a Coriolis tutorial, Part 2:

a rotating shallow water model and geostrophic adjustment

James F. Price
Woods Hole Oceanographic Institution
Woods Hole, Massachusetts, 02543
https://www2.whoi.edu/staff/jprice/ jprice@whoi.edu

Version 9 June 1, 2022

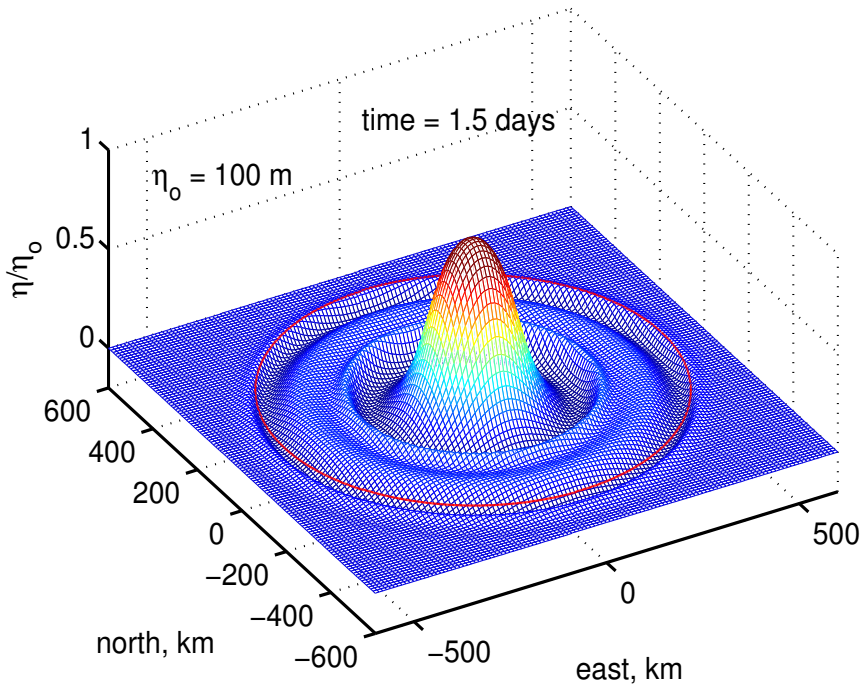

Figure 1: A snapshot taken 1.5 days after the start of a numerical experiment on geostrophic adjustment: a thickness anomaly of a dense fluid was released from rest onto an $f-$plane centered on 30^oN and allowed to evolve freely under the effects of gravity and rotation. The thickness anomaly slumps noticeably within the first half day, which excites inertia-gravity waves that expand isotropically. The leading edge propagates at the rate of a long, non-dispersive baroclinic gravity wave, roughly 275 km day^{-1} in this case. After about ten days the inertia-gravity wave radiation was largely completed, leaving most of the height anomaly seen here in a near geostrophic balance. An animation of this experiment is online at https://scienceweb.whoi.edu/PO/people/jprice/education/ga2d-h100-f.mp4 The goal of this Part 2 essay is to understand how the geostrophic adjustment process varies with the environment, the radius of the initial height anomaly, and the initial potential vorticity.

Abstract: This is the second of a four-part introduction to the effects of Earth's rotation on the fluid dynamics of the atmosphere and ocean. Part 1 derived the Coriolis force, $\propto -f\mathbf{k}\times\mathbf{V}$, where f is the Coriolis parameter, $\mathbf{k}$ is the vertical unit vector, and $\mathbf{V}$ is horizontal wind or current, and went on to analyze some of its basic properties in the context of a very simple, single parcel model. The goal and plan of this Part 2 is to develop further insight into the consequences of the Coriolis force by analyzing a sequence of experiments in which a thickness anomaly of horizontal scale L is released from rest and allowed to evolve under the influence of gravity and the Coriolis force. These geostrophic adjustment experiments are posed in a single layer fluid model, often called the shallow water model.

The initial gravitational slumping produces gravity waves (see the cover graphic). If there is no rotation, i.e., if $f = 0$, gravity waves will disperse the mass anomaly in a time L/C, where C is the gravity wave phase speed. When rotation is present, and if L exceeds several times the radius of deformation, $R_d = C/f$, and assuming that the eddy has a potential vorticity anomaly, then the Coriolis force will arrest the gravitational slumping and yield a geostrophically balanced eddy, an anti-cyclone if the thickness anomaly was a mass excess (high pressure). If f is presumed to be constant and if there is no friction, then a geostrophically balanced eddy could be exactly steady. Curvature effects on a two-dimensional eddy are measured by the Rossby number, U/fL, with U the azimuthal current. If the Rossby number is appreciable, then the actual, balanced current will exceed the geostrophic current in the case of an anticyclone.

These and other low frequency phenomenon are often best interpreted in the context of potential vorticity conservation, the geophysical fluid equivalent of angular momentum conservation. Earth's rotation contributes planetary vorticity = f, that is generally considerably larger than the relative vorticity of winds and currents. Small changes in the thickness of a fluid column may thus cause significant changes in the relative vorticity of winds and currents.

Contents

1 Large-scale flows of the atmosphere and ocean

This essay continues a four-part introduction to fluid dynamics on a rotating Earth. It is written for students who have some preparation in classical dynamics and applied mathematics and who are beginning a quantitative study of geophysical fluid dynamics. Part 1 examined the origin and the fundamental properties of the Coriolis force, and went on to consider a few of its consequences in the context of a single parcel model. The Coriolis force, often called simply 'rotation', admits two new modes of motion, a free oscillation usually called an inertial oscillation, and if there is a steady or time-mean external force, a steady or time-mean motion that is an analogue of geostrophic motion in the atmosphere and ocean.

1.1 The goal, and a plan to investigate the adjustment to geostrophic balance

The broad goal of this Part 2 is to develop insight for one of the most profound consequences of rotation, and specifically: **What circumstances lead to geostrophic balance?** It was noted in Part 1 that geostrophy was characteristic of most large scale and low frequency circulation, and now it is time to flesh out what we mean by 'large scale and low frequency' and to observe in some detail how geostrophy arises.

The plan is to solve and analyze a sequence of experiments in which a mass anomaly is released and allowed to evolve freely under the influence of gravity and rotation: gravity alone in Section 3, and gravity and rotation together in Section 4. As we will see first in Sec. 4, a nearly steady, nearly geostrophic state will arise spontaneously, provided that the mass anomaly has a sufficiently large horizontal scale, about 100 kilometers or greater at mid-latitudes. On the other hand, mass anomalies having a smaller horizontal scale, or having an initial current field that is characteristic of a gravity wave, will be dispersed into gravity waves. The essential results may be developed in a one-dimensional

domain, which makes it especially easy to conduct new experiments. However, oceanic and atmospheric eddies are more nearly circular than linear in plan view (Fig. 2) and will be modified somewhat by flow curvature, considered in Section 5. Results are summarized in Section 6, which then takes a look ahead to new phenomena that arise when the spatial variation of f is acknowledged.

1.2 Recurring themes

The discussion of the geostrophic adjustment experiments coming next will be based upon four noteworthy topics and themes that you will encounter over and over again in your study of the atmosphere and ocean:

Wave Dynamics As previewed in the cover graphic, many of the phenomena that arise in these experiments are wave-like. There are gravity waves that are more or less modified by rotation depending upon their wavelength and f, and nearly geostrophic eddies[1] that may also exhibit wave-like propagation (when the β-effect is included in Part 3). The language and concepts appropriate to waves — dispersion relations, phase speed, group speed — are thus very useful for describing the phenomena that arise in the these experiments.

Potential Vorticity Balance The dynamics of low frequency, quasi-geostrophic atmospheric and oceanic phenomena is revealed most effectively by analyzing the balance of potential vorticity, essentially the angular momentum of a rotating fluid, Sec. 2.2.3, rather than by the balance of linear momentum, which was all that was possible for the point-like parcel of Part 1. This use of potential vorticity balance was introduced by Rossby and colleagues in the 1930s[2] and remains an invaluable tool for the analysis and prediction of large scale and low frequency geophysical flows. Potential vorticity balance will be a central theme here and in Part 3.

Linear and Nonlinear, Finite Amplitude Waves that arise in linear systems, or in the limit that amplitude goes to zero in the solution of a nonlinear system, will have the most accessible description, e.g., by way of a dispersion relation. On the other hand, many geophysical phenomenon have an amplitude large enough that linear dynamics would not seem to be valid. Nevertheless, linear theory will generally be the starting point for an analyses that may then go on to include finite amplitude effects by way of more comprehensive (numerical) solution methods.

[1]The term 'eddy' is widely used in fluid dynamics, and with a wide range of meanings. Here eddy means a flow feature having a more or less circular planform, but with no particular dynamics or time-dependence implied. To characterize a movement as 'propagation' is suggestive of wave propagation, which is intended.

[2]Much of the pioneering research on the topics discussed in this essay appeared in a series of classic papers by Carl G. Rossby and colleagues published in the late 1930s. A scientific biography of Rossby and a bibliography of his classic papers is avilable at http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/rossby-carl-gustaf.pdf

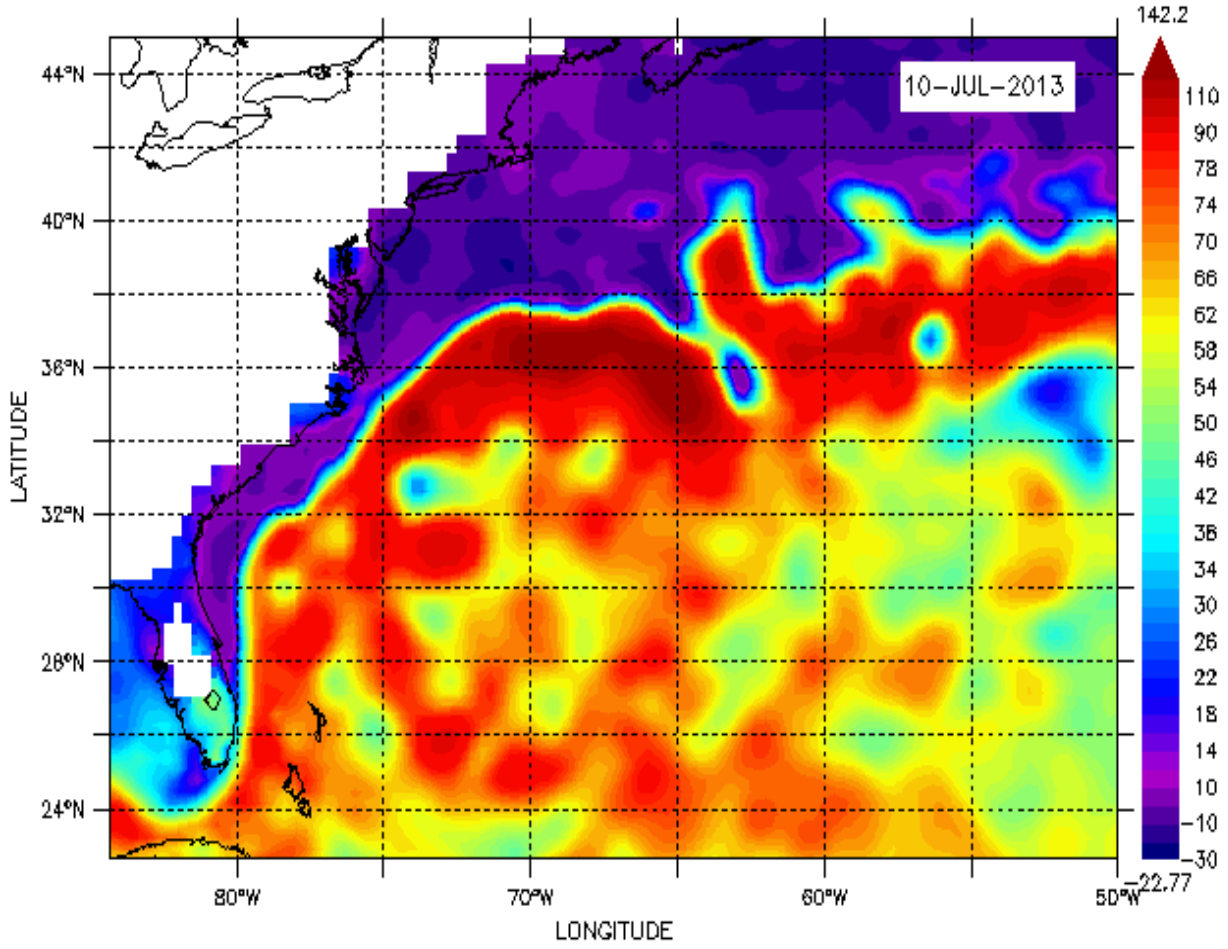

Figure 2: A snapshot of sea surface height, SSH, over the North Atlantic ocean from July, 2013 (thanks to the Aviso project, https://www.aviso.altimetry.fr/en/home.htm). Compared with the year-long mean of Fig. (1), Part 1, this field shows considerable variability on scales of several hundred kilometers, often termed the oceanic mesoscale (*meso* is Greek for middle) especially in subtropical and higher latitudes. These mesoscale eddies persist as identifiable features for many weeks or months and have an SSH amplitude of about $\pm$ 0.1 m. The corresponding thermocline displacements are about $\pm$ 50 m (not evident in this figure). The specific goal of this essay is to understand how Earth's rotation makes possible long-lived eddies of this kind.

Numerical Models and Solutions The shallow water system used here is kept as simple as possible but is nevertheless nonlinear, as are all more or less complete fluid models. Solving this nonlinear system, even for an idealized problem, usually requires more than just pencil and paper, *viz.*, a numerical code or model that serves to time-step a finite difference representation of the shallow water equations. In this and in many other ways, numerical models are an essential part of atmospheric and oceanic science, and for example, nearly all of the PhD students in the MIT/WHOI Joint Program (Physical Oceanography) use numerical models during their thesis research. For many of these students, that is also their first, hands-on experience. An implicit goal of this essay is to introduce some aspects of the practice of numerical modelling, especially hypothesis testing and solution interpretation. To that end, some of the homework assignments suggested here will require the generation of new numerical solutions. Source codes that are directly applicable to these problems are available on an anonymous ftp site linked in Sec. 6.3. By introducing some of the methods (and limitations) of numerical research, this essay is intended to supplement the excellent GFD texts available today by, among others, Cushman-Roisin, Gill and Pedlosky, that treat many of the same gesotrophic adjustment problems via mainly analytic solution methods.[3]

2 A shallow water model

This section derives the shallow water model, the conservation equations appropriate to a single fluid layer that varies in the horizontal, (x,y), but not in the vertical, z.[4] This layer is presumed to have a nominal thickness, H (Fig. 3). The vertical displacement of the upper surface (relative to a level surface) is $\eta(x,y,t)$ and will vary with horizontal position and with time. The thickness of the layer is thus imn general, $h(x,y,t) = H + b(x,y) + \eta(x,y,t)$ with b the variable lower boundary of the layer, or the sea floor. In this essay, only the case $b = 0$ (flat bottom) will be considered. The fluid above is presumed to be at rest, and to have a uniform density ρ_o, and pressure, P_a, that is uniform horizontally. The fluid within the lower layer has a greater density than the fluid above, $\rho_o + \delta\rho$ where $\delta\rho$ is a specified constant noted below.

Compared with the single parcel model of Part 1, this single layer model is a very big step toward a

[3] An introduction to geophysical fluid dynamics at about the level of this essay is by B. Cushman-Roisin, *Introduction to Geophysical Fluid Dynamics* (Prentice Hall, Engelwood Cliffs, New Jersey, 1994). Somewhat more advanced and highly recommended for the topic of geostrophic adjustment is A. E. Gill, *Atmosphere-Ocean Dynamics* (Academic Press, NY, 1982) and for waves generally, J. Pedlosky, *Waves in the Ocean and Atmosphere* (Springer, 2003).

[4] You may have noticed that this is a lengthy and mathematically dense section, and if so, the little voice in your head may be asking ... *do we really have to know all of this?* This is essential background for understanding the shallow water model and I hope that you will be sufficiently motivated to work through this material some day. If you are not ready for that task now, then jump ahead to Section 3 where discussion of the experiments begins in earnest.

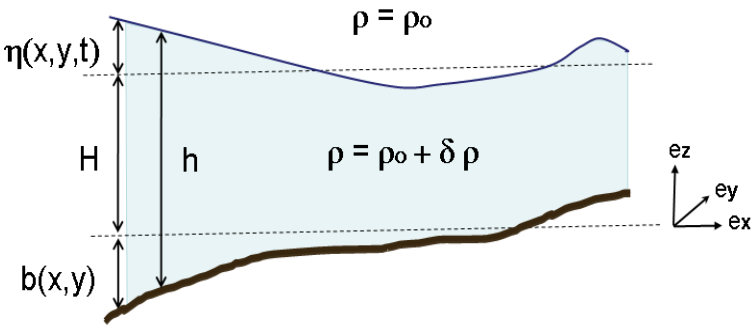

Figure 3: Schematic of a single fluid layer (light blue) sitting on a variable-depth lower, solid boundary, e.g., the sea floor. The horizontal dashed lines are level surfaces (perpendicular to gravity). This layer could be a homogeneous ocean, in which case the fluid above is the atmosphere and $\delta\rho \gg \rho_o$. In that case η is the displacement of the sea surface (as in Figs. 1 and 2). Or, it could be a relatively dense layer within a stratified ocean, in which case the fluid above is just slightly less dense, $\delta\rho \ll \rho_o$, and η is the displacement of a density surface. In either case, it is presumed that the fluid above the active layer does not impose a horizontally-varying pressure upon the active layer.

realistic model of the atmosphere or ocean. There are, however, two important idealizations and simplifcations that facilitate both the solution and its analysis, though at the expense of generality.

Single layer. The horizontal velocity $\mathbf{V}$ is implicitly depth-independent since it is represented by a two component vector field i.e., $\mathbf{V}(x,y,t) = [u(x,y,t), v(x,y,t)]$, vs. the true three-dimensional $\mathbf{V}(x,y,z,t)$, a bigger help than it may first appear. For this neglect of vertical structure to be a valid approximation, the nominal layer thickness, H, must be somewhat less than the horizontal scale of the motion. A single layer model of this kind is often and appropriately termed a shallow water model, even when H is taken to be the full depth of the ocean. The most convincing justification for the shallow water model comes from the analysis of a two layer model that includes both external (barotropic) and internal (baroclinic) normal modes (see the appendix to this section). The two layer model makes clear that a shallow water model is appropriate to a study of either barotropic or baroclinic normal modes, but not both at once. This is not a significant shortcoming given the broad goal of this study and, moreover, it wouldn't matter which normal mode was chosen. But given that a specific goal is to model the oceanic mesoscale eddies seen in Fig. 2, then as nominal values, $H = 500$ m and $\delta\rho = 2$ kg m^{-3}, are appropriate to the baroclinic motions of the ocean's main thermocline. Ocean mesoscale eddies have a radius $L \approx 100$ km, and hence this layer thickness is indeed 'shallow' in the sense that $H/L \ll 1$.

It is notable that oceanic eddies are mainly upper ocean phenomenon, i.e., strongest above the

main thermocline. They certainly are not resting on the sea floor, as the layer of Fig. (3) seems to be doing. The reduced gravity approximation discussed in Sec. 2.4 shows that such a single layer model will be most appropriate provided that the gravity wave speed of the model is made equal to the gravity wave speed of the baroclinic eddies of interest (and the sea floor is taken to be flat).

If the intent was to utilize the shallow water model to make a realistic simulation of a given, observed flow phenomenon, then three additional approximations or simplifications would have to be valid: 1) that the initial state was free of vertical shear, $\partial \mathbf{V}/\partial z = 0$, 2) that the fluid of interest was outside of frictional boundary layers, and 3) that the horizontal density variation within the layer was effectively zero. Most real geophysical flows violate all three of these conditions to some degree, especially 1) and 3). A realistic simulation is likely to require greater resolution in the vertical, e.g., some number of shallow layers stacked one on top of another to represent vertical shear, boundary layers, stratification, etc. However, consistent with the inductive method laid down in Part 1, one layer will be sufficient for now.

Ideal fluid. The physical processes allowed into this model are those of an ideal fluid, *viz.*, pressure and pressure horizontal variations, and transport by the fluid flow. The wide range of physical processes associated with the thermodynamic properties of a real fluid, e.g., compressibility, diffusion and viscosity, are omitted. This too may seem a bit high-handed, but is consistent with building the simplest model that will help us understand some of the consequences of Earth's rotation.

2.1 Conservation equations of an ideal fluid layer

2.1.1 Mass Conservation

The starting point is the assertion that mass (volume, really, since density is presumed constant) can be neither created nor destroyed by classical fluid dynamical processes. The positive assertion of this is that the mass of the layer at a given point can change only by virtue of mass fluxes associated with the (horizontal) fluid velocity within the layer. This is oftentimes referred to as the continuity requirement.

