to the statement that the three roots of the dispersion relation (7a) are distinct, which may be readily demonstrated. By making use of the fact that $\omega_{0,1} \neq \omega_{0,2}$, the remaining vectors needed, i.e., $\psi_0(1, 0, 1)$, $\psi_1(1, 0, 1)$ and $\psi_0(0, 1, 0)$, may be expanded in the vectors Φ_{-1}, $\Phi_{0,1}$, $\Phi_{0,2}$.

REFERENCES

Abramowitz, M. and I. A. Stegun. 1965. Handbook of Mathematical Functions, New York, Dover, 1046 pp.

Anderson, D. L. T. and A. E. Gill. 1975. Spin-up of a stratified ocean with applications to upwelling. Deep-Sea Res., *22*, 583–596.

Anderson, D. L. T. and P. B. Rowlands. 1976a. The role of inertia-gravity and planetary waves in the response of a tropical ocean to the incidence of an equatorial Kelvin wave on a meridional boundary. J. Mar. Res., *34*, 295–312.

—— 1976b. The Somali Current response to the Southwest Monsoon: the relative importance of local and remote forcing. J. Mar. Res., *34*, 395–417.

Blandford, R. 1966. Mixed Gravity-Rossby waves in the ocean. Deep Sea Res., *13*, 941–960.

Blumen, W. 1972. Geostrophic adjustment. Rev. Geophys., Space Phys., *10*, 485–528.

Brillouin, L. 1960. Wave Propagation and Group Velocity. New York, Academic Press. 154 pp.

Cane, M. 1974. Forced motions in a baroclinic equatorial ocean. GFD Notes, Woods Hole Oceanogr. Inst., Ref. No. 74–63.

—— 1975. A study of the wind-driven ocean circulation in an equatorial basin, Ph.D. Thesis, Massachusetts Institute of Technology, 372 pp.

Cane, M. and E. S. Sarachik. 1976. Forced baroclinic ocean motions. II. The linear equatorial bounded case. To appear.

Chester, C., B. Friedman and F. Ursell. 1957. An extension of the method of steepest descents, Proc. Camb. Phil Soc., *53*, 599–611.

Dickinson, R. E. 1969. Propagators of atmospheric motions 1 & 2, Revs. Geophys., *7*, 483–538.

Friedman, B. 1959. Stationary phase with neighboring critical points, J. Soc. Indust. Appl. Math., *7*, 280–289.

Leetmaa, A. 1973. The response of the Somali Current at 2°S to the southwest monsoon of 1971. Deep Sea Res., *20*, 397–400.

Lighthill, M. J. 1965. Group velocity, J. Inst. Maths. Applics., *1*, 1–28.

—— 1969. Dynamic response of the Indian Ocean to onset of the southwest monsoon. Phil. Trans. Roy. Soc., *A265*, 45–92.

Lindzen, R. 1967. Planetary waves on beta-planes. Mon. Wea. Rev., *95*, 441–451.

Longuet-Higgins, M. S. 1965. The response of a stratified ocean to stationary or moving wind stresses, Deep-Sea Res., *12*, 923–973.

—— 1968. The eigenfunctions of Laplace's Tidal Equation over a sphere. Phil. Trans. Roy. Soc., *A262*, 511–607.

Matsuno, T. 1966. Quasi-geostrophic motions in the equatorial area. J. Met. Soc. Japan, *44*, 25–43.

Moore, D. W. 1968. Planetary-gravity waves in an equatorial ocean. Ph.D. Thesis, Harvard University.

Moore, D. W. and S. G. H. Philander. 1976. Modelling of the equatorial oceanic circulation, *in* The Sea, Vol. 6, Goldberg, *et al.* eds., New York, Interscience, (to appear).

Munk, W. and N. Phillips. 1968. Coherence and band structure of inertial motion in the sea, Revs. Geophys., *6*, 447–472.

Rattray, M. 1964. Time dependent motion in an ocean; A unified two-layer beta-plane approximation. Studies on Oceanography, K. Yosida, editor, U. of Washington Press.

Rossby, C.-G. 1938. On the mutual adjustment of pressure and velocity distributions in certain simple current systems, 2, J. Mar. Res., *1*, 239–263.

—— 1945. On the propagation of frequencies and energy in certain types of oceanic and atmospheric waves, J. Met., *2*, 187–204.

Veronis, G. 1958. On the transient response of a β-plane ocean, J. Ocean. Soc. Japan, *14*, 1–5.

—— 1963a. On the approximations involved in transforming the equations of motion from a spherical to the β-plane, part I. Baroclinic systems, J. Mar. Res., *21*, 110–124.

—— 1963b. On the approximations involved in transforming the equations of motion from a spherical to the β-plane, part II. Barotropic systems, J. Mar. Res., *21*, 199–204.

Veronis, G. and H. Stommel. 1956. The action of variable wind stresses on a stratified ocean, J. Mar. Res., *15*, 43–75.

Williams, R. and C. Gibson. 1974. Direct measurement of turbulence in the Pacific Equatorial Undercurrent. J. Phys. Oceanogr., *4*, 104–108.

Yoshida, K. 1959. A theory of the Cromwell current and of the equatorial upwelling, J. Oceanogr. Soc. Japan, *15*, 154–170.

Received: *14 May, 1976;* revised: *28 September, 1976.*