To find the corresponding mathematical form of this physical assumption, the mass balance will be evaluated over a control volume that spans the full thickness of the layer, Fig. (4). The mass of the layer contained within the control volume is

$$M = \rho \bar{h} dc, \tag{1}$$

where ρ is the constant density within the layer, and $\bar{h}$ is the average thickness over the interval x to $x+d$ and c is the width (the dimension into the page). The mass flux due to fluid flow through a given

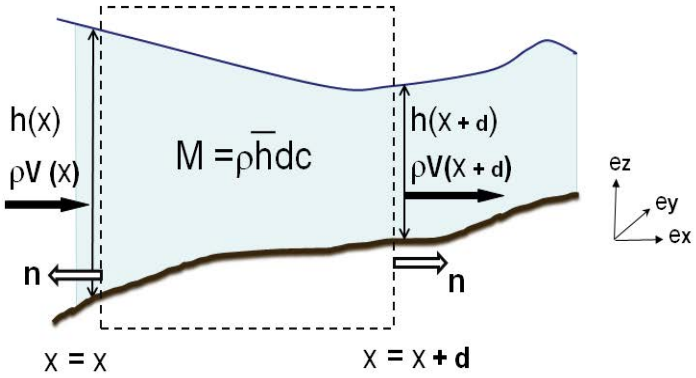

Figure 4: A control volume defined by thin dashed lines. The average thickness is $\bar{h}$, the length is d, and the width in the y-direction is c (y is normal to the page and is not shown). The statement of mass conservation appropriate to the problems considered here is that the mass of the active layer within the control volume, M, can change only because of mass fluxes due to fluid flow normal to the sides of the control volume, $-\rho A\mathbf{V}\cdot n$, where A is the area. The layer thickness variation shown here is exaggerated; changes of layer thickness of even a few percent can be significant.

surface of area A is just

$$\text{mass flux} = -\rho A\mathbf{V}\cdot\mathbf{n}, \tag{2}$$

where $\mathbf{n}$ is the outward unit normal of the surface. For the left-facing side of the control volume in Fig. (4), $\mathbf{n} = (-1,0,0)$ and the mass flux = ρhcu, where u is the x-component of the fluid velocity and hc is the area. These are evaluated at the x of the left-facing side. Noting that the unit normal on the right-facing side at $x+d$ is $\mathbf{n} = (1,0,0)$, then the sum of the mass fluxes through both x-facing sides is

$$\frac{\partial M}{\partial t} = \rho(h(x)u(x) - h(x+d)u(x+d))c. \tag{3}$$

By our construction of mass conservation, this net mass flux is equal to the time rate of change of mass within the control volume. The mass $M(x,y,t)$ varies with position and time, but notice that the time derivative of (3) is a partial derivative, since M is the mass of a control volume that is fixed in space. Inserting the definition of M, and dividing through by ρ, c and d gives

$$\frac{\partial \bar{h}}{\partial t} = \frac{1}{d}(h(x)u(x) - h(x+d)u(x+d)).$$

Now let d go to zero, and the right-hand side also becomes a partial derivative

$$\frac{\partial h}{\partial t} = -\frac{\partial (hu)}{\partial x}, \tag{4}$$

the one-dimensional, differential statement of mass conservation (though since the constant density and area are divided out, the units are length per time). Thus, the mass inside the control volume will change if and only if there is a *divergence* of the mass flux associated with the fluid flow; the mass flux alone is not relevant insofar as the M at a given position is concerned. Had variations in the y-direction (normal to the page) been considered, then there would arise an additional term, $-\partial(hv)/\partial y$, and the two-dimensional form of mass conservation for a shallow water model is

$$\frac{\partial h}{\partial t} = -\left(\frac{\partial (hu)}{\partial x} + \frac{\partial (hv)}{\partial y}\right) \tag{5}$$

$$= -\nabla\cdot(h\mathbf{V}). \tag{6}$$

This way of writing mass conservation is sometime said to be the flux form or conservative form, and is generally preferred for implementation in a numerical model. For purposes other than numerical integration it may be useful to expand the divergence operator and combine the partial time derivative and the advection term into the material derivative,

$$\frac{D(\ )}{Dt} \equiv \frac{\partial(\ )}{\partial t} + u\frac{\partial(\ )}{\partial x} + v\frac{\partial(\ )}{\partial y}, \tag{7}$$

about which more below. The mass balance (thickness balance) written using the material derivative is

$$\frac{Dh}{Dt} = \frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} = -h\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right). \tag{8}$$

2.1.2 Momentum conservation

At a given point, the momentum of the layer $\rho h\mathbf{V}$, can change only by the relative few prescribed processes, by momentum fluxes associated with the fluid velocity, by pressure variations, and by the Coriolis force, an inertial force that acts throughout the body of the fluid. There may also be an external force, $\mathbf{F}$, (not to be confused with the lower case f) that acts as a stress, a tangential force per area on the upper or lower boundaries.

Momentum flux by the flow. The x component of momentum (per unit area) is ρhu, and the momentum flux due to the x-component of the velocity is just

$$\text{momentum flux by the flow } = \rho hu^2,$$

which, notice, is proportional to kinetic energy.

Hydrostatic pressure and pressure flux. The phenomena of interest here are somewhat slowly varying, and the accelerations of the fluid are much, much less than the acceleration of gravity,

$$\frac{\partial^2 h}{\partial t^2} \ll g.$$

In that case the vertical component of momentum balance is the hydrostatic pressure relation,

$$\frac{\partial p}{\partial z} = -g\rho,$$

and the hydrostatic pressure within the layer ($z < h$) is

$$P(x,y,z) = P_a(z) - g\rho_o z + g\delta\rho(h(x,y,t) - z). \tag{9}$$

Thus the horizontally-varying pressure in this model is due solely to variation in the mass field, i.e., the layer thickness. $P_a(z)$ is the z-dependent but horizontally uniform pressure in the fluid above the active layer and the second term on the right is the hydrostatic pressure due to the constant density ρ_o. Neither of these z-dependent terms contribute to horizontally-varying momentum fluxes and may be ignored in what follows. The third term represents the pressure anomaly associated with the horizontally-varying layer thickness, h and notice that a relatively large layer thickness indicates a high pressure anomaly within the layer. The momentum flux due to the anomaly of hydrostatic pressure on an x-facing side of the control volume having width c is then

$$\text{pressure flux } = g\delta\rho c \int_0^h (h-z)dz \;=\; g\delta\rho c\frac{h^2}{2}.$$

The momentum flux. Summing these two terms gives the x-component momentum flux on an x-facing side of the control volume,

$$\text{momentum flux } \;=\; \rho chu^2 + g\delta\rho c\frac{h^2}{2}. \tag{10}$$

The first term on the right is the flux of u-momentum due to the u-component of the velocity and so is proportional to u^2, and the second term is the layer-averaged anomaly of the hydrostatic pressure $\propto h/2$ acting over the face of the control volume, ch. When the body force due to the Coriolis force is included, $-\mathbf{f}\times\mathbf{V} = (fv, -fu)$, and when the operations noted above to go from Eqn. (3) to Eqn. (6) are repeated here, the conservative (or flux) form of the shallow water momentum (per mass $\times$ area) equations are

$$\frac{\partial(hu)}{\partial t} = -\frac{\partial(huv)}{\partial y} - \frac{\partial(hu^2 + \frac{1}{2}g'h^2)}{\partial x} + fhv + F_x, \tag{11}$$

$$\frac{\partial(hv)}{\partial t} = -\frac{\partial(huv)}{\partial x} - \frac{\partial(hv^2 + \frac{1}{2}g'h^2)}{\partial y} - fhu + F_y, \tag{12}$$

where $g' = g\delta\rho/\rho$ is the reduced gravity (also buoyancy per unit volume). If the upper surface of the layer is the sea surface, then $\delta\rho \approx \rho$, and $g' \approx g$; if the upper surface is within the comparatively weakly stratified ocean, then $\delta\rho \ll \rho$ and so $g' \ll g$.

The external force shown here as $\mathbf{F} = (F_x,\ F_y)$ could represent wind stress, in which case we would probably say that $\mathbf{F} = \mathbf{F}(x,y)$ is a specified function of space only. Or, it could represent a bottom drag, in which case $\mathbf{F} \propto -r\mathbf{V}$ as in the single parcel model of Part 1. Notice that F must have dimensions of force/mass, though it will be referred to as 'force'. Our main interest is what the fluid flow does, rather than $\mathbf{F}$, and so unless it is noted otherwise, presume that $\mathbf{F}$ vanishes.

The shallow water mass and momentum balances written using the material derivative are then (repeating Eqn. 8),

$$\frac{Dh}{Dt} = \frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} = -h\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \tag{13}$$

$$\frac{Du}{Dt} = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -g'\frac{\partial h}{\partial x} + fv + \frac{F_x}{h}, \tag{14}$$

$$\frac{Dv}{Dt} = \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -g'\frac{\partial h}{\partial y} - fu + \frac{F_y}{h}. \tag{15}$$

The vector equivalents are

$$\boxed{\frac{Dh}{Dt} = \frac{\partial h}{\partial t} + \mathbf{V}\cdot\nabla h = -h\nabla\cdot\mathbf{V}} \tag{16}$$

$$\boxed{\frac{D\mathbf{V}}{Dt} = \frac{\partial \mathbf{V}}{\partial t} + \mathbf{V}\cdot\nabla\mathbf{V} = -g'\nabla h - f\mathbf{k}\times\mathbf{V} + \frac{\mathbf{F}}{h}} \tag{17}$$

The Coriolis force is written in the usual form, i.e., the horizontal component only. Moreover, here in Part 2, f will be taken as a constant, though we know that f varies with latitude. This choice to employ a constant f, or f-plane model, is purely for simplicity and convenience. Once we learn some of the consequences of a constant f here in Part 2, then in Parts 3 and 4 we can go on to learn the surprising and important consequences of f(latitude).

2.2 Shallow water model compared to a single parcel model

The shallow water model equations (16) and (17) define a complete if somewhat simplified model of an idealized, two-dimensionally varying fluid. Before plunging into their solutions it is useful to consider how these shallow water equations compare with the equation of motion appropriate to the single *parcel*

of Part 1, Sec. 5 (repeated here but omitting the bottom friction term from the latter and ignoring **F** of the former),

$$\frac{d\mathbf{V}}{dt} = \; g'\nabla b - f\mathbf{k}\times\mathbf{V}.$$

1) The Coriolis terms are identical in the two systems. The Coriolis force is an inertial force that depends solely upon the fluid velocity relative to the rotation vector, and makes no distinction of the physical properties of the material. And, of course, absent that fundamental property, it would not have been appropriate to use a single parcel model as the starting point in Part 1.

2) The pressure gradient term of the shallow water model, $-g'\nabla h$, has the form of the buoyancy force on a dense parcel sitting on a slope, $g'\nabla b$ (sign aside). A crucial difference is that the bottom slope of the single parcel model is prescribed and fixed, while the gradient of layer thickness in the shallow water model is a dependent variable, dependent upon space and time via the mass balance, Eqn. (16). Mass balance has no counterpart in the single parcel model. In an ideal fluid model such as this one, the variable pressure is the only way that a fluid parcel interacts with the rest of the fluid domain. For example, if fluid begins to pile up (converge) at a given location, then the layer thickness and the hydrostatic pressure will go up until the pressure gradient is sufficient to push fluid away from that location. The rate at which the fluid responds to a convergence determines the speed of gravity wave propagation (Sec. 3.1).

3) The time derivatives have quite different meanings implicit in the use of different symbols: the ordinary time derivative d/dt of the single parcel model is the time rate change of a specific, material parcel, which in a fluid dynamics context would often be termed a Lagrangian coordinate system. Different parcels have different initial conditions and subsequent forcing, and to solve for a fluid flow requires solving for many such (interacting) parcels. The material derivative D/Dt of the present shallow water model is equal to the time rate of change following a moving parcel at the instant the parcel is coincident with the spatial position where D/Dt is evaluated. The independent coordinates of our shallow water model are fixed spatial coordinates and time, which in fluid dynamics is often referred to as an Eulerian coordinate system. The fluid present at a fixed location changes due in part to the fluid flow via the advection process represented by the term $\mathbf{V}\cdot\nabla(\;)$. Advection is the product of two dependent variables, velocity times thickness gradient in Eqn. (16) or velocity times the velocity gradient in Eqn. (17) and so is nonlinear. Advection and this associated nonlinearity are right at the heart of fluid dynamics and are a significant reason why numerical methods and numerical models are essential tools of atmospheric and oceanic sciences.

2.3 Solving and diagnosing the shallow water system

Equations (16) - (35) are a coupled set of nonlinear, partial differential equations in three dependent variables, the two components of the horizontal velocity, $\mathbf{V} = (u,\ v)$, and the thickness, h. The parameters g and $\delta\rho$ are constants for any given problem, and the Coriolis parameter f is at most a specified function of the north-south coordinate, y, about which much more below.

The shallow water model is nonlinear because of the advection terms of the material derivative noted above, e.g., the east-west advection of north-south momentum is $u\ \partial v/\partial x$, which is the product of two dependent variables. Solutions of a shallow water system including these nonlinear terms are necessarily generated by numerical methods: the one-dimensional experiments shown here were solved by the Matlab script geoadj_1d.m, and the two-dimensional experiments by the Fortran code geoadj_2d.for, both of which are linked in Sec. 6.3.[5] In common with any numerical model, these models produce (putative) solutions in the form of very large data files, $u(x_i, y_i, t_j), v(x_i, y_i, t_j)$..., where x_i, t_i are discretized position and time. To verify that this mass of data is faithful to the model equations and to the boundary/initial conditions, and then to go on and learn something useful requires as much thought and effort as does generating the model and solution in the first place. To wit, we will seek to 1) construct useful visualizations of the solutions (this can be the artful side of numerical modelling), 2) diagnose the balances of energy and potential vorticity, 3) interpret the wave-like properties of the solution using the dispersion relation of the corresponding, linearized system, and then 4) form and test hypotheses regarding parameter dependence by conducting further numerical experiments (that will be the homework, actually).

2.3.1 Wave velocities and fluid velocities

A qualitative difference between a shallow water model and the single parcel model of Part 1 is that a shallow water model supports wave motions — relatively fast-moving gravity waves will arise in all of our experiments, and much lower frequency Rossby waves will arise when there is a north-south variation of f (deferred to Part 3). If the layer thickness anomaly is a simple harmonic motion varying with x and t only, $\eta(x,t) \propto \eta_o sin(kx - \omega t)$, then a constant phase propagates at the rate of the

$$\text{wave phase speed,}\ \ C_p = \frac{\omega}{k}, \tag{18}$$

where wave frequency is $\omega = 2\pi/(\text{wave period})$, or waves per time, and the wavenumber is $k = 2\pi/(\text{wave length})$, or waves per space interval. This wave propagation velocity is often readily

[5] The details of the numerical methods are all-important in setting the efficiency and the accuracy of the solutions, but not something that will be discussed here. An excellent, concise reference on numerical methods suitable for the shallow water model is https://www.mathworks.com/company/newsletters/articles/introducing-cleves-laboratory and look for waterwave

apparent (as in the cover graphic). In many cases the thickness anomaly will be the result of a superposition of waves, i.e., waves of different wave numbers and frequencies, and in that case the envelope of the superposition will move at the

$$\text{wave group speed, } C_g = \frac{\partial \omega}{\partial k}, \tag{19}$$

This makes clear that the dispersion relation, $\omega(k)$, is going to be very important. The dispersion relation is determined by the physics of the wave medium, specifically the relationship of restoring force to wavelength. This differs greatly between pure gravity waves (Sec. 3.1), gravity waves in the presence of rotation (Sec. 3.2), and Rossby waves (Part 3). The motion of the fluid is, in general, qualitatively different from the phase speed of waves that may be propagating through the fluid. Fluid velocity is usually much less than the gravity wave speed and is often much greater than the Rossby wave speed (Rossby waves in Part 3). To see the fluid motion one can simply plot the field of the instantaneous velocity vectors, which shows the

$$\text{Eulerian fluid velocity : } \mathbf{V}(x_i, y_i, t_j) = (u(x_i, y_i, t_j), v(x_i, y_i, t_j)), \tag{20}$$

at the fixed positions (x_i, y_i) and the times, t_j. This Eulerian velocity field is the direct output of the numerical model, useful in itself, and the starting point for much else. The fluid velocity is generally proportional to the amplitude of the motion; in the problems discussed here, $V \propto \eta_o$, where η_o is the initial interface displacement. Wave speeds, on the other hand, are independent of amplitude in linear approximation, and generally only weakly dependent upon amplitude in nonlinear models (more on this below).

To see the transport of fluid over a long term we can compute the evolution of a passive tracer, say s, that is carried along with the flow without in any way altering the flow. This tracer may be embedded in the fluid at the starting time, $s_o = s(x, y, t_o)$ in any way that will serve to highlight the features of interest. This tracer is then presumed to be conserved following the flow,

$$\frac{Ds}{Dt} = 0, \tag{21}$$

or, expanding the material derivative,

$$\frac{\partial s}{\partial t} = -(u\frac{\partial s}{\partial x} + v\frac{\partial s}{\partial y}) = -\mathbf{V} \cdot \nabla s. \tag{22}$$

The tracer value at a fixed location will thus change in time due solely to advection by the flow. Diffusion of tracer is omitted in this ideal fluid model, though some inadvertent 'numerical diffusion' will always be present in numerical solutions. Thus Eqn. (21) would hold exactly only in a perfect numerical solution (and you may never see one).

The trajectories of discrete fluid parcels, or 'floats', make a useful complement to the continuous tracer. The trajectory of the ith float, (x^i, y^i), is found by integrating the (Eulerian) fluid velocity at the moving location of the float,

$$x^i(t) = x^i_o + \int_{to}^{t} u(x^i, y^i, t)dt \quad \text{and} \quad y^i(t) = y^i_o + \int_{to}^{t} v(x^i, y^i, t)dt. \tag{23}$$

Which specific fluid parcel or float is being tracked in this way is given by the initial position, i.e., the ith float starts at $x^i_o = x^i(t = t_o)$. The float will likely be found between the discrete grid points of the numerical model's Eulerian velocity field and so in practice this will require some interpolation. The trajectory is the fundamental dependent variable of a Lagrangian description (rather than the velocity as in the Eulerian description). But if needed, the Lagrangian velocity may be computed from the trajectories via

$$\text{Lagrangian fluid velocity}: \quad \mathbf{U}^i(t) = \left(\frac{\partial x^i(t)}{\partial t}, \ \frac{\partial y^i(t)}{\partial t} \right). \tag{24}$$

Tracer evolution by advection Eqn. (21) and float trajectories by Eqn. (23) are very closely related, and, for example, the tracer concentration at the moving location of a float should be exactly conserved (numerical diffusion and tracking errors aside).

The relationship between the Eulerian and the Lagrangian velocities depends very much upon the length of time that a float is tracked, the $t = t_0 + \Delta t$ of Eqn. (23). If Δt is so brief that the float moves only an infinitesimal distance compared to the scale over which the Eulerian velocity varies in space, then the Lagrangian velocity will converge to the Eulerian velocity at the starting point. Much more interesting is the case that Δt is very long, in Part 3 up to a year, in which case any given float may wander into regions having very different Eulerian velocity, which is also likely changing in time. In that event, and aside from special cases, the Lagrangian velocity will likely bear no simple relationship to the Eulerian velocity.

2.3.2 Energy balance

Energy conservation is a fundamental physical law that can sometimes be of use in analyzing a fluid flow, especially if, as here, dissipation and thermal (internal) energy may be neglected. In that case the mechanical energy, $E = KE + PE$, is conserved, where E is the sum of kinetic energy (per unit mass), $KE = \frac{1}{2}h\mathbf{V}^2$, and potential energy (per unit mass), $PE = g' \int_0^\eta z dz = \frac{1}{2}g'\eta^2 + const$. To find the rate of change of KE, take the dot product of the momentum equation with the velocity times thickness, and the rate of change of PE is found by multiplying the continuity equation by the thickness anomaly times reduced gravity. The mechanical energy balance is

$$\boxed{\frac{DE}{Dt} = -g'\nabla\cdot(\eta h\mathbf{V}) + \mathbf{F}\cdot\mathbf{V}} \tag{25}$$

The flux term on the right-hand side is the product of pressure anomaly, thickness and velocity, i.e., pressure work (actually, pressure work rate, or power). For example, if the pressure anomaly is positive on the (imaginary) boundary of a control volume where the velocity is directed outward, say, then the fluid inside the control volume will do pressure work on the fluid outside of the control volume. The energy within the control volume will thus decline, while the energy outside the control volume will increase. The pressure-work term thus accounts for the outward energy transport associated with wave radiation through an open boundary, for example. The fluid flow can also transport energy at the rate $\mathbf{V}$, a process accounted by the advection term, $\mathbf{V}\cdot\nabla E$, of the total derivative. Finally, note that an external force $\mathbf{F}$ can either increase or decrease energy depending upon whether it has a component parallel or anti-parallel to the fluid velocity.

2.3.3 Potential vorticity balance

Throughout Part 1, our perspective on dynamics was through the *linear* momentum balance (linear here in the geometric sense), and rotation appeared by way of the Coriolis *force*. That was all that was possible given a single parcel model. Momentum balance is always relevant, but may not always be highly revealing. Given the present fluid model, there is another and complementary point of view, *angular* momentum balance, that has proven immensely fruitful for understanding some of the most important low frequency (frequencies less than f) phenomena of geophysical fluid dynamics. Two main reasons, to get a little ahead in this short story, are that 1) Earth's rotation provides a very large background angular momentum that is made visible by small changes in the thickness or latitude of a fluid column, and, 2) the angular momentum balance amounts to a kind of filter that eliminates high frequency gravity wave motions and so serves to highlight the processes that cause departures from geostrophic balance and that lead to low frequency currents and winds, i.e., large scale circulation.

Back a step or two..... to analyze the motion of a rotating, solid object, say a gyroscope, you might begin by computing the linear momentum balance of the component pieces. This would require an accounting of radial accelerations and internal stresses on each piece and would likely be a fairly arduous task. Assuming that the gyroscope is not at risk of breaking up, then at some point you might decide to take the structure for granted, and focus your effort toward analyzing the angular (azimuthal) momentum balance of the gyroscope as a whole. As a first step you would define a coordinate system that gave the most compact and least complex accounting of the moment of inertia of the gyroscope and then consider the processes that cause the angular momentum to change with time. The physical content of your angular momentum analysis would not be fundamentally different from the linear momentum description, but it would likely be a great deal simpler in the same way that choosing appropriate variables and an appropriate coordinate system can facilitate any mathematical analysis. The same considerations apply to an analysis of a fluid flow: as we will see in examples here, an angular momentum analysis and description will often (but not always!) be a good deal simpler and so

provide a great deal more insight than does the otherwise equivalent linear momentum analysis.

The fluid flow equivalent of angular velocity is the curl of the fluid velocity,

$$\boxed{\xi \equiv \nabla \times \mathbf{V}}$$

called the vorticity. If the fluid velocity is a three-dimensional vector, then the vorticity is also a three-dimensional vector. In the special case of the shallow water model, the velocity varies only in the two horizontal dimensions, and so the shallow water vorticity,

$$\xi = \nabla \times \mathbf{V} = (\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y})_z, \tag{26}$$

has a vertical component only and is effectively a scalar. You can visualize vorticity as the rotation of small (but not quite point-like) two-dimensional parcels, e.g., cylinders, that make up the fluid.[6] As we will see, however, vorticity in a plane wave will arise from horizontal shear in the direction normal to the velocity, e.g., $\xi = \partial v/\partial x$ or $\xi = -\partial u/\partial y$ are quite possible, and so spinning cylinders are not always apropos. If the direction of rotation is the same as Earth's rotation, the vorticity is said to be *cyclonic* (from the Greek *kyklon*, for circular motion). Cyclonic rotation is thus counterclockwise in the northern hemisphere and clockwise in the southern hemisphere. Anticyclonic is the reverse. Notice that vorticity has units of inverse time, or frequency, the same as the Coriolis parameter, f. There is an important sense in which ξ may be compared to and even added to f, as discussed below.

The vorticity of a fluid is unlike the angular momentum associated with a solid object in that it is defined at every point in a fluid, i.e., in principle there is a vorticity field that accompanies every fluid flow (although it could be zero), just as there is a velocity field and a thickness field in the shallow water model. The governing equation for this vorticity field may be found by taking the curl of the momentum equation, $\partial/\partial x$ of the y-component minus $\partial/\partial y$ of the x-component. A consequence of applying the curl operator is that all of the forces that are derivable from a potential are eliminated, most notably the pressure gradient in the shallow water model, i.e.,

$$\nabla \times \nabla(\eta(x,y)) = \frac{\partial^2 \eta}{\partial x \partial y} - \frac{\partial^2 \eta}{\partial y \partial x} = 0.$$

A divergence term $\partial u/\partial x + \partial v/\partial y$ will arise and may be eliminated using the thickness (continuity) equation, (16). After a little further rearrangement, the result is a balance equation for a scalar, q, called the potential vorticity,

$$\boxed{\frac{Dq}{Dt} = \frac{\partial q}{\partial t} + \mathbf{V} \cdot \nabla q = \frac{1}{h} \nabla \times \frac{\mathbf{F}}{h}} \tag{27}$$

[6] An essential resource for all students of fluid mechanics is the collection of fluid mechanics films made in the 1960s by Ascher Shapiro and colleagues and now available online at http://web.mit.edu/hml/ncfmf.html These films will seem a little old-fashioned today, but they are nevertheless very highly recommended in that they provide excellent visualizations of many key concepts of fluid mechanics, including vorticity and Eulerian and Lagrangian coordinate systems, that are timeless. A more modern film collection, many of which emphasize rotational effects, is https://spinlab.ess.ucla.edu/

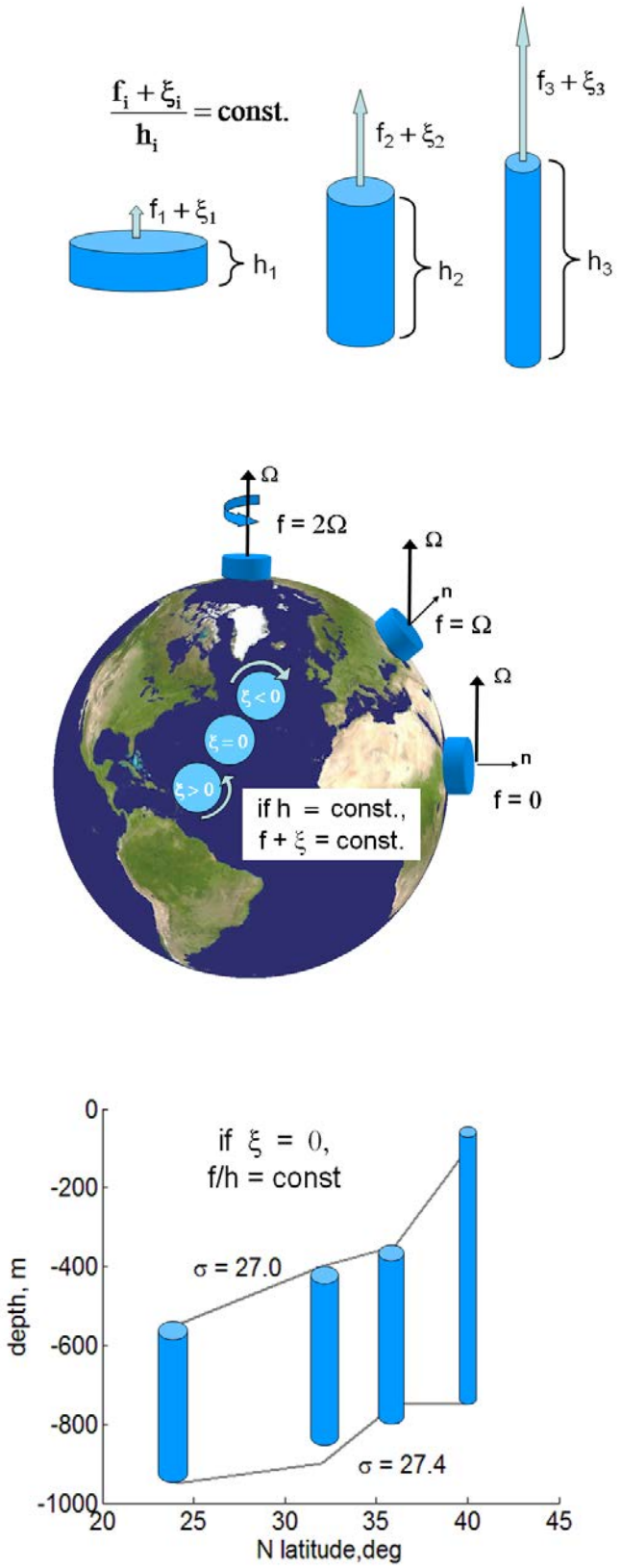

Figure 5: Some illustrations of q-conservation, $Dq/Dt = 0$, by fluid columns that have a constant volume but may have a variable thickness, h and that are free of external torques, $\nabla \times \mathbf{F} = 0$. **(upper)** If h is variable, then the absolute vorticity, $f + \xi$, will change in proportion to h. This is true whether the absolute vorticity $\approx \xi$, because $f \ll \xi$ as in small scale or engineering flows, or if $\xi \ll f$ and ξ is effectively zero, as in a gyre-scale flow, illustrated in the lower panel. **(middle)** The columns perched at upper and right on the Earth are intended to show that the planetary vorticity due solely to Earth's rotation is given by $\nabla \times \mathbf{V}_\Omega = 2\mathbf{\Omega} \cdot \mathbf{n} = 2\Omega \sin(\text{latitude}) = f$. The planetary velocity $\mathbf{V}_\Omega$ and thus the planetary vorticity f are apparent to an inertial observer, implicit in this view from space. We Earth-bound observers will see only the relative velocity $\mathbf{V}$, i.e., winds and ocean currents, and the relative vorticity, ξ. The planetary vorticity is not directly observable (though the stars turning overhead are a mirror image). The three columns in the North Atlantic are shown end-on to illustrate the sense of relative vorticity that would be acquired by a q-conserving column that was displaced north or south away from a reference site where $\xi = 0$ while holding h constant. This and the next mechanism involve $f(\text{latitude})$ that will be taken up in Part 3. **(lower)** If the relative vorticity ξ is much, much less than f and remains effectively zero, as in a gyre-scale flow, then a change in the latitude of a column will be accompanied by a change in the thickness of a column in the sense shown here, i.e., smaller h in the direction of smaller $|f|$. The solid lines are depths of the 27.0 and 27.4 potential density surfaces along a plausible trajectory through the main thermocline of the eastern subtropical North Atlantic. These data are after Fig. 4 of Luyten J. R. et al., 1983: 'The ventilated thermocline', *J. Phys. Oceanogr.* 13, 292 - 309.

where

$$\boxed{q = \frac{f+\xi}{h}} \tag{28}$$

Up until Part 3 our interest will be the fluid only, $\mathbf{F} = 0$, in which case we have a q-conservation law

$$\boxed{\frac{Dq}{Dt} = 0} \tag{29}$$

The terms in potential vorticity balance are:

$\xi = \nabla \times \mathbf{V}$ is often called the *relative* vorticity in this context, since it is the vorticity of the relative velocity, i.e., the winds and currents observed from an Earth-attached reference frame. In Part 1, Sec. 4 the relative velocity was denoted by $\mathbf{V}'$, though by now the prime superscript has been dropped.

$f = \nabla \times \mathbf{V}_\Omega$ is the *planetary* vorticity of the planetary velocity, $\mathbf{V}_\Omega = \mathbf{\Omega} \times \mathbf{X}$, due to Earth's rotation, Fig. (5), where $\mathbf{X}$ is the position vector (Sec. 4.3, Part 1). f may be written in several forms, $f = \nabla \times \mathbf{\Omega} \times \mathbf{X} = 2\mathbf{\Omega} \cdot \mathbf{n} = 2\Omega \sin(\text{latitude})$,

$f + \xi = \nabla \times (\mathbf{V}_\Omega + \mathbf{V})$ is appropriately termed the *absolute* vorticity, since it is indeed the vorticity of the absolute velocity, $\mathbf{V}_\Omega + \mathbf{V}$, i.e., the velocity that would be observed from an inertial reference frame, and finally,

h is the layer (or column) thickness.

The conservation equation (29) states that q is conserved following an ideal fluid parcel (or fluid column since this is a single layer model). By ideal we mean that there is no external torque due to wind stress or bottom friction. What is most important is to notice what is missing from (29): there is no process comparable to pressure work found in the energy balance, Eqn. (25), that transmits energy at the speed of gravity waves — q is transported with the fluid, exactly as a passive tracer, Eqn. (21), and not at the gravity wave velocity as is energy. Thus the potential vorticity of an ideal fluid has the conservation property of a tracer (or dye) that, once put into a fluid, remains with the fluid no matter how complex the flow may be. This includes during the process of geostrophic adjustment. But unlike a passive tracer, the potential vorticity includes an important part of the velocity field, the part having relative vorticity. In some cases it may be sufficient to calculate the evolution of a flow from q-conservation only, the steady state of an adjustment process is an important example in Sec 4.4, which

is a marked simplification over solving the full shallow water system. However, the pure gravity wave experiment in Sec. 3.1 will show that q conservation can be irrelevant if the dynamics is such that there is no vorticity.

Potential vorticity is a generalized angular momentum insofar as it accounts for a variable moment of inertia (variable thickness) as well as the planetary vorticity f due to Earth's rotation. The planetary vorticity is extremely important in this regard, because f is considerably larger than is the ξ of large scale flows, usually by a factor of 10 and often much more. A fluid column having potential vorticity will exhibit a kind of gyroscopic rigidity in the sense that it will respond to changes in the parcel configuration (thickness) or to changes in latitude and thus f. For example, suppose that a fluid column has absolute vorticity $(f_1 + \xi_1)$ and thickness h_1 in an initial state, and is then stretched from h_1 to h_2. Assuming that the stretching occurs without frictional or external torques, then the conservation law (29) applies and the absolute vorticity of the column will change to $(f_2 + \xi_2) = (f_1 + \xi_1)h_1/h_2$ (Fig. 5). Whether the change in absolute vorticity is due mainly to a change in the relative vorticity or due to a change in the planetary vorticity (latitude) can not be told without some additional information; sometimes knowing just the horizontal scale of the motion will suffice (Sec. 2, Part 3).

You may be wondering why we should make such a fuss over potential vorticity when the numerical model solves the more general shallow water equations. It is true that the winds or currents that might be inferred from a q-conservation argument can always be computed by the shallow water model and so q-conservation is redundant insofar as computation alone is concerned. The value of potential vorticity becomes evident when it is time to describe, interpret and understand a numerical model solution or a set of field observations: a potential vorticity-based description will often be simpler and yield far more insight than would the corresponding momentum plus continuity description. Of course, this presumes that potential vorticity concepts are a part of our working, fluid dynamics vocabulary. We can take a useful step in that direction by using potential vorticity balance to help describe and interpret the geostrophic adjustment experiments that follow in Sec. 4 and in Part 3.[7]

[7]Before ending this discussion it should be noted that the shallow water potential vorticity is not the most general form of potential vorticity, just as the shallow water model is not the most general fluid model. If the fluid velocity is a three component vector, which it generally is in the ocean and atmosphere, then so too is the vorticity. The three-dimensional vorticity equation includes a term that represents a change in the direction of the vorticity vector, e.g., from the horizontal into the vertical, often called 'tipping'. If the fluid flow is baroclinic, which the atmosphere and ocean generally are, then there is an additional term that arises from the cross-product of the density and pressure gradients, called the 'solenoidal' term. The corresponding three-dimensional and baroclinic potential vorticity that takes account of these additional processes is often called the Ertel potential vorticity; it reduces to the shallow water potential vorticity when the flow is appropriately two-dimensional and barotropic. By and large, the simpler shallow water potential vorticity will be adequate for an analysis of most large (horizontal) scale phenomenon, e.g., gyre-scale flows and even mesoscale eddies, once they are formed. However, it will not be adequate for analysis of small scale phenomenon, e.g., boundary layer roll vortices and three-dimensional turbulence. Knowing where this transition may occur requires a thorough understanding of Ertel potential vorticity. A very insightful discussion of some of these more advanced concepts is available online at http://www.damtp.cam.ac.uk/user/mem/

2.3.4 Finite amplitude effects

The phrases 'nonlinear' and 'finite amplitude' are often used interchangeably. Here, however, linear and nonlinear will be reserved for model equations, which are either one or the other with no gradation. The shallow water system is nonlinear because of the six terms that are the products of dependent variables, e.g., in the thickness equation (13) the terms $u\partial h/\partial x$ and $h\partial u/\partial x$ are nonlinear and so the model system is nonlinear when these are retained. [8]

Finite amplitude effects are the departures from linear dependence upon amplitude. Finite amplitude effects do not arise in the solution of a linear model, and are a matter of degree in the solution of a nonlinear model. For example, in the problems studied here, the amplitude is largely determined by η_o, the initial thickness anomaly. If the shallow water equations were linear (they aren't) then the solution $\eta(x,y,t)/\eta_o$ would be invariant to changes in η_o. Even in a nonlinear model, we can expect that $\eta(x,y,t)/\eta_o$ should be invariant to η_o in the limit of very small η_o (say $\eta_o = 1$ m with H = 500 m) simply because the nonlinear terms noted above will be very, very small compared to the dominant linear terms. The amplitude of the base case, $\eta_o = 50$ m, has been chosen to correspond with the observed SSH signature of mesoscale eddies (Fig. 1). At this amplitude, there are several interesting and important finite amplitude phenomena, including very large float (material) displacements given sufficient time. Nevertheless, a great deal of what we can see in a solution for $\eta(x,y,t)$ and $\mathbf{V}(x,y,t)$ can be attributed to linear wave dynamics and especially to the dispersion property of the waves supported by the linear shallow water system. (more on this below and in Part 3). The linear subset of the shallow water model results when $D(\ )/Dt$ is replaced everywhere by $\partial(\ )/\partial t$, and the variable layer thickness h is replaced by the constant initial thickness, H. In Cartesian components the linear shallow water model is

$$\begin{aligned}
\frac{\partial h}{\partial t} &= -H(\tfrac{\partial u}{\partial x} + \tfrac{\partial v}{\partial y}),\\
\frac{\partial u}{\partial t} &= -g'\tfrac{\partial h}{\partial x} + fv,\\
\frac{\partial v}{\partial t} &= -g'\tfrac{\partial h}{\partial y} - fu.
\end{aligned} \tag{30}$$

The stratification is represented in this one layer model by g' and H and assumed homogeneous. The stratification determines the phase speed of waves. The f is these equation is in general dependent

[8] Terms like $u\ \partial h/\partial x$ that are the product of two dependent variables are sometimes said to be 'bilinear' or 'semi-linear'. Model equations having such terms are nonlinear in the most important sense that a superposition of valid solutions does not yield a new valid solution. However, some solution methods are applicable to semi-linear systems, e.g., the method of characteristics, that are not applicable to more comprehensively nonlinear systems.

upon latitude (or north distance, y); the spatial variation of f determines the kinds of waves that are possible.

2.4 Boundary and initial conditions define the problem

The mass and momentum equations (16) and (17) could be used to model a wide range of phenomena. The definition of a specific, solvable problem follows from the specification of an initial condition on h and $\mathbf{V}$ throughout the model domain, and the definition of boundary conditions along the edges of the domain.

The problems studied here are variants on the classical problem of geostrophic adjustment[??] and specifically, geostrophic adjustment of a mesoscale eddy-like feature at mid-latitude. The stratification is chosen to be typical of the subtropical main thermocline,

$$H = 500 \text{ m} \text{ and } \delta\rho = 2 \text{ kg m}^{-3}, \tag{31}$$

and the resulting gravity wave speed,

$$C = \sqrt{g(\delta\rho/\rho_o)H} \approx 3.1 \text{ m s}^{-1},$$

is that of an internal or baroclinic gravity wave (Sec. 2.4.2). The initial condition is either a one-dimensional ridge with half-width L,

$$\eta(x,t=0) = \eta_o \text{ if } |x| \leq L, \text{ or else } \eta(x,t=0) = 0, \tag{32}$$

(here in Sec. 3) or a two-dimensional eddy with radius = L (in Sec. 4).[9] The half-width and the amplitude of the initial interface displacement are chosen to be comparable to observed mid-latitude mesoscale eddies (Fig. 1),

$$L = 100 \text{ km} \quad \text{and} \quad \eta_0 = 50 \text{ m}. \tag{33}$$

The initial velocity can be one of several forms, the default being a state of rest,

$$\mathbf{V}(x,t=0) = 0. \tag{34}$$

[9] The initial ridge given by Eqn. (32) has a very sharp edge, which in (numerical) practice is smoothed over a few horizontal grid points. Nevertheless, when released suddenly, this initial condition tends to produce energetic gravity waves that are not highly realistic of most natural phenomenon. Wind or tidal forcing acting on the ocean are by comparison rather slowly varying in time and space, and so less apt to produce energetic gravity waves. The initial condition may be easily changed to something smoother, and for that matter, the models may be readily configured to allow realistic wind or tidal forcing as in Part 3.

If the initial velocity and vorticity vanish, then the initial ridge has a potential vorticity $q = f/(H + \eta_o)$ compared with f/H in the outlying fluid. This potential vorticity anomaly is preserved during geostrophic adjustment, the full import of which will become clear in Sec. 4.5.

The computational domain will be either one-dimensional, having a width of several thousand kilometers, where that is appropriate (Secs. 3 and 4, where f is presumed to be a constant), or two-dimensional and 4000 km on a side (Sec. 5). There is no attempt to compute the fluid state outside of the computational domain, and so something other than the momentum and continuity equations will have to be imposed on the boundaries.

The only energy in the initial state is associated with the initial eddy which is placed near the center of the model domain. It is then reasonable that waves will be radiated outward only, i.e., that nothing will come into the model domain from the outside. A plausible and generally effective representation of this one-way, outward transfer is made by imposing a radiation boundary condition along the boundaries:

$$\frac{\partial \psi}{\partial t} = -U_{rad}\frac{\partial \psi}{\partial n}, \tag{35}$$

where ψ is thickness or a velocity component, n is the direction normal and outward from the boundary and U_{rad} is the appropriate velocity component normal to the boundary. This amounts to imposing a one-dimensional advection process normal to the boundary and at the speed U_{rad}, which is very important to the success of (35). The appropriate value of U_{rad} depends upon the dominant process local to the boundary, and that may change with time. Here, during the first 20 days of an experiment, U_{rad} is taken to be the gravity wave speed, $U_{rad} = C = \sqrt{g'H}$, the fastest wave speed in the shallow water system (Sec. 3.1). This serves well to usher along the gravity waves that first reach a distant boundary. But after that comes trouble ... very, very slowly ... in the form of low frequency Rossby waves (in Part 3).

2.5 Appendix to Sec. 2: Normal modes of a two layer model ocean

The shallow water model applied to a homogeneous (unstratified and so *barotropic*) ocean stands on its own. But if the shallow water model is applied to internal or *baroclinic* motions (defined below), then some discussion of the correspondence between observed and modelled layer thicknesses and phase speeds seems necessary.[10] In that vein it is useful to examine briefly the wave properties of a simple two

[10]The terms barotropic and baroclinic describe the dependence of density upon pressure. A barotropic fluid is one in which the density can be written as a function of the pressure only, $\rho = \rho(P)$. This would hold always in a fluid that was homogeneous, for example, in the (single layer) shallow water model. It would also hold in a fluid that was density stratified, provided that the density surfaces and pressure surfaces were everywhere parallel. This would likely be true in a resting state, but not when motion causes vertical displacements of density surfaces that will often greatly exceed the associated

layer model. For this purpose rotation and nonlinearity may be ignored, and the motion presumed to be in one horizontal dimension only (x,t). The stratification is represented by two homogeneous layers, Fig. (6), an upper layer 1, having an undisturbed thickness H_1 and density ρ_1, and a lower layer, 2, with undisturbed thickness H_2 and slightly greater density, $\rho_2 = \rho_1 + \delta\rho$. If this was meant to represent the open ocean, then the interface between the layers would correspond with the middle of the main thermocline and typical values would be $H_1 = 500$ m, $H_2 = 3500$ m, $\rho_1 = 1030$ kg m^{-3} and $\rho_2 = \rho_1 + \delta\rho$ with $\delta\rho = 2$ kg m^{-3}.

The interest here is in large scale, fairly slowly varying motions for which vertical accelerations are very, very gentle, $\ll g$, and the pressure is consequently hydrostatic, i.e., due to the weight of the fluid overhead. The bottom pressure is $P_b = g(\rho_1 h_1 + \rho_2 h_2)$, and the pressure within the layers varies as

$$P_1(x,z,t) = g\rho_1(h_1(x,t) + h_2(x,t) - z) \quad \text{and} \quad P_2(x,z,t) = g\rho_1 h_1(x,t) + g\rho_2(h_2(x,t) - z),$$

where $h(x,t)$ is the space and time-varying layer thickness. The pressure gradient divided by the density is then in each layer,

$$\frac{1}{\rho_1}\frac{\partial P_1}{\partial x} = g(\frac{\partial h_1}{\partial x} + \frac{\partial h_2}{\partial x}) \quad \text{and} \quad \frac{1}{\rho_2}\frac{\partial P_2}{\partial x} = g\frac{\rho_1}{\rho_2}\frac{\partial h_1}{\partial x} + g\frac{\partial h_2}{\partial x}. \tag{36}$$

With these results in hand, the momentum and continuity equations for the velocity and layer thicknesses are (anticipating a matrix format):

$$\begin{aligned}
\frac{\partial u_1}{\partial t} + g\frac{\partial h_1}{\partial x} + g\frac{\partial h_2}{\partial x} &= 0,\\
H_1\frac{\partial u_1}{\partial x} + \frac{\partial h_1}{\partial t} &= 0,\\
g\frac{\rho_1}{\rho_2}\frac{\partial h_1}{\partial x} + \frac{\partial u_2}{\partial t} + g\frac{\partial h_2}{\partial x} &= 0,\\
H_2\frac{\partial u_2}{\partial x} + \frac{\partial h_2}{\partial t} &= 0.
\end{aligned} \tag{37}$$

Notice that the thickness of a given layer can vary only if there is divergence within that layer, and, that the pressure gradient has a dependence upon both layer thicknesses. Thus the pressure gradient couples the layers together.

---

displacement of pressure surfaces. In that case, the fluid would be described as baroclinic, meaning that density and pressure surfaces intersect, and so density varied with more than the pressure, i.e., with horizontal position and time at a given pressure. Thus, barotropic is a special case in which density surfaces are always parallel with pressure surfaces, while baroclinic is all else. The distinction is important in that the pressure gradient can (will) generate vertical shear in a baroclinic fluid, but not in a barotropic fluid. Thus a shallow water model will be realistic model of a barotropic fluid and flow. A shallow water solution will require some interpretation if, as here, it is meant to represent a baroclinic phenomenon, e.g., ocean mesoscale eddies.

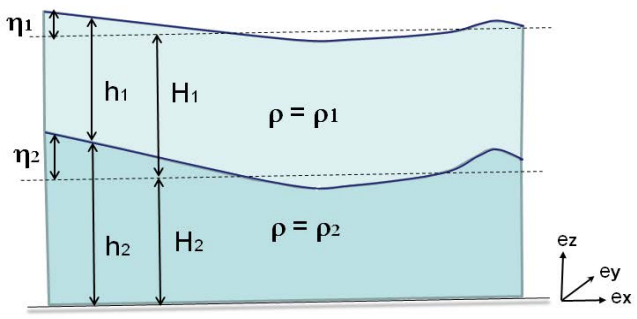

Figure 6: A two layer model appropriate to the stratified ocean. The dashed lines are level surfaces. The undisturbed layer thicknesses are H_1 and H_2, and the actual thicknesses are h_1 and h_2. The upper surfaces of the layers are displaced to $\eta_1 = h_1 + h_2 - (H_1 + H_2)$ and $\eta_2 = h_2 - H_2$. The pressure of the atmosphere above is presumed horizontally uniform and so negligible.

To find the wave properties of this system, presume that a wave may exist in both layers: $u_1(x,t) = U_1 cos(kx - \omega t)$, $u_2(x,t) = U_2 cos(kx - \omega t)$, $h_1(x,t) = H_1 + \Gamma_1 cos(kx - \omega t)$ and $h_2(x,t) = H_2 + \Gamma_2 cos(kx - \omega t)$. The U_1, Γ_1, U_2, Γ_2 are constant but to this point unknown amplitudes. Notice that the wave frequency and wavenumber are the same in the two layers, which is implicit in 'a wave', and are also unknown. After substitution of this wave form, the governing equations (37) may be written in matrices as

$$\begin{pmatrix} -\omega & gk & 0 & gk \\ H_1 k & -\omega & 0 & 0 \\ 0 & gk\frac{\rho_1}{\rho_2} & -\omega & gk \\ 0 & 0 & H_2 k & -\omega \end{pmatrix} \begin{pmatrix} U_1 \\ \Gamma_1 \\ U_2 \\ \Gamma_2 \end{pmatrix} = 0.$$

The simplest and most insightful description of a multi-part, linear system such as this will often be in terms of the normal mode frequencies and structure, also called the eigenvalues and eigenvectors. The normal modes are independent of one another, and hence the governing equations written in terms of the normal modes will be decoupled, which is not the case with Eqns. (37) written in the layer-wise thickness and velocity. To find the frequencies of the normal modes it is necessary to solve the characteristic equation of the coefficient matrix (same as saying that the determinant must vanish), which is a fourth order polynomial in ω

$$\omega^4 - gk^2(H_1 + H_2)\omega^2 + g^2 k^4 H_1 H_2 (\frac{\rho_2 - \rho_1}{\rho_2}) = 0.$$

This characteristic equation is bi-quadratic (the cubic and linear terms vanish) and can be readily solved

as a quadratic equation for ω^2,

$$\omega^2 = \frac{gk^2(H_1+H_2)}{2} \pm \frac{\sqrt{g^2k^4(H_1+H_2)^2 - 4g^2k^4H_1H_2(\frac{\rho_2-\rho_1}{\rho_2})}}{2}. \tag{38}$$

In the oceanic case in which $\frac{\rho_2-\rho_1}{\rho_2} \ll 1$ this may be simplified by factoring the term under the square root and then using the binomial theorem that $(1+\varepsilon)^{1/2} \approx 1+\varepsilon/2$ when $\varepsilon \ll 1$ as applies here,

$$\omega^2 = \frac{gk^2(H_1+H_2)}{2} \pm gk^2\frac{(H_1+H_2)}{2}\left(1 - 2\frac{H_1H_2}{(H_1+H_2)^2}(\frac{\rho_2-\rho_1}{\rho_2})\right). \tag{39}$$

2.5.1 Barotropic mode

The larger of the two roots of Eqn. (39) is the frequency squared of the external or *barotropic* mode,

$$\omega_{btr}^2 = gk^2(H_1+H_2) - gk^2\frac{H_1H_2}{(H_1+H_2)}(\frac{\rho_2-\rho_1}{\rho_2}). \tag{40}$$

Since the trailing factor involving the density difference is very, very small, ≈ 0.002, an excellent approximation of this frequency is

$$\omega_{btr} \approx \pm k\sqrt{g(H_1+H_2)}. \tag{41}$$

If the intent is to reproduce this dispersion relationship within a shallow water model, then the single-layer equivalent gravity and layer thickness are simply

$$g_e = g, \quad \text{and} \quad H_e = H_1+H_2, \tag{42}$$

which might have been guessed without help from a two layer model. This is the barotropic mode. The same kind of result is slightly less obvious for the baroclinic mode coming next.

By putting the appropriate frequency (41) back into the governing equations (37) we can solve for the ratio of the amplitude of the layer thicknesses changes,

$$\frac{\Gamma_1}{\Gamma_2} = \frac{\rho_2}{2\rho_1}\left((H_1 + (H_1^2+H_2^2-2H_1H_2+4H_1H_2\frac{\rho_1}{\rho_2})^{1/2})/H_2 - 1\right). \tag{43}$$

Noting that ρ_1/ρ_2 is very close to 1, then as a very good approximation,

$$\frac{\Gamma_1}{\Gamma_2} \approx \frac{H_1}{H_2}. \tag{44}$$

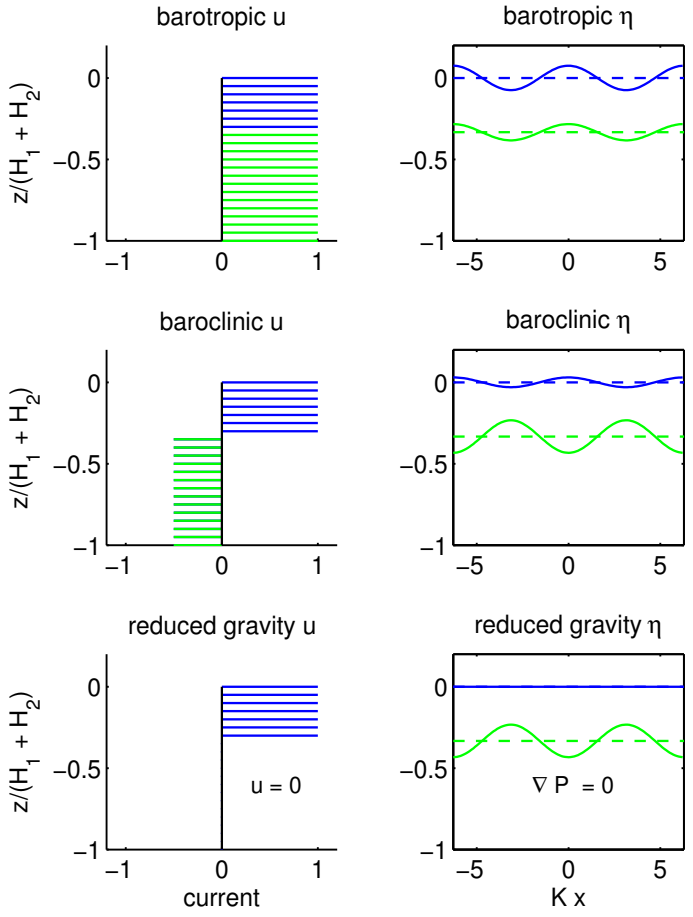

Figure 7: Normal modes of a two-layer ocean. **(upper)** The barotropic normal mode; current profile (upper left, arbitrary amplitude) and the sea surface and interface displacements (upper right, blue and green lines). **(middle)** The baroclinic normal mode. The currents and the sea surface and interface displacements have a self-consistent structure, but the thicknesses, $H_1 = 500$ m and $H_2 = 1000$ m, and the density difference, $\delta\rho = 300$ kg m^{-3}, were chosen to make this structure legible rather than realistic of the ocean thermocline. **(lower)** The reduced gravity approximation of the baroclinic mode uses only the density interface to compute the pressure gradient in the upper layer on the assumption that the lower layer currents and pressure gradient vanish. This would not be a good approximation for the stratification shown in this figure because H_1/H_2 is not $\ll 1$.

In this linear model the amplitude of either Γ_1 or Γ_2 is arbitrary, but the ratio Γ_1/Γ_2 is determined by the dynamics. Eqn. (44) indicates that the layer thicknesses oscillate in phase and with an amplitude that is proportional to the undisturbed layer thickness. For some purposes it is helpful to know the displacement of the upper surfaces of the layers; the sea surface, sometimes called the free surface, is displaced from its resting (level) height by

$$\eta_1 = (h_1 + h_2) - (H_1 + H_2),$$

and the density interface between the layers by

$$\eta_2 = h_2 - H_2.$$

In the barotropic mode, the amplitude of the vertical displacement of the free surface compared to the displacement of the interface is then, using Eqn. (44),

$$\frac{\Gamma_1 + \Gamma_2}{\Gamma_2} = \frac{H_1 + H_2}{H_2} \approx 1.14, \tag{45}$$

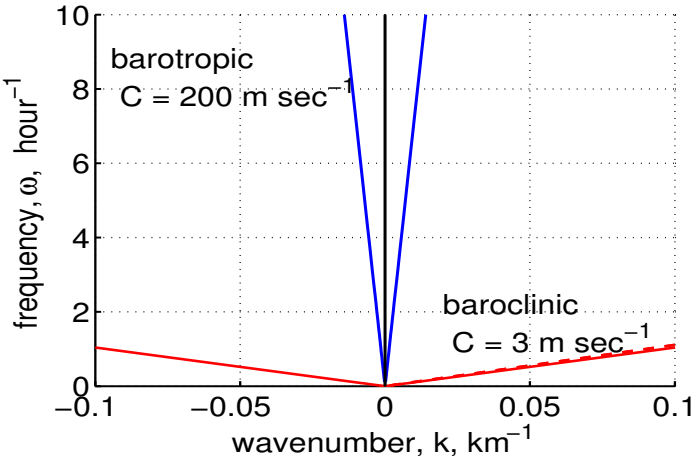

Figure 8: The dispersion relation for gravity waves of a two layer, nonrotating model of a stratified ocean. The solid lines are the full dispersion relation, and the dashed red line is an approximation discussed in the main text. At a common wavenumber, the barotropic and baroclinic modes have vastly different frequencies and phase speeds.

for our nominal, open ocean stratification. The vertical displacement is thus a maximum at the sea surface and decreases linearly to zero at the bottom (Fig. 7, upper). The displacements $\eta_1(x,t)$ and $\eta_2(x,t)$ have a $cos(kx-\omega t)$ time- and space-dependence that is common to both layers, which is what distinguishes a mode from an arbitrary motion. The pressure gradient is due almost entirely to the displacement of the free surface, and the currents are the essentially the same in the two layers (uniform with depth). The density interface η_2 moves up and down exactly as does the pressure at that level, and thus the density could be written as a function of the pressure in the initial state and during the subsequent motion. This kind of wave motion in which $\rho = \rho(P)$ everywhere in the fluid is, as here, termed barotropic.

The phase speed of a barotropic wave for a nominal, open ocean depth $H_1 + H_2 = 4000$ m is very fast,

$$C_{btr} = \frac{\omega_{btr}}{k} = \sqrt{g(H_1+H_2)} \approx 200 \text{ m sec}^{-1} \approx 680 \text{ km hour}^{-1},$$

or comparable to that of a jet transport. High frequency (5 to 20 min period) barotropic waves of this kind (in some circumstances called *tsunamis*, Japanese for harbor wave) are the primary oceanic response to a rapid vertical displacement of the sea floor (a few meters in a few minutes over a large horizontal scale). Because the fluid velocity associated with open ocean tsunami waves is comparatively gentle, a few centimeters per second, these waves undergo very little dissipation, and, aside from two-dimensional spreading, may arrive on a distant shore with a significant amplitude. Lower frequency barotropic waves (periods 1/2 to 1 day) are the principal components of the open ocean, astronomical tides. These near-daily frequency waves gravity waves are modified substantially by Earth's rotation and are often called inertia-gravity waves (Sec. 4, also called Poincare waves).[11]

[11] An excellent resource for tsunami waves is http://www.tsunami.noaa.gov/ A classic paper on open ocean tides is by http://articles.adsabs.harvard.edu//full/1944MNRAS.104..244P/0000254.000.html

2.5.2 Baroclinic mode

The smaller of the two roots of Eqn. (39) is labeled the internal or *baroclinic* normal mode, and

$$\omega_{bcl} = \pm k\sqrt{g\frac{H_1H_2}{(H_1+H_2)}(\frac{\rho_2-\rho_1}{\rho_2})}. \tag{46}$$

Baroclinic gravity waves have a much lower frequency and phase speed than does a barotropic wave of the same wavelength,

$$\frac{\omega_{bcl}}{\omega_{btr}} = \frac{C_{bcl}}{C_{btr}} \approx \sqrt{\frac{g'H_1H_2}{g(H_1+H_2)^2}} \approx \frac{1}{70},$$

mainly because the reduced gravity, $g' = g(\rho_2-\rho_1)/\rho_2 = g(2/1000)$, is very much less than the full gravity, g. If the intent is to model the baroclinic dispersion relation using a shallow water model, then

$$g_e = g' \text{ and } H_e = \frac{H_1H_2}{(H_1+H_2)} \tag{47}$$

will give an appropriate phase speed.[12] It is usually the case that H_2 exceeds H_1 by a factor of 5 - 10, and as a fair approximation, $H_e \approx H_1$. The baroclinic, long gravity wave phase speed is then, for the nominal stratification,

$$C_{bcl} = \frac{\omega_{bcl}}{k} = \sqrt{g'H_1} \approx 3.1 \text{ m sec}^{-1} \approx 270 \text{ km day}^{-1},$$

the red line of Fig. (8).

The ratio of layer thicknesses in the baroclinic mode is

$$\frac{\Gamma_1}{\Gamma_2} = \frac{\rho_2}{2\rho_1}\left((H_1 - (H_1^2+H_2^2-2H_1H_2+4H_1H_2\frac{\rho_1}{\rho_2})^{1/2})/H_2 - 1\right). \tag{48}$$

Making the same approximation noted earlier for the barotropic mode,

$$\frac{\Gamma_1}{\Gamma_2} \approx -\frac{\rho_2}{\rho_1} \approx -1, \tag{49}$$

and hence the layer thicknesses oscillate out of phase, and with nearly equal amplitude (Fig. 7, middle). There is a small but very important displacement of the free surface,

$$\frac{\Gamma_1+\Gamma_2}{\Gamma_2} \approx -\frac{\rho_2-\rho_1}{\rho_1} = -0.0017,$$

[12]Most authors define the equivalent depth from phase speed only, i.e., $H_e = C_{bcl}^2/g$ with g the full gravity. With that definition, the baroclinic $H_e \approx 1$ m (!). In other words, the baroclinic gravity wave mode has the phase speed of a 1 m thick, homogeneous (barotropic) layer. In a linear shallow water model and problem, the phase speed is all that matters and this can be a useful definition. However, if the shallow water model includes finite amplitude effects and is intended to simulate realistic amplitudes, then the actual layer thickness is relevant, and the equivalent values of Eqn. (47) seem more apt.

for the given stratification. An interface (thermocline) displacement of 50 m will thus be accompanied by a free surface displacement of about $50 * (\delta\rho/\rho_o) \approx 10$ cm, which is readily detectable to satellite-based, altimetric methods, as in Fig. (2). The currents in the upper and lower layers are exactly out of phase, and their ratio is such that the net transport vanishes.

2.5.3 Reduced gravity approximation of the baroclinic mode

The transport (depth integrated velocity) of the baroclinic normal mode = $u_1H_1 + u_2H_2 = 0$. Thus the upper and lower layer currents are in the ratio $u_1/u_2 = -H_2/H_1$. In the limit that $H_1/H_2 \to 0$, the lower layer current and pressure gradient are much, much less than in the upper layer. In that case, an approximation of Eqn. (36) is that

$$\frac{\partial h_2}{\partial x} = -\frac{\rho_1}{\rho_2}\frac{\partial h_1}{\partial x},$$

which may be used to eliminate the h_2 term from the upper layer pressure gradient,

$$\frac{\partial P_1}{\partial x} = g\left(\frac{\rho_2 - \rho_1}{\rho_2}\right)\frac{\partial h_1}{\partial x} = g'\frac{\partial h_1}{\partial x}.$$

The density interface can thus be used as a proxy for the sea surface insofar as the pressure gradient in the upper layer is concerned, provided that the full gravity that multiplies the sea surface slope is replaced by the much smaller reduced gravity, g', multiplying the much greater slope of the density interface. In that way all reference to the lower layer may be eliminated from the upper layer equations. This is sometimes referred to as a one and a half layer model, where the half refers to the deep, resting lower layer, and it is also called the reduced gravity approximation, (Fig. 7, lower). This is an appropriate interpretation of the shallow water model when applied to a simulation of baroclinic, mesoscale eddies which are mainly upper ocean (layer one) phenomena, and plausible for modelling some aspects of the wind-driven circulation of subtropical gyres (Fig. 9 and more in Part 4). The reduced gravity approximation may also be used in a multi-layered model, so long as it is appropriate to approximate the deep flow and pressure gradient as vanishing.

2.6 Problems

(1) What is the interpretation of $d/dt = 0$ in a Lagrangian system? How about $\partial/\partial t = 0$ and $D/Dt = 0$ in an Eulerian system? The material derivative and some aspects of advection are discussed in greater detail in 'Lagrangian and Eulerian representations...'. Price, James F., 12.808 Supplemental Material, Topics in Fluid Dynamics: Dimensional Analysis, the Coriolis Force, and Lagrangian and Eulerian Representations, https://ocw.mit.edu/resources/res-12-001-topics-in-fluid-dynamics-spring-2010/

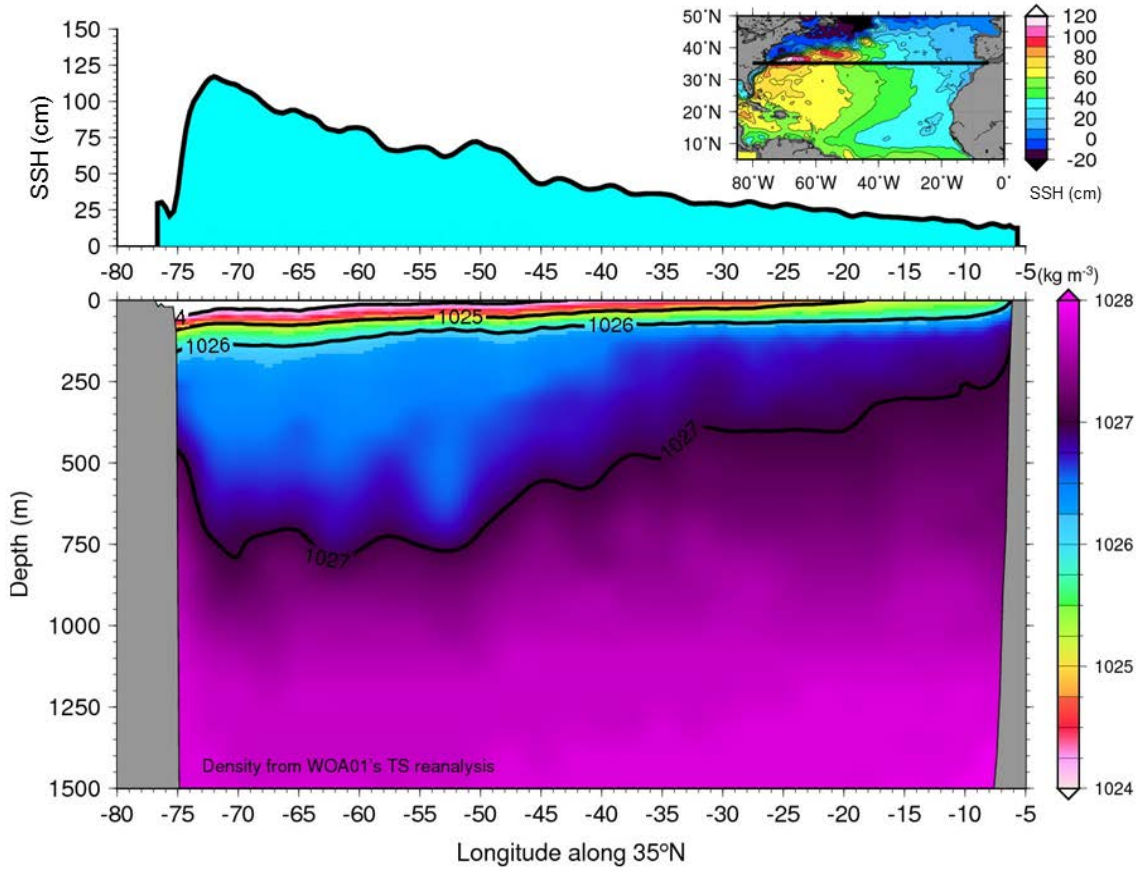

Figure 9: A cross section of the North Atlantic subtropical thermocline, sliced east-west along 35^oN, and viewed looking toward the north. The upper panel is SSH as in Fig.1, but here the monthly average for September over about twenty years of measurement. The lower panel is the long-term, September average of density along 35^oN from the World Ocean Atlas 2001 (http://www.nodc.noaa.gov/OC5/WOA01/pr_woa01.html). The tilt of the thermocline mirrors the tilt of the sea surface so that high SSH corresponds to a thick, warm upper layer. A result is that the pressure gradient and geostrophic velocity are comparatively small in the abyssal ocean, roughly 1500 - 2500 m depth, suggestive of a reduced gravity approximation (Section 2.4.3). Not shown in this figure are bottom-trapped density currents found at depths in excess of about 3000 m that make up the lower limb of the meridional overturning circulation. This figure was kindly provided by Iam-Fei Pun of WHOI.

(2) If phase speed and group speed are not already familiar to you, then the very brief discussion of Sec. 2.2.1 will probably not suffice. An excellent primary reference is Chapters 1 and 2 of Pedlosky's 'Waves in the Ocean and Atmosphere' (footnote 3). For a quick refresher, you might take a look at the script twowaves.m (Sec. 6.3), which allows you to define an arbitrary dispersion relationship between two waves that are then superimposed. The envelope of the superposition, and thus a wave form or pulse, propagates at the group speed, $C_g = \delta\omega/\delta k$, where δ is the difference between the two waves.

(3) You will feel much more at ease with the important energy and potential vorticity conservation laws if you will take the time to derive them yourself. There are no special techniques or approximations required, though the full shallow water equations entail a fair degree of manipulation. It will be helpful to start with the linear shallow water equations (see the next section) and then add just one of the advection terms, say $u\partial h/\partial x$. Performing this derivation does not aid applications, but it does help answer two important questions. 1) Does the linear shallow water system yield a self-consistent energy and vorticity balance? 2) Does the q conservation law include all of the phenomena of the shallow water system? How about energy conservation? (This second question will be answered in Sec. 3.1.)

(4) Some q-conservation problems for you: 1) A right cylinder having a radius r and height h has a moment of inertia $I = Mr^2/2$ where M is the mass of the cylinder. The angular momentum due to rotation at a rate ω about the central axis is $L = I\omega$. Show that conservation of angular momentum under changes of h and r that are mass (volume) conserving can be summarized with a vorticity conservation law like Eqn. (29). 2) Because $f \gg \xi$, generally, a change in thickness of 10% is often highly significant for the relative vorticity, as is a change in latitude of only a few degrees. Assume that a fluid column having a radius of 50 km and a thickness of 700 m is moved from 40^oN to 25^oN and that it conserves q. If all of the change in f is accounted by changes in ξ (no thickness change), estimate the magnitude of the resulting current. Now suppose that the change in f is accommodated entirely by a change in thickness.....how much?

(5) The algebra required for a three layer model is a bit tedious and is a good application for symbolic mathematics; see twolayer_eig.m linked in Sec. 6.3.

(6) Some modal questions for you: 1) Any arbitrary configuration of layer thicknesses may be decomposed into the normal modes. What would you infer is the two layer model composition of the offhand sketch of layer thicknesses in Fig. (6)? 2) In place of two active layers, suppose three layers of equal thickness. What would you guess for the modal structure of η? Check your answer against the (numerical) eigenvectors of twolayer_eig.m (Sec. 6.3). 3) The reduced gravity approximation is counterintuitive in that the density interface displacement is used to compute the hydrostatic pressure anomaly in the layer *above* the density interface. Rather than try to explain that while developing the shallow water model in Sec. 2.1, this valuable and sensible approximation (or interpretation) was deferred to this appendix. The essential, physical connection between pressure and mass or thickness anomaly is that a comparatively thick upper layer is a region of high pressure anomaly, as in Fig. (9). What do you think would happen if, due to a sign error in the code of a numerical model (we all make them) this relationship was reversed?

3 Gravitational adjustment

The sequence of experiments described here and in Part 3 differ mainly in the way that Earth's shape and rotation are represented. Of course, we know that the Earth is approximately spherical and that the Coriolis parameter varies significantly with latitude. Nevertheless, in the first experiments described here in Sec. 3, the model domain is presumed to be flat and not rotating, i.e.,

$$f = 0,$$

and there is no Coriolis force. In that case the shallow water model supports pure gravity waves only. Then, in Sections 4 and 5, the Earth will be approximated as rotating but flat, i.e.,

$$f = f_o = constant,$$

in which case there are gravity waves, inertial oscillations and geostrophic motions all at once. With an understanding of those experiments in hand, we will then be ready in Part 3 for the realistic case, spatially varying f,

$$f = 2\Omega sin(\text{latitude}),$$

first a mid-latitude case and then an equatorial case. These latter two experiments will include all of the phenomena that arose before, plus a low frequency wave called a planetary wave or Rossby wave.

3.1 Just gravity waves

The experiment starts at $t = 0$ when the thickness anomaly of dense water is released. Within the first tens of minutes the ridge begins to slump under the force of gravity, releasing potential energy and generating currents (kinetic energy). In this experiment, the motions take the form of two equal, outgoing solitary 'wave pulses' of amplitude $\eta_o/2$ and width $2L$. These pulses move at a steady speed that is very close to the (baroclinic) gravity wave speed, $C = \sqrt{g'H}$, and hence they run off of the model domain in just a few days. Other than these two discrete wave pulses, there is nothing. If you (like the author) had expected the response to look something like the waves excited on the surface of a pond by an errant golf ball, or tsunami waves generated on the surface of the ocean by a moving sea floor, then this solution will seem strange indeed. This raises a string of questions, some that should arise in any numerical study, and others that are specific to this case. In the first place, how can we know that this numerical solution is faithful to the model equations? [13] And then, these pulses look nothing like elementary gravity waves and so is it appropriate to call them 'wave' pulses?

[13] Looking at the plot of a numerical solution is a little bit like looking at Saturn through a new, homemade telescope (Fig. 11, Part 1). Barring catastrophic errors, the largest eddies and the rings will be obvious and unmistakable. However, at

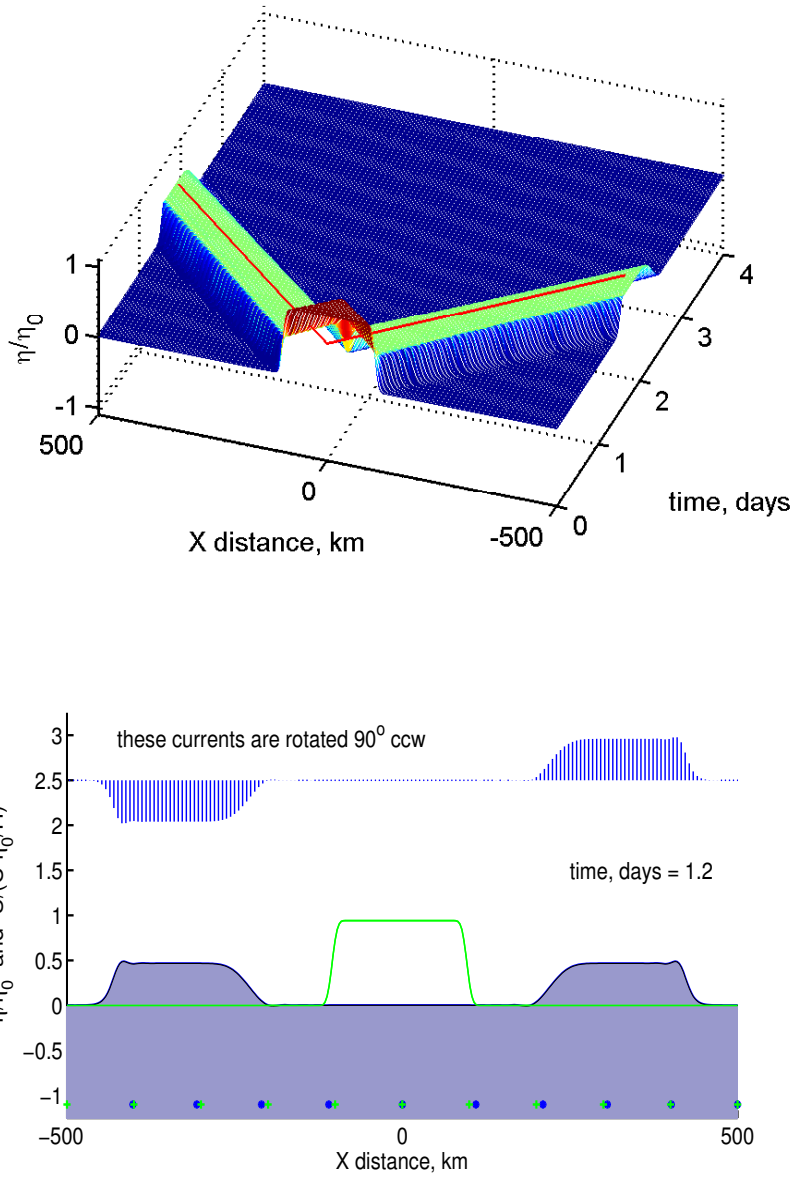

Figure 10: An experiment in gravitational adjustment (no rotation) solved by the numerical model geoadj_1d.m linked in Sec. 6.3. **(upper)** The nondimensional (scaled) thickness anomaly, $\eta(x,t)/\eta_0$, following the release of a rectangular ridge that had an initial amplitude $\eta_0 = 50$m. The thin red lines have a slope given by the gravity wave speed, $C = \sqrt{g'H} = 3.1$ m s^{-1} in this experiment. This solution was computed by the numerical model, geoadj_1d.m (Sec. 6.3). **(lower)** A snapshot of the solution at $t = 1.2$ days. The green line shows the initial thickness anomaly. The current, which is shown by the array of vectors plotted above, is in the x-direction only and has been rotated 90^o counterclockwise to be made visible here. The current is scaled with $C\eta_o/H$ (= 0.3 m s^{-1} for the case $\eta_o = 50$ m). The blue dots and green crosses at depth = -1.1 show the present positions and the initial positions of floats.

3.2 An exact solution of the linear, one-dimensional wave equation

There is a well-known, exact analytic solution for the initial value problem of the linear one-dimensional wave equation that can serve as a very useful reference for some aspects of this numerical solution. The wave equation of the linear, nonrotating shallow water system, Eqns. (30) with $f = 0$ and $v = 0$) is readily found by eliminating u in favor of η (recall that $\eta = H + h$ with H a constant and so $\partial h/\partial t = \partial \eta/\partial t$),

$$\frac{\partial^2 \eta}{\partial t^2} = g' H \frac{\partial^2 \eta}{\partial x^2}. \tag{50}$$

Given that the initial data is

$$\eta(x, t = 0) = \eta_o(x), \tag{51}$$

then the d'Alembert solution

$$\eta(x,t) = \frac{1}{2}\eta_o(kx - \omega t) + \frac{1}{2}\eta_o(kx + \omega t), \tag{52}$$

solves (50) and (51) provided that

$$\omega = k\sqrt{g'H}. \tag{53}$$

The relationship $\omega(k)$ given by Eqn. (53) is called the dispersion relation, and is the crucial and distinguishing property of a wave system. This is a particularly simple dispersion relation, $\omega(k)$ being a straight line with slope $\sqrt{g'H}$, Fig. (8). Waves of all wavenumber thus have the same phase speed,

$$\boxed{C_p = \frac{\omega}{k} = \sqrt{g'H} = C} \tag{54}$$

and the same group speed,

$$\boxed{C_g = \frac{\partial \omega}{\partial k} = \sqrt{g'H} = C} \tag{55}$$

Because all of the waves in this system (that is to say, waves of all wavenumbers) have the same phase speed, this system (shallow water, linear, nonrotating) is said to be nondispersive. As a consequence, the initial shape, $\eta_o(x)$, is retained in the propagating wave pulses. Examples of dispersive wave systems arise in later problems. Notice that the phase and group speed in this linear model are dependent upon the stratification g' and H only, and are independent of the amplitude of the motion.

some level of detail an active skepticism is healthy, even essential to avoid over-interpretation or outright error. For example, would you notice that Saturn is somewhat flattened if you didn't anticipate it? How could you discern genuine flatness from an optical distortion? One way would be to observe other, better-known phenomenon, e.g., the Moon, to calibrate the instrument's performance and extrapolate from there. Something like that can be done for numerical solutions by verifying that they follow adequately the conservation laws for energy and potential vorticity, among others, and that they reproduce known solutions well enough. The qualifications used here are deliberate; numerical solutions will never be perfect, and they don't need to be. You do have to verify that they are good enough for the problem at hand.

Gravity waves are the only possible nontrivial motion in this system. There is nothing that picks out a favored direction, and hence it is isotropic (from Greek *iso* + *tropos*, equal in all directions). Thus by symmetry half of the initial ridge propagates in one direction, and half goes the other.

The current that accompanies the wave pulses is in the direction of the pulse propagation, i.e., toward positive x within the right-traveling pulse. The velocity and the pressure gradient within the pulses is thus colinear (parallel or anti-parallel), a relationship often termed longitudinal in this context, and that is characteristic of pure gravity wave motion. On the other hand, geostrophic motion is transverse insofar as the velocity and the pressure gradient are approximately normal to one another (Part 1, Sec. 5 and other examples coming soon). The observed velocity/pressure gradient relationship thus provides a useful qualitative clue to dynamics at the level of the momentum balance.

Perhaps the two most important points are that 1) comparison with the exact d'Alembert solution shows that the numerical solution of Fig. (10) is qualitatively correct (more details below), and 2) the isolated pulses seen in the numerical solution are gravity waves in all respects save their appearance (and our prejudice that waves should look sinusoidal). Their structure is that of the initial condition and is retained because of the nondispersive property of shallow water, nonrotating gravity waves combined with the one-dimensional geometry of this specific problem.

3.3 The choice of scales

There is usually more than one way to plot a model solution (Fig. 10). It is highly desirable to use coordinates that will allow for use by the broadest possible audience. This generally requires the use of nondimensional coordinates, i.e., that independent variables, distance, time, etc. be reported in units that are intrinsic (natural) to the model system rather than in meters or seconds. In that respect, a curious feature of the linear, nonrotating shallow water model, Eq. (50), is that the gravity wave speed, $C = \sqrt{g'H}$, is the only intrinsic scale; there is no intrinsic horizontal length scale or time scale, as there will be when rotation is included in Sec. 3.1. Dimensional scales that are convenient for the specific initial condition, kilometers for horizontal distance and days for time, were therefore chosen for Fig. (10) and will be retained in all later plots to facilitate comparison with this first experiment.

The scales for the dependent variables, η and u, were chosen with the goal that the scaled (nondimensional) variables should be independent of the initial amplitude, η_0, in the limit of small amplitude. The obvious scale for the thickness anomaly is η_0, the initial value. Thus the thickness anomaly is plotted as η/η_o. (To preserve the important sign of η_o it is necessary to use the absolute value, $\eta/\mid \eta_o \mid$, though the plot legends do not show this.) The appropriate scale for the fluid speed is less obvious. The gravity wave speed $C = \sqrt{g'H}$ seems a promising candidate at first. However, C depends only upon the fluid properties and is independent of the amplitude, here η_o, while the fluid

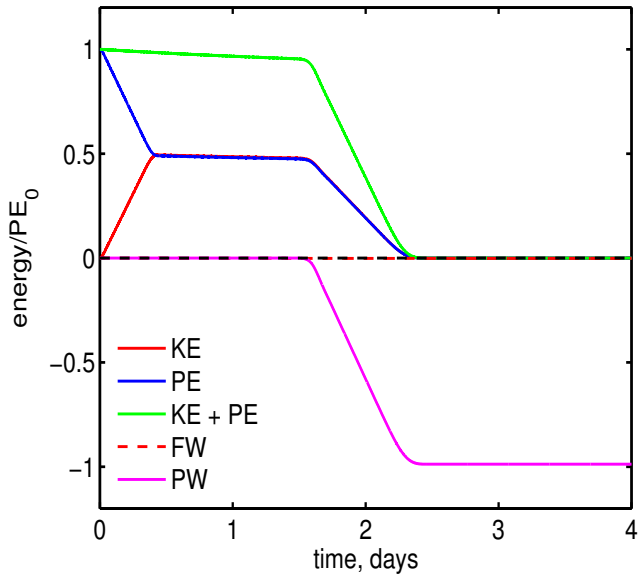

Figure 11: The energy balance for the gravitational adjustment problem (no rotation, Sec. 3.1) evaluated over a control volume $-500 \le x \le 500$ km. Energy is nondimensionalized with the initial potential energy, PE_0, and time is in days. The first four curves shown here have the same meaning as the curves with the same name (and color) in the parcel on a slope problem of Part 1, Sec. 5. Frictional work (FW) is vanishingly small since the frictional parameter r was effectively zero in this experiment.

speed should increase roughly linearly with η_o. Hence, C alone is evidently not an appropriate scale for the fluid speed. As a guess let's try instead $C\eta_o/H$, i.e., plot $u/(C\eta_o/H)$ (Fig. 10). Note that the current amplitude is about 0.5 in these nondimensional units. If this is indeed an appropriate scaling, then the nondimensional (or normalized) fluid speed should remain about 0.5 when η_o is varied within a small amplitude regime. At a large enough amplitude, the nondimensional speed will depart from this linear scale, which can be interpreted as a finite amplitude effect.

3.4 Energy and potential vorticity balances

The energy balance (Sec. 2.3.2) was evaluated over a domain $-500 \le x \le 500$ km (Fig. 11). Several features of this energy balance are in common with the parcel on a slope (Part 1, Sec. 5.3). First, the only source of energy is the potential energy, PE, stored in the initial ridge, as is true in all of the geostrophic adjustment problems that follow. Second, after the ridge is released and begins to slump, the decrease of PE is accompanied by a closely comparable increase of kinetic energy, KE. Thus the total energy, KE + PE, is approximately conserved in this ideal shallow water model, aside from small losses (a few percent per week) due to numerical viscosity and to a small but numerically necessary horizontal diffusion (discussed below).

Other aspects of this energy balance are quite different from that of the single parcel experiment. Notice that once the outward-going wave pulses are fully separated, $t \ge 2L/C = 0.8$ days, the kinetic energy and potential energy are thereafter equal. This kind of energy equipartition is a characteristic of

(nonrotating) linear gravity wave motion and so provides a useful check on the numerical solution in the much more common case that an exact analytic solution is not available. A second difference is the domain over which the energy balance was diagnosed. There was no choice about where to evaluate the energy balance for a single parcel; of course the balance was computed for the moving parcel. In fluid dynamics parlance that would likely be called a Lagrangian (material-following) balance. Given the present fluid model, there are a number of other ways that an energy balance could be evaluated. The balance could follow a wave pulse, or it could follow any specific fluid column. Here, the energy balance is evaluated over a fixed control volume, which is oftentimes called an Eulerian balance. As one consequence, energy could leave the control volume through the sides beginning at $t \approx 1.6$ days when the wave pulses reach the edges of the control volume. This energy flux is accounted by the pressure work term of Eqn (25).

Potential vorticity is interesting in quite a different way — there isn't any potential vorticity anomaly in this case! There is no planetary vorticity since $f = 0$, there is no relative vorticity in the initial condition since the ridge begins at rest, and there is no vorticity produced by the subsequent gravity wave processes because the pressure gradient in a shallow water model has zero curl.[14] Potential vorticity analysis is very often invaluable for the analysis of low frequency phenomenon (next sections), but it is not relevant to shallow water gravity waves.

3.5 Finer details of the solution; finite amplitude effects

On first seeing this solution the question came up — in what ways and to what degree can we trust this numerical solution? If we were attempting to make a realistic simulation of an observed, physical phenomenon (we aren't) then we would have to consider the assumptions made in formulating the shallow water model. But here the issue is much narrower .... how can we test that the plots in front of us represent an acceptably accurate solution to the problem posed? Comparison of the numerical solution with the exact D'Alembert solution suggests that wave propagation aspects are simulated reasonably well in the sense that the pulses move at nearly the phase speed C, and without significant dispersion, as they should in a small amplitude limit. Energy is also balanced fairly closely, though there is an unaccounted loss of a few percent in the first week. This implies that the amplitude generally must be slightly low as well.

A close look reveals two other, small but systematic differences between the numerical solution and the exact, linear d'Alembert solution insofar as the numerical wave pulses do not retain the exact

[14] In a shallow water model the fluid density is presumed constant in space and in time. If instead the density varies in space, then the pressure gradient divided by the variable density may have a curl, often called a solenoidal term. This term will produce vorticity having a component normal to the plane of the pressure and density gradients. This baroclinic process can be important in a more general three-dimensional fluid, but can not arise here.

shape of the initial ridge. First, the numerical $\eta(x,t)$ shows a small scale spatial oscillation (wavelength several times the numerical grid interval) that is characteristic of a finite-differencing error.[15] A second and more interesting difference is that the leading edge of a numerical wave pulse is steepened slightly, while the trailing edge is elongated or rarefied. This appears to be a genuine, finite amplitude effect due to nonlinear processes — horizontal advection and large variations in layer thickness — that are present in the nonlinear numerical model and solution, but not in the linear D'Alembert solution.

It is very useful to be able to make an *a priori* (before the event) estimate of the nonlinear terms and so to form a hypothesis regarding how finite amplitude effects may vary with the amplitude, η_o. For a longitudinal wave in which the wave and fluid velocity are colinear,
$u \propto u_o \cos(kx-\omega t)$, $h \propto h_o \cos(kx-\omega t+\phi)$, the ratio of the nonlinear thickness advection term to the linear local time rate of change may be estimated as

$$\frac{\text{nonlinear}}{\text{linear}} = \frac{u\partial h/\partial x}{\partial h/\partial t} = \frac{u_o k}{\omega} = \frac{u_o}{C}, \tag{56}$$

the ratio of the fluid speed to the wave phase speed. This ratio arises very widely and is often called the Froude number,

$$\boxed{F = \frac{u_o}{C}} \tag{57}$$

Given the (tentative) scale estimate of the current amplitude, $u_o \approx C\eta_o/H$ (Sec. 3.1), the Froude number can be written in external parameters and easily evaluated (recall that for the base case, $\eta_0 = 50$ m and $H = 500$ m) as

$$F = \frac{\eta_0}{H} = 0.1.$$

The other nonlinear term in the thickness equation may be readily factored into nonlinear and linear terms, $h(\partial u/\partial x) = (\eta + H)(\partial u/\partial x)$. Under the assumption that the variable η will be no larger than the initial value η_o, then again

$$\frac{\text{nonlinear}}{\text{linear}} = \frac{\eta(\partial u/\partial x)}{H(\partial u/\partial x)} = \frac{\eta}{H} \approx \frac{\eta_o}{H}.$$

The important ratio, η_o/H, called simply the 'amplitude', happens to be equal to the Froude number in this case, but not generally. These simple estimates indicate that nonlinear terms are small but not

[15] This error can be made larger or smaller depending upon the smoothness of the initial condition. If allowed to grow unchecked, this grid scale noise can readily swamp the 'physical' content of a numerical solution, i.e., the part that is faithful to the model equations. A simple cure is to include the smallest possible horizontal diffusion of thickness and momentum that will serve to damp high frequency variability. That has been done in these numerical solutions, though mention of it was left out of the previous discussion. The issue then becomes whether this *ad hoc* diffusion impacts significantly the aspects of the solution that are of interest. In this experiment, a thickness diffusion $D = 5$ m^2 sec^{-1} was sufficient to prevent runaway growth of the grid scale noise, though not enough to crush it completely. Over the short duration of this experiment, $T = 5$ days, this diffusion causes a small but acceptable spreading of the ridge judging from $\sqrt{DT^2} \approx 1.5$ km $\ll L$.

negligible compared to comparable linear terms in the base case in which $\eta_o/H \approx 0.1$. It is then plausible, but not certain from this analysis, that there should occur modest but detectable finite amplitude effects in the corresponding numerical solution.

A complementary estimate of finite amplitude effects comes from considering that the gravity wave speed likely depends upon the actual fluid thickness, $h = H + \eta$ and not just the nominal thickness as in the linear wave speed, Eqn. (54). A wave pulse, which has one sign, might then propagate at a speed given by the average thickness taking account of the pulse amplitude,

$$C_{pulse} = \sqrt{g'(H + \eta_o/4)} \approx C(1 + \eta_o/8H), \tag{58}$$

in the specific case considered here. Taking this thickness-dependent view of wave speed one step further, the observed steepening appears to be consistent with the idea that the thickest part of a wave pulse should propagate faster than in the undisturbed fluid, which should in consequence produce a steepened wave pulse front.

3.6 Problems

(1) Some important things went by rather quickly in Sec. 3.2. You should verify that 1) any function whose argument is $(kx \pm \omega t)$ satisfies the elementary wave equation provided that $\omega/k = \sqrt{g'H}$, and 2) the initial data is satisfied by Eqn. (52) for any $\eta_o(x)$. However, if $\eta_o \ll H$ does not hold, then the assumption that the system is linear would not be appropriate.

(2) The analysis of Sec. 3.2 lead to the (plausible) hypothesis that the wave pulses found in the numerical solution have the properties and parameter dependence of elementary gravity waves (if not their expected appearance). How can you test this, given the opportunity to specify the parameters of new numerical experiments via geoadj_1d.m (Sec. 6.3)? For now, exclude finite amplitude effects by keeping the amplitude small, $\eta_o/H \leq 0.1$, say, and omit friction, which will be discussed below.

(3) Friction was omitted from the discussion of energy balance in Sec. 3.4, but it is included in the numerical model via a linear damping of the velocity, $-r\mathbf{V}$, sometimes called Rayleigh damping. In the solution shown here r was set to a small enough value that frictional effects were negligible. Some friction questions: 1) What value of r is required to damp an outgoing wave pulse to half initial amplitude as it arrives at $x = 500$ km? 2) How does the required r vary with the nominal layer thickness, H? 3) Can you define a nondimensional number analogous to the Ekman number, $E = r/f$ (Part 1, Sec. 5) that serves to predict the amplitude of this damping effect?

(4) You can investigate finite amplitude effects systematically by experimenting with a much smaller and a much larger value for the initial thickness anomaly, say $\eta_o = 2$ and then $\eta_o = 200$ m (in the setup of the numerical model geoadj_1d.m). Some questions: 1) It is expected that finite amplitude effects should increase with η_o. But what if the nominal layer thickness, H, is increased apace with η_o so that the ratio, η_o/H remains constant? 2) Can you verify that wave front steepening (and rarefaction) is

large or small depending upon the Froude number, and conversely is independent of, e.g., the density difference, $\delta\rho$, when all else is held constant? 3) Does the scale for current speed, $C\eta_o/h$, account for the change in the actual (dimensional) current amplitude at small values of η_o/H? How about at very large values of η_o/H? 4) Test the estimated finite amplitude wave pulse speed Eqn. (58) using numerical solutions. 5) Instead of a ridge, suppose the initial interface displacement is a trough, $\eta_0 < 0$. Use the previous result 4) to explain the steepening and rarefaction seen in this case.

4 Geostrophic adjustment on an f-plane

The one new feature of the next experiment is that Earth's rotation is included, though in a simplified form, an f-plane, in which the Coriolis parameter is presumed to be constant. The latitude is taken to be 30^o N, and so

$$= 2\Omega \sin(\text{latitude}) \;=\; \Omega \;=\; 7.292 \times 10^{-5}\ \text{sec}^{-1}.$$

The inertial period is $2\pi/f \approx 1$ day (1 day less about 4 minutes). The one-dimensional domain and initial condition, including the initial state of rest, are unchanged from before.

We can compare the new experiment, Fig. (12), with the non-rotating experiment of the previous section, Fig. (10), and it is clear that the effects of rotation are very, very significant. We will discuss the details below, once we have a few more analysis tools in place. For now, note that there are still gravity waves that propagate away from the ridge, however their amplitude is much less than in the nonrotating case, and some wavelike motion persists for at least a week. The ridge subsides and spreads laterally in the first few days, but most of it remains in place, about 75% based upon a potential energy balance (Fig. 15). There are counter-flowing currents whose direction is such that the Coriolis force tends to oppose the pressure gradient associated with the sloping layer thickness, i.e., $0 \approx fv - g'\partial\eta/\partial x$, and after only a few days, the remaining ridge is in approximate geostrophic balance. In this f-plane experiment, an exact geostrophic balance is possible in principle (though not, in general, in a numerical solution).

4.1 Dispersion relation for waves on an f-plane

The waves in this rotating experiment are clearly very different from the pure gravity waves of the nonrotating case. The dispersion relationship that links the wave frequency ω and the wave number k makes a concise and very useful characterization of their properties. To find the dispersion relation, postulate a plane wave with wave vector, $\mathbf{K} = k_x\mathbf{e_x} + k_y\mathbf{e_y}$, which has a magnitude $K = \sqrt{k_x^2 + k_y^2} = 2\pi/\lambda$ with λ the wave length. A propagating plane wave having frequency ω is then $u(x,y,t) = U exp(i(k_x x + k_y y - \omega t))$ and similarly for v and η, where U, V, Γ are the constant but

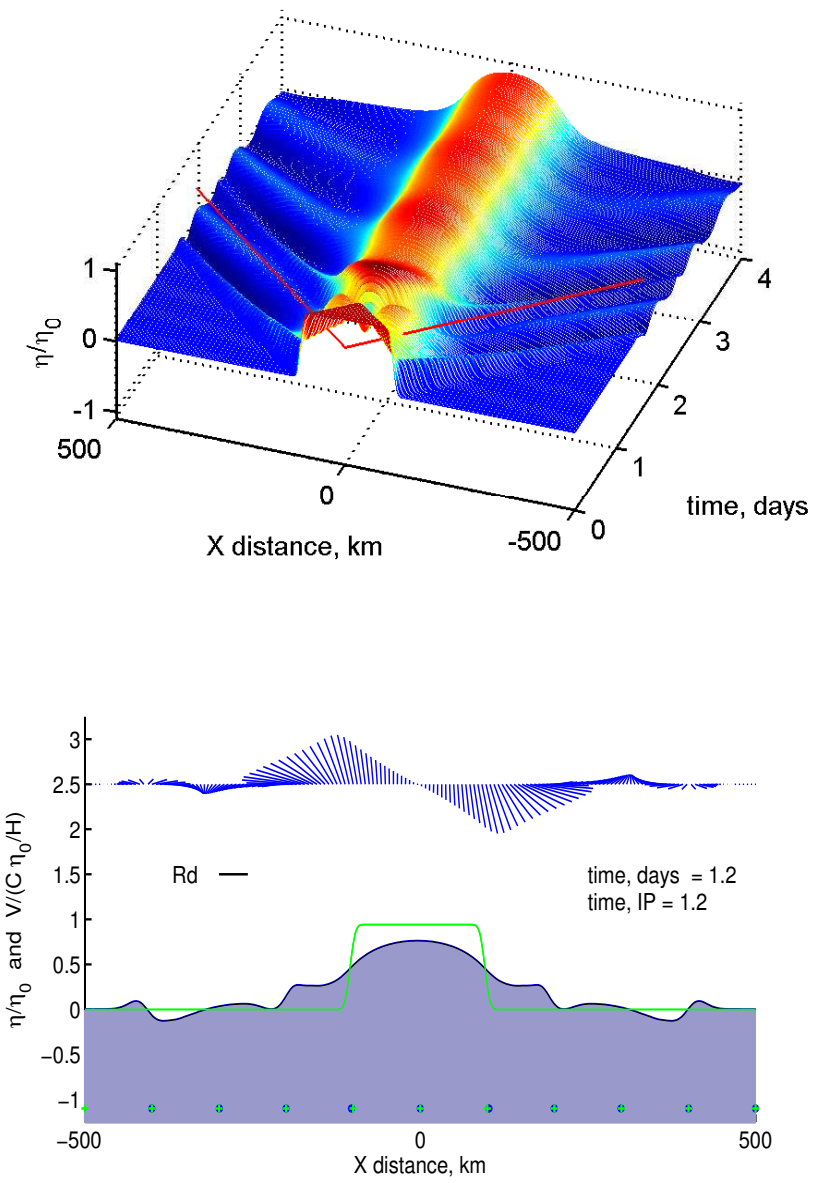

Figure 12: An experiment in geostrophic adjustment in one-dimension on an f-plane. **(upper)** The layer thickness anomaly, $\eta(x,t)/\eta_0$, of a rotating fluid layer following the release of a rectangular ridge having amplitude $\eta_0 = 50$ m. This solution was computed by the numerical model geoadj_1d.m Time and horizontal distance are shown in days and kilometers to make a direct comparison with the nonrotating experiment shown in Fig. (10). The thin red lines have a slope given by the gravity wave speed, $C = \sqrt{g'H} = 275$ km days^{-1} in this case. **(lower)** A snapshot of the solution at $t = 1.2$ days. The green line shows the initial displacement η_o. The current, which is shown by the array of vectors plotted above, is scaled with $C\eta_o/H = 0.30$ m s^{-1} for the parameters of this experiment. The blue dots and the green crosses at depth = -1.1 show the present and initial positions of floats. The floats do not move far in this experiment, and finite amplitude effects are small but detectable. Note the counter-flowing jets along the flanks of the partially subsided ridge. When the Coriolis force acting on these jets is sufficient to prevent the ridge from slumping further, the ridge is in a more or less steady, geostrophic balance.

unknown amplitudes of the velocity components and η. Substitution into the linear shallow water equations (30) and rewriting the resulting algebraic equations in a matrix format yields

$$\begin{bmatrix} -i\omega & -f & ig'k_x \\ f & -i\omega & ig'k_y \\ iHk_x & iHk_y & -i\omega \end{bmatrix} \begin{bmatrix} U \\ V \\ \Gamma \end{bmatrix} = 0. \tag{59}$$

This homogeneous system has nonzero solutions for $[U\ V\ \Gamma]$ only if it has a vanishing determinant,

$$\omega^3 - \omega(g'HK^2 + f^2) = 0, \tag{60}$$

which is to say, only if ω and K are related by a dispersion relation. The roots of (60) correspond to

$$\text{steady, geostrophic motion}: \quad \omega = 0, \text{ and} \tag{61}$$

$$\text{gravity wave motion modified by rotation}: \quad \omega = \pm\sqrt{g'HK^2 + f^2}. \tag{62}$$

Along with the dispersion relation, it is also helpful to find the polarization relation that relates the thickness and velocity components. Again we will presume a plane wave, but here there is no loss in assuming that propagation is in the x-direction only, i.e., $k_y = 0$. The thickness anomaly of a propagating plane wave can then be written $\eta(x,t) = \Gamma\cos(k_x x - \omega t)$. To find the corresponding $u(x,t)$ we can eliminate v from Eqn. (30),

$$\frac{\partial^2 u}{\partial t^2} = -f^2 u - f\frac{\partial \eta}{\partial y} - \frac{\partial^2 \eta}{\partial x \partial t}. \tag{63}$$

Substitution of the presumed $\eta(x,t)$ into (63) then gives,

$$u(x,t) = \frac{k_x \omega}{f^2 - \omega^2}\Gamma\cos(k_x x - \omega t). \tag{64}$$

Similar treatment for the other velocity component yields

$$v(x,t) = \frac{k_x f}{f^2 - \omega^2}\Gamma\sin(k_x x - \omega t). \tag{65}$$

For the wave solution, $\omega > f$, the velocity vector rotates clockwise around an ellipse whose major axis is in the direction of the wavevector, here x, and whose major/minor axes is in the ratio

$$\frac{U}{V} = \frac{\omega}{f}.$$

Since $\omega \geq f$, the ellipse is more or less elongated or polarized in the wave vector direction depending upon frequency.

4.2 Intrinsic scales of the f-plane

The rotating, shallow water model has just three external parameters: the Coriolis parameter, f, that provides a natural time scale, the rotation time, $1/f$, noted in Part 1. The reduced gravity, $g' = g\delta\rho/\rho_o$, is an acceleration, and the layer thickness, H, a degenerate length scale since there is no variation over the layer. The initial condition adds two lengths, η_o and L. The Coriolis parameter f is an obvious (inverse) time scale for normalizing the frequency of the dispersion relation Eqn. (62), and thus a given frequency can be said to be high or low compared to f. By that rearrangement to nondimensional form the dispersion relation (62) then reads

$$\frac{\omega}{f} = \pm\sqrt{\tfrac{g'H}{f^2}K^2 + 1}, \tag{66}$$

$$= \pm\sqrt{R_d^2K^2 + 1}, \tag{67}$$

in which the wavenumber K is automatically scaled by the very important length scale,

$$\boxed{R_d = \frac{C}{f}}$$

called the radius of deformation. Just as f^{-1} is the intrinsic time scale of a rotating shallow water model, R_d is the intrinsic horizontal length scale of this model. It is often the case that the initial condition or external forcing bring an external length scale to the problem, in this case the ridge width L, and the radius of deformation is the appropriate length scale for comparison. Thus a given ridge width is large or small as L/R_d is greater than or less than 1.[16] After this very simple but significant rearrangement to nondimensional coordinates, the dispersion relation (Figs. 13 and 14) holds for all values of f, g' and H, i.e., for all shallow water, f-plane models.

The value of R_d depends upon the stratification through C and the latitude through f. Over the open, subtropical oceans, $C \approx 3$ m sec^{-1} and not highly variable; $C \approx 2$ m sec^{-1} at subpolar latitudes. The more significant geographic variation of R_d comes from the latitudinal variation of $f^{-1} \propto 1/\sin(latitude)$. Notice that the equator is a special and an especially important case that will be considered in Part 3. In the ocean thermocline at 30^{o}N, $R_d \approx 40$ km. The R_d of the atmosphere at this latitude is much larger, $R_d \approx 1000$ km, because the atmosphere is both thicker than the ocean thermocline, and more strongly stratified. This very large disparity in R_d is reflected directly in the dominant horizontal scales of the variability seen in the atmosphere and ocean (cf. Figs. 1 of Part 1 and Part 2), i.e., much larger horizontal scales are observed in the atmosphere.

[16] Four remarks. 1) A more compelling physical interpretation of R_d as a length scale will come just below. 2) The important ratio of frequencies, ω/f, is sometimes referred to as the temporal Rossby number, for which there seems to be no widely accepted symbol. 3) The ratio KR_d is the layered model equivalent of the square root of the Burger number, $B = N^2/f^2$,

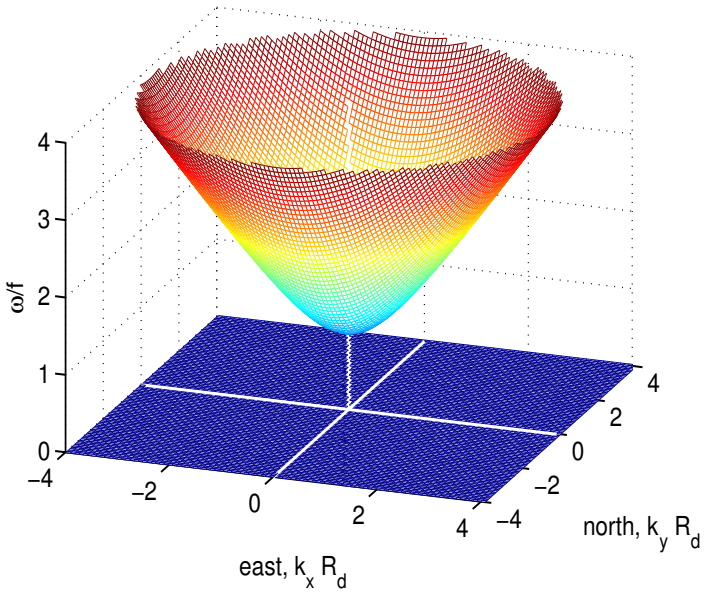

Figure 13: The dispersion relation for plane waves supported by the rotating, shallow water model. The frequency is normalized by the Coriolis parameter, f, and the wavenumber (an inverse length) by the inverse radius of deformation, R_d, discussed in the text. The colors are indicative of frequency only. A wavenumber vector $\mathbf{K}$ having an origin at $(0,0)$ and that terminates on either the bowl shaped surface or lies in the plane $\omega = 0$ is a solution, i.e., a possible free wave or geostrophic motion. This dispersion relation is symmetric about the ω axis (isotropic); the frequency of waves depends upon the magnitude of the wavenumber, $\mathbf{K}$, but not the direction of the wave vector.

4.3 Gravity, inertial and geostrophic motions

The non-zero ω/f root (the multi-colored parabola of Fig. 13 and the similar line of Fig. 14, upper) corresponds to gravity waves that are modified by rotation. There are three important limits in the dispersion relation that correspond with modes of the shallow water momentum equation.

High frequency, $\omega/f \gg 1$: In the short wave, high frequency limit, say $\omega/f \geq 5$, the dispersion relation asymptotes to pure gravity wave motion in which the phase and group speed are $C = \sqrt{g'H}$, as in the non-rotating, shallow water model (Sec. 3.1).[17] Pure gravity waves have the fastest group speed, $C_g \approx C$, and would be expected to arrive first in the far field, just as observed in the numerical solution. For these waves, the velocity is fairly strongly polarized in the direction of the wave propagation, as $U/V = \omega/f$, which may be called longitudinal.

Inertia-gravity waves: $1 < \omega/f < 5$ Intermediate scale waves fall somewhere between the limiting cases of pure gravity wave and pure inertial oscillation ($\omega/f = 1$, coming next), and are appropriately called inertia-gravity waves (or, rarely, gravity-inertia waves). Their phase speed depends upon K.

---

where $N^2 = (g/\rho)(\partial\rho/\partial z)$ is the static stability, times the aspect ratio H/λ. 4) There is no guarantee that the value 1 will always mark the boundary between large and small values of a given nondimensional number as factors of π or even $2\pi^2$ are sometimes relevant.

[17] If the wavelength is less than the layer thickness, H, then the wave motion will become a short gravity wave that is surface-trapped. These are the high frequency, wind-generated waves that make up most of the sea state and they are outside the domain of the shallow water model. Their frequency can be high enough that the pressure is nonhydrostatic.

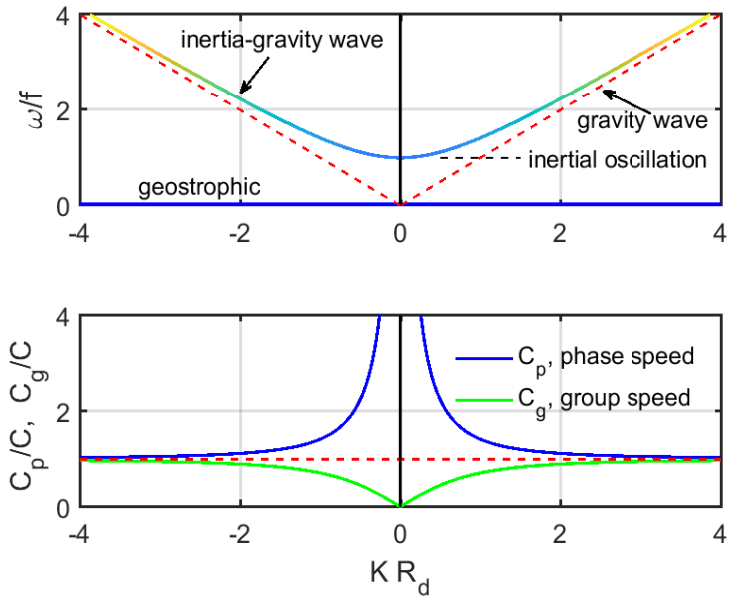

Figure 14: **(upper)** The dispersion relation for plane waves supported by the rotating, shallow water model. This is a slice through the isotropic surfaces of the previous figure. The blue line at zero frequency is the steady, geostrophic solution and the multi-colored parabola is the dispersion relation for inertia-gravity waves. The straight red dashed lines are for pure gravity waves ($f = 0$); in these nondimensional coordinates their slope is ± 1 for right- and left-going waves.**(lower)** The phase and group speeds of the inertia-gravity waves (blue and green lines) normalized by the gravity wave speed, C. Inertia-gravity waves are dispersive in that their phase speed varies with K. For short waves, $KR_d \gg 1$, the group speed asymptotes to that of pure gravity waves (the dashed red line) and in that limit $C_g = C_p = C$. The group speed goes to zero for very long waves, $KR_d \ll 1$, that have frequencies near f, i.e., near-inertial oscillations. The phase speed becomes infinite as $KR_d \to 0$.

Hence the inertia-gravity waves of an $f-$plane are dispersive, so that the shape of $\eta_0(x)$ is not, in general, preserved as these waves propagate away from a source (cf. the nonrotating, pure gravity wave experiment of Sec. 3.1). Instead, the initial form $\eta_0(x)$ will become more or less spread out, or 'dispersed', over time as waves of different wavelengths propagate at different C_g.[18]

Lowest possible wave frequency, $\omega/f \downarrow 1$: In the long wave limit, i.e., as KR_d becomes very small, inertia-gravity waves asymptote to inertial oscillations and the frequency approaches f. In this limit the phase speed is very large, infinite at $\omega/f = 1$. The entire field of motion consists of clockwise-rotating velocity having uniform phase, and thus no divergence and no associated η or pressure anomaly. It is appropriate to call this an inertial oscillation, to distinguish from vortical inertial motion that we will see in Sec. 5. Exact inertial oscillation, like exact geostrophic motion, amounts to a kind of frozen state that is unable to evolve. Consistent with this, the group speed in the limit $\omega/f \to 0$ is zero. Exact

[18]The linear initial value problem can also be solved usefully via Fourier transform. An example is carried out by the script ftransform.m (Sec. 6.3), which allows the choice of a dispersion relation. Somehow one must specify what fraction of the initial thickness anomaly will be dispersed into waves: in the non-rotating gravity wave case, all of the initial thickness anomaly η_o goes into propagating waves, but in the rotating case, a more or less significant fraction of η_o remains in the geostrophically balanced end-state as we will discuss shortly.

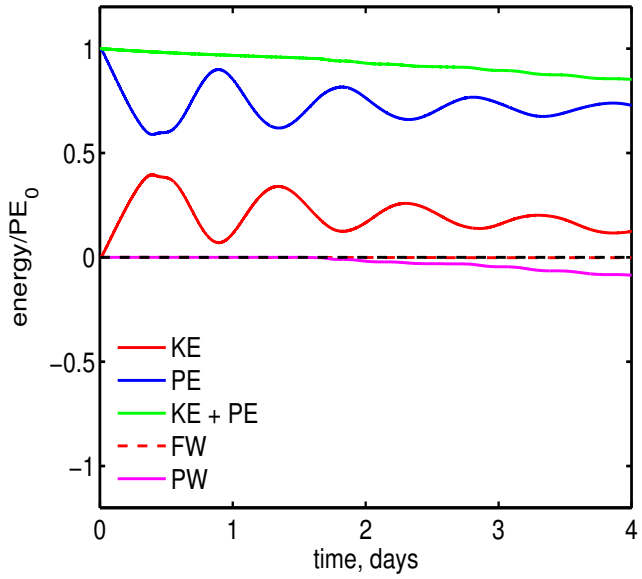

Figure 15: Energy balance of the f-plane geostrophic adjustment problem evaluated over $-500 \le x \le 500$ km. The energy and work terms are nondimensionalized with the initial potential energy, PE_0, and time is in days. At 30^o N, 1 day $\approx$ 1 inertial period. These data may be compared directly with that of the non-rotating experiment (Fig. 11) and are remarkably different. In this rotating experiment, very little energy was radiated out the sides of the domain, PW $\ll$ KE + PE, and was instead retained within the geostrophically balanced ridge as potential energy, mainly, plus significant kinetic energy. Notice that the decrease of total energy is slightly greater than the wave radiation through the sides, evidence of a small but likely growing numerical error.

inertial motion requires an infinite horizontal scale, and hence can not be realized in a finite ocean basin, or for any number of other reasons, e.g., the latitudinal variation of f and the existence of other ocean currents. While exact inertial oscillations are not expected in the ocean, rotary currents that turn in the direction of an inertial motion and that have frequencies within 5-10 percent of f — near-inertial oscillations — are very common, for example whenever there is a rapidly changing wind stress on the ocean surface (an example was noted in Section 5 of Part 1).

Notice that there is a low frequency range, $0 < \omega < f$, within which there are no free motions possible in this f-plane system.

Steady, geostrophic motion, $\omega = 0$: The root $\omega/f = 0$ corresponds to exactly steady, geostrophic motion (the blue plane of Fig. 14, upper). In this f-plane model, the corresponding wavenumber can have any orientation and any magnitude (this will not be true when the latitudinal variation of f is acknowledged in Part 3). At first this may seem a trivial solution, but it is instead the f-plane approximation of the slowly-varying and nearly geostrophic motion that makes up most of the atmosphere and ocean circulation. The polarization relation shows that the velocity associated with a plane wave thickness anomaly (not propagating) is solely in the y direction, parallel to constant thickness, and perpendicular (transverse) to the wave vector.

4.4 The steady state inferred from potential vorticity conservation

The dispersion relation is an invaluable guide to the properties of the waves that arise in a geostrophic adjustment experiment, but it can't tell us everything we might want to know. For example, the dispersion relation alone gives no hint to the amplitude of any process, i.e., it does not tell how much of the initial ridge will disperse into waves vs. survive into a geostrophic steady state. Notice too that a given wavenumber in Fig. (13) can be either a time-dependent wave, or, a steady geostrophic motion. What distinguishes these two very different kinds of motion? These important and fundamental questions lead to an analysis of potential vorticity conservation, described next.

A key result of this experiment is that most of the ridge survived the adjustment process, e.g., about 75% of the initial potential energy (Fig. 15). The conservation of potential vorticity provides real insight into why this was the case in this experiment, and why it will not be in others. In words, the $q-$conservation law Eqn. (29) states that q of a parcel (column) is unchanged by the process of geostrophic adjustment, or,

$$q(\alpha) = q_o(\alpha), \tag{68}$$

where α is a parcel tag. Assuming that q_o is known, then the task is to follow the parcels (all of the parcels) during the adjustment process. However, if the amplitude is small in the sense that parcel displacements are small compared to the scales over which the currents vary significantly, then a very helpful approximation of (68) is that

$$q(x) = q_0(x), \tag{69}$$

with x the usual (Eulerian) spatial coordinate. In other words, if the horizontal displacements during the adjustment process are small, then q will be conserved in place, approximately, and q conservation is linear. This may be checked by comparing the horizontal displacement of floats against the width of the adjusted portion of the ridge. In the base case (Fig. 12) there is a detectable, outward displacement of floats during the adjustment process, but the float displacement is small compared to the width of the adjusted portion of ridge, the radius of deformation, discussed below. As well, the geostrophic currents that form along the edge of the adjusted ridge are in a direction that is perpendicular to the thickness and velocity gradients and so cause no horizontal advection. The result is that q is indeed conserved in place, approximately, (Fig. 16, upper) as presumed in Eqn. (69). Note that the changes of v and η that occur during adjustment are significant so that linear q conservation is by no means trivial (Fig. 16).

The potential vorticity in this one-dimensional (x-dependent only) case is just

$$q = \frac{f + \partial v/\partial x}{H + \eta}.$$

The initial state included the ridge within which the layer thickness was $H + \eta_0$, and elsewhere the thickness was H. The ridge was presumed to be at rest, and so the initial relative vorticity vanishes. The

q distribution in the initial and in the final state is then, within the initial ridge,

$$\frac{f+\partial v/\partial x}{H+\eta} = \frac{f}{H+\eta_0} \text{ if } |x| \leq L, \tag{70}$$

and outside of the ridge

$$\frac{f+\partial v/\partial x}{H+\eta} = \frac{f}{H} \text{ if } |x| > L. \tag{71}$$

Note that both $v(x)$ and $\eta(x)$ are unknowns, but in addition, we have come to expect that the steady state should be in geostrophic balance, and hence we have also the geostrophic relationship between v and η, *viz.*

$$0 = fv_{geo} - g'\frac{\partial \eta_{geo}}{\partial x}.$$

When this v_{geo} is substituted into the q conservation law, say for the region $x \geq L$, Eqn. (71), there follows a second order, ordinary differential equation in $\eta_{geo}(x)$:

$$\frac{g'H}{f^2}\frac{d^2\eta_{geo}}{dx^2} - \eta_{geo} = 0.$$

The boundary conditions on η may be applied over four segments, $x \leq -L$, $-L < x \leq 0$, $0 < x \leq L$ and $x \geq L$. For example for the right-most segment, $x \geq L$, the boundary conditions are

$$\eta_{geo}(L) = \eta_0/2 \text{ and } \eta_{geo}(\infty) = 0.$$

The solution for this segment only is then

$$\eta_{geo}(x) = \frac{\eta_o}{2}\exp(-\frac{x-L}{R_d}). \tag{72}$$

By matching the solutions of the four segments, $\eta_{geo}(x)$ may be constructed across the entire ridge with a result that compares fairly well with the the numerical solution (Fig. 16, middle). Two notable features of this solution:

1) The horizontal length scale is $\mathbf{R_d}$: Compared with the initial ridge, the adjusted ridge is 'deformed' over the horizontal e-folding scale R_d. The width of the adjusted portion of the ridge is then about $3R_d$. This result suggests a clear physical interpretation of R_d. Recall from Part 1, Sec. 5.3 that $1/f$ is the time it takes for rotation to turn a parcel (started from rest) by one radian with respect to an impulsively-applied force. R_d is then the distance that a gravity wave will propagate in the time $1/f$, and hence R_d is proportional to the distance that the ridge will spread outward (as a gravity wave) before being arrested by the Coriolis force. Implicit in this is a definite meaning to the phrase 'large scale' as applied to a mass anomaly of half-width or radius L: the appropriate standard against which to compare (or measure) a horizontal length is R_d. The radius of deformation varies quite a lot with latitude, as noted earlier, and is very large (though not infinite) as f vanishes. The nonrotating

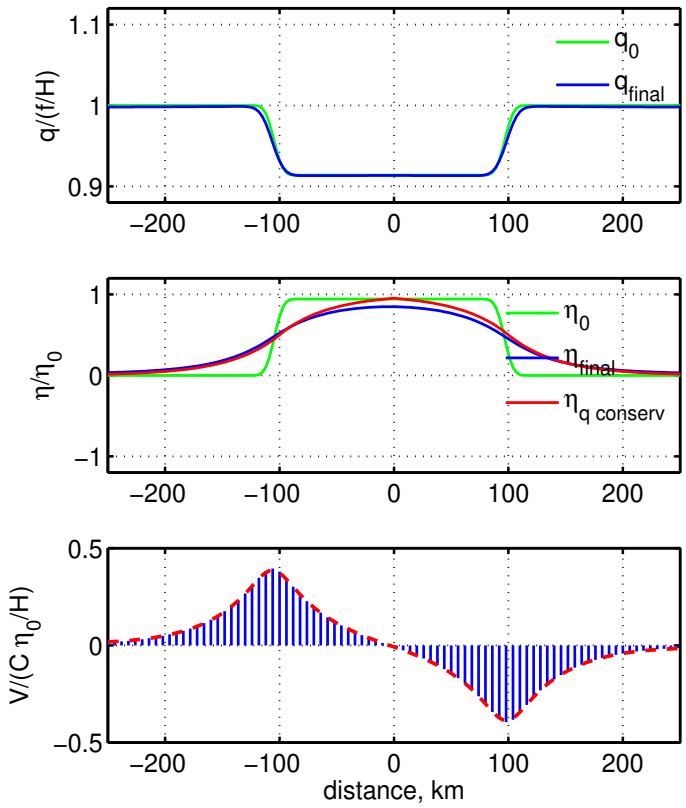

Figure 16: Elements of the potential vorticity balance for the $f-$plane geostrophic adjustment experiment. **(upper)** Initial (green) and final (blue) $q(x)$ (the latter is at $time = 6$ days). Note that q of the numerical solution is almost unchanged by the process of geostrophic adjustment, as should have been the case, despite that there are large changes in the thickness and the current. **(middle)** The initial and final thickness anomaly (green and blue lines). The red line, η_{geo}, is the steady state thickness anomaly computed from Eqn. (72) under the assumption of geostrophy and q conservation. **(lower)** Time-mean current (blue vectors). The velocity maxima, or jets, at $x = \pm$ 100 km = $\pm L$, coincide with the adjusted part of the initial ridge. These jets are in a near geostrophic balance as evidenced by a very close match with the geostrophic velocity, $(g'/f)\partial\eta_{final}/\partial x$ (dashed red line).

(gravitational) adjustment experiment of Sec. 3 can now be seen as the limit $R_d/L \to \infty$ in which rotational effects are expected to be negligible — literally zero in that case — compared to the spreading of an unbalanced thickness anomaly by gravity wave processes.

2) q-conservation and the geostrophic jets: The result Eqn. (72) gives an essentially complete account of the geostrophic jets. The jets have a relative vorticity $\partial v/\partial x$ (Fig. 16, lower) that follows from the effect of column thickness change and q conservation. For example, over the right-most segment, $x > L$, the layer thickness increased (the water column was stretched) from H to $H + \eta(x,t)$ during the adjustment process (Fig. 16, middle). The initial potential vorticity was just f/H since there was no initial velocity and no relative vorticity. To maintain constant q during the adjustment process, the relative vorticity $\xi = \partial v_{geo}/\partial x = f\eta/H > 0$ of the jet must thus be positive, or cyclonic, in this region. Stretching of the water column appears to produce positive relative vorticity by concentrating the planetary vorticity (Fig. 5, upper). To an Earth-bound, rotating observer, who sees stars turning overhead but not the rotation due to planetary velocity, it looks as if this relative vorticity has come out of nowhere. To an inertial observer for whom the same stars stand still, this looks like familiar angular momentum conservation, albeit of a deformable object.

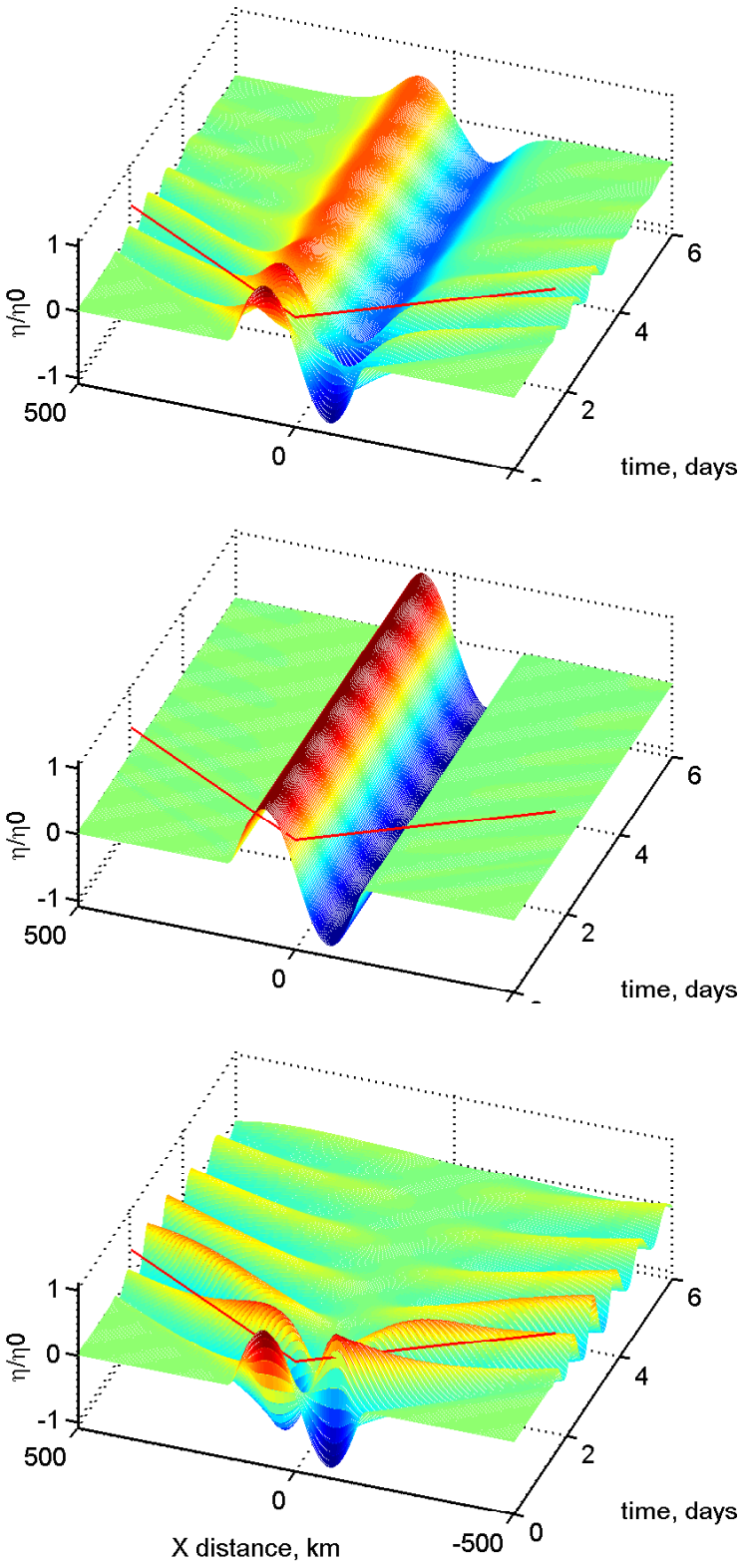

Figure 17: Three experiments in geostrophic adjustment in which the initial thickness anomaly was a single wave $\eta_0 sin(\pi x/L)$. The initial velocity was either **(top)** vanishing, as in the previous experiments, **(middle)** geostrophically balanced with the initial thickness anomaly, in which event nothing happens when the ridge is released, or **(bottom)** the velocity that gave vanishing q anomaly with respect to the outlying fluid. In this last case all of the initial ridge is dispersed by inertia-gravity waves; there is no geostrophic steady state. These experiments were carried out with geoadj_1d_uic.m, linked in Sec. 6.3.

4.5 If there is no potential vorticity anomaly

Note that if the initial q was spatially uniform (no q anomaly) the geostrophic, q-conserving solution (72) would be $\eta(x) = 0$, i.e., vanishing geostrophic steady state. In the experiment just described the initial ridge *was* a q anomaly because the initial velocity was assumed to vanish and hence the q of the initial ridge was $f/(\eta_o + H)$ compared with f/H outside the ridge. A vanishing initial velocity is the simplest initial state to define and a reasonable place to start experimenting, but it is not necessarily the most realistic or common initial condition. An experiment that allows a more general initial condition that includes velocity will make clear that insofar as the possible steady (adjusted) state is concerned, the initial ridge is defined by q its anomaly, generally, and not by η alone, as it may have seemed from this first rotating experiment.

For this more general case it is preferable to assume an initial $\eta_o(x)$ that has zero mean (though still called a ridge), e.g.,

$$\eta(x,t=0) = \eta_0 sin(\pi x/L) \text{ if } |x| < L, \text{ and } \eta(x,t=0) = 0 \text{ otherwise.}$$

Now consider one of three initial velocity fields that vanish outside of the thickness anomaly:

$$\text{vanishing velocity, } V(x,t=0) = 0, \text{ just as before,}$$

$$\text{geostrophic velocity, } V(x,t=0) = (g'/f)\partial\eta(x,t=0)/\partial x, \quad \text{or,}$$

$$\text{vanishing } q \text{ anomaly } V(x,t=0) = \int_{-L}^{x} f\eta(x,t=0)/H dx.$$

The first of these experiments (Fig. 17, upper) is very much like our previous experiment in that about two thirds of the initial ridge survives into a steady, geostrophic state. The second experiment is a non-event (Fig. 17, middle); nothing happens when a geostrophically-balanced ridge is released onto an f-plane (this would not be true if the latitudinal variation of f was retained, Part 3).

The third experiment is the important and perhaps most surprising one: a ridge having zero q anomaly is completely dispersed by inertia-gravity waves, with nothing surviving into a steady state (Fig. 17, lower).[19] This holds regardless of the width of the ridge. Evidently the width of a ridge is best defined from it's q anomaly, and not by the width of η alone, insofar as the possibility of a geostrophic end state is concerned.

The initial condition of this last experiment, vanishing initial q anomaly, may seem a special, contrived case, but it is a very, very common circumstance: tidal waves in the open ocean have large spatial scales, wavelengths of hundreds to thousands of kilometers, and yet tidal waves propagate as

[19] An animation of this third case is https://scienceweb.whoi.edu/PO/people/jprice/education/zeroq1d.mp4

almost free waves that show no sign of adjusting to the kind of geostrophic balance that characterizes mesoscale eddies and gyres. The reason is that tidal waves are generated by a small gravitational/centrifugal imbalance, the tidal force, that does not exert a torque on the ocean, and so do does not by itself generate a potential vorticity anomaly. Thus the difference between two identical wavenumbers of the dispersion relation, (Fig. 13), one being a steady geostrophic motion and the other being a wave, is evidently the association (or not) of a potential vorticity anomaly with the former but not the latter.

4.6 Problems

(1) Geostrophic balance is the final, steady state of the experiment in Fig. (12). How long would you say it takes for geostrophy to hold in this experiment? Recall from Part 1 that the time scale required for rotation to deflect a moving parcel significantly was $1/f \approx 1/4$ days at mid-latitudes. On the other hand, energetic near-inertial oscillations evidently persist for roughly five days. Are you thinking about an instantaneous geostrophic balance, or a time-mean geostrophic balance?

(2) Go back to the linear shallow water system Eqns. (30) and identify which of the terms are relevant in each of the three limits of the dispersion relation discussed in Sec. 4.3.

(3) Using the Matlab scripts geoadj_1d.m and geoadj_1d_uic.m to generate new solutions, 1) find experimentally how the width of a geostrophic jet varies with latitude and with H and g'. What do you expect if the latitude is from the southern hemisphere? Before trying these experiments in the numerical model, do a thought experiment and test your developing intuition against the numerical solution.

(4) Show that a plausible velocity scale for a geostrophic jet is $C\eta_o/H$. How/why does the jet speed vary with H? Show that a reasonable *a priori* estimate of parcel displacement is $C\eta_o/Hf$ and that the condition for linearity with respect to horizontal advection is then the familiar one, $\eta_o/H \ll 1$. Suppose that the initial thickness anomaly is made very large, say $\eta_o = 200$ m. How does this alter the potential vorticity balance during the adjustment process?

(5) The power of the potential vorticity analysis becomes evident when one considers different kinds of initial conditions, say $\eta_o = 0$, but with an unbalanced current, $u = u_o$ if $\mid x \mid \leq L$, and otherwise $u = 0$ (recall that u is the velocity component in the x-direction). Then assume the same profile but with the initial velocity directed in the y direction. Which of these initial states has vorticity, and thus a potential vorticity anomaly that should remain in the adjusted state? Solutions for these initial velocity distributions can be generated by geoadj_1d_uic.m.

(6) The solution shown here in Sec. 4 was made with a vanishingly small value of the linear damping coefficient. What happens to the wave response and to the geostrophic currents when damping is made significant? Connect the results you obtain here with the Ekman number developed in Part 1, Sec. 5.

5 In two horizontal dimensions

In this section we will consider briefly a two dimensional f-plane geostrophic adjustment experiment that introduces one new effect, flow curvature, and an important new nondimensional number, the Rossby number, that measures its magnitude. Aside from being two-dimensional, the configuration is unchanged from the basic case of Sec. 4: the initial thickness anomaly is taken to be a right cylinder having a radius $L = 100$ km that is typical of mesoscale eddies, and an amplitude 100 m. The stratification and central latitude are as before, $H = 500$ m, $g' = 2 \times 10^{-2}$ m sec^{-1} and hence $C = \sqrt{g'H} \approx 3.2$ m sec^{-1} (about 300 km day^{-1}), and $f = 7.29 \times 10^{-5}$ sec^{-1}.[20]

5.1 Adjustment is not greatly altered by the additional dimension

The slumping of the thickness anomaly produces inertia-gravity waves that propagate isotropically away from the center (cover graphic). The dispersion properties of these waves are familiar from the one-dimensional f-plane case. The wave front expands at a rate very close to the gravity wave speed, about 300 km day^{-1} (Fig. 1). The amplitude at the wave front is a small fraction of η_0, and decreases with time due to geometric spreading $\propto 1/r$ (not present in one dimension). The frequency at the wave front is roughly $2f$, and the current is at first polarized in the radial direction. With increasing time, the frequency decreases toward f, and the current vector rotates as a near-inertial motion. Near the initial eddy, weak near-inertial motion persists for several weeks.

Most of this eddy remained in place after the geostrophic adjustment process was essentially complete, several weeks after start. The resulting current was then nearly steady and anti-cyclonic (clockwise) around the eddy center, a high (Fig. 18). The initial condition was azimuthally symmetric, and nothing about the $f-$plane dynamics would change that (not the case when the β-effect is acknowledged, Part 3). Potential vorticity and energy balances are as expected from the one-dimensional cases of Sec. 4; e.g., potential vorticity stays with the fluid as prescribed in the initial condition.

5.2 Curvature and Rossby number

Once the motion reaches an azimuthally symmetric steady state we can diagnose some of its properties using the single parcel model of Part 1 (though, of course, we could not use the single parcel model to

[20]The numerical model used to solve these two dimensional problems is written in Fortran, geoadj_2d.for, linked in Sec. 6.3.

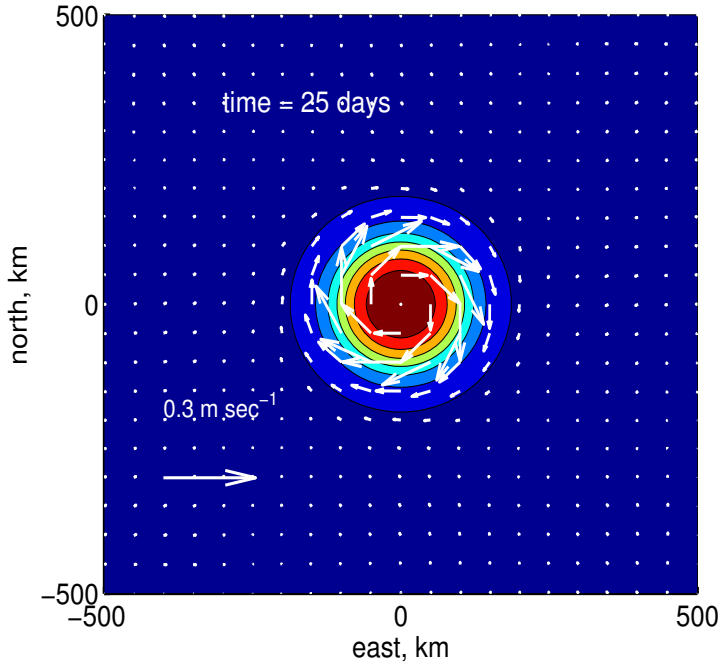

Figure 18: Velocity and interface displacement (color contours) from a two-dimensional f-plane experiment. This is a snapshot at 25 days. This eddy is a high in thickness (about 70 m at the center) and pressure and so the balanced flow is anti-cyclonic (clockwise). The big vector at lower left has a magnitude $0.5C\eta_0/H = 0.3$ m sec^{-1} and serves as a speed scale.

calculate the adjustment process in the first place). The radial component equation of motion for a rotating, polar system is (Part 1, Sec. 3.1, Eqn. 40)

$$\frac{d^2r}{dt^2} - r\omega^2 \;=\; r\Omega^2 + 2\Omega\omega r \;+\; F_r/M,$$

where we have dropped the primes used in Part 1 to indicate a rotating reference frame. The upper case Ω is the rotation rate of the reference frame, and the lower case ω is that of the parcel. Given that the flow is in steady state, the radial acceleration vanishes. The centrifugal force due to the reference frame rotation (first term on the right) also vanishes (as described in Part 1, Sec. 3). The azimuthal current is $U_\lambda \;=\; \omega r$, $f = 2\Omega$, and the external force F is the radial pressure gradient, $\propto \partial\eta/\partial r$. With these identifications, the equation of motion is

$$-\frac{U_\lambda^2}{r} \;=\; fU_\lambda \;-\; g'\frac{\partial\eta}{\partial r},$$

in which the centripetal acceleration equals the sum of the Coriolis force and the pressure gradient. It is easier to envision the steady force balance that would follow from adopting a co-rotating frame in which the parcel is not accelerating. In that case, the equation of motion is

$$0 \;=\; \frac{U_\lambda^2}{r} \;+\; fU_\lambda \;-\; g'\frac{\partial\eta}{\partial r}. \tag{73}$$

The Coriolis force will be positive (outward) if the current is cyclonic, or negative (inward) if the flow is anticyclonic. The pressure gradient force, $\propto -\partial\eta/\partial r$, may also have either sign, positive if a high, and

negative if a low. If the balance was just between these two terms, then the velocity would be the geostrophic velocity,

$$U_\lambda \;=\; U_{geo} \;=\; \frac{g'}{f}\frac{\partial \eta}{\partial r}.$$

The interesting term is the centrifugal force which is always positive (always outwards), with important consequences. The magnitude of the centrifugal force may be assessed by comparison to the Coriolis force,

$$\frac{centrifugal}{Coriolis} = \frac{U_\lambda^2/r}{fU_\lambda} = \frac{U_\lambda}{fr},$$

often written

$$\boxed{R_o = \frac{U_\lambda}{fr}} \tag{74}$$

an important nondimensional number called the Rossby number.[21] To estimate R_o for this eddy we could use the initial radius, $r = 100$ km, where the current is strongest (Fig. 18), and an estimate of the current, $U_\lambda \approx -0.2$ m s^{-1}, negative since clockwise or anticyclonic: $R_o \approx -0.03$. The centrifugal force is thus very small compared to the Coriolis force. In some applications the sign of R_o has no significance, but here it does if we use R_o to rewrite Eqn. (73),

$$0 = \; fU_\lambda(1+R_o) - g\frac{\partial \eta}{\partial r},$$

or,

$$U_\lambda \;=\; \frac{1}{1+R_o}U_{geo}.$$

In the case of a high pressure, anti-cyclone, Fig. (18), the centrifugal force adds constructively to the pressure gradient force, and in steady state, this combined centrifugal plus pressure gradient force must be balanced by a centripetal Coriolis force. The magnitude of the azimuthal velocity in a steady, high pressure anticyclone exceeds the geostrophic speed by a factor $1/(1+R_o) \approx 1.03$. Such a small effect of curvature is not apparent in Fig. (18) but evident on closer inspection, Fig. (19). A steady flow in which geostrophy is modified by curvature in this way is said to be in gradient balance or sometimes just 'balanced', with geostrophy modified by curvature being understood.

In the case of our nominal oceanic mesoscale eddy the effects of curvature are not very important or interesting. What is interesting is the insight that the gradient balance (73) can provide regarding

[21] This is a specific form of Rossby number. A more general form could be written $R_o = U/fL$, where U is a characteristic velocity (whose sign is usually not relevant), and L is the spatial scale over which the velocity varies significantly. For example, the ratio of momentum advection to the Coriolis force in the shallow water model is $U(\partial U/\partial x)/fU \approx (U^2/L)/fU = U/fL = R_o$, where equality here means same order of magnitude. Small R_o thus implies a small magnitude of momentum advection compared to the Coriolis force. An assertion of small R_o is often made at the start of a linear approximation.

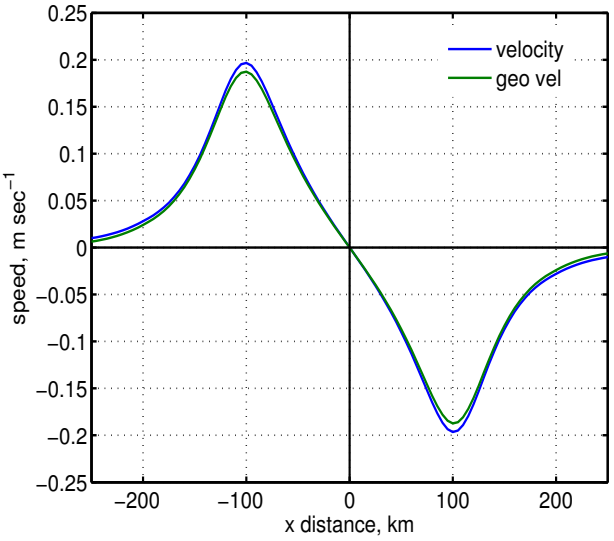

Figure 19: A slice across the adjusted eddy of Fig. (18). The geostrophic north-south velocity computed from the thickness field is the green line and the actual velocity is the blue line. The actual velocity of this high, anticyclone is only very slightly larger than the geostrophic velocity, just as expected from the small Rossby number, $R_o \approx -0.03$.

eddies generally. To that end we can treat Eqn. (73) as a quadratic equation for the unknown U_λ and with parameters f, r and $-g'\frac{\partial \eta}{\partial r}$;

$$U_\lambda = \frac{-f \pm \sqrt{f^2 + \frac{4g'}{r}\frac{\partial \eta}{\partial r}}}{2/r}. \tag{75}$$

This is a little difficult to interpret and can benefit from streamlining by dimensional analysis. There are four variables or parameters having two fundamental dimensions, length and time. Hence it should be possible to rewrite (75) in terms of one nondimensional dependent variable, and one nondimensional parameter. One possibility is

$$R_o = \frac{U_\lambda}{fr} = -\frac{1}{2} \pm \sqrt{\frac{1}{4} - \Pi}, \tag{76}$$

plotted as the blue and red lines in Fig. (20). The dependent variable is the Rossby number (74), and the independent variable is, from the second term in the radical of Eqn. (75),

$$\Pi = -\frac{g'}{rf^2}\frac{\partial \eta}{\partial r},$$

the nondimensional pressure gradient. The normalizing factor is rf^2, the centripetal acceleration of an inertial motion. The signs of R_o and Π are significant: $R_o > 0$ indicates a cyclonic (anticlockwise) flow, while $\Pi > 0$ indicates high pressure, which, if geostrophic, would be anticyclonic.

There are several noteworthy features of (76):

1) The upper left quadrant ($\Pi > 0$ and $R_o \geq 0$) of Fig. (20) is empty. There is no steady, gradient balance for a high pressure cyclone, since all three of the forces in such an eddy would be outward.

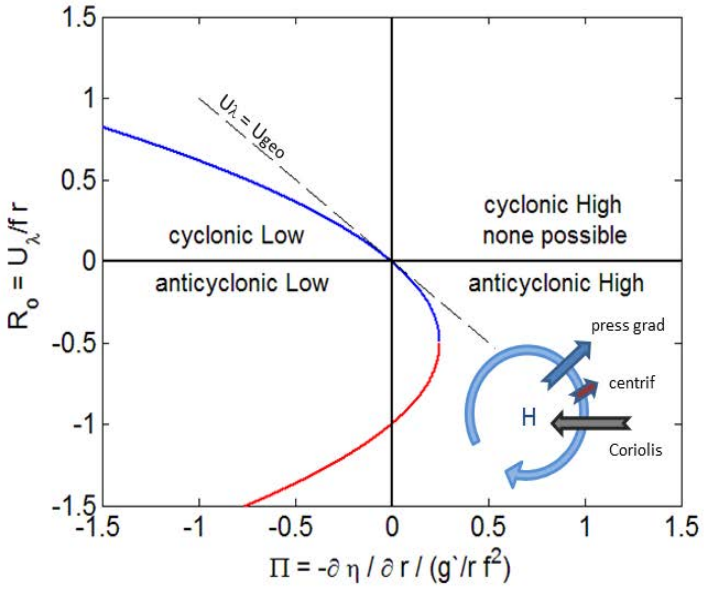

Figure 20: The blue and red line is the locus of possible solutions for gradient balanced flow over a large range of the nondimensional pressure gradient, Π. The blue line goes through the pure geostrophic case, (Π, R_o) = (0, 0), and includes the most commonly observed eddies. Red denotes high anticyclonic solutions that are uncommon and often called anomalous, but not impossible insofar as gradient balance alone is concerned. The schematic force balance for a high anticyclone is in the appropriate quadrant. The high anticyclone of Fig. (18) has a very small Rossby number, about -0.03, so that the centrifugal force is much smaller than the Coriolis force and hence it falls on the blue curve, close to the origin and close to the line $U_\lambda = U_{geo}$.

2) The large negative Π solutions, say $\Pi < 1$, are low pressure eddies for which the flow can be of either sign. For these intense eddies, the Coriolis force is of little importance, and the balance is mainly between the pressure gradient (inward) and the centrifugal force (outward), a so-called cyclostrophic balance. Tornadoes are the star examples from nature, but cyclostrophic eddies or vortices abound in everyday flows for which the spatial and time scales are small enough that the Coriolis force is unimportant; the vortices that spill off the edges of a paddle or the vortex that forms in the convergent flow above a drain are cyclostrophic eddies of this sort.

The cyclonic and anticyclonic branches differ importantly in that the cyclonic branch goes through the geostrophic state, (0, 0), while the anticyclonic branch does not. This has implications for the formation of such eddies. A cyclonic cyclostrophic eddy could result from the intensification (by vortex

stretching) of a larger scale, quasi-geostrophic flow whose absolute vorticity would be mainly planetary. Tornadoes are most often cyclonic. On the other hand, an anticyclonic eddy is more likely to be formed straightaway by intensification of whatever sign of relative vorticity might happen to be present at the site of a convergence (vortex stretching) event. Dustdevils, small scale versions of tornadoes formed by intense heating of the ground and subsequent convection, are about equally anticyclonic and cyclonic.

3) While the allowable negative (low) Π has no upper limit, the maximum possible (high) Π is 1/4 where $R_o = -1/2$. At that point the inward Coriolis force is balanced by the sum of outward pressure gradient and centrifugal forces, which are equal. Still larger positive Π implies larger centrifugal force which precludes a balance. This upper limit on Π has a real significance that can be seen in most synoptic weather maps: high pressure centers are invariably regions with very low pressure gradient and correspondingly low and often almost vanishing winds, i.e., high pressure centers give fair weather.

4) The anticyclonic high regime has two possible solutions for $0 < \Pi < -1/4$, a small magnitude R_o and a larger R_o solution ($R_o < 0$ in both cases). The former is predominantly geostrophic, with more or less minor effect of curvature. This is by far the most commonly encountered anticyclone, and our nominal oceanic eddy is of this sort. The larger magnitude R_o solution has stronger flow and the balance includes a significant contribution from the centrifugal force. This is often referred to as an anomalous anticyclonic high, on the basis that most winds or currents will generally be at least approximately geostrophic (small R_o), and indeed, large R_o anticyclones are very seldom observed in nature.[22]

5) An interesting case is anticyclonic flow at $(\Pi,\ R_o) = (0,\ -1)$ in which the balance is between the outward centrifugal force and inward Coriolis force (vanishing pressure gradient). This is one of the two kinds of inertial motion noted in Part 1, Sec. 5, vortical inertial motion, $U_\lambda = -fr$. Though possible as a steady gradient balance, such a flow would be highly unstable (see footnote above), and unlikely to persist in nature long enough to be observed.

5.3 Problems

(1) The discussion of curvature made allusion to centrifugal force acting on a parcel and is thus inherently Lagrangian. Our shallow water model is Eulerian, and there must be an equivalent

[22]This analysis considers only the possibility of a steady momentum balance, and on that basis an intense, high anticyclone is indeed possible. However, such eddies are rarely ever observed in nature. One contributing reason is that intense, high anticyclones tend to be vulnerable to centrifugal instability. Small perturbations in the azimuthally symmetric flow of an otherwise steady but unstable eddy may quickly escalate into chaotic flow that modifies the eddy away from its unstable configuration. Thus, even if an intense anticyclone was somehow formed, it would not survive. This goes beyond the present scope; see the excellent introduction by McWilliams (2006) and the review article by Chomaz, J. M, S. Ortiz, F. Gallaire and P. Billiant, 'Stability of quasi two-dimenional vortices', Lect. Notes Phys., **805**, 35-39 (2010), DOI 10.1007/978-3-642-11587-5_2

representation of this process in the Cartesian form of the Eulerian momentum equations, Eqns. (14) and (15). Can you discern what term(s) represent curvature? Hint: consider a control volume (in 2-d) set on the northern edge of the eddy and estimate the net east-west advection of north-south momentum into this control volume.

(2) Using Fig. (20), show that the actual velocity of a balanced anticyclone having $\Pi \leq 1$ exceeds the geostrophic velocity, and the reverse for cyclones. How about very intense, anticyclonic lows? What is the asymptote of R_o at very large positive Π? Why is the other sign, $-\Pi$, completely different?

(3) Sketch in the schematic force balance for the three quadrants of Fig. (20) not already noted. Describe the balance at (Π, R_o) = (0, 0) and (0, -1).

(4) Fill in the algebra needed to go from Eqn. (75) to Eqn. (76), which is one possible way to nondimensionalize the solution for balanced flow, in effect, $R_o(\Pi)$. Show that another equally efficient form is U_λ / U_{geo} is some function of Π.

6 Closing remarks

6.1 Summary of Part 2

In Section 1 we asked, **What circumstances lead to a geostrophic balance?** To develop some insight we analyzed several experiments on geostrophic adjustment posed in a numerical, shallow water model. The first experiment started with a thickness anomaly having a horizontal scale $L = 100$ km and amplitude, $\eta_0 = 50$ m, comparable to observed mesoscale eddies. Once released, the anomaly was free to evolve according to the physical processes allowed into the shallow water model.

1) When rotation is omitted by setting $f = 0$, the anomaly is quickly and completely dispersed in space by the propagation of shallow water gravity waves. These waves are nondispersive, having phase speed and group speed $C = \sqrt{g'H} \approx 3$ m sec^{-1} for the stratification typical of the mid-latitude oceans. These gravity waves are the only possible nontrivial free motions when $f = 0$. In the one-dimensional, gravitational adjustment problem of Sec. 3, the solution consists of discrete, propagating pulses that retain the shape of the initial ridge and so look nothing like elementary (sine) waves.

A given experiment is characterized also by the amplitude of the initial thickness anomaly, η_0/H. If the initial amplitude is small, $\eta_o/H \ll 1$, the pulses propagate at very nearly the expected gravity wave speed, C. For a larger initial amplitude, say $\eta_o/H \geq 0.2$, finite amplitude effects include distortions (from linear) in the shape of the pulses, and a systematic change in the propagation speed

depending upon the sign of η_o, greater when the layer is thicker.

2) When rotation is included by an f-plane approximation, the transient response includes gravity waves that are more or less modified by rotation, and a possible geostrophic steady state. The wave frequency depends upon wavelength compared to the radius of deformation, $R_d = C/f$, the intrinsic horizontal length scale of a rotating, shallow water system. The highest frequency and shortest waves are nondispersive gravity waves, while longer wavelengths are near-inertial gravity waves having very large phase speed but very small group speed. Aside from the short wave, high frequency limit, inertia-gravity waves are highly dispersive and so the radiating waves do not retain the structure of the initial thickness anomaly.

3) If the initial ridge is a potential vorticity anomaly with respect to the outlying fluid, and if the width of the initial ridge is comparable to or greater than R_d, then some fraction of the ridge will survive the adjustment process and reach geostrophic balance. In the case of a one-dimensional ridge, the geostrophic current takes the form of jets whose half-width is R_d. In an inviscid f-plane experiment, geostrophic balance can be exact, and exactly steady, i.e., geostrophy can persist forever on an $f-$plane. (As we will see in Part 3, this is not the case when the latitudinal variation of f is also considered.) In a two-dimensional experiment that begins with a cylindrical thickness anomaly (Fig. 1 and Sec. 5), a steady, balanced flow may result, and the azimuthal current U_λ deviate from geostrophic speed by a factor proportional to the Rossby number, U_λ/fL.

6.2 What comes next?

An understanding of geostrophy is an essential second step in a study of the atmosphere and ocean, but be assured that there is much, much more to learn about how such geostrophic features come to exist in the first place, and how they evolve in time. In Part 3 we will take up a fundamental question that comes from a closer look at the large scale mass (pressure) and circulation fields of the ocean and atmosphere, *viz.*, **How do small departures from geostrophic balance lead to time-dependence and to the marked east-west asymmetry of low frequency phenomenon?** The plan is again to solve experiments in geostrophic adjustment, but the next time taking account of the (surprisingly) important latitudinal variation of f.

6.3 Supplementary material

The most up-to-date version of this essay may be downloaded from the author's public access web site: https://www2.whoi.edu/staff/jprice/ and look under Education Projects. Matlab and Fortran source

codes useful for the homework are at https://www2.whoi.edu/staff/jprice/codes-aCt-P1P2/ which includes the following,

twowaves.m shows the result of superposing two sine waves whose dispersion relation may be specified arbitrarily.

twolayer_eig.m solves for eigenmodes of two and three layer models using symbolic math.

ftransform.m solves an initial value problem in one-dimension and without rotation. Shows the wave forms that result from dispersion that is normal, anomalous or none.

geoadj_1d.m is a shallow water model used for the one-dimensional geostrophic adjustment experiments of Secs. 3.1 and 3.2. The latitude, thickness anomaly, friction, etc. may be easily varied from experiment to experiment. These all assume that the initial velocity vanishes.

geoadj_1d_uic.m as above, but in this script the initial condition includes several choices for the initial current, including vanishing potential vorticity.

geoadj_2d.for is a two-dimensional shallow water model used to do the geostrophic adjustment experiment of Sec. 5. Written in a very simple Fortran. Solutions like those shown in Sec. 5 can be generated on a reasonably capable PC in a few minutes. Model parameters including latitude, the kind of f model, eddy amplitude, etc. are entered from the keyboard. Output is saved to disk.

galook.m is a Matlab script used to plot the data generated and saved by geoadj_2d.for.

Index

MIT OpenCourseWare
https://ocw.mit.edu

Resource: Topics in Fluid Dynamics

James Price

The following may not correspond to a particular course on MIT OpenCourseWare, but has been provided by the author as an individual learning resource.

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.